Diagnostic value of paraclinical tests in multiple sclerosis: relative sensitivities and specificities for reclassification according to the Poser committee criteria

S Beer, K M Rösler, C W Hess

Abstract
The yield of paraclinical tests was evaluated in a prospective study of 189 consecutive patients referred for suspected multiple sclerosis (142 patients with multiple sclerosis, 47 non-multiple sclerosis patients on discharge). Patients were first classified according to the Poser criteria by the clinical findings. Subsequently, the results of paraclinical tests (cranial MRI, visually evoked potentials (VEPs), somatosensory evoked potentials by tibial nerve stimulation (SSEPs), motor evoked potentials (MEPs), and analysis of CSF for oligoclonal banding and IgG-index (CSF)) were taken into account. The percentage of reclassified patients (recalification sensitivity, RS) was always lower than the percentage of abnormal results (diagnostic sensitivity, DS), and the divergence of RS v DS differed between the tests (60% v 84% in MRI, 31% v 77% in CSF, 29% v 37% in VEPs, 20% v 68% in MEPs, and 12% v 46% in SSEPs respectively). False reclassifications of non-multiple sclerosis patients to multiple sclerosis would have occurred with all tests (MRI: six of 47 patients, (recalification specificity 88%); CSF: one (98%); VEPs: two (96%); MEPs: two (96%); SSEPs: four (91%); P < 0.05). Although MRI had superior diagnostic capacity, 57 of the 142 patients with multiple sclerosis were not reclassified by the MRI result, 12 of whom were reclassified by CSF and 18 by one of the evoked potential (EP) studies. Of the 98 patients not reclassified by CSF, 53 were reclassified by MRI and 39 by EPs. The results suggest that for the evaluation of paraclinical tests in suspected multiple sclerosis, comparison of diagnostic sensitivities is inappropriate. In general, a cranial MRI contributes most to the diagnosis; however, due to its comparatively low specificity and its considerable number of negative results, EP or CSF studies are often useful to establish the diagnosis of multiple sclerosis.

Key words: multiple sclerosis; evoked potentials; magnetic brain stimulation; magnetic resonance imaging

The diagnosis of multiple sclerosis is based on the detection of multiple inflammatory demyelinating white matter lesions, which are disseminated in time and space. Clinically, dissemination in time is assessed by analysing the course of the disease or by serial examinations, and dissemination in space is suspected if clinical findings cannot be explained on the basis of a single lesion. In many patients, however, the clinical assessment is not sufficient to prove temporal and spatial dissemination, and paraclinical studies have to be performed. Clinically asymptomatic lesions may be detected by evoked potential studies or MRI of the brain and spinal cord. In addition, CSF analysis may provide evidence for the inflammatory nature of the disease. None of the clinical or paraclinical findings are, however, specific for the diagnosis of multiple sclerosis. Early classification criteria were based on clinical data only. The Poser Committee Criteria (PCC), published in 1983, include paraclinical tests as aids for the diagnostic classification of multiple sclerosis. Since then, newer techniques such as MRI and motor evoked potentials (MEPs) have been developed.

The aim of the present investigation was to analyse the diagnostic yield of paraclinical tests in multiple sclerosis. Most previous investigations on this topic analysed the frequency of an abnormal result in a given test (“sensitivity”) in patients with multiple sclerosis, yet high sensitivity of a test in multiple sclerosis does not equal high diagnostic yield. For instance, a paraclinical test that tends to confirm clinically detectable signs is not very likely to improve the diagnostic certainty in a patient, even though it might be very sensitive in terms of yielding a high proportion of abnormal results. Only a diagnostic test which, by the nature of its method and the neural system assessed, is apt to detect clinically silent lesions will be powerful in increasing diagnostic certainty. Thus a different approach is necessary to compare the diagnostic impact of paraclinical investigations in multiple sclerosis. Only a few studies have used a stratified approach to assess the increase of diagnostic accuracy from a clinical baseline by adding stepwise information obtained by paraclinical investigations. None of these included MEPs, even though the diagnostic sensitivity of MEPs in multiple sclerosis exceeds that of somatosensory evoked potentials (SEPs), visually evoked potentials (VEPs), and early acoustically evoked brainstem potentials. In this study, we analysed the number of patients with suspected multiple sclerosis who could be reclassified after the clinical examination
Diagnostic value of paraclinical tests in multiple sclerosis: relative sensitivities and specificities for reclassification according to the Poser committee criteria

