

eScholarship@UMassChan

Paramagnetic rim lesions are highly specific for multiple sclerosis in real-world data

Item Type	Journal Article
Authors	Hemond, Christopher C;Dundamadappa, Sathish K;Deshpande, Mugdha;Baek, Jonggyu;Brown, Robert H;Ionete, Carolina;Reich, Daniel S
Citation	Hemond CC, Dundamadappa SK, Deshpande M, Baek J, Brown RH Jr, Ionete C, Reich DS. Paramagnetic rim lesions are highly specific for multiple sclerosis in real-world data. Brain Commun. 2025 May 29;7(3):fcf211. doi: 10.1093/braincomms/fcaf211. PMID: 40496668; PMCID: PMC12150028.
DOI	10.1093/braincomms/fcaf211
Rights	Published by Oxford University Press on behalf of the Guarantors of Brain 2025. This work is written by (a) US Government employee(s) and is in the public domain in the US.
Download date	2026-04-15 09:30:49
Item License	http://creativecommons.org/publicdomain/zero/1.0/
Link to Item	https://hdl.handle.net/20.500.14038/54524

BRAIN COMMUNICATIONS

Paramagnetic rim lesions are highly specific for multiple sclerosis in real-world data

 Christopher C. Hemond,^{1,2} Sathish K. Dundamadappa,³ Mugdha Deshpande,¹ Jonggyu Baek,⁴ Robert H. Brown Jr,¹ Carolina Ionete¹ and  Daniel S. Reich²

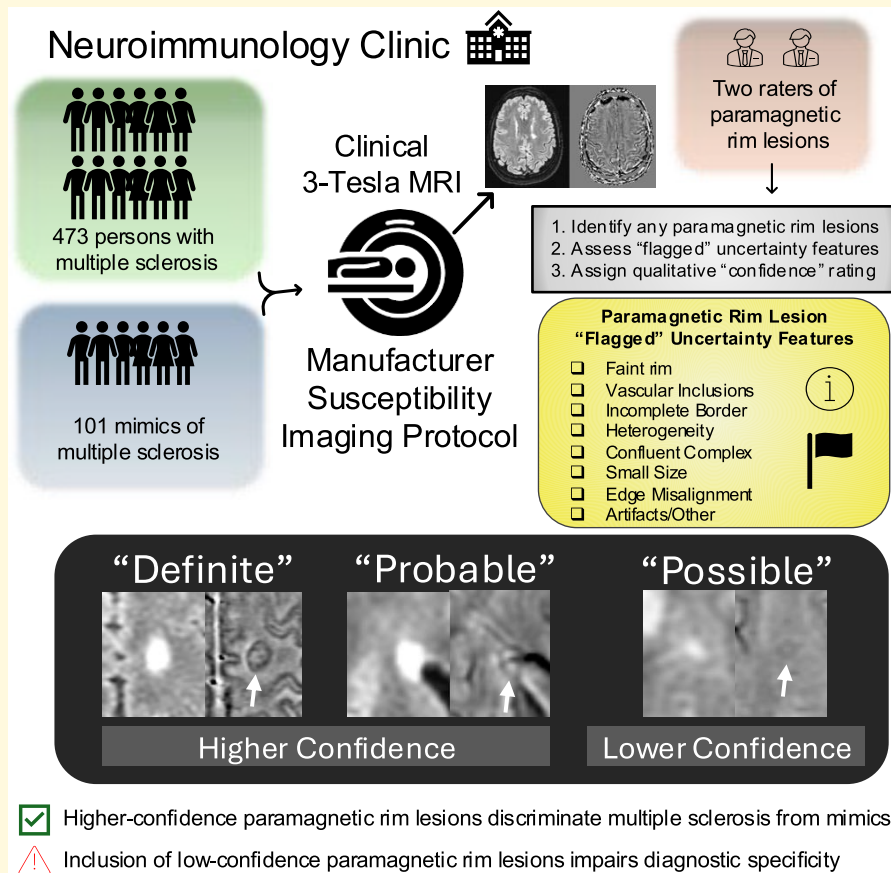
Paramagnetic rim lesions (PRLs) are an emerging biomarker for multiple sclerosis representing chronic, low-grade intraparenchymal brain inflammation. In addition to associating with greater disease severity, PRLs may be diagnostically supportive. Our aim in this study was to determine PRL specificity and sensitivity for discriminating multiple sclerosis from its diagnostic mimics using real-world clinical diagnostic and imaging data. This is a retrospective, cross-sectional analysis of a longitudinal cohort of patients with prospectively collected observational data. Patients were included if they underwent clinical evaluation in our academic neuroimmunology centre and had an available MRI scan from the same clinical 3-T magnet that included a T2*-weighted sequence with susceptibility post-processing (Susceptibility Weighted ANgiography protocol, General Electric). Susceptibility imaging-derived filtered phase maps and corresponding T2-fluid attenuated inversion recovery images were manually reviewed to determine PRLs. PRLs were categorized as ‘definite’, ‘probable’ or ‘possible’ based on modified, recent consensus criteria. We hypothesized that PRLs would convey a high specificity to discriminate multiple sclerosis from its MRI mimics. Five hundred seventy-four patients were evaluated in total: 473 with multiple sclerosis, 53 with non-inflammatory neurological disease and 48 with other inflammatory neurological disease. Identification of ‘definite’ or ‘probable’ PRL provided a specificity of 98% to discriminate multiple sclerosis from non-inflammatory neurological disease and other inflammatory neurological disease; sensitivity was 36%. Interrater agreement was almost perfect for definite/probable identification at a subject level. PRLs convey high specificity for multiple sclerosis and can aid in the diagnostic evaluation. Modest sensitivity limits their use as single diagnostic indicators. Including lesions with lower confidence (‘possible’) rapidly erodes specificity and should be interpreted with caution given the potential harms associated with misdiagnosis.

- 1 Department of Neurology, University of Massachusetts Memorial Medical Center, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA
- 2 Translational Neuroradiology Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA
- 3 Department of Radiology, University of Massachusetts Memorial Medical Center, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA
- 4 Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA

Correspondence to: Christopher C. Hemond, MD
Department of Neurology, University of Massachusetts Chan Medical Center
55 Lake Avenue North
Worcester, MA 01655, USA
E-mail: Christopher.hemond@umassmed.edu

Keywords: paramagnetic rim lesion; multiple sclerosis; specificity; diagnosis; mimics

Graphical Abstract



Introduction

Multiple sclerosis is a chronic inflammatory and degenerative disease of the central nervous system that benefits from early treatment with disease-modifying therapies.¹ Real-world data suggest concerning rates of misdiagnosis,² a problem that highlights an ongoing need for multiple sclerosis biomarkers with high specificity. The paramagnetic rim lesion (PRL) is a candidate MRI marker in this regard and has been shown in a series of studies to be associated with high ($\geq 95\%$) specificity for multiple sclerosis.^{3,4} PRLs additionally yield prognostic significance; higher counts are associated with greater neurological disability⁵ and portend accelerated disease progression.^{6,7} PRLs were recently announced as a new component of the upcoming revisions to the multiple sclerosis McDonald diagnostic criteria at the European Committee for Treatment and Research In Multiple Sclerosis congress in September 2024, underscoring the translational importance of this marker. However, most studies assessing PRL have used research protocols and cohorts that limit generalizability of the findings. Moreover, PRL categorization is a qualitative endeavour, and the pursuit of perfect interrater agreement has been challenging.⁸

Although recent North American Imaging in Multiple Sclerosis (NAIMS) consensus criteria for determination of PRL will aid in standardization of PRL identification,⁹ these metrics have yet to be systematically evaluated in the real world.

In this study, we aimed to determine the specificity and sensitivity of PRL through an analysis of clinical and MRI data from a single academic neuroimmunology clinic. These data reflect the wide variety of patients typically seen in clinical practice who are often referred for multiple sclerosis diagnostic consideration and management. We additionally aimed to evaluate how subjective ratings of PRL confidence could aid in guiding clinical relevance of lesion assessment features.

Materials and methods

Consent to participate

All participants were patients that underwent clinical evaluation and management at the University of Massachusetts neuroimmunology outpatient clinic in Worcester, MA,

USA. All patients provided explicit written informed consent to participate in one or more prospective observational studies that included a retrospective analysis of clinically acquired data. These studies were reviewed and approved by the University of Massachusetts Institutional Review Board (Protocols # H00016906 and 14143) in accordance with ethical principles outlined in the Declaration of Helsinki. Data collection, storage and access were in accordance with the Health Insurance Portability and Accountability Act to protect privacy and confidentiality.

