A longitudinal study of cognition in primary progressive multiple sclerosis

S. J. Camp,1 V. L. Stevenson,1 A. J. Thompson,1 G. T. Ingle,1 D. H. Miller,1 C. Borras,2 B. Brochet,3 V. Dousset,3 M. Falautano,4 M. Filippi,4 N. F. Kalkers,5 X. Montalban,2 C. H. Polman5 and D. W. Langdon 1,6

1Department of Clinical Neurology, Institute of Neurology, London, UK, 2Unitat de Neuroimmunologia Clinica, Hospitals Vall d’Hebron, Barcelona, Spain, 3Fédération des Neurosciences Cliniques de CHU de Bordeaux, Hôpital Pellegrin, Bordeaux, France, 4Department of Neuroscience, Scientific Institute Ospedale San Raffaele, Milan, Italy, 5University Hospital, Free University, Amsterdam, The Netherlands and 6Department of Psychology, Royal Holloway, University of London, Egham, UK

Correspondence to: A. J. Thompson, Institute of Neurology, Queen Square, London WC1N 3BG, UK
E-mail: a.thompson@ion.ucl.ac.uk

There are few longitudinal studies of cognition in patients with multiple sclerosis, and the results of these studies remain inconclusive. No serial neuropsychological data of an exclusively primary progressive series are available. Cross-sectional analyses have revealed significant correlations between cognition and magnetic resonance imaging (MRI) parameters in primary progressive multiple sclerosis (PPMS). This study investigated cognitive and MRI change in 99 PPMS patients from five European centres for 2 years. They were assessed at 12 month intervals using the Brief Repeatable Battery, a reasoning test, and a measure of depression. The MRI parameters of $T_1$ hypointensity load, $T_2$ lesion load, and partial brain volume were also calculated at each time point.

There were no significant differences between the mean cognitive scores of the patients at year 0 and year 2. However, one-third of the patients demonstrated absolute cognitive decline on individual test scores. Results indicated that initial cognitive status on entry into the study was a good predictor of cognitive ability at 2 years. There was only a small number of significant correlations between changes in cognition and changes on MRI, notably $T_1$ hypointensity load with the two attentional tasks ($r = -0.266, P = 0.017; r = -0.303, P = 0.012$). It is probable that multiple factors underlie this weak relation between the cognitive and MRI measures.

Keywords: cognitive function; longitudinal study; MRI; multiple sclerosis; primary progressive

Abbreviations: BRB = Brief Repeatable Battery; EDSS = Expanded Disability Status Scale; MADRS = Montgomery and Asberg Depression Rating Scale; MRI = magnetic resonance imaging; PPMS = primary progressive multiple sclerosis; SPMS = secondary progressive multiple sclerosis; RRMS = relapsing remitting multiple sclerosis; VESPAR = Verbal and Spatial Reasoning Test; SDMT = Symbol Digits Modalities Test; PASAT = Paced Auditory Serial Addition Task; WLG = Word List Generation Test


