PREDICTION OF A DIAGNOSIS OF MULTIPLE SCLEROSIS IN PATIENTS WITH CLINICALLY ISOLATED SYNDROME USING THE MAGNIMS 2016 AND MCDONALD 2010 CRITERIA: A RETROSPECTIVE STUDY

[1,2]Massimo Filippi, Prof, MD, [1,2]Paolo Preziosa, MD, [1]Alessandro Meani, MSc, [3]Olga Ciccarelli, Prof, PhD, [4]Sarlota Mesaros, PhD, [5]Alex Rovira, Prof, MD, [6]Jette Frederiksen, Prof, MD, [7]Christian Enzinger, Prof, MD, [3,8]Frederik. Barkhof, Prof, MD, [9]Claudio Gasperini, MD, [3]Wallace Brownlee, MD, [4]Jelena Drulovic, Prof, MD, [10]Xavier Montalban, Prof, MD [11]Stig P. Cramer, MD, [7]Alexander Pichler, MD, [12]Marloes Hagens, MSc, [9]Serena Ruggieri, MD, [2]Vittorio Martinelli, MD, [3]Katherine Miszkiel, MRCP, [10]Mar Tintorè, MD, [2]Giancarlo Comi, Prof, MD, [8,12]Iris Dekker, MSc, [12]Bernard Uitdehaag, Prof, MD, [4]Irena Dujmovic-Basuroski, MD, [1,2]Maria A. Rocca, MD.

[1]Neuroimaging Research Unit and [2]Department of Neurology, Institute of Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy; [3]Queen Square MS Centre, Department of Neuroinflammation, UCL Institute of Neurology, University College London (UCL), London, UK and NIHR UCLH BRC; [4]Clinic of Neurology, Faculty of Medicine, University of Belgrade, Serbia; [5]Section of Neuroradiology, Department of Radiology, Hospital Universitari Vall d'Hebron, Barcelona, Spain; [6]Clinic of Optic Neuritis and Clinic of Multiple Sclerosis, Department of Neurology, Rigshospitalet - Glostrup, University of Copenhagen, Denmark; [7]Department of Neurology, Medical University of Graz, Graz, Austria; [8]Department of Radiology and Nuclear Medicine, MS Center Amsterdam, VU University Medical Center, Amsterdam, The Netherlands; [9]Department of Neurosciences, San Camillo Forlanini Hospital, Rome, Italy; [10]Unitat de Neuroimmunologia Clinica, CEM-Cat, Hospital Universitari Vall d'Hebron, Barcelona, Spain; [11]Functional Imaging Unit, Dept. of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet, Copenhagen, Denmark; [12]Department of Neurology, MS Center Amsterdam, VU University Medical Center, Amsterdam, The Netherlands.

Running title: Validation of MAGNIMS 2016 criteria for MS diagnosis.

Key words: MRI, clinically isolated syndrome, multiple sclerosis, diagnostic criteria.

Character count - title: 166. **Character count - running title**: 52.
Word count - abstract: 386. **Word count - text**: 3525.
Number of tables: 3. **Number of figures**: 2.
Number of references: 30.
Supplemental Data: Supplementary Methods, Supplementary Table 1 and Table 2, Supplementary Figure 1 and Figure 2.

Correspondence should be addressed to: Prof. Massimo Filippi, Neuroimaging Research Unit,

Institute of Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute,

Vita-Salute San Raffaele University, Via Olgettina, 60, 20132 Milan, Italy. Telephone number:

#39-02-26433033; Fax number: #39-02-26433031; E-mail address: filippi.massimo@hsr.it

Abstract

Background. In 2016, the MAGNIMS network proposed modifications to the MRI criteria to define dissemination in space (DIS) and time (DIT) for the diagnosis of multiple sclerosis (MS) in clinically isolated syndrome (CIS) patients. Changes to DIS definition included removal of distinction between symptomatic and asymptomatic lesions; increasing to three the number of lesions needed to define periventricular (PV) involvement; combining cortical (CL) and juxtacortical lesions (JC); and adding the optic nerve (ON). For DIT, removing the distinction between symptomatic and asymptomatic lesions was suggested.

We compared the performance of the McDonald 2010 and MAGNIMS 2016 criteria for MS diagnosis in a large multicentre cohort of CIS patients to provide evidence to guide future revisions of MS diagnostic criteria.

Methods. Between 1995 and 2017, brain and spinal cord MRI and optic nerve assessments were collected from 368 patients with typical CIS suggestive of MS studied less than 3 months from clinical onset in eight European centres. Occurrence of a second clinical attack (clinically definite -CD- MS) was recorded during the follow-up. MRI criteria performance for DIS, DIT and DIS plus DIT was evaluated using a time-dependent receiver operating characteristic curve analysis, with CDMS at months (M) 36 and 60 as the outcome.

Findings. At the last evaluation (median=50.0 months, IQR=27.0-78.4), 189/368 (51%) patients developed CDMS. At M36, both DIS criteria showed high sensitivity (McDonald 2010=0.91, 95% CI=0.85-0.94; MAGNIMS 2016=0.93, 0.88-0.96), similar specificity (0.33, 95% CI 0.25-0.42; 0.32, 0.24-0.41), and similar area under the curve (AUC; 0.62, 95% CI 0.57-0.67; 0.63, 0.58-0.67). Performance of criteria was not affected by inclusion of symptomatic lesions (sensitivity 0.92, 0.87-0.96; specificity 0.31, 0.23-0.40; AUC 0.62, 0.57-0.66) or CL (sensitivity 0.92, 0.87-0.95; specificity 0.32, 0.24-0.41; AUC 0.62, 0.57-0.67). Our findings suggested that requiring three PV lesions might result in slightly lower sensitivity (0.85, 0.78-0.90), slightly higher specificity (0.40, 0.32-0.50) and similar AUC (0.63, 0.57-0.68). Our findings suggested that

inclusion of ON evaluation might result in similar sensitivity (0.92. 0.87-0.96) and slightly lower specificity (0.26, 0.18-0.34) and AUC (0.59, 0.55-0.64). AUCs for DIT (McDonald 2010=0.61, 0.55-0.67; MAGNIMS 2016=0.61, 0.55-0.66) and for DIS plus DIT (McDonald 2010=0.62, 0.56-0.67; MAGNIMS 2016=0.64, 0.58-0.69) were also similar.

