Abstract—Advancements in imaging technologies and newly evolving treatments offer the promise of more effective management strategies for MS. Until recently, confirmation of the diagnosis of MS has generally required the demonstration of clinical activity that is disseminated in both time and space. Nevertheless, with the advent of MRI techniques, occult disease activity can be demonstrated in 50 to 80% of patients at the time of the first clinical presentation. Prospective studies have shown that the presence of such lesions predicts future conversion to clinically definite (CD) MS. Indeed, in a young to middle-aged adult with a clinically isolated syndrome (CIS), once alternative diagnoses are excluded at baseline, the finding of three or more white matter lesions on a T2-weighted MRI scan (especially if one of these lesions is located in the periventricular region) is a very sensitive predictor (~80%) of the subsequent development of CDMS within the next 7 to 10 years. Moreover, the presence of two or more gadolinium (Gd)-enhancing lesions at baseline and the appearance of either new T2 lesions or new Gd enhancement on follow-up scans are also highly predictive of the subsequent development of CDMS in the near term. By contrast, normal results on MRI at the time of clinical presentation makes the future development of CDMS considerably less likely.

Overview. Multiple sclerosis (MS) affects up to 350,000 individuals in the United States. Different clinical courses of MS have been defined, including relapsing/remitting (RR) MS, secondary progressive (SP) MS, primary progressive (PP) MS, and progressive relapsing (PR) MS. RRMS is characterized by self-limited attacks of neurologic dysfunction followed by a variable degree of recovery. By contrast, PPMS is characterized by a steady decline in neurologic function from onset, without superimposed attacks. MS often presents at a time when the clinical extent of disease is apparently limited, although, even at this early stage of disease, substantial damage may have already occurred. Thus, approximately 50 to 80% of individuals who present with a clinically isolated syndrome (CIS) already have lesions on MRI, consistent with prior (occult) disease activity. Two randomized, controlled trials of interferon beta (IFNβ) have recently demonstrated a treatment benefit in patients with a CIS and MRI abnormalities suggestive of MS.
CSF analysis demonstrates the presence of oligoclonal bands or
Includes lesions detected by MRI or evoked potentials.

Symptoms lasting more than 24 hours would constitute an at-

Evidence of two separate lesions found on neurologic examina-

Laboratory-supported probable MS

Laboratory-supported definite MS

Clinically probable MS

Clinically definite MS

Table 1: Poser’s criteria for the diagnosis of multiple sclerosis (MS)

<table>
<thead>
<tr>
<th>Class</th>
<th>Attacks *</th>
<th>Lesions on Examination</th>
<th>CSF Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>2</td>
<td>2 or more</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>2</td>
<td>1</td>
<td>abnormal</td>
</tr>
<tr>
<td>B1</td>
<td>2</td>
<td>1 or more</td>
<td>+ abnormal</td>
</tr>
<tr>
<td>B2</td>
<td>1</td>
<td>2 or more</td>
<td>+ abnormal</td>
</tr>
<tr>
<td>B3</td>
<td>1</td>
<td>1 or more</td>
<td>+ abnormal</td>
</tr>
<tr>
<td>C1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>1</td>
<td>2 or more</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>1</td>
<td>1 or more</td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>2</td>
<td></td>
<td>+ abnormal</td>
</tr>
</tbody>
</table>

* Symptoms lasting more than 24 hours would constitute an attack even if they were “completely subjective and anamnestic.”
† Evidence of two separate lesions found on neurologic examination. Bilateral optic neuritis would constitute only one lesion provided that the episodes of optic neuritis were separated by less than 15 days.
‡ Includes lesions detected by MRI or evoked potentials.
§ CSF analysis demonstrates the presence of oligoclonal bands or an increased CNS synthesis of immunoglobulin G.

At least two distinct areas of the CNS (the so-called criteria of dissemination in time and space). Importantly, also, all of the existing diagnostic schemes require the exclusion of alternative diagnoses by appropriate laboratory and radiographic studies prior to the application of the diagnostic algorithm. As a result, the sensitivity, specificity, and diagnostic accuracy of each scheme must be considered in the context of a population of patients from which individuals with alternative diagnoses have been largely culled. The Poser criteria utilize both clinical and paraclinical information (including MRI and evoked potentials) to establish spatial dissemination (Table 1). For example, a CIS coupled with a single white matter MRI abnormality in an area unrelated to the clinical presentation establishes dissemination in space. Nevertheless, because the Poser scheme was developed at a time when MRI was in its infancy, and because white matter abnormalities are now known to be nonspecific (Table 2), there is concern that these criteria (at least insofar as they relate to the distribution of MRI lesions) may permit a diagnosis of MS in circumstances where it is unjustified. By contrast, certain other features of the Poser scheme seem overly restrictive and may, in some circumstances, prevent an appropriate diagnosis from being made. For example, there is no provision for the use of sequential paraclinical tests to establish dissemination in time, despite the fact that new lesions (presumably reflecting new disease activity) are commonly seen on follow-up MRI of patients with MS. Furthermore, these criteria have no provision for making a diagnosis of PPMS.

