

Are your MRI contrast agents cost-effective?

Learn more about generic Gadolinium-Based Contrast Agents.



**AJNR**

## **Development of a Standardized MRI Scoring Tool for CNS Demyelination In Children**

L.H. Verhey, H.M. Branson, S. Laughlin, M.M. Shroff, S.M. Benseler, B.M. Feldman, D.L. Streiner, J.G. Sled and B. Banwell

This information is current as of April 19, 2024.

*AJNR Am J Neuroradiol* published online 24 January 2013  
<http://www.ajnr.org/content/early/2013/01/24/ajnr.A3382>

# Development of a Standardized MRI Scoring Tool for CNS Demyelination In Children

L.H. Verhey, H.M. Branson, S. Laughlin, M.M. Shroff, S.M. Benseler, B.M. Feldman, D.L. Streiner, J.G. Sled, and B. Banwell



## ABSTRACT

**BACKGROUND AND PURPOSE:** The degree to which MR imaging is useful in the diagnosis of MS is predicated on standardized and reliable evaluation of MR imaging parameters. We aimed to devise items for an MR imaging scoring tool that would have high inter-rater agreement and would be straightforward to apply.

**MATERIALS AND METHODS:** On the basis of a literature search and consensus of an expert panel, we identified 48 parameters that describe acute CNS demyelination, predict MS diagnosis, or characterize demyelinating disorder mimics. MR images of children with clinically confirmed MS, monophasic ADEM, and angiography-negative biopsy-positive small-vessel primary angiitis of the CNS were scored by 2 neuroradiologists independently, using the preliminary 48-parameter tool. Parameters with Cohen  $\kappa \geq 0.6$  and deemed important in predicting diagnosis were retained. Parameters not visualized on routine clinical imaging or not important in differentiating MS, ADEM, and SV-cPACNS were discarded.

**RESULTS:** Of 65 eligible patients, 55 children were enrolled (16 with monophasic ADEM, 27 with MS, 12 with SV-cPACNS); 10 were excluded (6 had hard-copy films, 4 did not meet MR imaging quality requirements). Of the 48 parameters, 16 were retained in the final scoring tool. The remaining 28 parameters were discarded: 4 had  $\kappa < 0.6$  and were not deemed useful in predicting diagnosis; 9 were not visible on routinely acquired clinical images; and 15 had inter-rater agreement  $\geq 0.6$  but were not useful in differentiating monophasic ADEM, MS, and SV-cPACNS.

**CONCLUSIONS:** We propose a 16-parameter MR imaging scoring tool that is straightforward to apply in the clinical setting and demonstrates high inter-rater agreement.

**ABBREVIATIONS:** ADEM = acute disseminated encephalomyelitis; CI = confidence interval; ICC = intraclass correlation coefficient; SV-cPACNS = small-vessel childhood primary angiitis of the CNS

owing to its high sensitivity, MR imaging is an invaluable tool for the diagnosis and management of patients with MS. MR imaging plays a key role in confirming an MS diagnosis before a

second clinical attack in patients with an incident CNS demyelinating event<sup>1-3</sup>; in excluding alternate diagnoses; in monitoring response to MS-targeted therapy by evaluation of lesion accrual on serial imaging; and in monitoring disease progression seen by formation of confluent lesions and atrophy. The value of MR imaging in MS diagnosis and in informing clinical management largely rests on the extent to which standard acquisition protocols and consistent terminology are used. The Consortium of MS Centers published a consensus-based MR imaging acquisition protocol recommended for patients with MS.<sup>4</sup> A standard lexicon of the MR imaging features of CNS demyelination would further increase the consistency with which MR images are reported and would facilitate pooling of datasets in multicentered studies. We aimed to devise items for a standard MR imaging scoring tool for pediatric-onset CNS demyelination that demonstrates high inter-rater reliability and is straightforward to apply in the clinical setting.

Received July 10, 2012; accepted after revision August 13.

From the Program in Neuroscience and Mental Health (L.H.V., B.B.); Division of Pediatric Neuroradiology (H.M.B., S.L., M.M.S.), Department of Diagnostic Imaging; Divisions of Rheumatology (S.M.B., B.M.F.) and Neurology (B.B.), Department of Pediatrics, and Program in Physiology and Experimental Medicine (J.G.S.), The Hospital for Sick Children, Toronto, Ontario, Canada; Department of Psychiatry (D.L.S.), Department of Medical Biophysics (J.G.S.), Institute of Medical Science (L.H.V., B.B.), Department of Medical Imaging (H.M.B., S.L., M.M.S.), Institute of Health Policy, Management and Evaluation (S.M.B., B.M.F.); Department of Medicine (B.M.F.); and Dalla Lana School of Public Health (B.M.F.), University of Toronto, Toronto, Ontario, Canada; and Departments of Clinical Epidemiology and Biostatistics (D.L.S.) and Psychiatry and Behavioural Neurosciences (D.L.S.), McMaster University, Hamilton, Ontario, Canada.

This work was supported by the Canadian Multiple Sclerosis Scientific Research Foundation, Canadian Institutes of Health Research, and Multiple Sclerosis Society of Canada.