Interpretation. For CDMS development, MAGNIMS 2016 criteria have similar accuracy to McDonald 2010 criteria. Inclusion of symptomatic lesions is expected to simplify the clinical use of MRI criteria without reducing accuracy, while our findings suggest that needing three lesions to define PV involvement might slightly increase specificity. Future revisions of MS diagnostic criteria should consider these two factors.

Research in context

Evidence before this study

After the proposal of the MAGNIMS 2016 criteria, we searched PubMed for original research articles and reviews evaluating MRI criteria for the diagnosis of multiple sclerosis (MS) published from January 1, 1979 to October 31, 2017, with the terms "MRI", "clinically isolated syndrome", "multiple sclerosis", "second attack", "relapse", "conversion", "McDonald criteria", "diagnosis", "differential diagnosis", "white matter", "lesions", "symptomatic lesions", "periventricular", "brainstem", "cerebellar", "cortical lesions", "brain", "spinal cord", "optic nerve", "visual evoked potentials", "disease dissemination in space", "disease dissemination in time", and "high field". Only papers published in English were considered. We identified several studies which evaluated different MRI criteria, even though they were often limited by their small sample size, their retrospective design, their short follow-up, the presence of confounding factors (e.g., selection bias, treatment effects, etc.) and the fact that they assessed only single items of the criteria recently proposed.

Added value of this study

By evaluating a large multicentre cohort of patients with a clinically isolated syndrome (CIS) suggestive of MS, we compared the MAGNIMS 2016 criteria and the McDonald 2010 criteria for the development of clinically definite MS. Overall, our results suggest that the MAGNIMS 2016 criteria give similar performance compared to the McDonald 2010 criteria.

To our knowledge, this is the first complete study which evaluates the performance of the recently proposed MAGNIMS 2016 criteria in a large multicentre cohort of CIS patients. In addition to the global performance of the MAGNIMS 2016 criteria, we also investigated how the introduction of each of their proposed modifications into the McDonald 2010 criteria affected the performance.

Implications of all the available evidence

Accurate criteria in the diagnostic work-up of CIS patients are of critical importance not only to enable an early diagnosis of MS, thus allowing treatment to start sooner, but also to minimise the risk of misdiagnosis and overdiagnosis.

This study provides relevant pieces of information regarding the rationale for and performance of the MAGNIMS 2016 criteria. MAGNIMS 2016 criteria showed a similar overall diagnostic accuracy compared to McDonald 2010 criteria, and gave higher adjusted hazard ratio(aHR) for the development of clinically definite MS.

From the analyses of each proposed modification to the MRI criteria of the McDonald 2010 criteria, our results suggest that inclusion of symptomatic lesions and three lesions to define periventricular involvement could be considered for future revision of MS diagnostic criteria.

Introduction

In 2001, an international panel of multiple sclerosis (MS) specialists formally included MRI in the MS diagnostic criteria to provide objective evidence for disease dissemination in space (DIS) and time (DIT) and to exclude alternative diagnoses.[1] Since their introduction, these criteria have been modified to simplify them, to clarify specific aspects (e.g., inclusion of spinal cord [SC] MRI findings) and to enable earlier diagnosis of MS,[2-4] thus allowing treatment to start sooner. Since the last revision of these criteria (the McDonald 2010 criteria),[3] new data about the application of MRI for MS diagnosis have become available, and there have been improvements in MRI technology. This led members of the MAGNIMS network to organise an international workshop in 2015, to review recent imaging findings and to establish their contribution to MS diagnosis, proposing evidence-based modifications to the MRI diagnostic criteria. This, together with expert consensus, resulted in the MAGNIMS 2016 criteria.[5]

Modifications to the DIS criteria[5] were: 1) the removal of any distinction between symptomatic and asymptomatic lesions;[6, 7] 2) increasing from 1 to 3 in the number of lesions needed to establish periventricular (PV) involvement;[8-13] 3) combining cortical lesions (CL) (confined to the cortex without involving the underlying subcortical white matter [WM]) and juxtacortical lesions (JC) (contiguous to the gray matter [GM], and mixed WM/GM) to expand the concept of JC involvement;[14] and 4) including the optic nerve (ON) as an additional location for DIS definition.[15-17] Removing the distinction between symptomatic and asymptomatic lesions and inclusion of any new/active lesion were suggested for DIT criteria.

This study compares MAGNIMS 2016[5] and the McDonald 2010[3] criteria in a large multicentre cohort of patients with a clinically isolated syndrome (CIS) suggestive of MS, collected within the MAGNIMS network to evaluate their performance in the diagnosis of clinically definite (CD) MS. Each individual modification from the MAGNIMS criteria[5] was also assessed to investigate its influence on the diagnostic performance and thus its potential contribution to future modifications of MS diagnostic criteria.

Methods

Ethics committee approval. Approval was received from the local ethical standards committee and written informed consent was obtained from all subjects at the time of data acquisition. Because the study was based on existing data, with no additional burden on the patient, no further approval from the institutional review board was needed.

Patients. This project was run within the European MAGNIMS network (http://www.magnims.eu) and involved eight highly specialised MS centres with expertise in neuroimaging (see Supplementary Methods). The study design was similar to that of previous studies aimed at assessing the performance of MS diagnostic criteria.[14, 18, 19] Centres were asked to identify CIS patients recruited into prospective MRI and clinical follow-up (FU) studies from June 16, 1995 to January 27, 2017 with: (a) age between 16 and 60 years; (b) a first CIS suggestive of CNS demyelination;[20] (c) a clinical presentation typical for a relapsing-remitting MS;[21] (d) a complete neurological examination; (e) a baseline brain and SC MRI scan obtained <3 months from the clinical onset; and (f) a FU brain scan obtained <12 months from CIS onset. Development of CDMS was defined as the occurrence of a second clinical event attributable to demyelination lasting more than 24 hours and after an interval of at least one month from the first attack, with evidence of two separate lesions. Time to CDMS was calculated as the interval between the onset of the first and second events.

Appropriate investigations were carried out as necessary to exclude alternative diagnoses, the presence of comorbidities (psychiatric or other neurological disorders), and previous clinical events.

The following information was also collected: age at CIS onset, sex, type of onset, EDSS, presence of oligoclonal bands (OCB) (if cerebrospinal fluid examination had been performed),date of initiation and types of disease-modifying treatment (if applicable).