An international panel recently proposed new diagnostic criteria for both RRMS and PPMS. The new MRI criteria (Table 3) in some ways, are more stringent than those of Poser et al. For example, instead of a single lesion, these new criteria for spatial dissemination (MRI [a] criteria; see Table 3) can be fulfilled based on a minimum of two lesions provided that one is located juxtacortically, the other is located infratentorially, and, in addition, one of these two enhancements with gadolinium (Gd) administration (see Table 3). If no Gd-enhancing lesions are detected, a minimum of five MRI lesions in specific locations must be demonstrated (three periventricular, one juxtacortical, and one infratentorial) (see Table 3). By contrast, these new criteria are, in some ways, more permissive than those of Poser et al. Thus, MRI criteria for establishing dissemination in time and for diagnosing PPMS, topics not addressed in the Poser scheme, are proposed (see Table 1).

It is the purpose of this assessment to consider the evidence for the use of baseline and follow-up MRI in the diagnosis of patients with suspected MS. For the purpose of this assessment, the subsequent development of MS by purely clinical criteria was considered as the gold standard for comparison. Before reviewing the evidence, however, it is important to consider...
Table 3 Proposed McDonald criteria* for the diagnosis of multiple sclerosis (MS)

<table>
<thead>
<tr>
<th>Criteria for dissemination in space</th>
<th>Criteria for dissemination in time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: 2 Attacks† + 2 lesions on examination‡</td>
<td>1. ≥1 gadolinium (Gd)-enhancing lesion or ≥9 T2 lesions</td>
</tr>
<tr>
<td>2: 2 Attacks + 1 lesion on examination + [MRI (a)] or 2 MRI lesions + abnormal CSF¶</td>
<td>2. ≥1 infratentorial MRI lesion</td>
</tr>
<tr>
<td>3: 1 Attack + 2 lesions on examination + MRI (b)</td>
<td></td>
</tr>
<tr>
<td>4: 1 Attack + 1 lesion on examination + [MRI (a) or 2 MRI lesions + abnormal CSF] + MRI (b)</td>
<td></td>
</tr>
<tr>
<td>5: Progressive disease + abnormal CSF + MRI (b) + a) [≥9 T2 MRI lesions or ≥2 spinal cord lesions or 4–8 brain MRI lesions + 1 spinal cord lesion] OR b) Abnormal VEP† + [≥4 brain MRI lesions or &lt;4 MRI lesions + 1 spinal cord lesion]</td>
<td></td>
</tr>
</tbody>
</table>

* If the criteria are fulfilled and met, the diagnosis is MS. If the criteria are not completely met, the diagnosis is possible MS. If the criteria are fully explored and not met, the diagnosis is not MS.
† Symptoms lasting at least 24 hours would constitute an attack provided that the episode was not a pseudoademic such as due to fever or infection.
‡ Objective evidence of two lesions necessarily separated in space found on neurologic examination. A typical subclinical lesion detected by visual evoked potential testing (delayed response, well preserved waveform) can substitute for one clinical lesion.
§ MRI (a): Criteria for dissemination in space: three of the four following criteria need to be met:
1. ≥1 gadolinium (Gd)-enhancing lesion or ≥9 T2 lesions
2. ≥1 infratentorial MRI lesion
3. ≥1 juxtacortical lesion
4. ≥3 periventricular lesions
1 spinal cord lesion can substitute for 1 brain lesion.
¶ CSF analysis demonstrates the presence of oligoclonal bands different from any such bands in the serum or an increased immunoglobulin G index.
|| MRI (b): Criteria for dissemination in time: |
1. If original MRI was performed <3 months after the clinical attack, a follow-up MRI (>3 months after event) is necessary. A new Gd-enhancing lesion on this follow-up MRI provides evidence of dissemination in time. Otherwise, another follow-up MRI (>3 months after first MRI) is necessary. A new Gd-enhancing or a new T2 lesion on this second follow-up MRI provides evidence of dissemination in time.
2. If the original MRI was performed ≥3 months after the clinical attack, a Gd-enhancing lesion (at a site not implicated by the clinical event) provides evidence of dissemination in time. Otherwise, a follow-up MRI is necessary (suggested to be 3 months after original MRI). A new Gd-enhancing or a new T2 lesion on this MRI provides evidence of dissemination in time.