Introduction

Primary progressive multiple sclerosis (PPMS) occurs in 10–15% of patients with multiple sclerosis. By definition, these patients experience a progressive disease course from onset without any relapses or remission. Because of their comparative rarity, PPMS patients are less often studied (Thompson et al., 2000). The frequency of cognitive impairment in PPMS has been variously quoted as 7% (Comi et al., 1995) and 29% (Camp et al., 1999). The cognitive skills reported to be impaired in PPMS include attention and working memory (Camp et al., 1999; Foong et al., 2000; De Sonneville et al., 2002; Huijbregts et al., 2004), verbal memory (Camp et al., 1999; Gaudino et al., 2001; Blum et al., 2002; Huijbregts et al., 2004), spatial memory (Huijbregts et al., 2004; but not consistently, see Camp et al., 1999; Gaudino et al., 2001), spatial reasoning (Camp et al., 1999) and verbal fluency (Camp et al., 1999; Huijbregts et al., 2004). Spatial perception and word definition have been reported intact (Gaudino et al., 2001), as has verbal reasoning (Camp et al., 1999).
Reports comparing cognitive function in PPMS with other multiple sclerosis subtypes are equivocal, although emerging differences in the pathological characteristics maintain interest in the comparison (Bruck et al., 2002). PPMS patients were found to perform better than secondary progressive multiple sclerosis (SPMS) patients on a spatial working memory task involving planning (Foong et al., 2000) and a spatial memory task (Gaudino et al., 2001); but PPMS patients have also demonstrated poorer spatial recall and verbal fluency than SPMS patients (Huijbregts et al., 2004), a specific deficit in word fragment completion, that was not present in relapsing remitting multiple sclerosis (RRMS) or SPMS (Blum et al., 2002) and impaired complex attention skills, verbal memory and verbal fluency, compared with RRMS patients (Gaudino et al., 2001; Huijbregts et al., 2004). Cognitive dysfunction in PPMS correlates modestly with magnetic resonance imaging (MRI) parameters (Camp et al., 1999). Controlled, cross-sectional, neuropsychological studies of multiple sclerosis patients have demonstrated that cognitive deficits may occur early in the disease process, and worsen as the disease progresses (Heaton et al., 1985; Beatty et al., 1989; Ron et al., 1991). Patients with optic neuritis, and brainstem or spinal cord lesions, which are frequently the harbinger of multiple sclerosis (Francis et al., 1987; Miller et al., 1989), have been reported as demonstrating mild cognitive deficits (Lyon-Caen et al., 1986; Callanan et al., 1989; Feinstein et al., 1992a), while RR patients show mild to moderate forms of impairment (Heaton et al., 1985; Beatty et al., 1989), and SP patients exhibit more severe deficiencies (Heaton et al., 1985; Beatty et al., 1988; Feinstein et al., 1992b).

Less is known about progressive deterioration of neuropsychological skills in multiple sclerosis. Correlations between cognitive impairment and indicators of disease progression, such as disability and disease duration, are inconsistent, with some reports showing a significant correlation between cognitive impairment and physical disability (Beatty et al., 1990; McIntosh-Michaealis et al., 1991; Rao et al., 1991; Kessler et al., 1992; Basso et al., 1996; Troyer et al., 1996), while others have recorded no significant relationship (Jennekens-Schinkel et al., 1990; Minden et al., 1990; Mariani et al., 1991; Ron et al., 1991; Maurelli et al., 1992; Patti et al., 1995). The data regarding the relationship between cognitive function and disease duration are equally mixed, with most researchers reporting no correlation (Beatty et al., 1990; Jennekens-Schinkel et al., 1990; Minden et al., 1990; Rao et al., 1991; Maurelli et al., 1992; Patti et al., 1995), and others noting a significant relationship (McIntosh-Michaealis et al., 1991; Ron et al., 1991).

Longitudinal studies of cognitive function in multiple sclerosis do not provide definitive support for neuropsychological stability or vulnerability. A number of researchers have shown preservation of the cognitive skills of multiple sclerosis patients over time (Filley et al., 1990; Jennekens-Schinkel et al., 1990; Mariani et al., 1991; Mattioli et al., 1993; Amato et al., 1995; Hohol et al., 1997). In contrast, cognitive deterioration has been demonstrated (Canter, 1951; Feinstein et al., 1992b, 1993; Kujala et al., 1997; Amato et al., 2001).

The advent of large therapeutic trials for patients with multiple sclerosis provides the opportunity to utilize data from the untreated, control, patient group to examine the natural history of the disease, and the impact of serial assessment. The placebo group of Weinstein et al. (1999) comprised 126 RRMS patients, and showed significant improvement in verbal and spatial memory and attention over the 2 year study period. Weinstein et al. (1999) suggested that this was owing to practice effects, despite the use of parallel forms, and to the additional care the patients received. Fischer et al. (2000) also demonstrated practice effects in the neuropsychological performance of 74 placebo-treated, relapsing multiple sclerosis patients, followed up for 104 weeks. Apart from the new therapies, these studies provide valuable details of the change in cognitive ability over time, and with repeated assessment.

Recent longitudinal studies of cognition and MRI variables have struggled to demonstrate significant links over time. In a group of early RRMS patients, brain parenchymal volume decrease was the only independent predictor of cognitive decline, whereas neither $T_1$ and $T_2$ baseline values nor volume changes were significantly related to cognitive change (Zivadinov et al., 2001). Another study showed significant links between cognitive test performance and MRI parameters at both baseline and 4 year follow up, but no significant effects of change in MRI parameters on cognitive decline (Sperling et al., 2001).