According to the PCC by a given test result or by combining results of several tests. These paraclinical tests were: MEPs recorded from upper and lower extremity muscles, SEPs by stimulation of the tibial nerve (SSEPs), VEPs, cranial MRI, and CSF analysis. By its design, the study included some patients with initially suspected multiple sclerosis, in whom subsequently a different diagnosis was made. Comparison of results in these non-multiple sclerosis patients with the results in patients with multiple sclerosis additionally provided a measure for the specificity of these paraclinical tests. Some aspects of this study have been published previously in a preliminary form.15

Material and methods

Patients

We prospectively included all patients with suspected multiple sclerosis referred to our neurological department between 1989 and 1993. Only patients who agreed to be studied in hospital and who lacked a contraindication for one of the tests performed were included (presence of a cardiac pacemaker, history of epileptic fits or intracranial operations, clinical signs of increased intracranial pressure, age under 16 years). We did not include patients who presented with a clinically definite multiple sclerosis according to PCC because, by definition, paraclinical tests would not improve the diagnostic certainty in these patients. Patients with known diseases of the CNS other than multiple sclerosis were also excluded. All patients underwent a clinical examination, CSF analysis, cranial MRI, and multimodal evoked potentials (VEPs, SSEPs, MEPs).

Based on the diagnosis on discharge, we grouped the patients into non-multiple sclerosis patients (patients in whom a different diagnosis was established during the study) and patients with multiple sclerosis, comprising patients with definite, probable, and “possible” multiple sclerosis. Patients were classified as having possible multiple sclerosis if they did not meet the PCC criteria on discharge and no other CNS diseases were found, so that the suspicion of multiple sclerosis was maintained.

Paraclinical tests

All tests complied with the criteria of the local ethics committee. CSF was evaluated for the presence of oligoclonal banding by silver staining after isoelectric focusing on polyacrylamide gel,17 18 and intrathecal IgG production19 indicating an inflammation of the CNS. Normal values given by Reiber20 were applied.

The VEPs and SSEPs were measured in a standardised manner,9 and the latency of the P100 (VEP) and N22/P40 (SSEP) components were compared with our own normative data. The criteria for VEP abnormalities were an absence of P100, a prolonged P100 latency, or an interocular difference of P100 amplitude and latency beyond the mean ± 2.5 SD. Results for SSEPs were considered abnormal if the scalp recorded potential P40 was absent or if the central conduction time (N22-P40) or the side to side differences in latency and amplitude were beyond the mean ± 2.5 SD. We chose not to include early acoustic evoked brainstem potentials as their diagnostic value in multiple sclerosis has been shown to be limited.10 15 22 25

The MEPs were recorded bilaterally from the brachial biceps, the abductor digiti minimi, and the anterior tibial muscle. Transcranial magnetic brain stimulation was performed with a custom made magnetic stimulator and a circular coil of 9 cm diameter. Stimulator and coil placement have been described previously.26 Results for MEPs were considered as abnormal if the cortically evoked potentials were absent, or if the central motor conduction time (CMCT) or the side to side differences in CMCT were beyond 2.5 SD of the mean.

Cranial MRI was usually performed with a 1.5 Tesla General Electric scanner. In some patients, MRI had been performed in other institutions before referral; in these instances, the scan was only repeated if the quality of the external examination did not meet our standards. In all studies, T1 and T2 weighted sequences were available, as well as axial, coronal, and sagittal slices.