certain inherent limitations of the literature in this area. For example, consider a study looking at the predictive diagnostic validity of certain MRI features in patients with symptoms suggestive of MS. Often the design of such a study (as in this assessment) seeks to identify specific MRI features (present on the initial scan) and then to assess the ability of these features to predict the future development of MS. If the defined MRI features are both sensitive and specific for this outcome, then these baseline features are held to be useful in establishing the diagnosis of MS. Nevertheless, MRI features identified in this fashion (regardless of how strongly they are associated with future MS) do not constitute criteria for dissemination in space. Rather, they represent MRI features that help establish the diagnosis of MS at baseline (i.e., using these features to diagnose MS, there would be no need to await further developments). Moreover, any study of the predictive validity of MRI is dependent upon the gold standard used to establish the diagnosis of MS. Generally, this standard is the development of clinically definite MS (CDMS) by some criteria, after some period of follow-up. Nevertheless, in a disease like MS (where the development of CDMS can be delayed by decades from the onset of clinical symptoms) such a design has serious limitations, especially when the follow-up is either too short or too variable. Most studies of the predictive validity of MRI in MS are confounded by this limitation. For example, in a prospective study of patients with CIS followed for only 2 years, many of the patients classified as non-MS (because they have not yet developed CDMS) would still be expected to develop CDMS in the future. Moreover, if the period of follow-up is variable, each patient will not have had an equal chance to develop CDMS. In such circumstances, any calculated value for sensitivity, specificity, positive predictive value, or diagnostic accuracy will be invalid. The use of survival analysis methods might resolve, at least in part, some of these difficulties but, for some reason, these methods have not been widely employed in the clinical literature to date. In addition, it is not clear how best to measure specificity or predictive validity in these circumstances. Thus, it would be useful (from a therapeutic standpoint) to be able to distinguish a recurrent demyelinating disease (i.e., CDMS) from a nonrecurrent demyelinating disease (i.e., a monophasic syndrome). Just as crucial, however, is the important process of ascertaining which patients have another disease entirely (e.g., vitamin B12 deficiency, systemic lupus erythematosus, vasculitis, sarcoidosis). Therefore, it is the specificity, sensitivity, positive predictive value, and diagnostic accuracy of MRI in differentiating against these possibilities that is especially relevant clinically. In addition, essentially all of the MRI data relating to this topic are derived from young to middle-aged adults, so that the predictive value of MRI in children or elderly individuals in these contexts is not known.

Methods. In this assessment, we consider the evidence that MRI can prospectively predict future conversion to CDMS in patients presenting with a syndrome consistent with inflammatory demyelination (e.g., table 4). A large panel of physicians including MS investigators, MS clinicians, and MS neuroradiologists was assembled to analyze the evidence. A search was undertaken using the following terms: clinically isolated syndromes, multiple sclerosis, and magnetic resonance imaging.


We used the Medline database from 1966 to 2003. In addition, the reference lists of the articles identified were also reviewed to identify articles not found by the computer search. Using these methods we identified 46 articles. We reviewed the abstracts of these articles and further limited our assessment to English language studies that were prospective and utilized a well-defined gold standard for the development of CDMS. We only reviewed articles that studied at least 20 patients. There were 22 such studies identified. All of these articles considered the risk of developing CDMS in patients with CIS, based on the presence or absence of MRI lesions within the brain or spinal cord. The studies differed with respect to the number of patients included, the length of follow-up, and the definition of an abnormal scan (table E-5, available online at www.neurology.org). The predictive validity of CSF and evoked potential recordings in MS has also been studied, but is beyond the scope of the current assessment.

We used the Medline database from 1966 to 2003. In addition, the reference lists of the articles identified were also reviewed to identify articles not found by the computer search. Using these methods we identified 46 articles. We reviewed the abstracts of these articles and further limited our assessment to English language studies that were prospective and utilized a well-defined gold standard for the development of CDMS. We only reviewed articles that studied at least 20 patients. There were 22 such studies identified. All of these articles considered the risk of developing CDMS in patients with CIS, based on the presence or absence of MRI lesions within the brain or spinal cord. The studies differed with respect to the number of patients included, the length of follow-up, and the definition of an abnormal scan (table E-5, available online at www.neurology.org). The predictive validity of CSF and evoked potential recordings in MS has also been studied, but is beyond the scope of the current assessment.

We used the Medline database from 1966 to 2003. In addition, the reference lists of the articles identified were also reviewed to identify articles not found by the computer search. Using these methods we identified 46 articles. We reviewed the abstracts of these articles and further limited our assessment to English language studies that were prospective and utilized a well-defined gold standard for the development of CDMS. We only reviewed articles that studied at least 20 patients. There were 22 such studies identified. All of these articles considered the risk of developing CDMS in patients with CIS, based on the presence or absence of MRI lesions within the brain or spinal cord. The studies differed with respect to the number of patients included, the length of follow-up, and the definition of an abnormal scan (table E-5, available online at www.neurology.org). The predictive validity of CSF and evoked potential recordings in MS has also been studied, but is beyond the scope of the current assessment.

**Analysis of the evidence. Relationship of nonenhancing baseline MRI features to the risk of subsequently developing MS.** In 1988, the group from Queen Square reported the initial findings from a prospective series of patients (class II evidence) because of the large number of patients lost to follow-up that initially included 135 individuals with CIS (optic nerve, brainstem, and spinal cord) suggestive of MS who were first studied by MRI between 1984 and 1987. By the time of the first reports, only 109 patients remained in the cohort, apparently reflecting an early dropout of 26 patients. Also by the time of these initial reports, an additional three patients had received alternative diagnoses to MS (established at baseline) and were excluded from further study. The MRI findings in these three patients are not reported. For the purpose of this study, an MRI was defined as abnormal if it contained a single asymptomatic lesion at baseline.