To date, there are no longitudinal studies examining neuropsychological skills in PPMS patients exclusively. Therefore, the aims of this study were:

(i) To investigate cognitive change over time, in patients with PPMS.
(ii) To examine the relationship between cognitive change and MRI variables.

**Methods**

**Patients**

Of the 158 patients with PPMS recruited from six European centres (Amsterdam, Barcelona, Bordeaux, Lisbon, London and Milan) for the cross-sectional study (Camp et al., 1999; Stevenson et al., 1999), 147 PPMS patients from five of the centres were approached for repeat assessment (Lisbon was unable to continue participation in the study). These individuals were monitored for 2 years. All patients gave informed consent to participate in the study, which has been approved by the appropriate ethics committees, including the Institute of Neurology and National Hospital for Neurology and Neurosurgery, London, UK. All patients underwent a full neurological examination. Impairment and disability were measured in a standardized manner, using the Expanded Disability Status Scale (EDSS; Kurtzke, 1983). Wherever possible each patient was scored by the same observer at each examination.

**Neuropsychological tests**

The neuropsychological battery comprised tests of memory, attention, verbal fluency and reasoning. The Brief Repeatable Battery (BRB; Rao, 1990) is widely used (Hohol et al., 1997), and the
individual tasks have been detailed previously (Camp et al., 1999). In brief, the BRB comprises a list learning task to test verbal memory, the reproduction of an abstract spatial pattern with counters on a grid to test spatial memory, a timed coding task requiring numbers matched to shapes to be spoken aloud to test complex attention, the addition of pairs of spoken numbers to test working memory, and the generation of words from a given category to test executive skills. As the neuropsychological assessment was conducted yearly, parallel forms of the BRB were employed. Version B was used in year 0, Version A in year 1, and Version B again in year 2. In addition to the BRB, which has a narrow focus (Basso et al., 1996), the Verbal and Spatial Reasoning Test (VESPAR; Langdon and Warrington, 1995) was administered to patients in London, Amsterdam and Barcelona (i.e. the VESPAR was not used in Bordeaux and Milan). This test examines both verbal and spatial inductive reasoning skills. The Montgomery and Asberg Depression Rating Scale (MADRS; Montgomery and Asberg, 1979) was used to assess symptoms of depression.

MRI examination

MRI examinations were performed at each of the five centres (3 mm T1 weighted fast spin echo and T1 weighted spin echo images of the brain). A detailed account of the imaging protocol can be found in the paper by Stevenson et al. (2000). All electronic data were transferred to London for analysis by two observers (VLS and GTI). Total brain T1 lesion load and T1 hypointensity load were obtained using a semi-automated contour technique. The measure of partial cerebral volume, reflecting atrophy, was taken from a series of six 3 mm consecutive slices, with the most caudal at the level of the velum interpositum cerebri (Losseff et al., 1996). This site was chosen as it covers a large proportion of the lateral ventricles and cortical sulci, and the velum interpositum cerebri is thought to be a stable landmark despite ongoing atrophy, allowing repositioning for serial assessment.

Statistical analyses

Non-parametric statistics were employed for all analyses. The cognitive abilities of the patients at each time point, that is, year 0 and year 1, year 0 and year 2, and year 1 and year 2, were compared using a Wilcoxon signed-rank test, 2-tailed, with a Bonferroni correction for multiple tests (Bland and Altman, 1995). The impairment index used previously (Camp et al., 1999) was employed to provide a composite measure of cognitive dysfunction. The mean and standard deviation for each cognitive variable was derived from the matched control data collected at the start of the study. For each test, 0 was used to investigate the relationship between change in cognitive, mood and MRI parameters.

Results

Ninety-nine of the 147 patients approached returned for follow-up appointments, at year 1 and year 2, a return rate of 67.3%. Table 1 details the demographic, clinical, and MRI characteristics of these patients. Of the 158 PPMS patients originally enrolled in the study, 147 were approached for reassessment each year. The 48 patients who dropped out of the study were significantly older (Mann–Whitney with a Bonferroni correction) than those who returned for all assessments. Despite this, the clinical characteristics of disease duration and level of disability were comparable between the two groups. With respect to their cognitive skills at baseline, those who declined to complete the follow-up appointments scored significantly less (Mann–Whitney with a Bonferroni correction) than those who remained in the study on tests of information processing (SDMT; Symbol Digits Modalities Test) and spatial reasoning (spatial VESPAR).