Study design and subjects

This is a retrospective analysis of prospectively collected data from patients at a single, large, academic neuroimmunology clinic; the study timeframe was March 2016 to August 2022. All subjects were participating in at least one of two ongoing prospective, longitudinal observational studies of neuroimmunological disorders. Inclusion criteria were age greater than 18 but less than 80 years, having been evaluated in outpatient neuroimmunology clinic for neurological symptoms and having had at least one MRI performed on the same outpatient 3-T MRI scanner with a ‘multiple sclerosis’ protocol including susceptibility-sensitive imaging. Baseline was considered as the time of the first eligible MRI, with clinical measures calculated relative to that value. The total number of eligible patients between the observational studies at the time of data extraction was $n = 848$; 268 were excluded due to not having ≥ 1 eligible MRI scan performed on the same hardware platform using a generally harmonized software acquisition protocol; 6 patients were later excluded for not having any cerebral T2-hyperintense lesions with volumes $\geq 15\text{mm}^3$.

Patient neurological disability ratings and disease-modifying therapy use were procured by review of the patient electronic medical record and chosen based on proximity to the time of the baseline MRI. All Expanded Disability Status Scale scores and multiple sclerosis diagnoses/phenotypes extracted from the chart were rated by subspecialty-trained neuroimmunologists. The Multiple Sclerosis Severity Score¹⁰ and Age-Related Multiple Sclerosis Severity Score¹¹ were calculated based on tabulated values when possible; Multiple Sclerosis Severity Score was not calculated for disease durations of less than 6 months.

Patient diagnoses were extracted by chart review, using the latest information available. Diagnoses were made only by one of five expert neurologists, all of whom had dedicated subspecialty neuroimmunology clinical training and a minimum of 2 years post-fellowship clinical practice. All diagnoses for clinically isolated syndrome and multiple sclerosis met 2017 McDonald¹² criteria. Radiologically isolated syndromes were defined based on the updated (2023) criteria¹³ and given their low frequency, subsumed under ‘relapsing multiple sclerosis’ category for analysis purposes. Other inflammatory neurological disease (OIND) classification included the following diagnoses: (i) neuromyelitis optica spectrum disorder, (ii) myelin oligodendrocyte glycoprotein-associated

disease, (iii) proven or highly probable CNS infection or granulomatous disease and (iv) T2-hyperintense lesions and/or inflammatory cerebrospinal fluid profiles with concomitant diagnosis of a systemic inflammatory/autoimmune disease felt by their clinician to be causative or associated with CNS injury. ‘Unknown’ was defined as non-specific neurological symptoms (or unconfirmed retrospective symptoms) in the presence of at least one white matter brain lesion fulfilling dissemination in space candidate area(s), but without meeting revised radiologically isolated syndrome 2023 criteria or having a clear alternative diagnosis despite workup; these patients were subsumed under OIND. Non-inflammatory neurological disease (NIND) included all other non-multiple sclerosis patients not meeting OIND criteria. Neuromyelitis optica spectrum disorder and myelin oligodendrocyte glycoprotein-associated disease were diagnosed by expert neuroimmunologists using cell-based assays in conjunction with clinical presentation.

Image acquisition

All images were acquired on a 3-T MRI scanner (Signa Pioneer, General Electric Healthcare, WI, USA) using a 32-channel head and neck coil and a consistent acquisition protocol for multiple sclerosis patients. This standard protocol was implemented beginning in September 2016. There were no hardware upgrades that occurred during the observation period; three software updates occurred over the observation period of this study (PX25, PX26 and PX29), with two minor additional upgrades within each version. The protocol included sagittal 3D T2-FLAIR (FLuid Attenuated Inversion Recovery): TR/TE/flip angle/echo train length = 5400/maximum $\sim 133/90^\circ/140$, frequency/phase = 256/224, slice thickness = 0.8 at 50% resolution, reconstructed voxel size = $0.49 \times 0.49 \times 0.80$ mm, field of view = 250 mm, and scan duration = 3:34. Most susceptibility-sensitive scans (93.3%) were obtained using the following protocol: 3D SWAN: TR/TE/flip angle/echo train length = minimum ~ 42 ms/minimum ~ 24 ms/ $10^\circ/6$, frequency/phase = 320/224, slice thickness/overlap = 3.0 mm/1.5 mm, reconstructed voxel size = $0.47 \times 0.47 \times 3.0$ mm, field of view = 240 mm, and scan duration = 2:27; after November 2021, the following parameters with changed (affecting 6.7% of scans): flip angle/echo train length = $8^\circ/10$, frequency/phase = 256/256, slice thickness/overlap = 2.0 mm/1.0 mm, and field of view = 220 mm. If gadolinium was administered, an additional sagittal 3D T1-weighted fast spin echo (CUBE, no acronym) scan was obtained ~ 10 min after the intravenous infusion of gadoterate meglumine with the following parameters: TR/TE/flip angle 435 ms/16.1 ms/ 90° , frequency/phase = 256/256, slice thickness = 0.6, and reconstructed voxel size = $0.49 \times 0.49 \times 0.60$ mm.

Image analysis

MRI data underwent reconstruction with the scanner’s manufacturer software and were exported in Digital

Imaging and Communications in Medicine format. Images were converted to Neuroimaging Informatics Technology Initiative format using `dcm2niix` software (version 1.0.20201102).¹⁴ T1-weighted scans were resliced to a resolution of 0.80 mm isotropic. Manufacturer-reconstructed SWAN filtered phase images, adhering to the ‘right hand’ convention where paramagnetic substances appear dark, along with T2-FLAIR sequences, were aligned (using mutual information criteria and a rigid transformation with 6 degrees of freedom) with Statistical Parametric Mapping software (available at <https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). FLAIR and filtered-phase SWAN images were examined side by side in the axial view by one or both raters (C.C.H., a neurologist with 6 years of post-training experience with PRL specifically, and S.K.D., a neuroradiologist with 17 years of post-training experience; 6 with PRL) using ITK-SNAP software,¹⁵ with both evaluators blinded to the clinical details during the assessment. All PRLs were evaluated at a standardized contrast setting; intensity histograms for General Electric manufacturer phase images ranged between $-\pi$ and π , scaled by a factor of ~ 1000 . Our evaluation window was set between -300 and 300 with no adjustment of additional control points (linear mapping). Neither the software upgrades nor the small number of susceptibility scans (6.7%) affected by the SWAN sequence upgrade substantially affected image quality or interpretation.

PRL determination

White matter hyperintensities ≥ 3 mm in the long axis were identified on FLAIR images, and potential PRLs were defined as having a colocalized paramagnetic rim as seen on the filtered-phase SWAN in the axial plane. Raters were free to visualize PRL in the (lower resolution) orthogonal planes at their discretion; in practice, this occurred only rarely. The paramagnetic rim was required to be contiguous through at least 2/3 of the border (excepting ventricular and cortical margins) and must be visualized in at least two axial slices. The core of the PRL must be similarly isointense (with small variations accepted) to normal appearing white matter. These criteria are similar to the recent NAIMS consensus criteria,⁹ with a primary difference being that we did not immediately exclude lesions with complex or edge-like venous features. We additionally implemented a categorical ‘confidence’ assessment system for raters based on judgements of lesion features that contributed uncertainty. These ‘flag’ categories were defined as (i) heterogeneity in core or rim; (ii) incomplete paramagnetic border; (iii) lesion within a confluent complex; (iv) faint paramagnetic signal; (v) vascular complexity near lesion rim; (vi) incomplete edge colocalization between paramagnetic rim and FLAIR lesion edge; (vii) small size ($\sim < 5$ mm long-axis diameter); and (viii) susceptibility artefacts (proximate to dental hardware or air/bone interfaces or major dipole phase inversion). Milder dipole edge artefacts (phase inversion outside the paramagnetic rim) were not recorded. Raters marked their

level of confidence based on the number and severity of these categories. Up to two of the most contributory conditions that reduced confidence were recorded. Because the severity of these confidence-reducing features was variable, we ultimately allowed rater judgement as the final arbitrating factor, akin to a real-world radiology setup.