In 1991, the experience in 89 of these 109 patients who had been followed for 5 years was reported. This cohort included 45 patients of the original 61 patients with brainstem and spinal cord syndromes, plus an additional 44 patients with isolated optic neuritis derived from a cohort reported earlier. By the time of this report, an additional 20 patients had apparently been lost to follow-up. Three of the original 135 patients (2.2%) were found (by the time of this report) to have diseases other than MS (cerebrovascular disease, myasthenia gravis, and HIV). The MRI findings in these patients, however, are not reported and these patients were excluded from further analysis (leaving 132 patients in the study cohort). Forty-six of these were apparently lost to follow-up by the time of this report. Of these 89 patients, 57 (64%) had abnormal MRI at baseline. Of these, 72% developed CDMS at follow-up (Poser criteria) compared to only 6% in the group of patients without MRI abnormalities.

In 1994, a study of 89 of these patients with CIS followed prospectively for 43 to 84 months after their initial MRI (class II evidence) was reported. This cohort consisted of the same patients as those reported earlier by Morrissey et al. These authors found that patients with T2-weighted lesion burdens in excess of 1.23 cm² had a 90% risk of developing CDMS during follow-up compared to only 6% in patients with a normal baseline MRI.

In 1998, a further report detailed their experience with 81 patients with CIS (these patients were derived from the same initial cohort of 135 patients reported earlier) who were followed for 10 years after their initial clinically isolated event (class II evidence). Apparently, of the 89 patients reported earlier, 8 had been lost to follow-up (2 of whom apparently died from causes unrelated to MS). Fifty-one of these patients (63%) had two or more T2-weighted white matter lesions on their baseline scans and, of these, 44 (86%) developed CDMS at follow-up compared to only 13% of patients without such MRI abnormalities. One additional patient was excluded from study because systemic lupus erythematosus had been diagnosed rather than MS. As in the other patients with alternative diagnoses, the findings (if any) on the baseline MRI of this patient are not reported (making a total of four patients from the initial cohort with alternate diagnoses).
In 2002, a 14-year follow-up study of 71 of these patients with CIS (class II evidence) was published.\textsuperscript{24} This cohort represented the remaining group from the initial Queen Square cohort of 135 patients reported earlier.\textsuperscript{19-25} Between this report\textsuperscript{24} and the previous one,\textsuperscript{23} an additional eight were lost to follow-up, one refused to participate, and another was diagnosed as having a cerebrovascular accident. Of these 71 patients, 50 (70\%) had an abnormal MRI at baseline and, of these, 88\% subsequently developed CDMS. Moreover, 48 of these 50 patients (96\%) had developed either CDMS or probable MS using only the clinical criteria of Poser.\textsuperscript{17} Documentation of temporal dissemination by either a new clinical exacerbation or the radiologic appearance of new lesions consistent with MS was observed in 98\% of those individuals with abnormal baseline MRI.\textsuperscript{24} In those patients with CIS who had a normal MRI at baseline, only 19\% of 21 patients went on to develop CDMS over the same period of follow-up. Nevertheless, over the 14-year period, 40\% of those with a normal baseline MRI exhibited either new clinical activity or evidence of new MRI lesions. Interestingly, higher MRI lesion numbers and volumes at baseline did not seem to predict higher conversion rates to CDMS. Thus, the low MRI burden (1 to 3 lesions and mean lesion volume of 0.6 cm\(^3\)), medium burden (4 to 10 lesions and mean lesion volume of 0.9 cm\(^3\)), and high burden (>10 lesions and mean lesion volume of 5.6 cm\(^3\)) were all associated with similar rates of conversion to CDMS (89\%, 87\%, and 88\%). This suggests that the presence of even a few (perhaps even one) characteristic MRI abnormalities in patients with a CIS is strongly predictive of the ultimate development of CDMS.

Unfortunately, however, the authors of this Queen Square cohort study did not utilize the statistical method of survival analysis for their data and the actual survival function cannot be constructed for the available reports. In addition, a large number of patients have been lost to follow-up and the MRI findings in those patients who developed alternative diagnoses were omitted from consideration in the available reports. These difficulties complicate interpretation of this data set. Nevertheless, the reported findings suggest that the presence of even a few MRI abnormalities (one to three lesions at baseline) are just as predictive for future MS as are larger lesion loads. In addition, after excluding alternative diagnoses at baseline, only a small number of patients\textsuperscript{5} from the original cohort were known to have received alternative diagnoses to MS at follow-up.

In 1991, a study of 200 suspected MS patients followed prospectively for a mean of 2.1 years (class I evidence) was reported.\textsuperscript{25} MRI scans were classified as strongly supportive of MS at baseline in 94 patients (47\%) on the basis of having either four or more T2-weighted lesions in the cerebral white matter or three lesions, one of which was situated periventricularly. Of these patients, CDMS (Schumacher criteria) developed in 46 (49\%) compared to only 5\% of patients who had normal baseline MRI. Three patients (all over age 60 years at initial presentation) with MRI strongly supportive of MS had received alternative diagnoses to MS at follow-up.