Cognitive results

Table 2 details the scores at year 0, year 1 and year 2, of the 99 patients who returned for all three neuropsychological assessments. There were no significant differences (with one exception) between the scores of the patients at year 0, compared with year 1, year 0 relative to year 2, or between year 1 and year 2 (Wilcoxon signed-rank test, with a Bonferroni correction for multiple tests (Bland and Altman, 1995; Table 2). There was a significant difference only between the mean MADRS score at year 1 and year 2 (P = 0.012).

Although the mean scores at the three time points were not significantly different for the patients taken together as one group, it is possible that the results of the patients who showed
no cognitive deterioration were masking the results of those who did exhibit decline. An individual cognitive change index was derived for each patient, by comparing their BRB impairment index at year 0 with that at year 1 and year 2, respectively (see Statistical analyses above). The change index gave an indication of the amount and direction of change in cognition over the 2 years (Table 3). Broadly similar numbers experienced deterioration, stability and improvement. To examine the number of cognitive skills affected, patients were categorized according to the number of tests that they failed. The pattern of cognitive competence and impairment for a large number of patients remained stable during the 2 year period (Table 4). Of the 73 patients with complete BRB data, 35 patients failed the same number of tests, while 15 patients failed fewer and 23 patients failed more. Of the 67 patients with complete VESPAR data, 54 patients remained unchanged, nine patients showed an improved performance, and four deteriorated. In terms of general cognitive performance, 52 patients were classed as intact at baseline (BRB), with only four of these cognitively impaired at year 2. Of the 21 patients impaired at baseline (BRB), that is failing 3 or more tests, six who were minimally impaired returned to the normal range, while the 15 with moderate to severe impairment remained in the impaired range [either improved slightly, but were still classed as impaired (N = 3), deteriorated (N = 6) or remained unchanged (N = 6)].

Mean scores on the MADRS (Montgomery and Asberg, 1979) did not differ significantly when comparing year 0 and year 2 (the means for the 3 years were 7.6, 8.0 and 6.5). The majority of participants was categorized as not depressed (53, 55 and 65% for the 3 years scoring ≤6). A significant proportion was mildly depressed (43, 40 and 28% scoring 7–19). A small minority was moderately depressed (4, 5 and 7% scoring 20–34). No participant scored within the severely depressed range (35–60). There were no significant correlations between the cognitive impairment index and disability, as assessed by the EDSS (Kurtzke, 1983). The cognitive impairment index was significantly related to depression, but only for 1 year (Spearman’s rho: year 0, 0.218, P = 0.072; year 1, 0.437, P = 0.042; year 2, 0.105, P = 0.434).

### Cognitive and MRI parameters

Using Spearman’s rank correlation coefficient, the relations reported at baseline between MRI parameters and cognitive impairment remained, that is, there were moderate correlations between cognitive impairment and T1 hypointensity load (r = 0.373; P = 0.005), T2 lesion load (r = 0.383; P = 0.003), and partial cerebral volume (r = −0.243; P = 0.066). With respect to absolute change in cognitive scores and absolute change in MRI parameters, the only significant correlations were between: change in T1 hypointensity load

### Table 2 Mean (SD) cognitive scores at baseline, and at years 1 and 2

<table>
<thead>
<tr>
<th>Table 2 Mean (SD) cognitive scores at baseline, and at years 1 and 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verbal memory</strong></td>
</tr>
<tr>
<td>SRT: long-term storage</td>
</tr>
<tr>
<td>SRT: consistent long-term retrieval</td>
</tr>
<tr>
<td>SRT: delayed recall</td>
</tr>
<tr>
<td>Spatial memory</td>
</tr>
<tr>
<td>10/36: total immediate recall</td>
</tr>
<tr>
<td>10/36: delayed recall</td>
</tr>
<tr>
<td>Attention</td>
</tr>
<tr>
<td>SDMT: total</td>
</tr>
<tr>
<td>PASAT: three second rate</td>
</tr>
<tr>
<td>PASAT: two second rate</td>
</tr>
<tr>
<td>Verbal fluency: WLG</td>
</tr>
<tr>
<td>Reasoning</td>
</tr>
<tr>
<td>VESPAR: verbal total</td>
</tr>
<tr>
<td>VESPAR: spatial total</td>
</tr>
<tr>
<td>VESPAR: total</td>
</tr>
</tbody>
</table>