Please address correspondence to Brenda Banwell, MD, Division of Neurology, Department of Pediatrics, The Children's Hospital of Philadelphia, Rm. 10018 Colket Research Building, 3501 Civic Center Blvd., Philadelphia, PA, 19104; e-mail: banwellb@email.chop.edu

Indicates open access to non-subscribers at [www.ajnr.org](http://www.ajnr.org)

Indicates article with supplemental on-line appendix and tables

<http://dx.doi.org/10.3174/ajnr.A3382>

## MATERIALS AND METHODS

### Participants and Definitions

Children and adolescents younger than 18 years of age with MS and monophasic ADEM were identified from the Pediatric Demyelinating Disease Registry (The Hospital for Sick Children–SickKids, Toronto, Ontario, Canada), and those with SV-cPACNS were identified from a single-center prospective cohort of children followed at SickKids.<sup>5,6</sup> Participant selection was based on the following inclusion criteria: 1) availability of an MR imaging scan acquired within 30 days of the initial clinical presentation; 2) children with MS diagnosed on the basis of at least 2 demyelinating episodes<sup>7</sup> and followed from the first attack for a minimum of 2 years; 3) children with monophasic ADEM (defined as polyfocal neurologic deficits and encephalopathy)<sup>8</sup> followed for at least 3 years without further clinical or MR imaging evidence of new or recurrent demyelination; and 4) children with SV-cPACNS having met the Calabrese diagnostic criteria (no evidence of systemic vasculitis, normal cerebral angiography findings, and brain biopsy confirmation of isolated small-vessel vasculitis).<sup>9</sup> Participants with SV-cPACNS were included in this study because their presenting clinical and MR imaging features often overlap those of acute CNS demyelination and thus they were thought to be informative in delineating MR imaging features suited to future studies in which the MR imaging tool will be evaluated for specificity across different CNS disorders. Ethics approval was obtained for this study.

### MR Imaging Scoring Tool

A comprehensive PubMed search for articles published between January 1, 1980, and January 1, 2012, was performed by using combinations of the following search terms: “magnetic resonance imaging” OR “MRI,” “multiple sclerosis” OR “MS,” “pediatric,” “inflammatory demyelination,” “acute disseminated encephalomyelitis” OR “ADEM,” and “small-vessel CNS vasculitis.” The search was restricted to English language publications. References cited in original and review articles were also assessed. Fifty-one articles were reviewed, from which 48 MR imaging parameters were identified. On-line Table 1 lists the MR imaging parameters that formed the initial iteration of the scoring tool. A panel of experts in pediatric CNS demyelinating and inflammatory disease that included a neurologist (B.B.), a rheumatologist (S.M.B.), and 3 neuroradiologists (H.M.B., S.L., M.M.S.) was convened (moderated by L.H.V.) to create working definitions for each of the 48 parameters on the basis of the literature and panel consensus. As detailed in On-line Table 1, all parameters with the exception of lesion count were binary (present/absent). Following the panel session, a dictionary was created to document the definition of each parameter. A second follow-up panel meeting was held to further revise the parameters and definitions.

Two investigators (H.M.B., S.L.) independently applied the 48-parameter scoring tool, blinded to clinical information, to a training set of 9 randomly ordered MR imaging scans (3 with MS, 3 with ADEM, 3 with SV-cPACNS). These 9 scans were not included in the test set described subsequently.

### MR Imaging Analysis

MR images were acquired at 1.5T according to clinical protocol and archived at SickKids. At minimum, T1- and T2-weighted or FLAIR images were required for each patient. Postcontrast T1-weighted and diffusion-weighted images were evaluated when available. MR images were reviewed for image quality by a pediatric neuroradiologist (H.M.B. or S.L.) blinded to clinical information; hard-copy films and scans degraded by dental hardware or patient motion artifacts were not evaluated. All MR images were copied from the PACS, anonymized, and subsequently analyzed on an eFilm (Version 1.5.3; <https://estore.merge.com/na/index.aspx>) DICOM viewer workstation.