MRI analysis. Details of the MRI acquisition are reported in the Supplementary Methods. All images were assessed by consensus by two experienced observers (PP and MAR), blinded to the

patients' identity and MS status at the Neuroimaging Research Unit (Milan, Italy). Brain WM lesions were identified on DE/FLAIR images and were defined as hyperintensities involving at least 3 voxels, present on at least two slices and visible on two different sequences (e.g., FLAIR and T2 or proton density and T2). Total number of WM lesions, number of PV (in direct contact with the ventricular system), JC, and posterior fossa (PF) lesions (located in the brainstem and/or cerebellum) were evaluated. CL and JC lesions identified from double inversion recovery (DIR) (when available) and T2/FLAIR sequences, where combined to expand the definition of JC involvement. Gd-enhancing lesions were identified on post-contrast T1-weighted scans. Hyperintense SC lesions were counted.

To evaluate the effects of symptomatic lesions, if a subject had a brainstem or SC syndrome, we counted lesions both including and excluding those present in symptomatic regions.

ON involvement was evaluated using MRI (presence of a ON lesion) and/or visual evoked potentials (VEP) (presence of an increased latency and/or a decreased amplitude and/or an asymmetry between the eyes).

From the FU MRI scans, the numbers of new T2-hyperintense and Gd-enhancing lesions were quantified.

DIS and DIT criteria. On baseline MRI scans, the following DIS criteria were assessed (see Supplementary Table 1): 1) McDonald 2010 criteria;[3] 2) MAGNIMS 2016 criteria;[5] 3) modified DIS criteria 1: McDonald 2010 criteria modified to include symptomatic lesions; 4) modified DIS criteria 2: McDonald 2010 criteria modified to change to 3 the minimum number of lesions necessary to define PV involvement; 5) modified DIS criteria 3: McDonald 2010 criteria modified to combine CL and JC; 6) modified DIS criteria 4: McDonald 2010 criteria modified to include ON involvement as an additional location for the definition of DIS.

On baseline and FU MRI, DIT was defined according to 1) the McDonald 2010 criteria[3] and 2) the MAGNIMS 2016 criteria[5] (Supplementary Table 1).

The fulfilment of DIS plus DIT criteria for all DIS criteria was also assessed.

Statistical analysis. Cumulative/dynamic time-dependent receiver operating characteristic (ROC) curve analysis[22] for censored survival data was applied to assess the performance of the MRI criteria for DIS, DIT and DIS plus DIT, using the clinical status (CDMS or CIS) over time as outcome. Sensitivity, specificity, accuracy, positive and negative predictive values at months 36 (M36) and 60 (M60) were calculated. Bias-corrected and accelerated bootstrap method[23] was used to estimate 95% confidence intervals (CIs).

The cumulative risk of CDMS development from the first clinical event up to the last available FU was represented using Kaplan-Meier survival curves (patients censored according to their FU). Extended Cox regression models using time to CDMS as the outcome and adjusted for age (continuous), sex (binary), treatment (binary, time-dependent, i.e., treatment effects were modeled considering the time when a patient started any treatment), disease onset type (optic neuritis *vs* others), and presence of OCB (binary) were performed to obtain adjusted hazard ratios (aHRs) and 95% CIs. A shared gamma-frailty term was also included to address centre effects, accounting for unobserved heterogeneity and statistical dependence between clustered time-to-event data.[24] Based on a bootstrap resampling technique, a test for the hypothesis of no difference between aHRs for McDonald 2010 criteria and other MRI criteria was performed. Similar models including a specific interaction term were estimated to explore possible interaction between MRI criteria and treatment, type of onset or the presence of OCB.

Role of the funding source. The funding sources had no role in study design, collection, analysis, and interpretation of data; writing of the report; or the decision to submit the paper for publication. The corresponding author had full access to all data in the study and had final responsibility for the decision to submit the paper for publication.

Results

From an initial selection of 571 CIS patients, the final cohort comprised 368 CIS patients fulfilling all study inclusion criteria (see Figure 1 for study flow chart).

Table 1 summarises the main baseline demographic, clinical and MRI findings of the CIS patients included.

For the final cohort, MRI had been obtained at 1.0 Tesla (15 [4%] patients), 1.5 Tesla (243 [66%] patients) or 3.0 Tesla (110 [30%] patients).

Assessment of ON involvement was performed in 241 (65%) CIS patients: 216 with a monofocal onset (97 optic neuritis, 42 brainstem/cerebellar syndrome, 57 SC syndrome and 20 hemispheric) and 25 with multifocal onset. This assessment was based on VEP in 219/241 (91%) patients, ON MRI in 3/241 (1%) patients and both VEP and ON MRI in 19/241 (8%) patients.

At baseline, 41 (11%) CIS patients showed no lesions on brain MRI (6 of them presented lesions at SC MRI), while 195 (53%) had no SC lesions.

One-hundred and twenty-eight (35%) patients had a brainstem/cerebellar or SC syndrome at onset, and at least one T2-hyperintense lesion in the symptomatic region. ON involvement was found in 130/241 (54%) patients.

At the last evaluation (median=50.0 months; IQR27.0-78.4 months), 189 (51%) CIS patients had a second clinical episode (median time to conversion=10.9 months, IQR=5.0-24.6) and 184 (50%) developed new T2 or Gd-enhancing lesions. During the FU, 157 (43%) patients started disease-modifying drugs and 50 (32%) of them did not develop CDMS.

The performance of the different combined diagnostic criteria is shown in Table 2, while Figure 2a and Supplementary Figure 1a show the area under the curve (AUC) over time, up to 10 years, from disease onset of the McDonald 2010 and MAGNIMS 2016 criteria according to the development of CDMS. The main results from these analyses at M36 are:

1) considering DIS alone, both the McDonald 2010 and MAGNIMS 2016 criteria had high sensitivity (0.91 and 0.93), similar specificity (0.33 and 0.32) and similar AUC (0.62 and 0.63);

2) all modified DIS criteria showed similar performance compared to the McDonald 2010 criteria (sensitivity ranging from 0.85 to 0.92, specificity from 0.26 to 0.40, AUC from 0.59 to

0.63). Our results suggest that the use of 3 PV lesions resulted in slightly lower sensitivity (0.85) and slightly higher specificity (0.40), while ON evaluation had the lowest specificity (0.26);

3) the two DIT criteria investigated showed similar performance (sensitivity=0.78 and 0.80, specificity=0.44 and 0.42, AUC=0.61 for both);

4) for both McDonald 2010 and MAGNIMS 2016 criteria, fulfilling DIS plus DIT gave similar sensitivity (0.73 and 0.77), specificity (0.50 for both) and AUC (0.62 and 0.64);

5) DIS plus DIT with all modified DIS criteria gave similar performance compared to the McDonald 2010 criteria (sensitivity ranging from 0.70 to 0.76, specificity from 0.48 to 0.55, AUC from 0.61 to 0.63).