Diagnostic classification

For each patient a baseline PCC classification was established based on the clinical data only: clinically probable multiple sclerosis (C1, C2: abbreviations according to Poser et al2), and “possible” multiple sclerosis for patients not fulfilling the PCC. Patients with clinically definite multiple sclerosis (A1) where excluded for the reasons mentioned. Subsequently, the findings of each single paraclinical test and the results of various combinations of the tests were additionally taken into account and the patient was reclassified. These classifications were done by two of us (SB, KMR) independently, and in the few cases of disagreement were jointly discussed and classified. The reclassified patients were then counted for each test and each combination of tests.

From the clinical baseline classification, abnormal paraclinical test results were only considered for reclassification if they proved a separate lesion in a different functional system. In particular, MEP findings were only considered for reclassification if the clinical examination did not show brainstem dysfunction, pyramidal signs, or motor deficits of the respective spinal segment. For instance, if clinically a patient had a paraparesis with exaggerated tendon jerks of the legs, an abnormal MEP to the tibialis anterior was not considered for reclassification, but an abnormal MEP to the biceps was considered as evidence of a second lesion. Results for SSEPs were interpreted similarly—that is, only if there was no brainstem disturbance and no sensory deficit. Abnormal VEP findings were only considered for reclassification if no visual disturbance of the same side had influenced...
the baseline PCC classification. When the yield of multimodal EPs was assessed, an abnormal result for both MEPs and SEP was not considered as evidence of two lesions if both could be attributed to the same spinal or brainstem location. Findings from MRI were taken into account in two different ways. Firstly, all MRIs demonstrating multiple areas of increased signal (multifocality) or single areas that could not account for the clinical symptoms of multiple sclerosis, were accepted for reclassification. Secondly, MRIs were only accepted as indicative of dissemination in space if they fulfilled the MRI criteria proposed by Fazekas et al.\textsuperscript{27} Three or more areas of increased signal and the presence of at least two of the following characteristics: (a) lesion size > 5 mm, (b) lesion abutting the ventricular bodies, and (c) infratentorial lesion location. These criteria had the greatest specificity for multiple sclerosis in a comparative analysis.\textsuperscript{28}

**Analysis of Results**

Five parameters were calculated for each paraclinical test. These were: (a) the diagnostic sensitivity, which was defined as the percentage of patients with an abnormal result in a given test or combination of tests; (b) the reclassification sensitivity, which was defined as the percentage of patients with multiple sclerosis that could be reclassified according to PCC by results of a given test or a combination of tests; (c) the reclassification index, which was defined as the ratio of reclassification sensitivity and diagnostic sensitivity; (d) the diagnostic specificity, which was defined as the percentage of non-multiple sclerosis patients with a normal finding in a given test; and (e) the reclassification specificity, which was defined as the percentage of non-multiple sclerosis patients that were correctly not reclassified by a given test. To assess significant differences of these parameters between the paraclinical tests, McNemar’s ∼ 2 test was applied. The level of significance was set at P = 0.05.

**Results**

**Patients**

Inclusion criteria were fulfilled by 189 patients (107 women, 82 men). Their mean age was 38 (range 16–67) years, and the mean duration of the illness before admission to hospital was 2–9 (0–25) years. These 189 patients were split into two groups based on the discharge diagnosis: the multiple sclerosis group consisted of 142 patients (80 women, 52 men), containing 98 patients with definite, 15 with probable, and 29 with possible multiple sclerosis. Of the 29 patients with possible multiple sclerosis, 14 had a chronic progressive course, and 15 had a first attack with consecutive remission. Of these 15 patients, nine experienced further bouts after a mean follow up of two years, allowing the later classification as probable or definite multiple sclerosis. Fifty five patients presented with a spinal cord syndrome, 40 had a brainstem symptomatology, 12 had an optic neuritis, and 35 had other CNS symptoms of multiple sclerosis. The mean age in the multiple sclerosis group was 37 (range 16–66) years. Table 1 shows the classification of the patients with multiple sclerosis to the diagnostic categories using the PCC and including all clinical and paraclinical data.