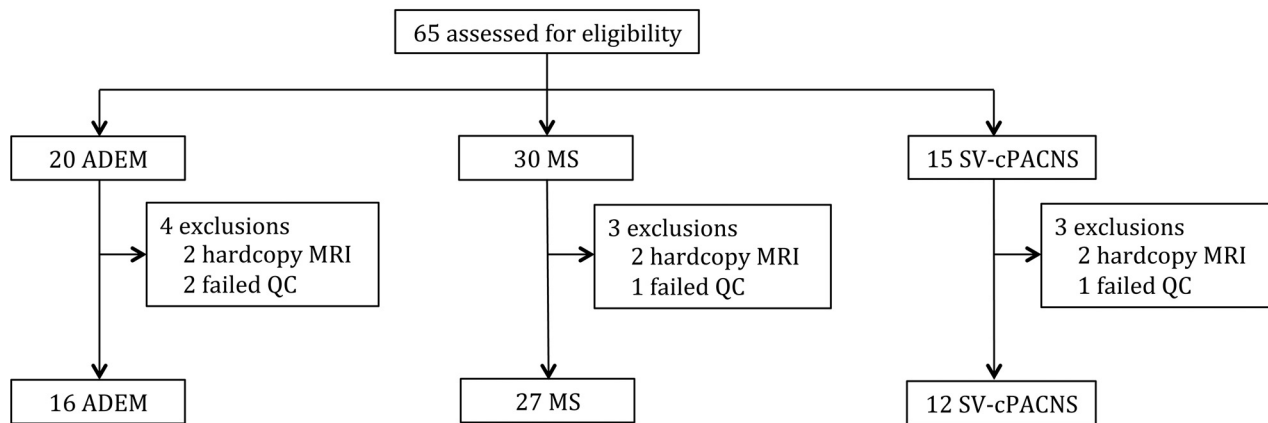
The MR image acquired at presentation was evaluated by the 2 trained raters (H.M.B., S.L.) independently and blinded to presenting symptoms and diagnosis. One individual (L.H.V.) was present during all scoring sessions to ensure consistent use of the parameters by the 2 trained raters and to perform data entry. A dictionary of the 48 parameters was available to both raters during the scoring sessions. An Access (Version 2003; Microsoft, Redmond, Washington) database was created, into which the 2 trained raters' responses to the 48 parameters were entered.

A lesion was defined as a T2-weighted or FLAIR hyperintensity with a minimum diameter of 3 mm in either the axial, sagittal, or coronal plane. Adjacent lesions were classified as distinct when separated by at least 1 mm of normal-appearing tissue. T1 hypointense lesions were defined as hypointense to cortical gray matter on T1-weighted imaging and were correlated with hyperintense lesions on T2-weighted imaging. T1 hypointense lesions were confirmed as nonenhancing on postcontrast T1-weighted imaging.

### Statistical Analysis

The level of inter-rater agreement for each of the 48 parameters was determined by calculating Cohen  $\kappa$ <sup>10</sup> or the ICC<sup>11</sup> as appropriate. Strength of agreement was arbitrated as  $\leq 0$  = poor, 0.01–0.20 = slight, 0.21–0.40 = fair, 0.41–0.60 = moderate, 0.61–0.80 = substantial, 0.81–1 = almost perfect.<sup>12</sup> For determining whether parameters were retained or discarded, the expert panel proposed the following a priori rules: 1) Parameters with inter-rater agreement  $\geq 0.6$  and the lower limit of the 95% CI  $\geq 0.5$  would be retained, provided the parameters were deemed diagnostically useful on the basis of the panel expertise and relevant literature. 2) Parameters deemed by the panel as important on the basis of the literature in discriminating monophasic ADEM, MS, and SV-cPACNS for which  $\kappa$  or ICC was  $< 0.6$  or the lower limit of the 95% CI was  $< 0.5$  would be re-evaluated independently by both raters on all scans after refinement of the parameter definition; if inter-rater agreement increased to  $\geq 0.6$  after re-scoring, the parameter would be retained. 3) Parameters not visualized on routine clinical MR imaging sequences would be discarded, irrespective of the level of inter-rater agreement. 4) Parameters not contributory to differentiating ADEM, MS, and SV-cPACNS would be discarded.

Following completion of MR imaging scoring by both raters, the panel was reconvened to review the inter-rater agreement of each parameter and decide which parameters were retained on the basis of the a priori rules defined previously. When there was disagreement among the panel members, 1 individual (M.M.S.), expert in neuroimaging and not involved in the scoring, served as an arbitrator.



**FIG 1.** Description of participants. QC indicates quality control.

**Table 1: Inter-rater agreement of parameters retained**

Parameter	$\kappa$	95% CI
Lesion count	0.81 <sup>a</sup>	0.68–0.89
Gyral projection <sup>b</sup>	0.74	0.54–0.93
Diffusion restriction of lesion	1.0	1.0–1.0
Lesion enhancement	0.92	0.81–1.0
Leptomeningeal enhancement	0.85	0.55–1.0
Black hole	0.76	0.54–0.98
Periventricular lesion	0.80	0.66–0.94
Internal capsular lesion	0.74	0.58–0.91
Brain stem lesion <sup>c</sup>	0.72	0.57–0.88
Cerebellar lesion	0.73	0.56–0.91

<sup>a</sup> ICC reported.

<sup>b</sup> Formerly fingerlike projection.

<sup>c</sup> Brain stem collapsed from right brain stem ( $\kappa = 0.71$ ; 95% CI, 0.53–0.88), midline brain stem ( $\kappa = 0.78$ ; 95% CI, 0.61–0.95), and left brain stem ( $\kappa = 0.78$ ; 95% CI, 0.61–0.95).

A 1-way ANOVA with post hoc Bonferroni pair-wise comparisons was used to compare age at first attack in children with monophasic ADEM, MS, and SV-cPACNS.

Statistical analyses were performed by using STATA version 12 (StataCorp, College Station, Texas) and the Statistical Package for the Social Sciences, Version 20.0 (IBM SPSS Statistics for Windows, Armonk, New York).