Similar results were found considering CDMS at M60 as the outcome (Table2, Figure 2a and Supplementary Figure 1a). Given the high percentage of patients with optic neuritis at onset, the performance of the McDonald 2010 and MAGNIMS 2016 criteria according to type of onset (optic neuritis *vs* other types of onset) was also investigated (Supplementary Table 2 and Supplementary Figure 2). Both sets of criteria were more sensitive (CIS with optic neuritis: McDonald 2010=0.87, 95% CI=0.77-0.94 and MAGNIMS 2016=0.90, 95% CI=0.80-0.95; CIS with other types of onset McDonald 2010=0.93, 95% CI=0.86-0.97 and MAGNIMS 2016=0.95, 95% CI=0.89-0.99) and less specific (CIS with optic neuritis: McDonald 2010=0.44, 95% CI=0.32-0.56 and MAGNIMS 2016=0.39, 95% CI=0.27-0.52; CIS with other types of onset McDonald 2010=0.21, 95% CI=0.12-0.33 and MAGNIMS 2016=0.25, 95% CI=0.14-0.37) in patients without optic neuritis, but this difference was less pronounced for the MAGNIMS 2016 criteria.

Figure 2b and Supplementary Figure 1b show the Kaplan-Meier curves used to estimate the cumulative risk of CDMS development using DIS only, DIT only or DIS plus DIT from the McDonald 2010 and MAGNIMS 2016 criteria, while the aHRs from the extended Cox regression models are shown in Table 3. The cumulative risk of CDMS development was similar for the McDonald 2010 and MAGNIMS 2016 criteria. A lack of fulfilment of the MAGNIMS 2016 criteria

was associated with a higher conversion-free survival than was the case for the McDonald 2010 criteria (Figure 2b).

aHRs were higher for the MAGNIMS 2016 compared to the McDonald 2010 criteria,(aHR=4.43 and 3.48, for DIS only; aHR=2.95 and 2.52, for DIS plus DIT, p<0.0001 for all), but not significantly different (p=0.12 for DIS only; p=0.08 for DIS plus DIT) (Table 3). Similar significant aHRs, not different from those of the McDonald 2010 criteria, were found for the other modified criteria evaluated (aHRs ranging from 3.13 to 3.66 for DIS only and from 2.54 to 2.60 for DIS plus DIT, p<0.0001 for all) (Table 3).

The aHRs of the criteria were not affected by disease onset type, presence of OCB or treatment (Supplementary Table 3).

Discussion

Thanks to improvements in MRI technology and its increasing application in MS, new evidence for the use of MRI in the diagnostic work-up of patients with a suspicion of MS has emerged since the publication of the McDonald 2010 criteria,[3] and was considered in the MAGNIMS 2016 criteria.[5] Since their publication, different aspects proposed in the MAGNIMS 2016 consensus guidelines have been examined. A recent study of 170 CIS patients[25] compared the performance of the McDonald 2010 and MAGNIMS 2016 criteria for DIS only, considering the clinical status at heterogeneous FU durations (at least two years) as the outcome. Additionally, SC and ON evaluations were performed only in those patients with clinical involvement of these regions. As a consequence, a proper validation of these criteria has so far been lacking.

By evaluating a large multicentre cohort of patients experiencing a typical CIS, we found that the performance of the McDonald 2010 and MAGNIMS 2016 criteria was similar considering CDMS conversion at M36 and M60. Specifically, both criteria showed high sensitivity and accuracy and similar specificity. For both sets of criteria, specificity was lower than that of previous studies that evaluated the diagnostic performance of the McDonald 2010 criteria.[6, 7, 13, 19] Several

factors may help to explain the current findings, including the different FU durations,[6, 13] the methods used for the statistical analysis (using a time-to-event analysis in the current study), and the influence of treatment, which may have delayed or prevented the occurrence of the second attack during the study period.

Survival probability analyses confirmed that the two sets of criteria had similar performance. Compared to the McDonald 2010, the MAGNIMS 2016 DIS criteria had a higher aHR, although not significantly different, which agrees with the higher conversion-free survival in CIS patients not fulfilling these criteria and the suggestion of slightly higher negative predictive values (0.79, 95% CI 0.67-0.87, *vs* 0.83, 0.71-0.91, at M36).

Quite similar results were obtained when the combination of DIS plus DIT for the two sets of criteria were tested.

In addition to the global performance of the MAGNIMS 2016 criteria, we also investigated how the introduction of each of their proposed modifications into the McDonald 2010 criteria changed the final performance. The purpose of this approach was to see whether the overall specificity of the McDonald 2010 criteria could be increased, thus minimising overdiagnosis.[26]

Consistent with some recent studies,[6, 7] our study provides further support for the inclusion of lesions in the symptomatic region in CIS patients with a brainstem or SC onset, since this did not affect the performance of DIS and DIT diagnostic criteria. The definition of what is symptomatic may be extremely challenging, and thus the removal of the distinction between symptomatic and asymptomatic lesions is likely to simplify the clinical implementation of these criteria. Moreover, since CIS patients with a single symptomatic lesion have a high risk of CDMS conversion,[6, 7] the inclusion of these lesions may have clinical relevance in predicting a diagnosis of MS in these patients.

According to the MAGNIMS 2016 criteria, a single PV lesion is not deemed sufficient to indicate involvement of the PV region. Indeed, PV lesions have been described in many other neurological conditions, which often enter the differential diagnosis of MS, including migraine,[9, 10]

cerebrovascular diseases and other inflammatory disorders.[12] Moreover, a threshold of three or more PV lesions was found to be the most accurate for predicting CDMS conversion,[8] and the presence of at least 3 PV lesions predicted CDMS in large cohorts of CIS patients,[11] and in a CIS cohort with SC onset. In a recent study, 3 PV lesions improved specificity in older CIS patients.[27] In line with this evidence and with that of a recent 15-year FU study,[13] our results suggested that the use of 3 lesions to define PV involvement slightly reduced sensitivity (0.85 *vs* 0.91 at M36), but slightly increased specificity (0.40 *vs* 0.33 at M36), without affecting diagnostic accuracy. Given the increased risk of misdiagnosis due to an oversimplification of MS diagnostic criteria, there is the need to consider features which could be more distinctive for this condition compared to other diseases.[26] Our findings suggest that this criterion might improve the specificity of the MRI diagnostic criteria, reducing misdiagnosis and also representing a possible prognostic factor.