Forty seven patients (27 women, 20 men) had other neurological diseases or functional disturbances (non-multiple sclerosis group). Their mean age was 41 (range 19–67) years. The diagnoses in this group were psychogenic deficits (15), cerebrovascular diseases (13), non-inflammatory myelopathy (five), other inflammatory diseases of the CNS (four), and various other neurological diseases (10).

**Table 1** 
Classification of patients with multiple sclerosis (n = 142) based on the Poser committee criteria (PCC) before and after paraclinical testing

<table>
<thead>
<tr>
<th>Classification</th>
<th>Attacks</th>
<th>Clinical evidence</th>
<th>Paraclinical evidence</th>
<th>CSF</th>
<th>Baseline classification (n)</th>
<th>Discharge classification (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinically definite MS</td>
<td>A1</td>
<td>2</td>
<td>2</td>
<td>and</td>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>2</td>
<td>1</td>
<td></td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Laboratory supported definite MS</td>
<td>B1</td>
<td>1</td>
<td>2</td>
<td>or</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>2</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B3</td>
<td>1</td>
<td>1</td>
<td>and</td>
<td>1</td>
<td>42</td>
</tr>
<tr>
<td>Clinically probable MS</td>
<td>C1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>1</td>
<td>2</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>1</td>
<td>1</td>
<td>and</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Laboratory supported probable MS</td>
<td>D1</td>
<td>2</td>
<td></td>
<td></td>
<td>82</td>
<td>29</td>
</tr>
</tbody>
</table>

Baseline classification from clinical data only; discharge classification based on clinical data and results of all paraclinical tests; MS = multiple sclerosis; for definition of possible multiple sclerosis see text.
The diagnostic impact of an abnormal result of a paraclinical test is expressed by the reclassification index (in reclassified/in abnormal). MRI results were interpreted with and without taking into account the criteria by Fazekas et al.2. Brackets indicate statistically significant differences (McNemar's χ² test).

Table 2  Number and frequency of abnormal findings (diagnostic sensitivity) and of reclassified patients (reclassification sensitivity) for the different paraclinical tests in the multiple sclerosis group (n = 142)

<table>
<thead>
<tr>
<th>Test</th>
<th>No abnormal findings (%)</th>
<th>Reclassified sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI (Fazekas et al)</td>
<td>119 (84)</td>
<td>112 (79)</td>
</tr>
<tr>
<td>CSF</td>
<td>85 (60)</td>
<td>85 (60)</td>
</tr>
<tr>
<td>MEP</td>
<td>110 (77)</td>
<td>44 (31)</td>
</tr>
<tr>
<td>SSEP</td>
<td>96 (68)</td>
<td>26 (20)</td>
</tr>
<tr>
<td>VEP</td>
<td>53 (37)</td>
<td>17 (12)</td>
</tr>
</tbody>
</table>

The diagnostic impact of an abnormal result of a paraclinical test is expressed by the reclassification index (n reclassified/in abnormal). MRI results were interpreted with and without taking into account the criteria by Fazekas et al. Brackets indicate statistically significant differences (McNemar's χ² test).

Reclassification to definite MS ( ):

Reclassification to probable or definite MS ( ):

Table 3  Patients in the multiple sclerosis group (n = 142) that could not be reclassified by a given paraclinical test (-) and patients that would have been reclassified by additional paraclinical tests (+) if the initial single test was normal.