Lesions that were part of confluent complexes were included for assessment. PRL could not be considered ‘definite’ or ‘probable’ if they did not share a significant edge portion with the FLAIR signal, however. All lesions were assessed for contrast enhancement and excluded if they exhibited this property. Scans that did not include contrast were either (i) compared with a prior available scan to ensure lesion presence and chronicity or (ii) compared with a subsequent susceptibility-sensitive scan to ensure PRL chronicity. Cases in which these conditions were not met ($N = 8$) were few, but remained included in the analysis after chart review revealed that none of these patients were having new symptoms at the time of the scan, and the lesions in question did not exhibit any restricted diffusion or signs of oedema that might be suggestive of active demyelination. In cases where PRL could not be confirmed at follow-up (and no available prior, $N = 2$), patients had at least one definite/probable PRL that was already confirmed, such that this would not change patient category of ≥ 1 versus 0.

Interrater assessment and lesion segmentation

We randomly assessed 43 patients without confluent lesions to determine interrater agreement. No attempts were made to reconcile findings, and the analysed results reflect the original rater only ($N = 465$ rated by C.C.H., $N = 115$ rated by S.K.D. with 43 overlap). Cohen’s kappa was used to determine interrater agreement measures as implemented in the R toolbox ‘IRR’ (interrater agreement) at the patient level (≥ 1 PRL with and between confidence categories). To coordinate interrater comparisons, T2-hyperintense lesion maps were segmented from T2-FLAIR images using the lesion prediction algorithm¹⁶ as implemented in the Lesion Segmentation Toolbox version 3.0.0 (www.statistical-modelling.de/lst). A lesion index was created using the ‘cluster’ function in FSL, to label individual lesions. Lesion segmentation failures ($N = 29$) or scans with substantial confluent lesions ($N = 44$) were excluded from interrater comparison. Lesions were later segmented using the ‘LST-AT’ tool,¹⁷ which exhibited greater qualitative accuracy and no failures. Lesions < 15 mm³ in volume were excluded and lesion topography automatically determined using this same tool. Lesion counts and volumes reported here are based on this latter tool.

Statistical analysis

We assessed normality using a combination of skew/kurtosis metrics and histogram inspection. All statistical analyses were performed in R-studio (www.r-studio.com) version

4.02, with libraries including Tidyverse,¹⁸ yardstick (sensitivity and specificity)¹⁹ and IRR.²⁰

Timed analysis

One rater (C.C.H.) timed their analysis of randomly assessed patients and determined the number of seconds to (i) first PRL detection and (ii) total count of PRL detected.

Results

Cohort description

A total of 574 patients met inclusion criteria. Of these, 473 were diagnosed with multiple sclerosis [relapsing = 375 including clinically isolated syndrome ($n = 8$) and radiologically isolated syndrome ($n = 3$), progressive = 98], 53 with NIND, and 48 with OIND. A summary of these patient populations can be seen in Table 1, with detailed lesion segmentation characteristics in Supplementary Table 3. Relapsing multiple sclerosis patients were notably younger (45 ± 12 years) compared with the progressive (58 ± 9 years), NIND (51 ± 13 years) and OIND (51 ± 11 years) cohorts.

Of patients with NIND and OIND, diagnoses were highly variable. OIND included unknown inflammatory or demyelinating abnormal/inflammatory CSF results and MRI but no identified aetiology (30%), neurosarcoid (15%), neurolupus (13%), NMOSD (6%), neuro-Sjogren's disease (6%) and others <5% each: MOGAD, primary angiitis of the CNS, CLIPPERS, viral and autoimmune encephalitis, steroid-responsive encephalopathy associated with autoimmune thyroiditis, other rheumatological-associated, other suspected autoimmune-associated, probable CNS lyme, other probable CNS infectious. NIND was composed of small vessel disease or leukoaraiosis (34%), migraine/headache (19%), neurodegenerative (15%), nonspecific white matter abnormalities in functional neurological disorder or

benign disease (9%) and other (23%), a heterogeneous category that included <5% each of various genetic leukodystrophies or epilepsies, toxic-metabolic conditions, fibromyalgia and idiopathic intracranial hypertension with white matter lesions.

PRL analysis

Among patients meeting 2017 McDonald criteria for multiple sclerosis ($N = 473$), we observed at least one 'definite' PRL in 22% of relapsing patients and 15% of progressive patients. No 'definite' PRLs were identified in either the NIND or OIND categories, corresponding to a 100% specificity and 21% sensitivity when limited to the category of highest confidence (see Table 2). When the confidence threshold was lowered to include 'probable' (e.g. 'definite' or 'probable'), PRLs were seen in 37% of relapsing multiple sclerosis and 29% of progressive multiple sclerosis, as well as 2% ($n = 1$) of both NIND and OIND. This combined category yielded a specificity of 98% and sensitivity of 36%. When reducing confidence to include 'possible' lesions (i.e. all categories), 37% of relapsing and 56% of progressive patients had at least one lesion meeting these criteria; 16% of NIND and 14% of OIND were also identified as having at least one PRL in this combined category; discrimination specificity was reduced to 85%, and sensitivity was increased to 54%. Figure 1 exhibits PRL from different confidence categories. One 'probable' PRL was identified in both NIND and OIND. The one NIND patient with an identified PRL had been previously treated with glatiramer acetate for suspicion of multiple sclerosis, but this diagnosis was later felt to be unlikely based on further workup including a lack of oligoclonal bands and spinal cord lesions, as well as a history of lacunar strokes and multiple vascular risk factors ultimately felt to be most consistent with (non-inflammatory) small vessel ischemic disease (Fig. 2A). The diagnosis of the OIND patient was neurolupus, although their clinician

Table 1 Summary of the patient cohorts

Variable	Relapsing multiple sclerosis, N = 375	Progressive multiple sclerosis, N = 98	NIND, N = 53	OIND, N = 48	P-value
Age, years	45 ± 12	58 ± 9	51 ± 13	51 ± 11	<0.001 ^a
Sex, female (%)	313 (83)	58 (59)	43 (81)	37 (77)	<0.001 ^b
Disease duration, years	10.4 ± 8.9	18.9 ± 11.8			<0.001 ^c
EDSS	1.5 (1.5,2.5)	5.3 (3.5,6.5)			<0.001 ^c
MSSS	2.8 ± 2.2	5.2 ± 2.5			<0.001 ^c
ARMSS	3.1 ± 2.1	5.1 ± 2.5			<0.001 ^c
DMT at scan	237 (64%)	59 (61%)			0.831 ^d
Total T2 lesion count, N	28 ± 21	36 ± 19	22 ± 19	18 ± 19	<0.001 ^e
Total T2 lesion volume, mL	6.4 ± 7.9	15.1 ± 17.2	5.0 ± 7.1	3.9 ± 9.9	<0.001 ^f

Scores are reported as N (%), mean ± SD or median (IQR). ARMSS, Age Related Multiple Sclerosis Severity Score; DMT, disease-modifying therapy; EDSS, Expanded Disability Status Scale; MSSS, Multiple Sclerosis Severity Score; NIND, non-inflammatory neurological disease; OIND, other inflammatory neurological disease; relapsing multiple sclerosis includes diagnoses of clinically isolated syndrome ($n = 8$) and radiologically isolated syndrome ($n = 3$). ^aANOVA; all pairwise *post hoc* tests using Tukey's honest significant difference were significant ($P < 0.05$), except NIND versus OIND. ^b χ^2 test; for sex, significant ($P < 0.05$) pairwise tests were between relapsing multiple sclerosis and progressive multiple sclerosis. ^ct-test. ^d χ^2 test. ^eANOVA; significant ($P < 0.05$) *post hoc* pairwise comparisons using Tukey's honest significant difference are between relapsing multiple sclerosis versus progressive multiple sclerosis, NIND versus progressive multiple sclerosis and OIND versus progressive multiple sclerosis. ^fANOVA; significant ($P < 0.05$) *post hoc* pairwise comparisons using Tukey's honest significant difference are between relapsing multiple sclerosis versus progressive multiple sclerosis, NIND versus progressive multiple sclerosis, OIND versus progressive multiple sclerosis and OIND versus relapsing multiple sclerosis.