In 1991, the findings in 60 patients with a CIS (class I evidence), 24 of whom had three or more white matter lesions, at least one of which was periventricular in location, were reported.\textsuperscript{26} Of these 24 patients, 71\% progressed to CDMS after an average follow-up of 14.3 months compared to only 24\% in those without lesions. No alternative diagnosis to MS or CIS was made or mentioned at follow-up in this group of patients.

In 1993, the experience in a study prospectively following 303 patients with demyelinating syndromes who did not meet Poser criteria for MS at baseline (class I evidence) was reported.\textsuperscript{27} Of those patients with three or more T2-weighted white matter MRI lesions (or two lesions with one being periventricular in location), 75\% developed CDMS after an average follow-up of only about 8 months compared to only 27\% without these MRI findings. No alternative diagnosis to MS was made or mentioned in this group of patients.

In 1996, the findings in 44 patients with clinically isolated brainstem or spinal cord syndromes, followed prospectively (class I evidence), were presented. After 7 years, 18 of the 30 patients (60\%) who had abnormal MRI (three or more lesions) developed CDMS (McAlpine criteria\textsuperscript{29}) compared to only 4 patients (29\%) without such MRI findings. No alternative diagnosis to MS or CIS was mentioned in this report.

In 1997, the final results of a prospective study of optic neuritis (class I evidence) were reported.\textsuperscript{41} This study involved 74 patients with isolated optic neuritis followed for an average of 5.6 years (range of 4 months to 19.5 years). MRI were obtained an average of 16.4 months after onset and the findings were abnormal (one or more white matter lesions consistent with MS) in 42 patients (57\%). Of these, 16 patients (38\%) developed CDMS compared to only 5 patients (16\%) without such lesions. No alternative diagnosis to MS or optic neuritis was made or mentioned in this group of patients.

In 1997, the final results of the prospective Optic Neuritis Treatment Trial\textsuperscript{31} were published (class I evidence). In the 89-patient cohort who initially had isolated optic neuritis and three or more T2-weighted white matter lesions on their baseline MRI, the cumulative probability of developing CDMS after 5 years of follow-up was 51\%, compared to only 16\% in patients with normal baseline brain MRI.\textsuperscript{31} In 1997, the Amsterdam group\textsuperscript{32} reported their experience in prospectively following 74 patients with a CIS suggestive of MS (class I evidence). Patients were subsequently classified as either CDMS or not MS based on Poser criteria\textsuperscript{17} (excluding paraclinical tests) during follow-up. In this study, patients with an abnormal MRI by Paty et al.'s criteria\textsuperscript{33} (either four or more T2-weighted lesions, or
three lesions, one of which was situated periventricularly) had a 60% chance of developing CDMS during a follow-up period ranging from 23 to 96 months (median = 39 months), compared to only 16% in patients with normal brain MRI at baseline. Using a logistic regression analysis, with dichotomized MRI parameters, in addition to the dichotomized clinical outcome, these authors arrived at the MRI (a) criteria (see table 3), which they reported had the best performance of any MRI criteria with a diagnostic accuracy of 80%. By contrast, Paty et al.’s criteria, despite being more sensitive, had a diagnostic accuracy (by their calculation) of only 69%. Ten patients who were originally in this study (all from one site) were subsequently diagnosed as having a disease other than MS. These patients, however, were excluded from the analysis and their MRI findings are not reported.

In 1999, the findings in 102 patients with isolated optic neuritis followed for an average of 2.3 years using survival analysis methods (class I evidence) were reported. After 10 years, the cumulative probability of developing CDMS in patients with abnormal MRI (one or more white matter lesions consistent with MS) was approximately 55% after 10 years compared to 0% in patients without such lesions. No alternative diagnosis to MS or optic neuritis is mentioned in this report.

The results of a 7-year prospective study consisting of 163 patients followed for 13 to 84 months after a first episode of neurologic symptoms and with at least three typical demyelinating lesions on brain MRI (class IV evidence because there were no MRI-negative controls) were presented in 2000. These authors reported that of these patients, 136 (83%) had an abnormal MRI at baseline. Of these, 81% developed CDMS after 5 years compared to only 22% of patients without such lesions. Of the patients with normal baseline MRI, 31% developed new MRI lesions at the 1-year mark. Of these, 20% developed MS after 5 years compared to none of the patients with persistently normal scans.

In 2000, the Queen Square group reported on the likelihood of developing CDMS at 12 to 19 months in 50 patients with CIS with MRI scans done both at the time of initial presentation and 3 months later (class I evidence). Both the appearance of new T2 lesions (sensitivity 92%; specificity 81%) and the appearance of new Gd enhancement (sensitivity 54%; specificity 89%) at 3 months follow-up were highly predictive of the subsequent development of CDMS in the near term.