<table>
<thead>
<tr>
<th>Test</th>
<th>MRI (Fazekas et al)</th>
<th>CSF</th>
<th>MEP</th>
<th>SSEP</th>
<th>VEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI (Fazekas et al)</td>
<td>57</td>
<td>12</td>
<td>10</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>CSF</td>
<td>98</td>
<td>39</td>
<td>19</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>MEP</td>
<td>77</td>
<td>18</td>
<td>0</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td>SSEP</td>
<td>125</td>
<td>38</td>
<td>25</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>VEP</td>
<td>101</td>
<td>55</td>
<td>28</td>
<td>18</td>
<td>7</td>
</tr>
</tbody>
</table>

Number of patients that would have been reclassified to definite multiple sclerosis are in parentheses. MRI was always interpreted according to Fazekas et al. Multimodal EP = combination of MEPs, SSEPs, and VEPs.
Table 4 Results for the 47 patients of the non-multiple sclerosis group

<table>
<thead>
<tr>
<th>Test</th>
<th>No of abnormal results (%) (sensitivity)</th>
<th>No falsely reclassified by PCC to MS (%) (diagnosis)</th>
<th>Diagnostic specificity (%)</th>
<th>Reclassification specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
<td>18 (38)</td>
<td>15 (32) (8 CVI, 1 Lyme disease, 1 mitochondrial disease, 1 spinal tumour, 4 functional deficits)</td>
<td>62</td>
<td>68</td>
</tr>
<tr>
<td>MRI (Fazekas et al)</td>
<td>6 (13)</td>
<td>6 (13) (4 CVI, 1 Lyme disease, 1 mitochondrial disease)</td>
<td>87</td>
<td>87</td>
</tr>
<tr>
<td>CSF</td>
<td>7 (15)</td>
<td>1 (2) (1 unverified CNS inflammation)</td>
<td>85</td>
<td>98</td>
</tr>
<tr>
<td>MEP</td>
<td>11 (23)</td>
<td>2 (4) (1 spinocerebral atrophy, 1 thoracic myelopathy)</td>
<td>77</td>
<td>96</td>
</tr>
<tr>
<td>SSEP</td>
<td>7 (11)</td>
<td>4 (9) (1 CVI, 1 cervical myelopathy)</td>
<td>85</td>
<td>91</td>
</tr>
<tr>
<td>VEP</td>
<td>5 (11)</td>
<td>2 (4) (1 CVI, 1 cervical myelopathy)</td>
<td>89</td>
<td>96</td>
</tr>
</tbody>
</table>

MS = multiple sclerosis; diagnostic specificity = percentage of correct negative results; reclassification specificity = percentage of correctly unreclassified patients. MRI was interpreted with criteria given by Fazekas et al. CVI = cerebrovascular infarction.

**Diagnostic and Reclassification Specificity**

In the non-multiple sclerosis group (n = 47), 5 patients would have been falsely classified as having probable multiple sclerosis on the basis of the clinical examination alone. This corresponds to a specificity of the clinical examination of 89%. All paraclinical tests yielded abnormal results in some non-multiple sclerosis patients, and some of these abnormal results would have led to incorrect reclassification of these patients to probable or definite multiple sclerosis (table 4). The number of falsely reclassified patients was greatest for MRI, followed by SSEPs, MEPs, VEPs, and CSF (MRI v CSF, P < 0.05). Application of the criteria of Fazekas et al. to the analysis of MRIs reduced the number of falsely reclassified patients considerably (table 4). Most of the falsely reclassified patients had cerebrovascular disorders (table 4). The incidence of non-specific MRI abnormalities is greater after the age of 50, and our six falsely reclassified patients were somewhat older (mean 45, range 20–69 years) than the non–multiple sclerosis group as a whole (41 years). Nevertheless, four of the six patients were younger than 50 years, suggesting that false reclassifications by MRI may not only be a problem of older age. These data allowed calculation of diagnostic and reclassification specificities for each of the paraclinical tests (table 4). The reclassification specificity decreased considerably if more than one paraclinical test was considered: multimodal EPs had a reclassification specificity of 87%, and if all paraclinical tests were combined (and the criteria of Fazekas et al. were applied to MRI interpretation), the reclassification specificity was only 72%.