Table 2 Summary of PRL frequencies observed across different diagnostic categories and their associated sensitivity/specificity and interrater agreement

Variable	Relapsing multiple sclerosis, N = 375 (%)	Progressive multiple sclerosis, N = 98 (%)	NIND, N = 53 (%)	OIND, N = 48 (%)	Specificity (%)	Sensitivity (%)	Cohen's κ
Definite PRL (≥ 1)	82 (22)	15 (15)	0 (0)	0 (0)	100	21	0.91
Probable PRL (≥ 1)	103 (27)	21 (21)	1 (1.8)	1 (2.0)	n/a	n/a	0.90
Possible PRL (≥ 1)	137 (37)	41 (42)	8 (14)	7 (14)	n/a	n/a	0.53
Definite or probable PRL (≥ 1)	140 (37)	28 (29)	1 (1.9)	1 (2.1)	98	36	0.94
Definite, probable or possible PRL (≥ 1)	201 (54)	55 (56)	9 (17)	7 (15)	85	54	0.74

n/a, not applicable for calculation, as these categories would naturally be combined with categories of higher confidence in practice; NIND, non-inflammatory neurological disease; OIND, other inflammatory neurological disease; progressive multiple sclerosis includes primary and secondary progressive forms; PRL, paramagnetic rim lesion; relapsing MS includes clinically isolated syndrome, $N = 8$, and radiologically isolated syndrome, $N = 3$).

commented that they suspected co-morbid multiple sclerosis as well (Fig. 2B).

Interrater analysis

Forty-three patients were evaluated by both raters. At the subject level, patients having ≥ 1 'definite' PRL were agreed upon with a κ of 0.91; similarly, patients with ≥ 1 'probable' PRL showed a κ of 0.90. This agreement was substantially reduced when considering a category of ≥ 1 'possible' PRL, in which there was a $\kappa = 0.53$. Combination categories yielded $\kappa = 0.94$ ('definite'/'probable') and $\kappa = 0.74$ ('definite'/'probable'/'possible'). Examples of interrater disagreements can be seen in Fig. 3.

PRL confidence analysis

Raters identified up to two of the most significant factors that influenced their perceived confidence in PRL determination. Among the total 563 PRL, raters stated that the following factors played a role: 'faint' (33%), 'vascular complexity' (25%), 'incomplete border' (23%), 'heterogeneity' (22%), (being part of a) 'confluent (lesion) complex' (17%), 'small' (9%), 'misalignment' between FLAIR and paramagnetic edges (4%), 'artefact' (2%) or 'other' (2%). Examples of many of these features can be seen in Figs 1–3; Fig. 4 depicts explicit examples of reasons why caution should be advised in lesions that are ultimately very unlikely to represent PRL. A breakdown of 'flags' visualized in possible or probable PRL by diagnostic category is presented in Supplementary Table 4, and a supplementary visual guide with additional examples of each category is presented in Supplementary Fig. 1.

Timed analysis

Thirty patients were randomly evaluated in an optimized setting by one rater (C.C.H.). The mean time to identify the first PRL (in cases when 'definite' or 'probable' PRL ≥ 1 , $N = 9$) was 38 ± 37 s (median = 22; min = 4, max = 103). Mean time including cases without 'definite' or 'probable' PRL

($N = 21$) was 55 ± 31 s (median = 54; min = 4, max = 128). Mean time to individually count and summate all PRL, among all included patients, was 79 ± 53 s (median = 66; min = 24, max = 279). The median number of lesions in this cohort was 28 (minimum = 1, maximum = 101). The Pearson correlation coefficient between lesion number and total evaluation time was $r = 0.64$ and between lesion number and time-to-first PRL was $r = 0.21$.

Sensitivity analysis for early diagnoses

The most urgent diagnostic need in discriminating multiple sclerosis from its radiological mimics typically occurs within the first months to years of symptom onset. We therefore performed a sensitivity analysis examining only patients with early disease, defined as having had symptom onset within the prior year. All patients from OIND and NIND cohorts were included as the evaluation MRI was typically their index scan. The results of this analysis are presented in Supplementary Tables 1 and 2. This smaller ($N = 44$) cohort had total lesion volumes and numbers in similar proportion to the OIND and NIND comparators, with definite and probable PRL exhibiting the same specificity but with improved sensitivities (definite: 32% versus 21%; definite/probable: 48% versus 36%; definite/probable/possible: 59% versus 54%) due to a higher observed overall frequency of PRL in this early-multiple sclerosis sub-cohort.

Discussion

PRLs are a candidate diagnostic biomarker for multiple sclerosis, given their high demonstrated specificity for multiple sclerosis versus other MRI-mimicking diseases. Here, we confirm high discriminative specificity in the largest clinical cohort to date. Strengths of this study include the use of a real-world neuroimmunology clinical cohort and a manufacturer-provided susceptibility-sensitive sequence, as well as the exploration of a novel rater confidence system to enhance clinical translation in the determination of a PRL.

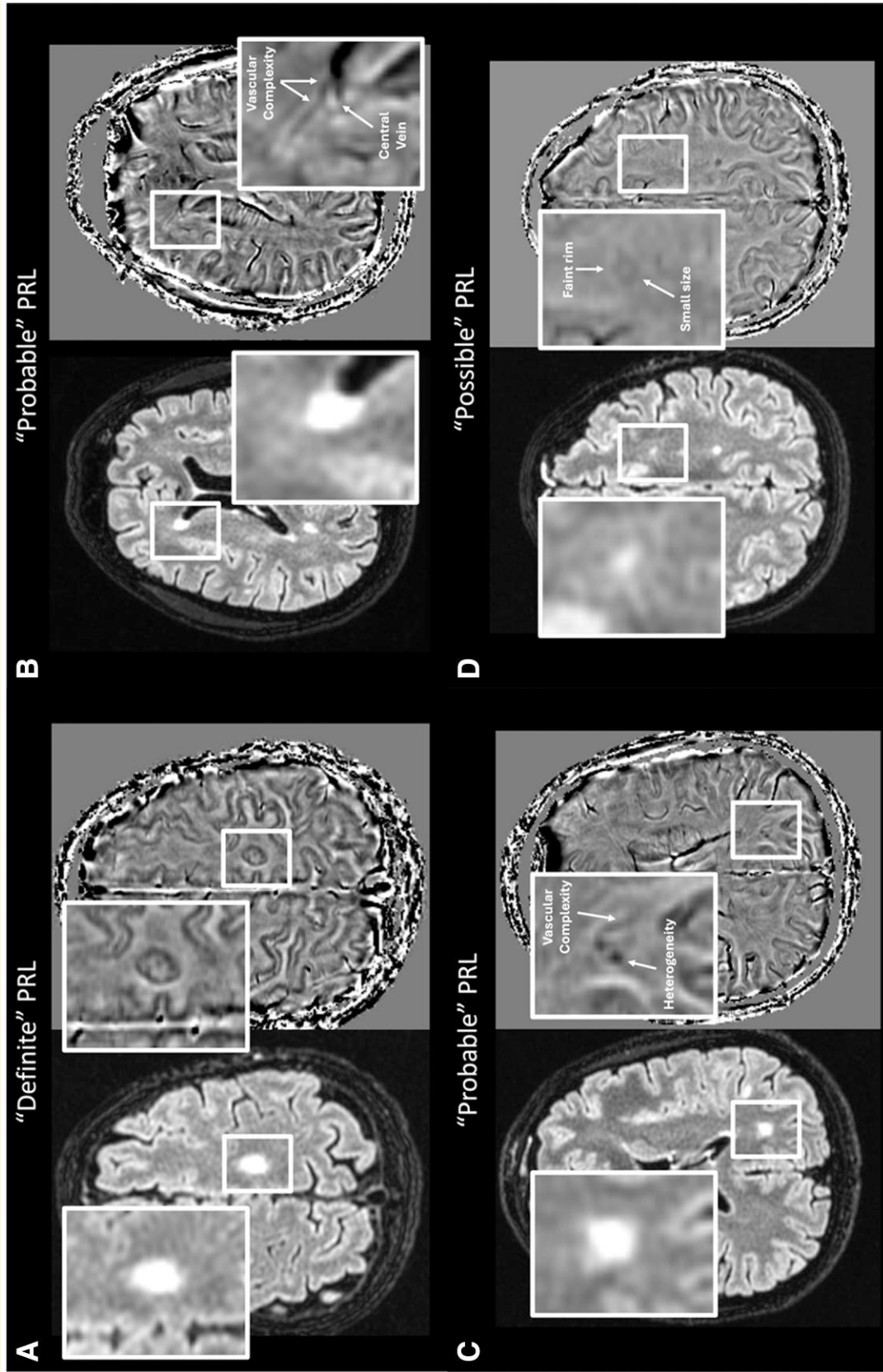


Figure 1 Examples of PRL ratings of different confidence categories in relapsing-remitting multiple sclerosis. Axial, fluid attenuated inversion recovery images (left) are coregistered to high-pass filtered phase susceptibility images (susceptibility weighted angiography, General Electric; right) in four different examples demonstrating ratings of (A) 'definite', (B) and (C) 'probable' and (D) 'possible' PRLs. Small white arrows identify some of the many features that often reduce rater confidence in assessment.