## RESULTS

### Participants

Sixty-five children and adolescents were assessed for eligibility. MR images for 10 participants were excluded because either they were of insufficient quality ( $n = 4$ ) or the images were on hardcopy film ( $n = 6$ ). MR images were evaluated for the 55 children (16 with ADEM, 27 with MS, and 12 with SV-cPACNS) included (Fig 1). Children with ADEM were younger (mean age,  $6.2 \pm 4.4$  years) at the time of first attack than those with MS (mean age,  $12.8 \pm 3.5$  years;  $P < .0001$ ) and SV-cPACNS (mean age,  $10.3 \pm 4.3$  years;  $P = .028$ ). Onset age did not differ between children with MS and those with SV-cPACNS ( $P = .209$ ).

### Retained Parameters

Of the 48 parameters, 10 (demonstrating an inter-rater agreement statistic of  $\geq 0.6$  with a lower limit of the 95% CI  $\geq 0.5$ ) could be readily evaluated on routine clinical sequences and were deemed diagnostically useful on the basis of panel consensus and the lit-

erature (Table 1). A post hoc decision was to collapse “midline brain stem,” “left brain stem,” and “right brain stem” into 1 parameter (brain stem) because any 1 lesion commonly co-occurred in all 3 locations. The panel agreed to rename the parameter “fingerlike projections” as “gyral projections,” a more accurate descriptor of the feature. Of the 10 parameters, 2 that were not deemed by the panel to be MR imaging features of acute CNS demyelination were retained for their perceived utility in distinguishing CNS demyelination and SV-cPACNS or other mimics: 1) leptomeningeal enhancement ( $\kappa = 0.85$ ; 95% CI, 0.55–1.0) was deemed by the panel as an important feature in differentiating CNS inflammatory demyelination from SV-cPACNS in cases in which presenting symptoms were nondiscriminatory; and 2) diffusion restriction visible on DWI ( $\kappa = 1.0$ ) was deemed by the panel as a specificity parameter, given its reported sensitivity for arterial ischemia.

The panel attributed the low inter-rater agreement of 3 of the 48 parameters (caudate:  $\kappa = 0.49$ , 95% CI, 0.25–0.73; putamen:  $\kappa = 0.44$ , 95% CI, 0.16–0.71; globus pallidus:  $\kappa = 0.39$ , 95% CI, 0.13–0.65) to the challenge of precisely delineating the borders of the lentiform nuclei and caudate on conventional MR imaging. The panel also deemed that the radiologic distinction among the nuclei was not relevant to acute CNS demyelination. Therefore, the panel agreed to collapse these 3 parameters into 1, basal ganglia.

Of the 48 parameters, 6 (1 of which included the collapsed parameter “basal ganglia”) were deemed by the panel as important in predicting chronic demyelination as opposed to a monophasic demyelinating illness but had inter-rater reliability statistics  $< 0.6$  following first-pass scoring (Table 2). The panel agreed that ambiguity in the definitions of 5 of the 6 parameters hindered their reliable application across raters. Revisions were made to the definition of the 5 parameters as follows: 1) “bilateral lesion distribution,” refers to both the supratentorial and infratentorial regions; 2) “juxtacortical,” a lesion must involve the subcortical U-fibers to be scored as juxtacortical; 3) “intracallosal,” provision of anatomic landmarks to define borders of the corpus callosum in the transverse plane (On-line Appendix, Parameter 1) and specification that 1 mm of normal-appearing white matter surrounding a lesion required to define a lesion as being “intracallosal” (to distinguish such lesions from periventricular lesions);

**Table 2: Inter-rater agreement of parameters retained following rescoring**

Parameter	First Pass		Second Pass	
	$\kappa$	95% CI	$\kappa$	95% CI
Bilateral distribution	0.66	0.45–0.87	0.73	0.50–0.95
Intracallosal	0.65	0.42–0.87	0.90	0.76–1.00
Thalamic	0.60	0.40–0.80	0.85	0.71–0.99
Basal ganglia <sup>a</sup>	0.37	0.16–0.58	0.77	0.60–0.94
Cerebral white matter-other <sup>b</sup>	0.38	0.16–0.59	0.69	0.46–0.91
Juxtacortical	0.46	0.28–0.63	0.64	0.41–0.86

<sup>a</sup> Basal ganglia collapsed from caudate ( $\kappa = 0.49$ ; 95% CI, 0.25–0.73), putamen ( $\kappa = 0.44$ ; 95% CI, 0.16–0.71), and globus pallidus ( $\kappa = 0.39$ ; 95% CI, 0.13–0.65).

<sup>b</sup> Formerly subcortical lesion.