In a recent study (class IV evidence) of 383 patients with a CIS and characteristic MRI abnormalities, 82% of the 190 placebo-treated patients had new subclinical MRI lesions by 18 months and 50% of these (using survival analysis methods) had developed CDMS after 3 years. In those who demonstrated two or more Gd-enhancing lesions at baseline, 52% developed CDMS by 18 months and 84% had either developed CDMS or exhibited temporal dissemination of disease activity on MRI after 18 months. Contrast, in patients with fewer than two enhancing lesions at baseline, only 24% developed CDMS at 18 months (class I evidence). In this study, 60% of the placebo patients not meeting MRI (a) criteria (see table 3) went on to develop temporal dissemination of disease activity by either experiencing another exacerbation or developing new MRI lesions.

In a recent study of the positive predictive value, sensitivity, and specificity of MRI abnormalities for the development of CDMS, 68 patients with CIS were evaluated with an MRI both at baseline and at 3 months and 1 year post-onset (class I evidence). Of the 48 (71%) patients with an abnormal baseline MRI (class IV evidence) criteria, which they reported had the best performance of any MRI criteria with a diagnostic accuracy (by their calculation) of 73%. By contrast, Paty et al.’s criteria, despite being more sensitive, had a diagnostic accuracy of 64%. Apparently, none of the patients in this study were diagnosed as having a disease other than MS during follow-up.

Results of a recent phase III clinical trial of IFN-β-1a (class IV evidence because there were no MRI-negative controls) in 308 patients with a CIS and MRI findings suggestive of MS (four white matter lesions on T2, or three lesions, one of which was infratentorial or enhancing after Gd infusion) have been reported. Some of the patients were polysymptomatic at onset although the exact number cannot be derived from the text. Of these patients, 45% of placebo-treated patients had developed CDMS after 2 years. In this study no alternative diagnosis to MS was made or mentioned during the follow-up interval.

Relationship of baseline Gd enhancement and subsequent MRI features to the risk of developing MS. In the study from Queen Square discussed earlier, a new MRI lesion was found (after 1 year) in 68% of patients who had an abnormal MRI at baseline. Of these, 81% developed CDMS after 5 years compared to only 22% of patients without such lesions. Of the patients with normal baseline MRI, 31% developed new MRI lesions at the 1-year mark. Of these, 20% developed MS after 5 years compared to none of the patients with persistently normal scans.
MRI, 18 (21%) developed CDMS and 4 (6%) developed probable MS after the period of follow-up. Using the baseline MRI alone, the most useful imaging correlate for predicting conversion to CDMS was the presence of one or more enhancing lesions (specificity 80%; sensitivity 61%). Moreover, the identification of even a single lesion on T2-weighted scans at baseline was found to be very sensitive in predicting the development of CDMS (89%) at the expense of lower specificity (36%). Integrating the baseline MRI findings with the subsequent emergence of new T2 abnormalities at follow-up appeared to be the most predictive for the development of CDMS (sensitivity 83%; specificity 76%). By contrast, in those patients who did not develop new lesions on the 3-month follow-up scan, CDMS developed in only 5% by 1 year.

The diagnostic value of the new McDonald criteria was recently compared with that of the Poser criteria in a group of patients with CIS followed prospectively for up to 3 years. The clinical diagnosis of MS was made by unblinded physicians who were aware of the (nonquantified) MRI findings (class III evidence). No alternative diagnosis to MS was made or mentioned in the article, although one patient died of an asthmatic attack. All of the patients who developed CDMS at the 3-year follow-up had at least one lesion on brain MRI at baseline. By contrast, only 63% of these patients met McDonald criteria for dissemination in space on baseline brain MRI. After 3 years, a diagnosis of MS was made more often using McDonald criteria than using Poser criteria (58% vs 38%). This difference, however, was due almost entirely to the fact that, after 3 years, 58% of patients with CIS met McDonald MRI criteria for dissemination in time (see table 3), compared to only 38% who met purely clinical criteria for this outcome. Such a result is hardly surprising. As noted earlier, Poser criteria do not consider evidence from sequential paraclinical tests in the diagnostic scheme and it has already been well established that MRI evidence of disease activity (either new lesions or Gd enhancement) is a considerably more frequent occurrence than clinical attacks. Nevertheless, these results are consistent with the notion that new clinical MRI activity is associated with future development of MS.

In 2003, the diagnostic value of the new McDonald criteria was studied retrospectively in a cohort of 139 patients with CIS recruited since 1995 (class III evidence) and who were followed for at least 12 months (mean = 37 months). Only some of these patients received Gd and these data are not reported. These authors found that, in the first year, using the new McDonald criteria, the diagnosis of CDMS was able to be established in 40% of patients compared to only 11% using Poser criteria. Once again, however, this difference in performance between diagnostic schemes is due to the added criteria for dissemination in time in the McDonald scheme, an area not addressed in the Poser scheme. In fact, just considering the baseline information, the Poser scheme is considerably more sensitive than the McDonald scheme for identifying patients who will ultimately develop CDMS. Thus, of the 38 patients in this cohort who developed CDMS (defined by a second clinical attack) at follow-up, the baseline MRI met the McDonald criteria in 27 (71% sensitivity), the Paty criteria in 33 (87% sensitivity), and the Poser criteria in 35 (92% sensitivity). Moreover, no patient developed any alternative diagnosis during follow-up, so the criteria seem equally specific against such an outcome. For example, if the Poser criteria for spatial dissemination were accepted, together with the temporal dissemination criteria of McDonald (see table 3), an additional 11 patients with CDMS (a 22% increase in diagnostic yield) would have been identified without a single patient being misidentified as having MS when, in fact, they had an alternative diagnosis.