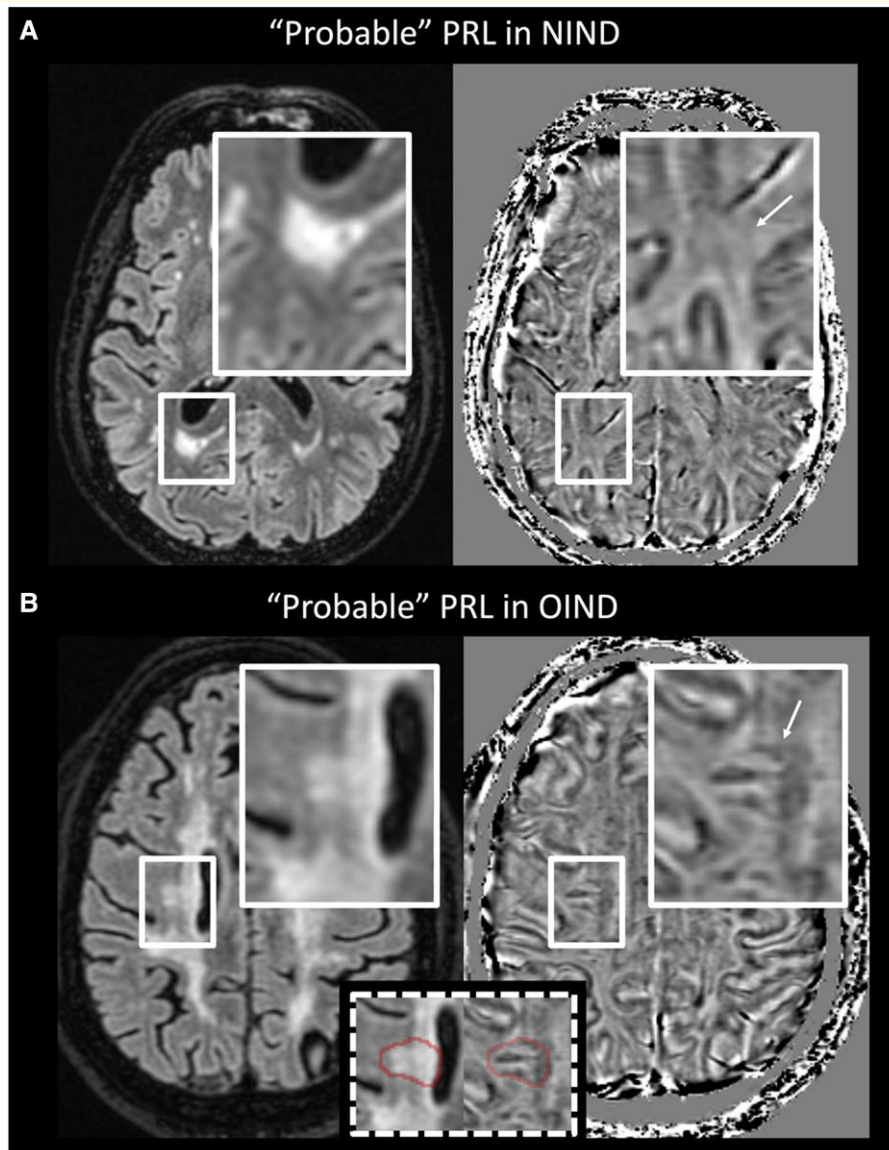


Figure 2 Probable PRLs in non-multiple sclerosis. PRLs were rated as ‘probable’ in two patients from the cohort, diagnosed with **(A)** small vessel disease (NIND) and **(B)** neurolyupus (OIND). Axial, fluid attenuated inversion recovery images are on the left and high-pass filtered phase susceptibility images are on the right. White arrows point to an edge of the probable PRL.

The rationale for a confidence system in this study is based on imperfect interrater agreement and challenges that the raters in this study have previously encountered. This system is based on recent consensus criteria determined by the NAIMS Cooperative,⁹ but without the universal exclusion of lesions with veins that could mimic a border. We also added consideration of a reduced susceptibility signal (i.e. ‘faint’) as a ‘flagged’ factor that may reduce PRL confidence and did not have a formal category for ‘challenges in reaching an agreement’ for PRL, since raters assessed in isolation. However, despite intentions to create a formal system of ‘flags’ to evaluate PRL, raters rapidly agreed that the number of factors (‘flags’) that could reduce certainty did not necessarily correspond to a linear decline in subjective confidence.

As a simple example, many potential PRLs were simply faint (low contrast-to-noise ratio), and at a certain threshold of ‘faintness’, one begins to question whether a lesion truly represents a faded PRL or simply is part of the background. Thus, a PRL with only one flag (‘faint’) might be classified as ‘possible’ (the relevance of this ‘faint’ category is underscored by how commonly it was reported to affect PRL assessments, with raters noting 46% of suspected PRL in progressive multiple sclerosis and 32% in relapsing multiple sclerosis being ‘faint’; see [Supplementary Table 4](#)). Conversely, a (large) lesion could contain 3+ flags including vascular inclusions, paramagnetic heterogeneities and incomplete border and yet still remain ‘definite’ confidence based on strong signal hypointensity and close congruence

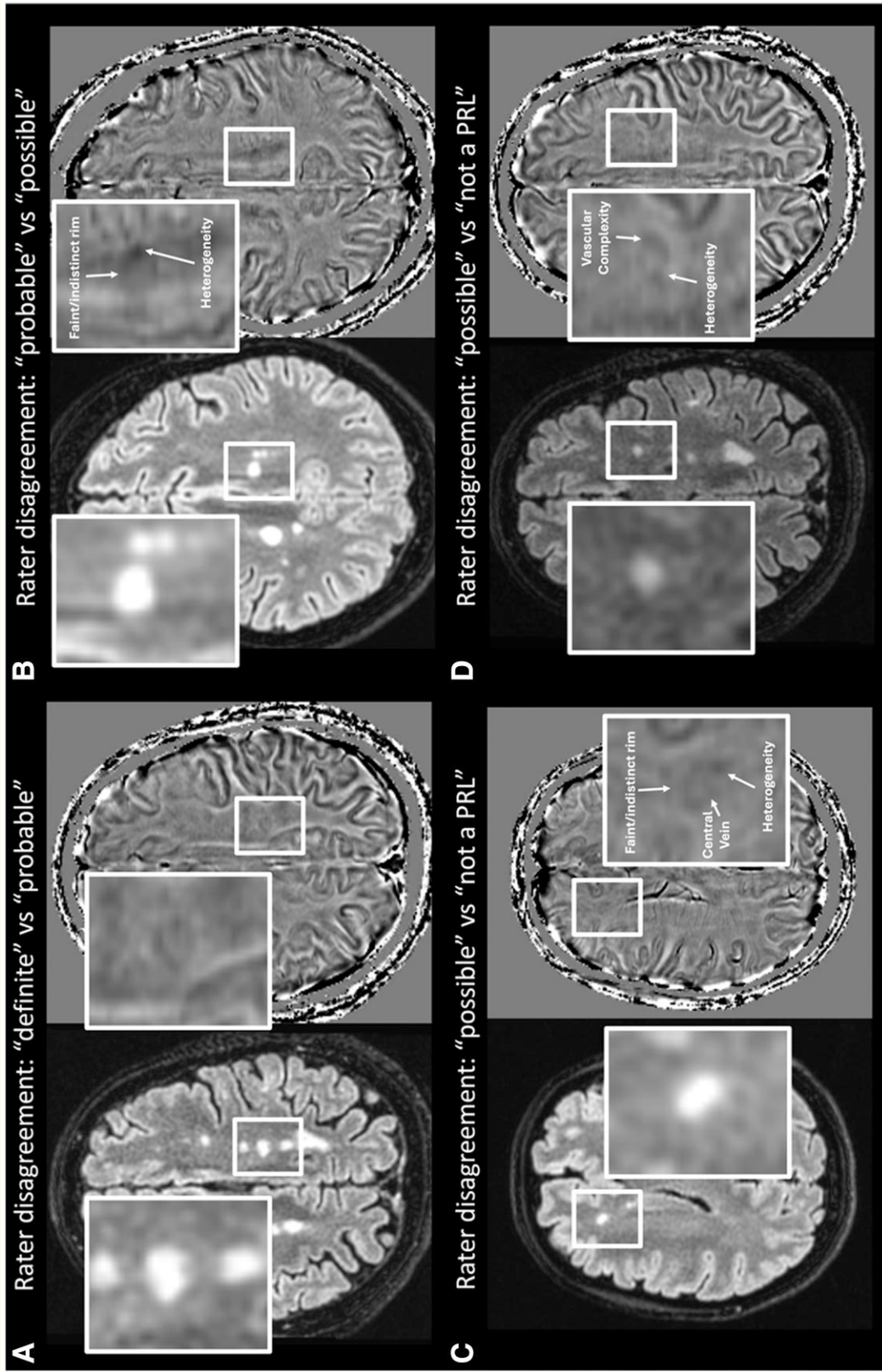


Figure 3 Examples of rater disagreement at different confidence levels for PRL ratings. Frequent sources of disagreement occurred due to faint signal, paramagnetic heterogeneity within the lesion or lesion border, as well as the potential for venous structures to mimic the PRL edge. Diagnoses are relapsing multiple sclerosis (A, B, D) and clinically isolated syndrome (C).