**Discussion**

Paraclinical tests are performed to increase the certainty of a clinically suspected diagnosis. This is of particular interest in a disease such as multiple sclerosis, for which a specific test does not exist. In this study, the efficacy of paraclinical testing in multiple sclerosis was analysed by focusing on the diagnostic power of various methods in a clinical setting. It shows that mere evaluation of diagnostic sensitivities may be inappropriate. By using a stratified approach in unselected patients with suspected multiple sclerosis we analysed the impact of different paraclinical tests on PCC classification. As expected, the reclassification sensitivity (the percentage of patients who could be reclassified by an abnormal result) was lower than the diagnostic sensitivity (the percentage of patients with an abnormal result) in all of the performed tests. Because this reclassification index (table 2) differed largely between the tests, comparison of these tests on the basis of diagnostic sensitivity alone may be misleading. When taken as single tests, MRI had the greatest reclassification sensitivity (60%), followed by CSF analysis (31%), and VEPs (28%), MEPs (19%), and SSEPs (12%).

By including patients with suspected multiple sclerosis in whom subsequently a diagnosis was made, we were able to calculate specificities of the paraclinical tests for reclassification according to the PCC (table 4). The reclassification specificities were relatively high for all performed tests. It is, however, important to note and of considerable clinical interest that by increasing the number of paraclinical examinations, the number of falsely reclassified non-multiple sclerosis patients increased as well. The diagnostic accuracy in a given patient may thus decrease rather than increase if a great many paraclinical examinations are performed. Although the few non-multiple sclerosis patients in our study did not allow a thorough assessment, they suggest a lower specificity of MRI compared with the other paraclinical tests. The MRI results would have led to an incorrect reclassification to probable or definite multiple sclerosis in six of 47 non-multiple sclerosis patients, even if the restrictions described by Fazekas et al. were used for the interpretation. Contrary to MRI, only one patient would have been falsely reclassified by the result of the CSF analysis (table 3). Summarised, our data support our clinical impression that sometimes MRI findings that are highly suggestive of multiple sclerosis fail to exclude other neurological conditions, as stated previously by others. Hence, the high reclassification sensitivity of the MRI may, to some extent, be at the expense of a comparatively low reclassification specificity; and diagnostic criteria such as those proposed by Fazekas et al. should be used for the interpretation of MRI abnormalities.

A possible reason for performing paraclinical tests is to confirm and characterise clinically doubtful or suspected findings. This aspect was not assessed in the present investigation, and applies especially to EP studies. For example, in a patient with visual disturbances it may be helpful to perform VEPs, as a prolonged latency supports the presence of an
optic neuritis, whereas a decreased amplitude of the cortical response may suggest an ophthalmological problem or an axonal lesion to the optic nerve.\textsuperscript{21} Such a patient would not have been reclassified by an abnormal VEP result. Also the diagnostic values of MEPs and SSEPs is probably underrated by our analysis, because they may be helpful in ascertaining doubtful clinical findings such as brisk reflexes without spasticity, equivocal extensor plantar responses, or ambiguous sensory symptoms of multiple sclerosis.

It is difficult to compare our data with previously published studies. Firstly, most investigations only reported diagnostic sensitivities of various paraclinical tests without evaluating the impact of these abnormal results on the diagnostic appraisal. Secondly, the diagnostic sensitivity of paraclinical testing is known to increase during the course of the disease.\textsuperscript{22,29} The average duration of the disease of our cohort of patients was probably shorter than in most previous studies, as we included cases with possible and probable multiple sclerosis. Thirdly, by excluding patients with possible multiple sclerosis, Espay et al.\textsuperscript{29} could not yet be classified according to the PCC, previous studies excluded a most important group of patients who only later developed probable or definite multiple sclerosis. We are aware that our inclusion of patients with possible multiple sclerosis with monophasic syndromes bears the risk of misinterpretation of an acute disseminated encephalomyelitis (ADEM) as multiple sclerosis. As ADEM is much less common than multiple sclerosis, this risk is of little clinical relevance, however. In a pathomorphological study, Izquierdo et al.\textsuperscript{30} showed that, based on the PCC, the average delay for diagnosing clinically definite multiple sclerosis was longer than eight years. Hence, when assessing the diagnostic efficacy of a test, exclusion of cases with possible multiple sclerosis from the study fails to meet the clinical reality. Finally, the specificity of paraclinical studies to reclassify patients with suspected multiple sclerosis according to the PCC criteria has not been evaluated so far.