As a result, the findings from both of these studies do not support the use of McDonald MRI criteria for dissemination in space (see table 3), which have been consistently found to be less sensitive than either the Paty criteria or the Poser criteria. Rather, the improved diagnostic ability of the new scheme seems to rest entirely upon its use of temporal dissemination criteria. Indeed, the predictive validity of such criteria for temporal dissemination is established both by these studies and by several pieces of class I evidence, as discussed earlier. Moreover, because no alternative diagnosis to MS was made at follow-up in either study, all of the sets of criteria would seem to be equally specific against such a possibility. Therefore, it seems unnecessary to sacrifice sensitivity by requiring more stringent spatial MRI criteria in the development of a diagnostic algorithm for CDMS.

**Discussion.** Newly proposed diagnostic guidelines base their MRI recommendations on the criteria suggested by Barkhof et al. and Tintore et al., stating that “these criteria provide an acceptable degree of sensitivity while providing greater specificity and accuracy than the criteria proposed by Paty et al. and Fazekas et al.” Although these new MRI criteria (a criteria is used to establish dissemination in space, neither of the studies from which these criteria were derived actually addresses the question of spatial dissemination. Rather, as discussed earlier, these studies assess the ability of MRI (measured once at baseline) to predict subsequent CDMS (dissemination in both time and space). Also, in both studies, the follow-up intervals were too short and too variable to draw strong conclusions from the results. Thus, patients were dichotomized into CDMS or not MS groups, as if the not MS group were no longer at risk to develop MS. Despite this dichotomization, however, the cumulative probability of developing MS after a CIS increases steadily with time. Therefore, such a dichotomization does not constitute a valid stan-
standard. For example, the probability of CDMS in high-risk patients with CIS may be as high as 90% after 14 years,\textsuperscript{24} compared to the 31% found in the 2-year study of Tintore et al.\textsuperscript{36} Indeed, the presence of even a small number of lesions at baseline (perhaps as few as one to three) in patients with CIS has been associated with a 98% chance of ultimately developing either probable or definite MS.\textsuperscript{24} In addition, because these studies used a wide range of follow-up intervals, individual patients in these studies were not given an equal chance to develop CDMS. All of these considerations raise serious concerns about the validity of focusing on these observations\textsuperscript{32,36} in the development of diagnostic recommendations.

A striking feature of the literature in patients with CIS with typical MRI lesions is the apparent low risk of any alternative diagnosis being made at follow-up, even when the MRI changes are minimal.\textsuperscript{24} The existing studies have, in general, not fully addressed this point but nonetheless suggest that, in patients with CIS, even a few characteristic MRI lesions are specific for demyelinating disease once appropriate alternative diagnoses (e.g., table 2) have been excluded at baseline. As a result, the principal diagnostic consideration is whether the patient has a nonrecurrent demyelinating disease (CIS or ADEM) or will ultimately prove to have the recurrent disease process we designate as MS. In such a circumstance, it seems appropriate to focus on using MRI to establish recurrent disease activity, rather than developing more stringent MRI criteria to document dissemination in space.

Moreover, because the existing studies have included only patients with monophasic syndromes, their findings cannot be validly applied to patients with progressive disease without attacks (e.g., PPMS). Indeed, there have been no systematic longitudinal studies concerning the predictive validity of any MRI pattern in patients with progressive neurologic syndromes. Thus, no evidence-based diagnostic recommendations for MRI in this clinical setting can be made.

Recommendations. MRI changes seen in MS are known to be nonspecific (e.g., table 2). Therefore, the information derived from imaging investigations must always be considered in the context of the specific clinical situation presented by an individual patient. As a result, the following recommendations are predicated on the exclusion, at baseline, of appropriate alternative conditions that can mimic MS or can mimic the radiographic findings seen in MS.

1. On the basis of consistent class I, II, and III evidence, in patients with CIS, the finding of three or more white matter lesions on a T2-weighted MRI scan is a very sensitive predictor (>80%) of the subsequent development of CDMS within the next 7 to 10 years (Type A recommendation). It is possible that the presence of even a smaller number of white matter lesions (e.g., one to three) may be equally predictive of future MS although this relationship requires better clarification.

2. The presence of two or more Gd-enhancing lesions at baseline is highly predictive of the future development of CDMS (Type B recommendation).

3. The appearance of new T2 lesions or new Gd enhancement 3 or more months after a clinically isolated demyelinating episode (and after a baseline MRI assessment) is highly predictive of the subsequent development of CDMS in the near term (Type A recommendation).

4. The probability of making a diagnosis other than MS in patients with CIS with any of the above MRI abnormalities is quite low, once alternative diagnoses that can mimic MS or can mimic the radiographic findings of MS have been excluded (Type A recommendation).