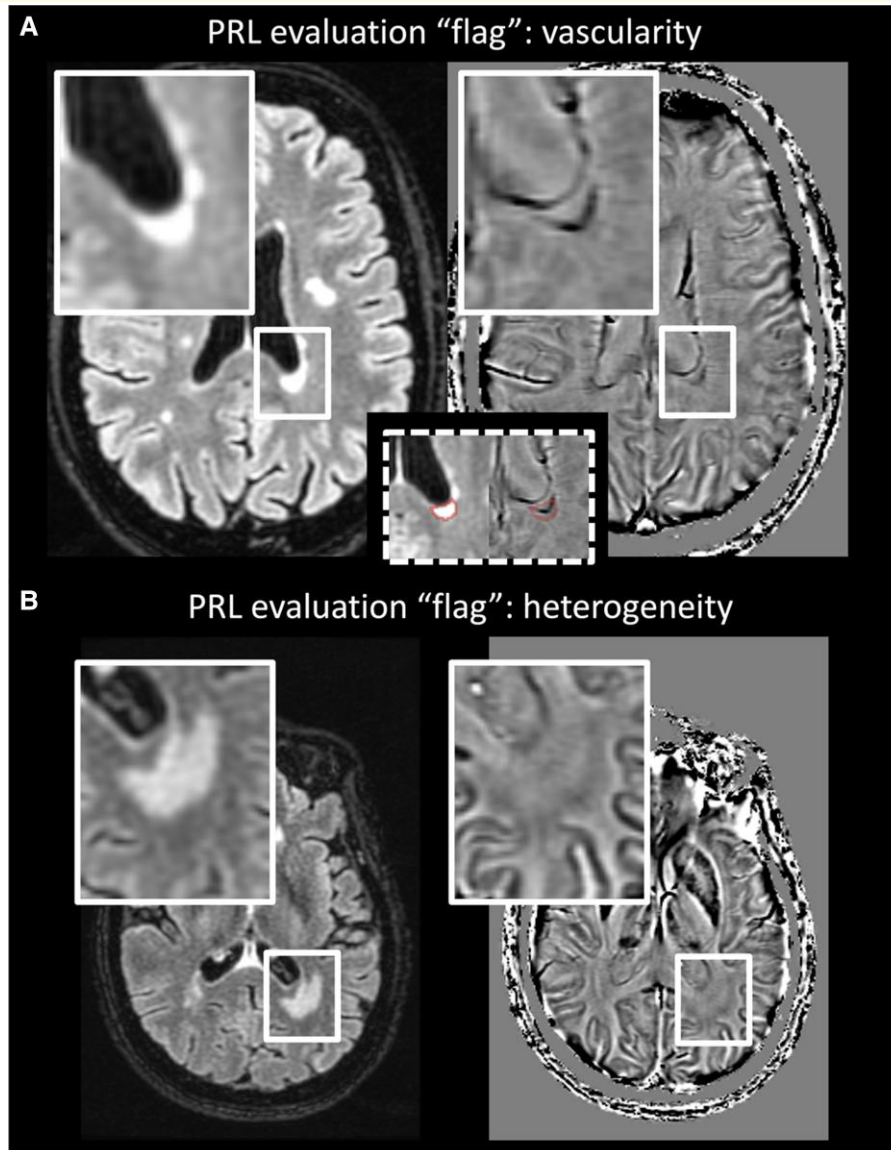


Figure 4 Examples of vascular and heterogeneity ‘flags’ in PRL evaluation that could lead to false positives. **(A)** An enlarged central vein that could readily be mistaken for a paramagnetic rim in a patient with relapsing multiple sclerosis; an inlay (dotted box) depicts the coregistered edges of the fluid attenuated inversion recovery lesion, which depict an important misalignment. Because the ependymal surface of the ventricle is also paramagnetic, periventricular lesions could be graded with caution. Example **(B)** shows a faint, thickened and incomplete paramagnetic signal proximate to the edges of a fluid attenuated inversion recovery hyperintense lesion in a patient with non-inflammatory neurological disease (vascular risk factors, hyperhomocysteinemia and history of moderate traumatic brain injuries); additional ‘flags’ include dense vascularity traversing the edges. Neither rater identified this as a possible PRL.

to the edges of the FLAIR signal. For these reasons, we believe a formal confidence rating system should be a topic of further study and expert consensus. Note that our use of ‘possible’ PRL here should not be confused with the NAIMS consensus⁹ statement’s use of ‘possible’, which was defined as visualization of PRL in the context of a lack of contrast administration.

In general, our data support the idea—in context of goals maximizing diagnostic specificity and reducing harms from misdiagnosis—that PRL of questionable conviction

(‘possible’) should be excluded from diagnostic considerations. Raters frequently disagreed on what constituted a ‘possible’ PRL, and although the inclusion of these lesions did improve sensitivity from 36% to 54%, we observed an unacceptably high decrease in specificity, from 98% to 85%. Interestingly, a diagnosis of progressive multiple sclerosis notably benefited from inclusion of low-confidence ‘possible’ lesions, with ≥ 1 PRL prevalences concordant with histological work²¹; we hypothesize that effective treatment and older age (both reducing lesion incidence), combined

with longer observation times, allow PRL to fade into 'possible' lesions by the time of cross-sectional evaluation in this population. PRL fading has been observed previously and estimated to contribute to PRL disappearance over roughly 7 years.^{22,23} Broad inclusion of susceptibility-sensitive sequences with phase reconstructions across neuroradiology protocols may allow interrogation of historical scans in some cases to better clarify whether a lesion should be categorized as a PRL.

Interrater agreement was nearly perfect for the determination of ≥ 1 PRL of 'definite' or 'probable' confidence, which is reassuring when consideration is given to implementation of PRL as a potential diagnostic biomarker. To our knowledge, only one other study has implemented a PRL confidence rating system²⁴; in this study, Stötling *et al.* assigned lesions a PRL confidence rating between 1 and 5, with higher confidence lesions (4 'certain' or 5 'very certain') defined as meeting NAIMS consensus criteria.⁹ The authors report an interrater kappa of 0.74 for per-subject PRLs presence, defined as a confidence cut-off of 4 or 5. Here, we find the same interrater kappa (0.74) when collapsing across all categories, including 'possible'. We hypothesize that our (conservative) 'possible' definition is similar to these authors' 'certain' definition, in that most 'possible' PRL identified in this study would still meet NAIMS consensus criteria. However, given the differences in the rating definitions, a direct comparison is not possible. Stötling *et al.* also used a different, higher-resolution, isotropic MRI protocol, and raters perhaps had different training backgrounds/protocols. Another potential difference is that our raters had worked together previously and generally formed a 'local' consensus for what constitutes a PRL using our specific imaging protocol. Although not necessarily dissimilar from real-world neuroradiologists comparing notes, this prior consensus nonetheless could have artificially inflated agreement. However, arguing against this bias, we (C.C.H. and S.K.D.) have previously reported an unremarkable interrater kappa of 0.70 using the same local consensus²⁵; this kappa is similar to multiple other studies as noted in a recent review on this topic.⁸ Rather, we interpret the high interrater agreement here to suggest that 'probable' and 'definite' PRLs were simply examples of high-clarity PRL to trained raters. Indeed, given this local consensus, it is notable that raters continued to disagree on what constitutes a 'possible' PRL, underscoring the importance of a cautious interpretation for low-confidence lesions even among experts. Nonetheless, as a single-centre research effort, the lack of external validation represents a limitation. An important area of future research is the need for the development and assessment of standardized rater training across different datasets, to enhance both internal and external research validity of PRL research and clinical translation.