**Table 3: Inter-rater agreement of parameters not included in the scoring tool**

Parameter	$\kappa$	95% CI
Inter-rater agreement <0.6 or lower limit of 95% CI, <.5		
Symmetric pattern	–0.04	–0.08–0.01
Ependymal enhancement	0.38	–0.15–0.91
Fingerlike + projection	0.46	0.10–0.82
Proportion of discrete lesions	0.39 <sup>a</sup>	0.24–0.54
Not visualized on routine clinical brain imaging		
Cervical spinal cord	0.72	0.42–1.0
Dot-dash sign	0.71	0.46–0.97
Cortical gray matter	0.36	0.15–0.58
Optic nerve lesion	0.48	–0.03–0.98
Optic nerve enhancement	0.82	0.50–1.0
Optic nerve sheath enhancement	–	–
Extraoptic fat enhancement	–	–
Extraoptic muscle enhancement	–	–
Perineural enhancement	–	–
Not contributory to differentiating ADEM, MS, and SV-pACNS		
Frontal lobar location	0.60	0.36–0.85
Temporal lobar location	0.81	0.67–0.94
Parietal lobar location	0.80	0.66–0.94
Occipital lobar location	0.70	0.53–0.86
ACA vascular territory	0.56	0.37–0.75
MCA vascular territory	0.80	0.58–1.0
PCA vascular territory	0.64	0.47–0.82
Vertebrobasilar vascular territory	0.73	0.57–0.89
Cerebellar peduncle	0.76	0.59–0.93
Target lesion	0.78	0.62–0.94
Nodular leptomeningeal enhancement	–	–
Linear leptomeningeal enhancement	–	–
Dural enhancement	–	–
Supratentorial lesion enhancement	0.76	0.59–0.94
Infratentorial lesion enhancement	0.78	0.58–0.98

**Note:**—ACA indicates anterior cerebral artery; PCA, posterior cerebral artery; –,  $\kappa$  invalid due to low frequency of occurrence in  $\geq 1$  cell of the  $2 \times 2$  table.

<sup>a</sup> Weighted  $\kappa$  reported.

and 4) “thalamic” and “basal ganglia,” encompass lesions that involve the thalamus or basal ganglia even if the lesions are not entirely contained within these regions.

The poor inter-rater agreement of the sixth parameter, “subcortical lesions,” was due to a difference in opinion among raters on what represented subcortical white matter. One rater referred to all supratentorial white matter extending between the cortical ribbon and the lateral ventricles as subcortical; therefore, a lesion located in the supratentorial white matter that did not abut the cortex or lateral ventricle was scored as subcortical. The other rater viewed the supratentorial white matter as divided into deep (adjacent to the lateral ventricles) and superficial (adjacent to the cerebral cortex) white matter and therefore scored only those lesions in the white matter that were adjacent to, but not contiguous

with, the cortex as “subcortical lesions.” To ensure that the subcortical lesion parameter was interpreted as involving all supratentorial (nonjuxtacortical and nonperiventricular) white matter, the definition was revised and the parameter was renamed “cerebral white matter-other.”

After the definitions were revised to eliminate ambiguity, the 6 parameters were independently re-evaluated on all scans by both trained raters. Inter-rater agreement increased to  $\geq 0.6$ , permitting their retention in the final tool (Table 2).

In total, 16 parameters were retained in the final tool (Tables 1 and 2). A textual and pictographic atlas of the final 16-parameter scoring tool was created and published in our recent work (On-line Appendix).<sup>13</sup> In the atlas, anatomic landmarks have been delineated for parameters that rely on accurate identification of anatomic structures.

### Parameters Discarded

As shown in Table 3, of the 48 parameters, we excluded 28: Four due to low inter-rater agreement, 9 due to poor visualization of the parameter on routine clinical imaging, and 15 that the expert panel deemed to be not diagnostically valuable.

Specifically, “symmetric pattern,” “ependymal enhancement,” “fingerlike + projection,” and “proportion of discrete lesions” had  $\kappa < 0.6$  (or the lower limits of the 95% CI  $< 0.5$ ) and, even if redefined, were not deemed by the panel to be diagnostically useful.

Nine parameters could not be accurately scored on routine clinical sequences. Although well-recognized as features of MS, the evaluation of “intracortical,” “cervical spinal cord,” and “optic nerve lesions,” requires targeted cortical, spinal cord, or orbital imaging sequences—all of

which are not routinely acquired in a clinical brain MR imaging protocol. One parameter (the “dot-dash” sign<sup>34</sup>) has been described in 1 study as an early MR imaging feature of MS; however, scoring the parameter requires thin sagittal T2-weighted or FLAIR imaging through the midline. The remaining 5 parameters (“optic nerve enhancement,” “optic nerve sheath enhancement,” “extraoptic fat enhancement,” “extraoptic muscle enhancement,” and “perineural enhancement”) require specialized fat-suppressed orbital imaging.

Finally, 15 parameters, despite demonstrating acceptable inter-rater agreement, did not aid in differentiating ADEM, MS, and SV-cPACNS (On-line Table 2) and were, therefore, excluded. The formation of demyelinating and SV-cPACNS lesions does not respect lobar (4 parameters) or vascular territory (4 param-



ters) boundaries, and the location of contrast-enhancing lesions (ie, supratentorial and infratentorial) was not discriminatory. While the presence of leptomeningeal enhancement was retained in the scoring tool, the more specific parameters (“nodular leptomeningeal enhancement,” “linear leptomeningeal enhancement,” and “dural enhancement”) were too infrequently noted to permit computation of inter-rater reliability and were not deemed by the panel to be discriminatory. The parameter called “cerebellar peduncle lesions” was discarded due to its co-occurrence with brain stem lesions and cerebellar lesions and the challenge of deciphering the margins of the cerebellar peduncles. “Target lesions” was discarded because the panel could not reach consensus on a consistent definition.