CLs are a distinctive feature of MS patients from the beginning of the disease,[14] and have not been detected in patients with other neurological conditions, including neuromyelitis optica and migraine. As a consequence, including CL assessment in the diagnostic algorithm of CIS patients, is likely to increase specificity, as suggested by a single-centre study[14] and recently confirmed by a multi-centre investigation.[28] Despite this, acquiring MR images that depict CLs is technically challenging, and there are no standardised sequences for this task. Consequently, MAGNIMS guidelines[5] recommended using the combined term CL/JC to expand the concept of JC involvement. In our analysis, inclusion of CLs evaluation in the subgroup of CIS patients (45%) with DIR acquisition did not significantly influence DIS criteria performance, and only four additional patients fulfilled the DIS criteria.

The effect of the inclusion of the ON as an additional site for DIS fulfillment was also investigated in those patients who had VEP or optic nerve MRI evaluations. Clearly, since this study was not pre-planned to include proper investigation of this region, the data were available for only about half of the patients, with a relatively heterogeneous assessment of neurophysiology, MRI, or both. Optic neuritis is the first manifestation of MS in up to 30% of CIS patients,[20] who

have generally been considered to have a lower risk of developing MS. Despite this, recent evidence has demonstrated that the concomitant presence of WM lesions,[15, 17] early new lesion formation[16] or cerebrospinal fluid OCB[17] increases the risk of developing MS in patients with this type of presentation. Our findings suggested that the inclusion of ON assessment in the definition of DIS might slightly decrease specificity (0.26 *vs* 0.33 at M36), which is expected whenever additional criteria are included. An assessment of whether the combination of ON involvement with different DIS criteria (e.g., presence of 3 PV lesions) or a change in the number of sites needed to define DIS (e.g., from 2 to 3) modifies performance of diagnostic criteria was beyond the scope of the present investigation. Consequently, studies aimed at validating MRI and neurophysiological measures of ON involvement in fulfilling diagnostic criteria in support of MS should be undertaken.

Given the high percentage of patients with optic neuritis, we also explored whether this type of presentation gives different performance of the criteria. Both sets of criteria were more sensitive and less specific in patients without than those with optic neuritis. However, the difference of performance was less pronounced for MAGNIMS 2016 criteria.

This study has some limitations. First, our CIS patients were recruited in highly specialised centres, possibly resulting in the selection of patients at higher risk of conversion to CDMS (51%), as also suggested by their high median lesion number (median=11). Moreover, our cohort included only patients with a typical syndrome suggestive of MS and without comorbidities, and thus the performance of MRI criteria in CIS patients with atypical features, with other diseases mimicking MS or with other concurrent diseases was not evaluated. However, it should be borne in mind that this is the type of patients for whom the criteria have been proposed and that application of MS criteria outside this typical scenario is discouraged by the International Panel.[1-4]

Secondly, because we needed a long enough FU for the analysis, the data were collected retrospectively. However, all patients included are part of ongoing studies run by the participating centres, and are therefore representative of the current population of CIS patients in Europe, including the type of presentation and policies for early treatment initiation (which, however,

showed no effect in the interaction analysis). As a consequence, we believe that our results are generalizable to CIS patients attending the majority of MS clinics. Thirdly, due to the multicentre setting, MR images were acquired using different field strengths (from 1.0 to 3.0 Tesla) and MRI parameters (e.g. slice thickness). This allowed us to evaluate the MRI criteria in a situation close to a clinical setting. Although high-field MRI allows the detection of higher number of WM lesions in CIS and MS patients,[29] it has been demonstrated that field strength does not influence significantly fulfilment of DIS and DIT criteria.[30] Nevertheless, to take all these differences into account, survival analyses were performed after correcting for centre.

Overall, this study provides important pieces of information regarding the application of the recently proposed MAGNIMS 2016 criteria for MS diagnosis. Our findings suggest that the MAGNIMS 2016 criteria perform similarly to the McDonald 2010 criteria. Among the different modifications proposed, our results support removal of the distinction between symptomatic and asymptomatic lesions, which simplifies the clinical use of MRI criteria, and suggests that further consideration of increasing to 3 the number of lesions needed to define PV involvement is warranted, as this seemed to slightly increase specificity. Further effort is still necessary to improve CL assessment and more studies should be performed to better evaluate the influence of including ON assessment as an additional DIS criterion.

Author contributions
MF study concept, and analysis and interpretation of the data. He also acted as study supervisor.
PP patient enrollment and analysis and interpretation of the data.
AM statistical analysis and interpretation of the data.
OC patient enrollment and analysis of the data.
SM patient enrollment and analysis of the data.
AR patient enrollment and analysis of the data.
JF patient enrollment and analysis of the data.
CE patient enrollment and analysis of the data.
FB patient enrollment and analysis of the data.
CG patient enrollment and analysis of the data.
WB patient enrollment and analysis of the data.
JD patient enrollment and analysis of the data.
XM patient enrollment and analysis of the data.
SPC patient enrollment and analysis of the data.
AP patient enrollment and analysis of the data.
MH patient enrollment and analysis of the data.
SR patient enrollment and analysis of the data.
VM patient enrollment and analysis of the data.
KM patient enrollment and analysis of the data.
MT patient enrollment and analysis of the data.
GC contributed to study concept.
ID patient enrollment and analysis of the data.
BU patient enrollment and analysis of the data.
IDB patient enrollment and analysis of the data.
MAR study concept, and analysis and interpretation of the data.
All the authors contributed to drafting/revising the manuscript and gave their approval to the current version of the manuscript.

Declaration of interests
MF is Editor-in-Chief of the Journal of Neurology; serves on a scientific advisory board for Teva Pharmaceutical Industries; has received compensation for consulting services and/or speaking activities from Biogen Idec, Merk-Serono, Novartis, and Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Teva Pharmaceutical Industries, Novartis, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla (FISM), Cure PSP, Alzheimer's Drug Discovery Foundation (ADDF), the Jacques and Gloria Gossweiler Foundation (Switzerland), and ARiSLA (Fondazione Italiana di Ricerca per la SLA).
PP received speakers honoraria from Biogen Idec, Novartis and ExceMED.
AM, CG, WB, JD, AP, MH, SR, KM and IDB have nothing to disclose.
OC serves as consultant for Novartis, Biogen Idec, Genzyme, and TEVA.
SM has received funding for travels from Merck, Bayer Schering, Medis and Genzyme Sanofi, and honoraria for speaking from Merck Serono and Novartis. She also received research grant support from the Ministry of Education and Science, Republic of Serbia (project no. 175031).
AR reports personal fees from Novartis, Sanofi-Genzyme, OLEA Medical, Bayer, Bracco, Merck-Serono, Teva Pharmaceutical Industries Ltd, Roche and Biogen Idec, grants from Siemens AG and Icometrix, outside the submitted work..
JF reports personal fees from Almirall, Biogen, Genzyme, Merck Serono, Novartis, Sanofi-Aventis, Teva, and Takeda , outside of the submitted work.
CE reports personal fees from Biogen, Bayer Schering, Merck Serono, Novartis, Shire, Genzyme and Teva Pharmaceutical Industries Ltd./sanofi-aventis, and Roche , grants from Merck Serono, Biogen, and Teva Pharmaceutical Industries Ltd./sanofi-aventis, outside the submitted work.