The superiority of MRI in our investigation was due to its known high sensitivity in detecting clinically silent lesions.\textsuperscript{29,31} Dissemination in space can be directly demonstrated by MRI whereas EPs suggest spatial dissemination only in the context of clinical findings. Paty et al.,\textsuperscript{1} analysing a similar patient group (n = 200), found a similarly high diagnostic sensitivity for MRI (62\%) compared with sensitivities between 46\% and 49\% for CSF, VEPs, and SSEPs. In a follow up study of the same patients,\textsuperscript{32} the initial MRI had the greatest predictive value for later evolution to clinically definite multiple sclerosis, followed by CSF/VEPs and SSEPs. Similar results were reported by Filippini et al.,\textsuperscript{31} who found a diagnostic sensitivity of 70\% for MRI in 34 cases of suspected multiple sclerosis. In these studies, however, the impact of the abnormal test results on the initial diagnostic assessment was not evaluated.

It has been suggested that multimodal EP recordings may increase the diagnostic yield of electrophysiological testing in multiple sclerosis.\textsuperscript{33,34} This was confirmed in the present study, where multimodal EPs allowed reclassification in a greater number of patients than single EP studies. Even if all three EP modalities were combined, however, MRI had a superior reclassification sensitivity (60\% for MRI \& 46\% for multimodal EPs; figure). The low reclassification sensitivity of MEPs and SSEPs in our study may be due to our restrictive interpretation based on the Poser committee recommendations.\textsuperscript{35} An abnormal result in EPs was only considered for reclassification if none of the clinical findings could be associated with a lesion along the known EP pathways. This is reflected in the high percentage of abnormal multimodal EP and SSEPs results and their relatively low reclassification sensitivity (table 2). Nevertheless, EP testing may be useful in patients in whom the MRI or CSP testing is not diagnostic, as was the case in 18 or 39 of our patients with non-reclassify- ing MRI or CSP respectively and abnormal EPs (table 3). In 39 patients with suspected multiple sclerosis, Farlow et al.\textsuperscript{36} found multiple MRI lesions in 72\%, but abnormal multimodal EPs only in 41\%. Similarly, Cutler et al.\textsuperscript{37} found abnormal MRIs in 21 of 27 patients with definite or probable multiple sclerosis (78\%), but abnormal multimodal EPs in only 14 of these patients (58\%). By contrast, Giesser et al.\textsuperscript{38} found a higher sensitivity for multimodal EPs (78\%) than for MRI (65\%) in a group of 23 patients with possible multiple sclerosis. It should be noted that in these studies multimodal EPs included VEPs, SSEPs, and acoustic evoked brainstem potentials.

In our cohort of patients, VEPs had a greater chance to detect subclinical lesions than had MEPs and SSEPs, because fewer patients complained of visual disturbances (12 patients in the multiple sclerosis group). Thus, even though VEP abnormalities were considerably less frequent than MEP or SSEP abnormalities, they led to a greater relative number of reclassifications (table 2). Previous studies have also shown that VEPs may have a superior sensitivity to detect clinically silent lesions\textsuperscript{39} even though their diagnostic sensitiv- ity for multiple sclerosis may be lower than EP modalities.\textsuperscript{35}