## DISCUSSION

We created an MR imaging scoring tool consisting of 16 parameters that demonstrate substantial inter-rater reliability. The rationale for each of the parameters included in the tool is detailed in the Supplementary Panel of the On-line Appendix and is based on evidence for their utility in characterizing the MR imaging features of acute demyelination and for their utility in discriminating acute CNS demyelination and SV-cPACNS. For a clinically useful tool to be generally accepted, it must be practical to use in the clinical setting without the need for rigorous training. Thus, we intentionally created the tool to be binary-response (with the exception of lesion count, which has an upper limit lesion count of 15) to minimize the quantitative requirement and maximize its efficient use in clinical practice.

The first component of developing the MR imaging tool was to devise the parameters themselves. We used established methods of item identification, including a literature search and expert consensus,<sup>14</sup> to formulate a comprehensive list of potential parameters. The rationale for performing a literature search was so that the tool would comprise parameters that have been empirically demonstrated to be MR imaging characteristics of acute CNS demyelination and relapsing-remitting MS or features that might discriminate demyelination and SV-cPACNS. This method of devising tool parameters has been used in other instances, such as in a scale for differentiating irritable bowel syndrome and organic bowel disease, in which the parameters represented clinical features and laboratory values that were found to distinguish the 2 patient groups.<sup>15</sup>

We also used expert consensus as a second method for devising potential parameters. We selected the expert panel from neuroradiology, neurology, and rheumatology staff working in our established pediatric demyelinating disease and CNS vasculitis programs, ensuring extensive clinical and radiologic experience with pediatric-onset MS and SV-cPACNS. The expert panel played a key role in defining the comprehensive list of parameters identified from the literature and in devising definitions for each parameter that were objective and were straightforward to apply. Similar utility of expert opinion has been reported elsewhere, such as the Patient-Reported Outcomes Measurement Information System initiative funded by the National Institutes of Health, in which experts contribute potential parameters to be used in evaluating patient-reported outcomes across multiple health conditions.<sup>16</sup>

Following identification of relevant MR imaging parameters for the diseases under evaluation, we created definitions for each parameter to enable consistency in application of the parameters among raters.<sup>17,18</sup> Therefore, a key component of our work was to assess the inter-rater reliability of the potential parameters identified by the literature search and expert opinion; only parameters that demonstrated substantial or good inter-rater agreement ( $\kappa \geq 0.6$ )<sup>10,12</sup> were retained in the final tool. Several MR imaging parameters demonstrated poor inter-rater agreement on first evaluation but were deemed important discriminatory parameters by the panel. These parameters were redefined; then, all scans were rescored, and only when the inter-rater agreement increased to  $\geq 0.6$  was the parameter retained.

A second noteworthy aspect is that we focused on the inter-rater reliability of our parameters and not intrarater reliability. Because the error contributing to intrarater reliability is considered contained within inter-rater reliability, demonstration of high inter-rater reliability is sufficient.<sup>17</sup> Of the 16 parameters retained in the final tool, 14 demonstrated inter-rater agreement that was considerably higher ( $\geq 0.72$ ) than our a priori cut-point of 0.6. The use of inter-rater agreement to guide the selection of parameters included in a scoring tool is not unique to our study. Similar methodology was used during the creation of the well-established Glasgow Coma Scale, a clinical scale for evaluating the depth and duration of impaired consciousness and coma,<sup>19</sup> and the International Standards for Neurologic Classification of Spinal Cord Injury of the American Spinal Injury Association, a practice guideline for classifying the degree of neurologic impairment due to spinal cord injury.<sup>18,20,21</sup>

A key aspect of the diagnostic criteria for multiple sclerosis is the exclusion of mimics of CNS demyelination.<sup>22,23</sup> In creating our proposed tool, due to the clinical challenge of distinguishing SV-cPACNS from MS, we included MR images of children with SV-cPACNS because this disease was thought to have a similar but potentially distinct inflammatory pattern on MR imaging. We also included parameters that have been reported in CNS infection and malignancy. Specifically, leptomeningeal enhancement has been described in children with SV-cPACNS<sup>6</sup> and is well-documented in infectious<sup>24,25</sup> and neoplastic<sup>26-28</sup> processes. Leptomeningeal enhancement is not a feature of MS, and its presence was highlighted by an international consensus panel as a “red flag” to consider other nondemyelinating etiologies.<sup>23</sup> In contrast, the MR imaging parameter of “acute diffusion restriction,” while highly characteristic of vascular occlusion such as stroke,<sup>29-31</sup> has also been a feature recently reported in acute demyelinating lesions, particularly tumefactive lesions,<sup>32,33</sup> and therefore was retained in the final tool.

In developing the MR imaging scoring tool, we considered the challenges of MR imaging acquisition in the pediatric context. Parameters that required sequences beyond a standard clinical brain protocol for accurate scoring were not included in the final tool. For example, while optic nerve and spinal cord lesions are well-documented features of MS, focused spinal cord or fat-suppressed orbital imaging is required to adequately resolve these lesions. Adding these sequences to a standard clinical brain MR imaging protocol significantly increases scan time, rendering such acquisitions intolerable in children who are not sedated and in-

curing higher cost in the context of research studies. The yield of orbital or spinal imaging in the absence of clinical evidence of acute or remote optic nerve or spinal cord involvement has not been established in children subsequently diagnosed with MS; further studies will be required, with appropriate consideration of scanning time.