What could be done to improve confidence within this problematically broad 'possible' category? Examination of the 'flags' that contributed to the most uncertainty (Supplemental Table 3) provides some insight. PRL 'faintness' was a frequent concern: this is a challenging problem to correct as it relates to physiological rim iron concentrations. The hypointensity of the

filtered phase signal depends on both the magnetic field strength as well as the echo time. To improve signal, radiologists could consider protocoling multiple sclerosis patients on an MRI with the highest available field strength and increasing protocol echo time, albeit at the expense of increased scan times. Other common flags like small lesions, susceptibility heterogeneity and vascular complexity would all likely be improved by increasing the acquisition resolution to better visualize and differentiate these fine structures. For example, the in-plane (axial) acquisition voxel size for our susceptibility sequence is 0.80×1.07 mm before reconstruction, and we find a definite/probable PRL prevalence of 29% (progressive multiple sclerosis) to 37% (relapsing multiple sclerosis); as an extreme comparison, a recent study employing 7-T resolution at 0.15×0.15 mm in-plane resolution reported a PRL prevalence of 87%.²⁶

Nonetheless, even using our unoptimized clinical susceptibility protocol, we find few examples of higher confidence PRL in non-multiple sclerosis conditions. These results are comparable with the several other large ($N > 100$) studies. An early 2012 study by Hagemeyer *et al.*²⁷ assessed 135 multiple sclerosis patients compared with 49 controls with incidental small vessel disease, finding a PRL specificity of 98% and sensitivity of 22%. Clarke *et al.* used the Siemens SWI sequence at 3 T to determine PRL by local consensus in 112 patients with clinically isolated syndrome compared with 35 other non-multiple sclerosis mimics and report 100% specificity for PRL and 59% sensitivity. Maggi *et al.*³ employed high-resolution T2* segmented EPI sequence at 3 T to examine 329 multiple sclerosis patients compared with 83 mimics and find a specificity of 93% and a sensitivity of 52%; follow-up work showed similarly with 93% specificity and 60% sensitivity.²⁸ Last, Meaton *et al.*⁴ assessed PRL using heterogeneous mix of manufacturer- and custom-susceptibility protocols at 3 T across several clinical sites; using blocks of the brain to improve rater blinding, investigators in this study report 98% specificity and 24% sensitivity for PRL in 254 multiple sclerosis patients versus 91 MRI mimics and 217 healthy controls. Across all these studies, few false positives occurred; sensitivity is likely most related to the comparison population and the technical aspects of the susceptibility protocol as discussed above. An exception may be Susac disease, in which PRLs were commonly observed,^{3,29} although the number of reported cases of this rare disease is small. Unfortunately, our cohort did not include any patients with this diagnosis for comparison, and this remains an area of need for clinical research.

This study was designed with clinical translation in mind. PRL evaluations were rapid (55 s), especially when only considering the time-to-first PRL identification. Determining total PRL counts is more laborious but likely to be clinically relevant, as higher PRL count predicts clinical worsening.⁶ Promising efforts are emerging for computer-assisted PRL identification using advanced statistics or deep learning methods,³⁰⁻³² and these methods are poised to facilitate rapid PRL counts and reduce or eliminate rater inconsistencies. Sensitivity could likely be improved using a higher in-plane resolution, at the expense of longer scan times and

consequent increase in risk of motion artefacts. Determining an ‘optimized’ resolution and protocol to detect PRL with each manufacturer remains an area of high clinical need. For example, the use of a single General Electric ‘SWAN’ sequence here may not directly compare to other manufacturer sequences (such as Philips ‘SWIp’ or Siemens ‘SWI’) or research protocols (such as the NIH T2* Segmented Echo-Planar Imaging sequence); additional comparative research is needed in the context of upcoming McDonald criteria revisions as well as multicentre research and large clinical trials. We also implemented a single-rater system for lesion assessment, as would be reflected in practice. FLAIR and SWAN sequences were coregistered as part of our preprocessing, an important step that is included in many modern PACS software and does not often represent a translational limitation in neuroradiology practice. We did not use the T1 sequence in PRL determination, which may aid in cases of poor FLAIR-susceptibility colocalization (here affecting 4% of PRL assessments). Although raters were allowed to view the lower-resolution multiplanar reformats here, this occurred only rarely as the 3 mm out-of-plane thickness adds little perceived confidence and slows evaluations. Acquiring thinner slice thicknesses would facilitate orthogonal evaluations, likely at the expense of increased scan times. Notably, much research with PRL is based on a high-resolution, isotropic, segmented-EPI acquisition (optimized for high contrast between lesions and central veins³³) in which orthogonal reformats provide more detailed information; this sequence is unfortunately not yet widely available as a manufacturer product sequence.

Given the design of this study as a retrospective analysis of prospectively collected data, other potential biases exist. Notably, there is an imbalance between the number of multiple sclerosis and non-multiple sclerosis diagnoses, with enrichment in multiple sclerosis. This selection and information bias could be related to a clinician not ordering the multiple sclerosis-specific 3-T MRI protocol, which includes the SWAN sequence for evaluation of PRL and CVS. Presumably, in these patients not receiving a 3-T MRI, multiple sclerosis could be ruled out with their existing imaging and without the need for a more specialized, dedicated 3-T scan. This selection bias would thus favour inclusion of patients with significant suspicion of multiple sclerosis, a bias that would decrease the observed effect and is therefore unlikely to affect the conclusions of this study. An additional bias could arise from raters’ knowledge of global MRI features including T2 lesion distribution and morphology, potentially influencing PRL interpretation. One group aimed to reduce this bias by presenting raters with randomized sections of the brain, initially created by splitting each MRI into eight blocks⁴; this approach largely obscures global features. Here, however, we chose to pursue a ‘real-world’ method better reflecting clinical practice, in which neuroradiologists and neurologists would be making PRL assessments using all available information. Non-multiple sclerosis patients were included in the study when they had at least one T2-hyperintense lesion $\geq 15\text{mm}^3$ in volume, but other lesion characteristics and ‘green flag’ morphology that could help identify multiple sclerosis³⁴ were not specifically

considered. We reasoned that if the patient’s neuroimmunology-trained clinician was concerned enough to order an ‘multiple sclerosis protocol’ 3-T brain scan, this clinical decision was sufficient to represent a pragmatic mimic, even if specificity could likely be increased with a more detailed morphological assessment. Last, although we provide a suggested PRL evaluation outline here, we emphasize that this ‘confidence categorization’ assessment is ultimately qualitative in nature, remaining untested in wider neuroradiological practice.

Finally, care should be made to find an appropriate contrast and window for viewing PRL. In our experience with GE (and Philips) susceptibility protocols, PRL histogram data are normally distributed around 0 (with a range that should fall between π and negative π , sometimes with a scaling factor), and we suggest a window covering $\sim\pm 0.7$ SD around the mean of the data to facilitate high PRL contrast. However, we did not systematically explore this issue, which is an important topic for future research directions.

In summary, our data support the use of PRL to discriminate multiple sclerosis from other common neurological mimics seen in routine neuroimmunology care. We anticipate these imaging markers providing an important diagnostic perspective for clinicians, especially in conjunction with other susceptibility-based markers such as the central vein sign. Limited research examining the combination of these two imaging biomarkers suggests an augmentation of utility compared with either one alone.²⁸ Additionally, as markers of chronic active lesion activity, PRLs also provide important biological insight into multiple sclerosis disease phenotype. We anticipate a forthcoming emergence of a more comprehensive, non-invasive approach to the assessment of compartmentalized parenchymal inflammation with susceptibility-sensitive imaging. Complementary perspectives on chronic active lesion activity can also be obtained from other non-invasive methods such as chronic lesion activity such as slowly expanding lesions^{35,36} or TSPO-PET+ activity³⁷; these approaches, however, remain challenging to implement in routine clinical practice.

Conclusions

The observation of at least one PRL of ‘probable’ or ‘definite’ confidence is associated with very high discriminative specificity for multiple sclerosis versus other common MRI mimics. Raters should remain conservative in their PRL assessments for diagnostic purposes and be familiar with the many features that degrade confidence in lesion assessments. PRLs are readily and rapidly visualized using a manufacturer protocol in the clinical routine. These data, using a large cohort of real-world neuroimmunology cases, support proposals for PRL integration into future diagnostic multiple sclerosis criteria.

Supplementary material

Supplementary material is available at *Brain Communications* online.