SRT, Selective Reminding Test; 10/36, 10/36 Spatial Recall Test; SDMT, Symbol Digits Modalities Test; PASAT, Paced Auditory Serial Addition Task; WLG, Word List Generation Test; VESPAR, Verbal and Spatial Reasoning Test; MADRS, Montgomery and Asberg Depression Rating Scale (Montgomery and Asberg, 1979).

### Table 3 Individual cognitive change, derived by calculating change on the BRB cognitive impairment index for each participant between year 0 and year 1, and year 0 and year 2

<table>
<thead>
<tr>
<th>Table 3 Individual cognitive change, derived by calculating change on the BRB cognitive impairment index for each participant between year 0 and year 1, and year 0 and year 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRB index</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Year 0–I</td>
</tr>
<tr>
<td>Breakdown</td>
</tr>
<tr>
<td>Breakdown</td>
</tr>
<tr>
<td>Year 0–2</td>
</tr>
<tr>
<td>Breakdown</td>
</tr>
</tbody>
</table>
| Only patients completing the entire BRB on both occasions are included in this table.
Longitudinal study of cognition in multiple sclerosis

Table 4 Number of cognitive tasks failed at baseline and at year 2 (only patients completing all tasks in the battery were included in these analyses)

<table>
<thead>
<tr>
<th>BRB</th>
<th>At year 0 no. of tasks failed</th>
</tr>
</thead>
<tbody>
<tr>
<td>At year 2 no. of tasks failed</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VESPAR</th>
<th>At year 0 no. of tasks failed</th>
</tr>
</thead>
<tbody>
<tr>
<td>At year 2 no. of tasks failed</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

and (i) delayed spatial recall memory ($r = -0.260; P = 0.020$), (ii) SDMT ($r = -0.266; P = 0.017$) and (iii) change in 3 s Paced Auditory Serial Addition Task (PASAT) ($r = -0.303; P = 0.012$), and change in partial brain volume and change in spatial VESPAR ($r = -0.242; P = 0.047$).

When comparing the MRI characteristics of those cognitively impaired and intact on BRB at baseline, there were significant differences (Mann–Whitney with a Bonferroni correction) between the mean values for all MRI parameters, at both year 0 and year 2. Patients with cognitive impairment at year 0 not only exhibited significantly more pathology at baseline, but these differences were maintained at 2 years. They did not, however, demonstrate significantly more change in pathology, that is, greater MRI change values.

There were no significant correlations between a composite global measure of cognitive change and absolute or percentage change in MRI parameters. Categorizing patients according to whether their baseline pathology, as evaluated using MRI, was less than, greater than, or equal to the median value for all MRI parameters, at both year 0 and year 2. Patients with cognitive impairment at year 0 performed worse on the cognitive tasks at year 2 than patients with less pathology. Applying the categorization to baseline brain volume values, patient performance differed significantly at year 2 on the spatial recall test, SDMT, and both levels of the PASAT (Mann–Whitney with a Bonferroni correction). Using the categorization with year 0 values for $T_1$ lesion load, there were significant differences between the mean scores of the patients at year 2 on the tests of verbal and spatial memory, attention and verbal fluency (Mann–Whitney with a Bonferroni correction). Applying the classification to baseline $T_2$ lesion load values, mean scores were significantly different at year 2 on the SDMT, and the word list generation (WLG) Task (Mann–Whitney with a Bonferroni correction). When categorizing change in MRI measures, at or above median change in MRI or below the median change, there were significant differences only between year 2 scores on the SDMT (change in $T_1$ hypointensity load, change in $T_2$ lesion load) (Mann–Whitney with a Bonferroni correction).