We created a manual that contains an atlas and definition for each parameter and highlights important anatomic delineations to ensure that the tool can be accurately used by radiologists and clinicians who were not part of the creation of the tool. Not only will this aid in use of the tool across clinical centers, it also serves as a model for MR imaging scoring tools to be applied in pediatric MS clinical trials in which inclusion in the trial will be predicated on accurate MS diagnosis.

The intended future application of our proposed MR imaging tool is multifaceted and requires validation. The first priority will be to determine the ability of the MR imaging tool to identify distinct features of different CNS inflammatory disorders, such as monophasic CNS demyelination, MS, ADEM, and SV-cPACNS because such disorders share clinical features rendering accurate diagnosis challenging at onset. The MR imaging tool will then be evaluated for specificity in identifying noninflammatory CNS disorders, such as inherited or metabolic diseases that also have onset during childhood and impact CNS white matter. Finally, we need to evaluate user satisfaction with the tool because application in a busy clinical environment will require endorsement of utility and applicability.

## CONCLUSIONS

We developed an MR imaging scoring tool consisting of 16 items that demonstrate substantial inter-rater agreement. Using established methods of item identification, including a literature search and expert consensus, the parameters are based on evidence for their utility in characterizing the MR imaging features of acute CNS inflammation and demyelination. The binary-response nature of the parameters and the tool manual will facilitate utility of the tool in clinical practice without the need for rigorous training. The scoring tool will inform the creation of structured reporting that is increasingly being proposed for use in diagnostic radiology.

## ACKNOWLEDGMENTS

We thank Drs Adrian Fawcett and Sunita Venkateswaran for their contributions to the initial phase of this work, the staff of the Pediatric Demyelinating Disease Program at The Hospital for Sick Children (Toronto, Ontario, Canada), and the participating children and their families for their commitment to our research program.

Disclosures: Leonard H. Verhey—*RELATED: Grant:* Canadian Multiple Sclerosis Scientific Research Foundation,\* *Comments:* This study was funded under a research grant from the Canadian Multiple Sclerosis Scientific Research Foundation. L.H. Verhey receives scholarship support from the Canadian Institutes of Health Research, Multiple Sclerosis Society of Canada, University of Toronto, and The Hospital for Sick Children, *Support for Travel to Meetings for the Study or Other Purposes:* Canadian Multiple Sclerosis Scientific Research Foundation, *Comments:* Travel to present at scientific meetings was covered under a research grant from the Canadian Multiple Sclerosis Scientific Research Foundation. Helen M. Branson—*RELATED: Grant:* Canadian Multiple Sclerosis Scientific Research Foundation,\* *Support for Travel to Meetings for the Study or Other Purposes:* Canadian Multiple Sclerosis

Scientific Research Foundation,\* *Comments:* I did not personally receive any funding but other authors received funding for travel to meetings. Suzanne Laughlin—*RELATED: Grant:* Canadian Multiple Sclerosis Scientific Research Foundation,\* Manohar M. Shroff—*RELATED: Grant:* Canadian Multiple Sclerosis Scientific Research Foundation,\* *Comments:* This study was conducted under a research grant (paid to SickKids) from the Canadian Multiple Sclerosis Scientific Research Foundation, *Support for Travel to Meetings for the Study or Other Purposes:* Canadian Multiple Sclerosis Scientific Research Foundation,\* *Comments:* The grant covered travel expenses for fellows or postdoctorates with whom I worked, when presenting this work at scientific meetings. I personally was never paid out of this grant for travel expenses to present at any meeting. Brenda Banwell—*RELATED: Grant:* Canadian Multiple Sclerosis Scientific Foundation,\* *UNRELATED: Consultancy:* Biogen-IDEc, Merck-Serono, *Comments:* no relevance to present work, *Grants/Grants Pending:* Canadian Institute of Health Research,\* MS Society of Canada,\* *Payment for Lectures (including service on Speakers Bureaus):* Novartis, Biogen-IDEc, Serono, *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed:* visiting professorships. \*Money paid to the institution.