Acknowledgements

We thank colleagues including Drs. Roberto Bomprezzi, Idanis Berrios-Morales and Raffaella Umeton as well as Celia Gomes-McGillivray, Nimmy Francis, Aurelie Martin-Puig and Mariana Kurban at the University of Massachusetts for contributions to patient recruitment in ongoing, longitudinal observational studies. We thank the family of Mrs. Jane Gagne, in memory of Edward Gagne, for their ongoing research support.

Funding

C.C.H. receives extramural research funding from the National Institutes of Health, National Institute of Neurological Disorders and Stroke, grant K23NS126718. D.S.R. is funded by the National Institute of Neurological Disorders and Stroke Intramural Research Program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests

The authors declare no conflicts of interest with the study. D.S.R. has received research funding from Abata and Sanofi for separate projects related to therapeutic targeting of PRL.

Data availability

Deidentified data are available to qualified collaborators, following a signed institutional data sharing agreement. No custom R code was developed for this analysis. All analyses were performed using standard functions from publicly available R packages, as described in the Materials and methods.

References

1. Reich DS, Lucchinetti CF, Calabresi PA. Multiple sclerosis. *N Engl J Med*. 2018;378(2):169-180.
2. Kaisey M, Solomon AJ. Multiple sclerosis diagnostic delay and misdiagnosis. *Neurol Clin*. 2024;42(1):1-13.
3. Maggi P, Sati P, Nair G, et al. Paramagnetic rim lesions are specific to multiple sclerosis: An international multicenter 3T MRI study. *Ann Neurol*. 2020;88(5):1034-1042.
4. Meaton I, Altokhis A, Allen CM, et al. Paramagnetic rims are a promising diagnostic imaging biomarker in multiple sclerosis. *Mult Scler J*. 2022;28(14):2212-2220.
5. Absinta M, Sati P, Masuzzo F, et al. Association of chronic active multiple sclerosis lesions with disability *in vivo*. *JAMA Neurol*. 2019;76(12):1474.
6. Treaba CA, Conti A, Klawiter EC, et al. Cortical and phase rim lesions on 7 T MRI as markers of multiple sclerosis disease progression. *Brain Commun*. 2021;3(3):fcab134.
7. Reeves JA, Mohebbi M, Wicks T, et al. Paramagnetic rim lesions predict greater long-term relapse rates and clinical progression over 10 years. *Mult Scler J*. 2024;30(4-5):535-545.

8. Reeves JA, Mohebbi M, Zivadinov R, et al. Reliability of paramagnetic rim lesion classification on quantitative susceptibility mapping (QSM) in people with multiple sclerosis: Single-site experience and systematic review. *Mult Scler Relat Disord*. 2023;79:104968.
9. Bagnato F, Sati P, Hemond CC, et al. Imaging chronic active lesions in multiple sclerosis: A consensus statement. *Brain*. 2024;147(9):2913-2933.
10. Roxburgh R, Seaman SR, Masterman T, et al. Multiple sclerosis severity score: Using disability and disease duration to rate disease severity. *Neurology*. 2005;64(7):1144-1151.
11. Manouchehrinia A, Westerlind H, Kingwell E, et al. Age related multiple sclerosis severity score: Disability ranked by age. *Mult Scler J*. 2017;23(14):1938-1946.
12. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17(2):162-173.
13. Lebrun-Fréney C, Okuda DT, Siva A, et al. The radiologically isolated syndrome: Revised diagnostic criteria. *Brain*. 2023;146(8):3431-3443.
14. Li X, Morgan PS, Ashburner J, Smith J, Rorden C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J Neurosci Methods*. 2016;264:47-56.
15. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*. 2006;31(3):1116-1128.
16. Schmidt P, Pongratz V, Küster P, et al. Automated segmentation of changes in FLAIR-hyperintense white matter lesions in multiple sclerosis on serial magnetic resonance imaging. *NeuroImage Clin*. 2019;23:101849.
17. Wiltgen T, McGinnis J, Schlaeger S, et al. LST-AI: A deep learning ensemble for accurate MS lesion segmentation. *NeuroImage Clin*. 2024;42:103611.
18. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *J Open Source Softw*. 2019;4(43):1686.
19. Kuhn M, Vaughan D, Hvitfeldt E. yardstick: Tidy characterizations of model performance. Published online 2024. Accessed February 2024. <https://yardstick.tidymodels.org>
20. Gamer M, Lemon J, Fellows I, Singh P. irr: Various coefficients of interrater reliability and agreement. Published online 2019. Accessed February 2024.
21. Luchetti S, Franssen NL, Van Eden CG, Ramaglia V, Mason M, Huitinga I. Progressive multiple sclerosis patients show substantial lesion activity that correlates with clinical disease severity and sex: A retrospective autopsy cohort analysis. *Acta Neuropathol*. 2018;135(4):511-528.
22. Absinta M, Maric D, Gharagozloo M, et al. A lymphocyte-microglia-astrocyte axis in chronic active multiple sclerosis. *Nature*. 2021;597(7878):709-714.
23. Zhang S, Nguyen TD, Rúa SMH, et al. Quantitative susceptibility mapping of time-dependent susceptibility changes in multiple sclerosis lesions. *AJNR Am J Neuroradiol*. 2019;40(6):987-993.
24. Stölting A, Vanden Bulcke C, Borrelli S, et al. Clinical relevance of paramagnetic rim lesion heterogeneity in multiple sclerosis. *Ann Clin Transl Neurol*. 2024;11(12):3137-3151.
25. Hemond CC, Reich DS, Dundamadappa SK. Paramagnetic rim lesions in multiple sclerosis: Comparison of visualization at 1.5-T and 3-T MRI. *AJR Am J Roentgenol*. 2022;219(1):120-131.
26. Su L, Zhang Z, Gao C, et al. Brain lesion characteristics in Chinese multiple sclerosis patients: A 7-T MRI cohort study. *Ann Clin Transl Neurol*. 2025;12(2):300-310.
27. Hagemeyer J, Heininen-Brown M, Poloni GU, et al. Iron deposition in multiple sclerosis lesions measured by susceptibility-weighted imaging filtered phase: A case control study. *J Magn Reson Imaging*. 2012;36(1):73-83.
28. Borrelli S, Martire MS, Stölting A, et al. Central vein sign, cortical lesions, and paramagnetic rim lesions for the diagnostic and prognostic workup of multiple sclerosis. *Neurol Neuroimmunol Neuroinflammation*. 2024;11(4):e200253.

29. Wuerfel J, Sinnecker T, Ringelstein EB, et al. Lesion morphology at 7 Tesla MRI differentiates Susac syndrome from multiple sclerosis. *Mult Scler J*. 2012;18(11):1592-1599.
30. Barquero G, Rosa FL, Kebiri H, et al. RimNet: A deep 3D multimodal MRI architecture for paramagnetic rim lesion assessment in multiple sclerosis. *NeuroImage Clin*. 2020;28:102412.
31. La Rosa F, Wynen M, Al-Louzi O, et al. Cortical lesions, central vein sign, and paramagnetic rim lesions in multiple sclerosis: Emerging machine learning techniques and future avenues. *NeuroImage Clin*. 2022;36:103205.
32. Lou C, Sati P, Absinta M, et al. Fully automated detection of paramagnetic rims in multiple sclerosis lesions on 3T susceptibility-based MR imaging. *NeuroImage Clin*. 2021;32:102796.
33. Sati P, Oh J, Constable RT, et al. The central vein sign and its clinical evaluation for the diagnosis of multiple sclerosis: A consensus statement from the North American Imaging in Multiple Sclerosis Cooperative. *Nat Rev Neurol*. 2016;12(12):714-722.
34. Filippi M, Preziosa P, Banwell BL, et al. Assessment of lesions on magnetic resonance imaging in multiple sclerosis: Practical guidelines. *Brain*. 2019;142(7):1858-1875.
35. Calvi A, Clarke MA, Prados F, et al. Relationship between paramagnetic rim lesions and slowly expanding lesions in multiple sclerosis. *Mult Scler J*. 2023;29(3):352-362.
36. Elliott C, Rudko DA, Arnold DL, et al. Lesion-level correspondence and longitudinal properties of paramagnetic rim and slowly expanding lesions in multiple sclerosis. *Mult Scler J*. 2023;29(6):680-690.
37. Kaunzner UW, Kang Y, Zhang S, et al. Quantitative susceptibility mapping identifies inflammation in a subset of chronic multiple sclerosis lesions. *Brain*. 2019;142(1):133-145.