Discussion

Cognitive results

There were no significant differences between the mean scores of neuropsychological test performance of patients with PPMS at year 0 and year 2 (Table 2). This would appear to suggest preservation of cognitive skills over this period, as reported in other subtypes of multiple sclerosis (Filley et al., 1990; Jennekens-Schinkel et al., 1990; Mariani et al., 1991; Mattioli et al., 1993; Amato et al., 1995; Hohol et al., 1997). However, the typical course of cognitive function over this relatively short time scale (2 years) may remain unchanged, perhaps especially so in PPMS, where cerebral lesions have been reported to sustain less inflammation than in SPMS (Revesz et al., 1994). It may be that the observation period of 2 years was too short to identify significant change and extensive links to clinical and MRI parameters (Ingle et al., 2003). It is possible that the lack of significant deterioration in cognitive ability between year 0 and 2 may have been partly attributable to bias in the sample. At recruitment to the study, the cohort had been diagnosed for an average of 10 years and, therefore, it is possible that significant cognitive decline may have started and slowed in some individuals. The cohort demonstrated significant cognitive impairment on eight of the eleven cognitive measures at baseline (Camp et al., 1999). The relatively large group of patients who did not remain over the 2 years of the study was more impaired on two of the eleven cognitive measures at baseline, than the group who remained. Those patients who were more cognitively impaired at year 0, and who did not return, may have continued to decline (Kujala et al., 1997). In addition, patients who were cognitively intact at baseline but experienced cognitive decline during the course of the study may have been more likely to withdraw from the study. Cognitive impairment has been demonstrated to be a major predictor of subject attrition in other neurological samples (Levin et al., 2000).

Therefore, the results of the analysis may be conservative, as data from those individuals most likely to show a marked result were unavailable.

On further inspection, variability in cognitive performance over time was noted (Tables 3 and 4), in line with previous longitudinal studies in RRMS and SPMS patients (Jennekens-Schinkel et al., 1990; Feinstein et al., 1993). Approximately one-third of the patients showed a numerical decline in test scores and a similar proportion failed more tests after 2 years. If the possibility of practice effects raising scores and a drop out of more cognitively impaired individuals are accepted as likely to have reduced the incidence of cognitive impairment in this experimental cohort, then the identification of one-third of the individuals demonstrating individual cognitive decline could be taken as an indication of a marked...
occurrence and progression of cognitive impairment in middle stage PPMS. It is noteworthy that in a cohort of early RRMS patients, about a third worsened cognitively over 2 years (Zivadinov et al., 2001).

Cognitive level at entry into the study appeared to be a relatively good indicator of cognitive status at 2 years. Those patients classed as cognitively intact on entry demonstrated preservation of these skills at 2 years. Those with minimal impairment at baseline returned to the normal range. In contrast, those who were categorized as moderately or severely impaired at baseline continued to exhibit cognitive deterioration. These findings provide some support for the study by Kujala et al. (1997) using a heterogeneous multiple sclerosis sample. They reported cognitive stability at 3 years in those patients intact at year 0, although almost all of the patients impaired at baseline had suffered further cognitive deterioration.

When comparing patient performance in the current study with that of the placebo group of Weinstein et al. (1999), the PPMS patients did not exhibit comparable practice effects. Whereas the PPMS group was significantly below matched healthy controls at the start of the study, Weinstein et al.’s RRMS group was not and therefore possibly more likely to exhibit a group practice effect. That the scores of the PPMS patients appeared to remain stable may even indicate that the patients were in fact performing worse at year 2, but that the practice effects were artificially elevating their scores to a comparable level with that at baseline. PPMS patients performed worse than the normative sample of Boriniga et al. (2001) on both version B, completed at year 0, and A, completed at year 1, of the BRB. The poor performance of the patients may be less marked on version A as the effects of practice in the patient group have not been controlled.

There are a number of problems associated with accurately evaluating the results obtained from a longitudinal neuropsychological study. Standardized administration of tasks and scoring procedures across centres, and the administration of the tests by the same examiner within a centre, reduce some of the artefactual sources of variability. Practice effects, that is, the enhanced scoring of a subject owing to familiarity with the task and/or the stimuli, may also bias performance. This study attempted to reduce item familiarity by employing alternate forms of the BRB. These forms have been shown to be equivalent in some studies (Rao, 1990; Bever et al., 1995), although Boriniga et al. (2001) reported significant differences between the mean scores obtained on versions A and B on the 10/36 Selective Reminding Test, SDMT, and WLG in a normative sample. However, the current study utilized the same version for baseline and final assessment, which minimizes the bias of varying difficulty of ‘parallel’ forms. Task familiarity is more difficult to address, and can only really be solved by the use of controls assessed serially, so that any difference in the two groups’ serial performance can be examined. Longitudinal control data were not collected in the current study. The sensitivity of the neuropsychological test battery to detect significant change in ability over time must also be considered. Bever et al. (1995) suggested that the BRB may be useful in serial studies of cognitive ability. However, they reported practice effects on the PASAT between trials 1 and 2, and commented on the need for testing of large control and multiple sclerosis patient samples.