## REFERENCES

1. McDonald WI, Compston A, Edan G, et al. **Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the Diagnosis of Multiple Sclerosis.** *Ann Neurol* 2001;50:121–27
2. Polman CH, Reingold SC, Edan G, et al. **Diagnostic criteria for multiple sclerosis: 2005 revisions to the “McDonald Criteria.”** *Ann Neurol* 2005;58:840–46
3. Polman CH, Reingold SC, Banwell B, et al. **Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria.** *Ann Neurol* 2011;69:292–302
4. Simon JH, Li D, Traboulsee A, et al. **Standardized MR imaging protocol for multiple sclerosis: Consortium of MS Centers consensus guidelines.** *AJNR Am J Neuroradiol* 2006;27:455–61
5. Hutchinson C, Elbers J, Halliday W, et al. **Treatment of small vessel primary CNS vasculitis in children: an open-label cohort study.** *Lancet Neurol* 2010;9:1078–84
6. Elbers J, Halliday W, Hawkins C, et al. **Brain biopsy in children with primary small-vessel central nervous system vasculitis.** *Ann Neurol* 2010;68:602–10
7. Poser CM, Paty DW, Scheinberg L, et al. **New diagnostic criteria for multiple sclerosis: guidelines for research protocols.** *Ann Neurol* 1983;13:227–31
8. Krupp LB, Banwell B, and Tenenbaum S. **Consensus definitions proposed for pediatric multiple sclerosis and related disorders.** *Neurology* 2007;68:S7–12
9. Cellucci T, Benseler SM. **Diagnosing central nervous system vasculitis in children.** *Curr Opin Pediatr* 2010;22:731–38
10. Sim J, Wright CC. **The kappa statistic in reliability studies: use, interpretation, and sample size requirements.** *Phys Ther* 2005;85:257–68
11. Shrout PE, Fleiss JL. **Intraclass correlations: uses in assessing rater reliability.** *Psychol Bull* 1979;86:420–28
12. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74
13. Verhey LH, Branson HM, Shroff MM, et al. **MRI parameters for prediction of multiple sclerosis diagnosis in children with acute CNS demyelination: a prospective national cohort study.** *Lancet Neurol* 2011;10:1065–73
14. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use.* New York: Oxford University Press; 2008:17–29
15. Kruis W, Rindfleisch GE, Weinzierl M. **Zinc deficiency as a problem in patients with Crohn’s disease and fistula formation.** *Hepatology* 1985;32:133–34
16. Cella D, Yount S, Rothrock N, et al. **The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years.** *Med Care* 2007;45:S3–11
17. Streiner DL, Norman GR. *Health Measurement Scales: A Practical*

*Guide to their Development and Use*. New York: Oxford University Press; 2008:167–207

18. Marino RJ, Jones L, Kirshblum S, et al. **Reliability and repeatability of the motor and sensory examination of the international standards for neurological classification of spinal cord injury.** *J Spinal Cord Med* 2008;31:166–70
19. Teasdale G, Jennett B. **Assessment of coma and impaired consciousness: a practical scale.** *Lancet* 1974;2:81–84
20. Jonsson M, Tollback A, Gonzales H, et al. **Inter-rater reliability of the 1992 international standards for neurological and functional classification of incomplete spinal cord injury.** *Spinal Cord* 2000;38:675–79
21. Priebe MM, Waring WP. **The interobserver reliability of the revised American Spinal Injury Association standards for neurological classification of spinal injury patients.** *Am J Phys Med Rehabil* 1991;70:268–70
22. Charil A, Yousry TA, Rovaris M, et al. **MRI and the diagnosis of multiple sclerosis: expanding the concept of “no better explanation.”** *Lancet Neurol* 2006;5:841–52
23. Miller DH, Weinshenker BG, Filippi M, et al. **Differential diagnosis of suspected multiple sclerosis: a consensus approach.** *Mult Scler* 2008;14:1157–74
24. Baumann M, Birnbacher R, Koch J, et al. **Uncommon manifestations of neuroborreliosis in children.** *Eur J Paediatr Neurol* 2010;14:274–77
25. Jang S, Suh SI, Ha SM, et al. **Enterovirus 71-related encephalomyelitis: usual and unusual magnetic resonance imaging findings.** *Neuroradiology* 2012;54:239–45
26. Singh SK, Leeds NE, Ginsberg LE. **MR imaging of leptomeningeal metastases: comparison of three sequences.** *AJNR Am J Neuroradiol* 2002;23:817–21
27. Pang J, Banerjee A, Tihan T. **The value of tandem CSF/MRI evaluation for predicting disseminated disease in childhood central nervous system neoplasms.** *J Neurooncol* 2008;87:97–102
28. Chamberlain MC. **A review of leptomeningeal metastases in pediatrics.** *J Child Neurol* 1995;10:191–99
29. Warach S, Gaa J, Siewert B, et al. **Acute human stroke studied by whole brain echo planar diffusion-weighted magnetic resonance imaging.** *Ann Neurol* 1995;37:231–41
30. De Vries LS, Van der Grond J, Van Haastert IC, et al. **Prediction of outcome in new-born infants with arterial ischaemic stroke using diffusion-weighted magnetic resonance imaging.** *Neuropediatrics* 2005;36:12–20
31. Kirton A, Shroff M, Visvanathan T, et al. **Quantified corticospinal tract diffusion restriction predicts neonatal stroke outcome.** *Stroke* 2007;38:974–80
32. Kiriya T, Kataoka H, Taoka T, et al. **Characteristic neuroimaging in patients with tumefactive demyelinating lesions exceeding 30 mm.** *J Neuroimaging* 2011;21:e69–e77
33. Malhotra HS, Jain KK, Agarwal A, et al. **Characterization of tumefactive demyelinating lesions using MR imaging and in-vivo proton MR spectroscopy.** *Mult Scler* 2009;15:193–203