Acknowledgments

References

1.   McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001;**50**:121-7.
2.   Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* 2005;**58**:840-6.
3.   Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011;**69**:292-302.
4.   Montalban X, Tintore M, Swanton J, Barkhof F, Fazekas F, Filippi M, et al. MRI criteria for MS in patients with clinically isolated syndromes. *Neurology* 2010;**74**:427-34.
5.   Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *The Lancet Neurology* 2016;**15**:292-303.
6.   Brownlee WJ, Swanton JK, Miszkiel KA, Miller DH, Ciccarelli O. Should the symptomatic region be included in dissemination in space in MRI criteria for MS? *Neurology* 2016;**87**:680-3.
7.   Tintore M, Otero-Romero S, Rio J, Arrambide G, Pujal B, Tur C, et al. Contribution of the symptomatic lesion in establishing MS diagnosis and prognosis. *Neurology* 2016;**87**:1368-74.
8.   Barkhof F, Filippi M, Miller DH, Scheltens P, Campi A, Polman CH, et al. Comparison of MRI criteria at first presentation to predict conversion to clinically definite multiple sclerosis. *Brain* 1997;**120 (Pt 11)**:2059-69.
9.   Absinta M, Rocca MA, Colombo B, Copetti M, De Feo D, Falini A, et al. Patients with migraine do not have MRI-visible cortical lesions. *Journal of neurology* 2012;**259**:2695-8.
10.   Liu S, Kullnat J, Bourdette D, Simon J, Kraemer DF, Murchison C, et al. Prevalence of brain magnetic resonance imaging meeting Barkhof and McDonald criteria for dissemination in space among headache patients. *Multiple sclerosis* 2013;**19**:1101-5.
11.   Ruet A, Arrambide G, Brochet B, Auger C, Simon E, Rovira A, et al. Early predictors of multiple sclerosis after a typical clinically isolated syndrome. *Multiple sclerosis* 2014;**20**:1721-6.
12.   Kim SS, Richman DP, Johnson WO, Hald JK, Agius MA. Limited utility of current MRI criteria for distinguishing multiple sclerosis from common mimickers: primary and secondary CNS vasculitis, lupus and Sjogren's syndrome. *Multiple sclerosis* 2014;**20**:57-63.
13.   Brownlee WJ, Miszkiel KA, Altmann DR, Ciccarelli O, Miller DH. Periventricular lesions and MS diagnostic criteria in young adults with typical clinically isolated syndromes. *Multiple sclerosis* 2016:1352458516667565.
14.   Filippi M, Rocca MA, Calabrese M, Sormani MP, Rinaldi F, Perini P, et al. Intracortical lesions: relevance for new MRI diagnostic criteria for multiple sclerosis. *Neurology* 2010;**75**:1988-94.
15.   Optic Neuritis Study G. Multiple sclerosis risk after optic neuritis: final optic neuritis treatment trial follow-up. *Archives of neurology* 2008;**65**:727-32.
16.   Swanton JK, Fernando KT, Dalton CM, Miszkiel KA, Altmann DR, Plant GT, et al. Early MRI in optic neuritis: the risk for clinically definite multiple sclerosis. *Multiple sclerosis* 2010;**16**:156-65.
17.   Tintore M, Rovira A, Rio J, Otero-Romero S, Arrambide G, Tur C, et al. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. *Brain* 2015;**138**:1863-74.
18.   Rovira A, Swanton J, Tintore M, Huerga E, Barkhof F, Filippi M, et al. A single, early magnetic resonance imaging study in the diagnosis of multiple sclerosis. *Archives of neurology* 2009;**66**:587-92.

19.   Swanton JK, Rovira A, Tintore M, Altmann DR, Barkhof F, Filippi M, et al. MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study. *The Lancet Neurology* 2007;**6**:677-86.

20.   Miller D, Barkhof F, Montalban X, Thompson A, Filippi M. Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. *The Lancet Neurology* 2005;**4**:281-8.

21.   Brownlee WJ, Hardy TA, Fazekas F, Miller DH. Diagnosis of multiple sclerosis: progress and challenges. *Lancet* 2017;**389**:1336-46.

22.   Hung H, Chiang C. Estimation models for time-dependent AUC with survival data. *Canadian Journal of Statistics* 2010;**38**:8-26.

23.   Efron B. Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association* 1987;**82**:171-85.

24.   Glidden DV, Vittinghoff E. Modelling clustered survival data from multicentre clinical trials. *Stat Med* 2004;**23**:369-88.

25.   Hyun JW, Huh SY, Kim W, Park MS, Ahn SW, Cho JY, et al. Evaluation of 2016 MAGNIMS MRI criteria for dissemination in space in patients with a clinically isolated syndrome. *Multiple sclerosis* 2017:1352458517706744.

26.   Solomon AJ, Bourdette DN, Cross AH, Applebee A, Skidd PM, Howard DB, et al. The contemporary spectrum of multiple sclerosis misdiagnosis: A multicenter study. *Neurology* 2016;**87**:1393-9.

27.   Arrambide G, Tintore M, Auger C, Río J, Castilló J, Vidal-Jordana A, et al. Lesion topographies in multiple sclerosis diagnosis: a reappraisal. *Neurology* 2017, in press.

28.   Preziosa P, Rocca MA, Mesaros S, Meani A, Montalban X, Drulovic J, et al. Diagnosis of multiple sclerosis: a multicenter study to compare revised McDonald-2010 and Filippi-2010 criteria. *J Neurol Neurosurg Psychiatry* 2017; July 19) [epub ahead of print July 19, 2017; doi: 10.1136/jnnp-2017-315863].