Performance on cognitive tasks may also be affected by mood (Anastasi, 1997). Scores on the MADRS (Montgomery and Asberg, 1979) were unremarkable, and depression was only significantly related to cognitive impairment in year 1. The small numbers with moderate depression, taken with the absence of severe depression, support Vleugels et al.’s (1998) finding that PPMS patients were significantly less depressed than the SPMS patients, suggesting this may be owing to the more predictable nature of PPMS.

**Cognitive and MRI parameters**

There were few significant correlations between change in neuropsychological performance and change in the MRI parameters. There were no corrections for multiple comparisons in the MRI analysis, which raises the possibility of spurious significant results. However, these findings support other cognitive and MRI longitudinal studies in different multiple sclerosis subtypes (Mariani et al., 1991; Mattioli et al., 1993). The few significant correlations between cognitive change and absolute change in MRI parameters suggest that the relationship between change in pathology and cognitive ability is complex, involving multiple factors. Although measurable change in several MRI parameters was demonstrated in the PPMS and transitional progressive multiple sclerosis (TPMS) cohort within just 1 year (Stevenson et al., 2000), this did not correlate with definite clinical change (EDSS status). Filippi et al. (1998) reported that changes in MRI parameters are much more sensitive than clinical changes, and perhaps this is also applicable to cognitive skills. More sophisticated MRI techniques might have demonstrated more significant links.

The correlations between cognitive ability and MRI parameters in cross-sectional studies of PPMS patients are moderate (Camp et al., 1999); therefore, it is unsurprising that there is limited significance in the relationship between absolute change in MRI parameters and change in cognitive performance. The limitations of conventional MRI techniques and those of the serial administration of neuropsychological tests are magnified when examining change over a 2 year time period. Changes in normal appearing brain tissue have been reported to contribute significantly to cognitive deterioration in patients with multiple sclerosis (Filippi et al., 2000). Thus, the poor demonstration of microscopic changes in normal appearing white matter, and the failure to address the integrity of cortical tracts may account for the current findings, although the measure of atrophy may be influenced by events in these tissues. The lack of change on cognitive measures over the 2 years may also contribute to the low correlation with change in imaging measures.
Patients with a higher lesion load at year 0 exhibited more cognitive impairment at 2 years than those with a lower lesion load at baseline. However, categorizing patients according to their change in pathology between year 0 and year 2 did not reveal any separation with respect to cognitive abilities. These findings perhaps suggest that it is the initial amount of pathology, at entry into the study, which impacts on cognitive function over time, rather than the absolute change during the study period.

The current study is the only one to date to examine serially cognitive function in PPMS patients. Despite the relatively short follow-up period results indicate that initial neuropsychological status may predict cognitive ability at 2 years. The relationship status change in cognition and change in MRI variables involves many factors. Further assessments may elucidate these complexities.

Acknowledgements
This work was part of MAGNIMS (Magnetic Resonance Networks in Multiple Sclerosis) funded by an EC initiative. The Institute of Neurology NMR Research Unit and the posts held by SJC and VLS were funded by the Multiple Sclerosis Society of Great Britain and Northern Ireland. The work in Amsterdam was supported by the Foundation Friends of MS Research, in Barcelona was supported by Fundación Esclerosis Múltiple (FEM), in Bordeaux was supported by a grant from Ligue Française Contre la sclérose en plaques, and in Milan by the Fondazione Italiana Sclerosi Multipla (FISM).

References
Huijbregts SCJ, Kalkers NF, de Sonneville LMJ, de Groot V, Reuling IEW, Polman CH. Differences in cognitive impairment of relapsing remitting, secondary, and primary progressive MS. Neurol 2004; 63: 335–9.