The present direct comparison of MRI/EPs and CSF may be inappropriate from a theoretical standpoint, because of their different appreciation in the PCC classification. These tests are not directly competitive, because CSF furnishes information about possible CNS inflammation, and MRI/EP about possible dissemination in space. The direct compar- ison in the present study was chosen from a clinical standpoint, where the diagnostic impact (sensitivity and specificity) of an examination is the main interest and not necessarily the pathophysiologlcal nature of abnormal findings or the way by which they are achieved. Even though CSF analysis had a similar diagnostic sensitivity as MRI, the reclassification sensitivity of CSF was only
half of that of MRI (table 2). By its special appreciation within the PCC, however, CSF analysis was particularly powerful in allowing reclassification to definite multiple sclerosis (figure). In previous studies, the weight of CSF analysis for the diagnosis of multiple sclerosis has been a contentious issue. Morrisey et al 14 found MRI more powerful than CSF in predicting the evolution of clinically isolated syndromes in multiple sclerosis. Also, others 15 16 suggested that CSF analysis does not add substantially to diagnostic certainty. In our investigation, however, 12 of 57 patients (21%) with normal or non-diagnostic MRI could be reclassified by the CSF finding (table 3). In 90 patients with multiple sclerosis with a cranial MRI not "strongly suggestive of multiple sclerosis", Lee et al 32 found 25 patients (28%) with positive CSF oligoclonal bands, four of them developing clinically definite multiple sclerosis in the follow up. Sharief and Thompson 28 found a greater predictive value of CSF (including IgM and IgG analysis) for the later development to definite multiple sclerosis than MRI. Five of their 45 patients with normal MRI findings showed an evolution to multiple sclerosis after a mean follow up of 18 months.

To our knowledge, only one previous investigation used a similar analysis of reclassification sensitivity as ours. Gilmore et al 10 analysed the diagnostic value of paraclinical testing in a group of 40 patients with possible or probable multiple sclerosis. The diagnostic reclassification sensitivity of all tests (CSF, MRI, and multimodal EPs) was 63% in their patients, compared to 78% in the present study. Analysis of CSF had a reclassification sensitivity of 25%, compared with 31% in our investigation, and multimodal EPs combined with CSF 53% (59% in this study). On its own, MRI was indicative of a silent lesion in 45% only. The findings for MRI and EPs, however, cannot be directly compared with ours, because Gilmore et al 10 included the results of the CSF analysis for the baseline classification and subsequently calculated the diagnostic impact of the other paraclinical tests. Also, the performed EPs were different (VEPs, SSEPs, and BAEPs).

In conclusion, our data suggest that the clinical appraisal of suspected multiple sclerosis should start with a cranial MRI and a lumbar puncture for CSF testing. The combination of these two examinations seems particularly useful as they assess two different aspects of the disease—namely, dissemination of lesions in space (MRI) and CNS inflammation (CSF). Moreover, it is the combination of the examinations that yields the greatest specificity (CSF) and the greatest sensitivity (MRI) for reclassifying patients with multiple sclerosis according to the PCC. If MRI and CSF do not allow the attainment of a probable or definite diagnosis (or a lumbar puncture is considered too invasive by the patient), further testing can follow, depending on the clinical situation, by the use of VEPs, MEPs, or SSEPs, or a combination of these with the aim of disclosing clinically silent lesions. Studies on EPs remain helpful especially in evaluating doubtful clinical signs, and, due to their high sensitivity and easy performance, MEPs seem particularly suitable for screening patients with equivocal symptoms. 13 15 37 One should be aware, however, that performing additional tests may bear the risk of false reclassifications of non-multiple sclerosis patients to multiple sclerosis. A greater number of paraclinical examinations may thus decrease rather than increase the diagnostic accuracy in multiple sclerosis.

We thank our colleagues of the Department of Neuroradiology (Head: Professor G Schnitker) for their collaboration. We also thank Dr J Kesselring (Valens) for helpful comments. This work was supported by the Swiss National Science Foundation Grant No 3.552-1.86.

Diagnostic value of paraclinical tests in multiple sclerosis: relative sensitivities and specificities for reclassification according to the Poser committee criteria