29.   Wattjes MP, Lutterbey GG, Harzheim M, Gieseke J, Traber F, Klotz L, et al. Higher sensitivity in the detection of inflammatory brain lesions in patients with clinically isolated syndromes suggestive of multiple sclerosis using high field MRI: an intraindividual comparison of 1.5 T with 3.0 T. *European radiology* 2006;**16**:2067-73.

30.   Wattjes MP, Harzheim M, Lutterbey GG, Hojati F, Simon B, Schmidt S, et al. Does high field MRI allow an earlier diagnosis of multiple sclerosis? *Journal of neurology* 2008;**255**:1159-63.

Table 1. Main baseline demographic, clinical, and MRI findings from the final cohort of clinically isolated syndrome (CIS) patients.

	All cases (n=368)
Demographic details	
Number (%) of	
• Men	126 (34%)
• Women	242 (66%)
Median age at onset (IQR) [years]	32.5 (26.0-39.2)
Clinical details	
Median disease duration at baseline MRI (IQR) [months]	1.8 (0.8-2.8)
Median EDSS at baseline (IQR)	1.5 (1.0-2.0)
Clinical presenting symptom(s) (%): Monofocal	340 (92%)
• Optic neuritis	• 169/340 (50%)
• Brainstem/cerebellar syndrome	• 61/340 (18.0%)
• Spinal cord syndrome	• 79/340 (23%)
• Hemispheric syndrome	• 31/340 (9%)
Multifocal	28 (8%)
Number (%) of patients with CSF analysis	256 (70%)
Number (%) of patients with oligoclonal bands	176/256 (69%)
Number (%) of patients receiving treatment at FU	157 (43%)
CDMS at 12 months from onset (%)	99 (27%)
CDMS at FU (%)	189 (51%)
Median time to CDMS (IQR) [months]	10.9 (5.0-24.6)
Median FU duration in not converters (IQR) [months]	36.2 (15.1-60.5)
MRI details	
Baseline number (%) of patients with lesions (brain and	333 (90%)

cord)	
Median lesion number (IQR)	11 (4-25)
Median time to FU MRI (IQR) [months]	6.4 (5.1-11.3)
MRI criteria	
Baseline number (%) of patients with ≥ 1 PV lesion	292 (79%)
Baseline number (%) of patients with ≥ 3 PV lesions	231 (63%)
Baseline number (%) of patients with ≥ 1 JC lesion	259 (70%)
Baseline number (%) of patients with ≥ 1 CL//JC	266 (72%)
Baseline number (%) of patients with ≥ 1 PF lesion	165 (45%)
Baseline number (%) of patients with ≥ 1 SC lesion	173 (47%)
Baseline number (%) of patients with symptomatic lesions*	128 (35%)
Baseline number (%) of patients with ≥ 1 Gd-enhancing lesion	150 (41%)
Number (%) of patients with ≥ 1 new T2/Gd-enhancing lesion at FU MRI	184 (50%)

*for patients with a brainstem or spinal cord syndrome.

Abbreviations: CIS=clinically isolated syndrome; MRI=magnetic resonance imaging;

IQR=interquartile range; EDSS=Expanded Disability Status Scale; CSF=cerebrospinal fluid;

FU=follow-up; CDMS=clinically definite multiple sclerosis; PV=periventricular; JC=juxtacortical;

CL=cortical lesion; PF=posterior fossa; SC=spinal cord; Gd=gadolinium. See text for further

details.

Table 2. Performance of the different combined MRI criteria for dissemination in space (DIS), dissemination in time (DIT) and DIS plus DIT for development of clinically definite multiple sclerosis (CDMS) in the final cohort (n=368).

Criteria	Timepoint	Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)	PPV (95% CI)	NPV (95% CI)
DIS only						
McDonald 2010[3]	M36	0.91 (0.85-0.94)	0.33 (0.25-0.42)	0.62 (0.57-0.67)	0.56 (0.49-0.62)	0.79 (0.67-0.87)
	M60	0.87 (0.80-0.91)	0.33 (0.21-0.46)	0.60 (0.53-0.67)	0.65 (0.57-0.72)	0.63 (0.47-0.76)
MAGNIMS 2016[5]	M36	0.93 (0.88-0.96)	0.32 (0.24-0.41)	0.63 (0.58-0.67)	0.56 (0.50-0.63)	0.83 (0.71-0.91)
	M60	0.90 (0.83-0.94)	0.34 (0.23-0.48)	0.62 (0.56-0.69)	0.66 (0.59-0.73)	0.70 (0.54-0.82)
Modified DIS criteria 1 (symptomatic)	M36	0.92 (0.87-0.96)	0.31 (0.23-0.40)	0.62 (0.57-0.66)	0.56 (0.49-0.62)	0.80 (0.69-0.89)
	M60	0.88 (0.81-0.92)	0.33 (0.22-0.46)	0.60 (0.54-0.67)	0.65 (0.58-0.72)	0.65 (0.49-0.79)
Modified DIS criteria 2 (3PV)	M36	0.85 (0.78-0.90)	0.40 (0.32-0.50)	0.63 (0.57-0.68)	0.57 (0.51-0.64)	0.74 (0.63-0.82)

	M60	0.82 (0.75-0.87)	0.41 (0.29-0.55)	0.62 (0.55-0.69)	0.66 (0.59-0.74)	0.61 (0.48-0.73)
Modified DIS criteria 3 (CL/JC)	M36	0.92 (0.87-0.95)	0.32 (0.24-0.41)	0.62 (0.57-0.67)	0.56 (0.50-0.62)	0.81 (0.69-0.89)
	M60	0.88 (0.81-0.92)	0.31 (0.20-0.44)	0.59 (0.53-0.66)	0.64 (0.57-0.71)	0.64 (0.47-0.77)
Modified DIS criteria 4 (ON)	M36	0.92 (0.87-0.96)	0.26 (0.18-0.34)	0.59 (0.55-0.64)	0.54 (0.48-0.60)	0.78 (0.64-0.88)
	M60	0.90 (0.84-0.94)	0.26 (0.16-0.38)	0.58 (0.52-0.65)	0.63 (0.56-0.70)	0.64 (0.46-0.79)
DIT only						
McDonald 2010[3]	M36	0.78 (0.71-0.84)	0.44 (0.35-0.53)	0.61 (0.55-0.67)	0.58 (0.51-0.65)	0.67 (0.57-0.76)
	M60	0.78 (0.71-0.84)	0.49 (0.37-0.62)	0.63 (0.56-0.71)	0.69 (0.61-0.77)	0.60 (0.48-0.70)
MAGNIMS 2016[5]	M36	0.80 (0.73-0.85)	0.42 (0.33-0.50)	0.61 (0.55-0.66)	0.57 (0.50-0.64)	0.68 (0.57-0.77)
	M60	0.79 (0.72-0.85)	0.46 (0.33-0.59)	0.62 (0.55-0.69)	0.68 (0.60-0.76)	0.59 (0.46-0.70)