5. The MRI features helpful in the diagnosis of PPMS cannot be determined from the existing evidence (Type U recommendation).

Recommendations for future research.

1. Most of the articles that deal with the relationship of MRI findings to long-term development of CDMS have utilized older MRI technology. Today's scanners have higher field strength, a more complete coverage area, more sensitive techniques, and utilize better postacquisition processing than was available in these early MRI studies. As a result, this topic will need frequent re-review to assess the impact of these improved imaging techniques.

2. In general, the technical aspects of MRI recordings have been poorly described, documented, or analyzed in the existing literature. Nevertheless, these considerations are probably important and should be specifically detailed and explored in future reports. For example, the size that an abnormality needs to be in order to be classified as a lesion has varied (e.g., some authors use 2.5 mm, others 6 mm, and others do not indicate), and the optimal definition requires clarification. Whether lesions should be identified by fluid attenuated inversion recovery (FLAIR) images, proton density (PD) images, or routine T2-weighted images needs clarification. Whether a slice thickness of 3 mm is necessary or whether 5 mm is adequate needs investigation. Presumably, MRI scans without gaps are preferable to scans with gaps (which may miss lesions), but this potential pitfall is often overlooked.

3. At the moment, individual radiologists or institutions perform MRI scans in a very idiosyncratic manner (e.g., some prefer either axial or sagittal PD images whereas others do not record PD images at all; some use FLAIR, others do not; some intentionally wait between Gd administration and T1-weighted image acquisition, and others do not). These wide variations in technique make it...
extremely difficult for practicing neurologists to compare one scan to the next or to consistently evaluate their patients. The field would be greatly benefited by development of a standardized (nationwide) protocol for the technical aspects of MRI recordings to be used in all patients with suspected or known MS.

4. New imaging technologies, such as magnetization transfer ratios (MTR), magnetic resonance spectroscopy, diffusion tensor imaging, tractography, and brain atrophy measurements will undoubtedly facilitate a better understanding of the extent and dynamic aspects of disease pathology in MS. T1-black hole volume may be a better surrogate measure of disability than T2-lesion volume, although defining lesions by a specific MTR value or by the apparent diffusion coefficient may prove to be more quantitative and less subjective than trying to identify T1-black hole by eye. In addition, three-dimensional imaging techniques, or using specialty head coils, may permit the detection of smaller lesions in areas that are currently called normal-appearing white matter. Each of these new MRI techniques will need to be evaluated for sensitivity and specificity in detecting tissue injury in MS and for predicting the development of MS in the future.

5. Future studies should assess the specificity, sensitivity, and predictive value of specific lesion characteristics (e.g., configuration of the lesion profile, number, and location) that may confer greater or lesser susceptibility to the development of CDMS in patients with a CIS.

6. Studies should be undertaken to assess the use of surveillance MRI as a means of monitoring the disease activity in order to assess the response to different therapies.

7. Studies to define the usefulness of MRI in the diagnosis of PPMS are needed.

**Disclaimer.** This statement is provided as an educational service of the American Academy of Neurology. It is based on an assessment of current scientific and clinical information. It is not intended to include all possible proper methods of care for a particular neurologic problem or all legitimate criteria for choosing a specific procedure. Neither is it intended to exclude any reasonable alternative methodologies. The AAN recognizes that specific patient care decisions are the prerogative of the patient and the physician caring for the patient, based on all of the circumstances involved.

### Appendix 2: Classification of evidence definitions

<table>
<thead>
<tr>
<th>Rating of recommendation</th>
<th>Translation of evidence to recommendations</th>
<th>Rating of diagnostic article</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = Established as useful/predictive or not useful/predictive for the given condition in the specified population</td>
<td>Level A rating requires at least one convincing class I study or at least two consistent, convincing class II studies</td>
<td>Class I: Evidence provided by a prospective study in a broad spectrum of persons with the suspected condition, using a “gold standard” for case definition, where test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy. In addition, there must be adequate accounting for drop-outs with numbers sufficiently low to have minimal potential for bias.</td>
</tr>
<tr>
<td>B = Probably useful/predictive or not useful/predictive for the given condition in the specified population</td>
<td>Level B rating requires at least one convincing class II study or at least three consistent class III studies</td>
<td>Class II: Evidence provided by a prospective study of a narrow spectrum of persons with the suspected condition, or a well designed retrospective study of a broad spectrum of persons with an established condition (by “gold standard”) compared to a broad spectrum of controls, where test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy.</td>
</tr>
<tr>
<td>C = Possibly useful/predictive or not useful/predictive for the given condition in the specified population</td>
<td>Level C rating requires at least two convincing and consistent class III studies</td>
<td>Class III: Evidence provided by a retrospective study where either persons with the established condition or controls are of a narrow spectrum, and where test is applied in a blinded evaluation.</td>
</tr>
<tr>
<td>U = Data inadequate or conflicting. Given current knowledge, test/predictor is unproven</td>
<td></td>
<td>Class IV: Any design where test is not applied in blinded evaluation OR evidence provided by expert opinion alone or in descriptive case series (without controls).</td>
</tr>
</tbody>
</table>
References