DIS plus DIT						
McDonald 2010[3]	M36	0.73 (0.66-0.80)	0.50 (0.42-0.59)	0.62 (0.56-0.67)	0.58 (0.51-0.65)	0.67 (0.58-0.75)
	M60	0.72 (0.64-0.78)	0.52 (0.39-0.65)	0.62 (0.54-0.69)	0.68 (0.59-0.76)	0.56 (0.45-0.67)
MAGNIMS 2016[5]	M36	0.77 (0.70-0.83)	0.50 (0.41-0.59)	0.64 (0.58-0.69)	0.60 (0.52-0.67)	0.70 (0.61-0.78)
	M60	0.76 (0.68-0.82)	0.52 (0.39-0.65)	0.64 (0.57-0.71)	0.69 (0.61-0.77)	0.60 (0.49-0.70)
Modified DIS criteria 1 (symptomatic)	M36	0.76 (0.69-0.83)	0.49 (0.40-0.58)	0.62 (0.57-0.68)	0.58 (0.51-0.65)	0.68 (0.60-0.77)
	M60	0.74 (0.66-0.80)	0.50 (0.36-0.62)	0.62 (0.55-0.69)	0.68 (0.59-0.75)	0.57 (0.46-0.68)
Modified DIS criteria 2 (3PV)	M36	0.70 (0.63-0.77)	0.55 (0.46-0.63)	0.62 (0.56-0.68)	0.59 (0.51-0.67)	0.66 (0.57-0.74)
	M60	0.69 (0.61-0.76)	0.55 (0.42-0.68)	0.62 (0.54-0.69)	0.69 (0.60-0.77)	0.56 (0.45-0.66)
Modified DIS criteria 3 (CL/JC)	M36	0.75 (0.67-0.81)	0.50 (0.42-0.60)	0.63 (0.57-0.68)	0.59 (0.51-0.66)	0.68 (0.59-0.76)

	M60	0.73 (0.65-0.79)	0.52 (0.39-0.64)	0.62 (0.55-0.69)	0.68 (0.59-0.76)	0.57 (0.45-0.67)
Modified DIS criteria 4 (ON)	M36	0.75 (0.67-0.81)	0.48 (0.39-0.57)	0.61 (0.56-0.67)	0.57 (0.50-0.65)	0.67 (0.57-0.75)
	M60	0.74 (0.66-0.80)	0.50 (0.36-0.62)	0.62 (0.54-0.69)	0.68 (0.60-0.75)	0.57 (0.46-0.68)

Abbreviations: MRI=magnetic resonance imaging; CI=confidence interval; AUC=area under the curve; PPV=positive predictive value; NPV=negative predictive value; DIS=dissemination in space; DIT=dissemination in time; PV=periventricular; CL=cortical lesion; JC=juxtacortical; ON=optic nerve; M=month.

Modified DIS criteria 1: McDonald 2010 criteria modified to include symptomatic lesions; modified DIS criteria 2: McDonald 2010 criteria modified to change to 3 the minimum number of lesions necessary to define PV involvement; modified DIS criteria 3: McDonald 2010 criteria modified to combine CL and JC; modified DIS criteria 4: McDonald 2010 criteria modified to include ON involvement as an additional location for the definition of DIS.

Table 3. Adjusted hazard ratios (aHRs) with their corresponding 95% confidence intervals (CI) and a bootstrap-based comparison with the McDonald 2010 criteria. aHRs were obtained from extended Cox regression models using time to clinically definite multiple sclerosis (CDMS) as the outcome in the final cohort.

Criteria	aHR* (95% CI)	p value	p value vs McDonald 2010[3]
DIS only			
McDonald 2010[3]	3.48 (2.16-5.62)	<0.0001	-
MAGNIMS 2016[5]	4.43 (2.59-7.56)	<0.0001	0.12
Modified DIS criteria 1 (symptomatic)	3.59 (2.18-5.93)	<0.0001	0.74
Modified DIS criteria 2 (3PV)	3.13 (2.06-4.76)	<0.0001	0.51
Modified DIS criteria 3 (CL/JC)	3.66 (2.24-6.00)	<0.0001	0.51
Modified DIS criteria 4 (ON)	3.34 (1.98-5.64)	<0.0001	0.81
DIT only			
McDonald 2010[3]	2.63 (1.81-3.82)	<0.0001	-
MAGNIMS 2016[5]	2.47 (1.69-3.61)	<0.0001	0.39
DIS plus DIT			

McDonald 2010[3]	2.52 (1.78-3.58)	<0.0001	-
MAGNIMS 2016[5]	2.95 (2.04-4.26)	<0.0001	0.08
Modified DIS criteria 1 (symptomatic)	2.54 (1.77-3.65)	<0.0001	0.89
Modified DIS criteria 2 (3PV)	2.54 (1.80-3.58)	<0.0001	0.92
Modified DIS criteria 3 (CL/JC)	2.60 (1.83-3.71)	<0.0001	0.27
Modified DIS criteria 4 (ON)	2.58 (1.81-3.67)	<0.0001	0.65

Abbreviations: aHR=adjusted hazard ratios; CI=confidence interval; DIS=dissemination in space; DIT=dissemination in time; PV=periventricular; CL=cortical lesion; JC=juxtacortical; ON=optic nerve.

*=adjusted for age, sex, centre, treatment, type of onset and oligoclonal bands.

Figure legends

Figure 1. Study flow chart. See text for further details. Abbreviations: NMOSD=neuromyelitis optica spectrum disorder.

Figure 2. (a) Overall accuracy of the McDonald 2010 (red line) and MAGNIMS 2016 (blue line) criteria determined by the area under the curve (AUC) over time, up to 10 years, from disease onset, to the development of clinically definite multiple sclerosis (CDMS), considering dissemination in space (DIS) only or DIS plus dissemination in time (DIT). (b) Kaplan-Meier curves showing the survival probability estimates of not developing CDMS up to 10 years from disease onset considering DIS only or DIS plus DIT according to the McDonald 2010 and MAGNIMS 2016 criteria. See text for further details. Abbreviations: CI=confidence interval.