Interferon Beta-1b for the Treatment of Primary Progressive Multiple Sclerosis

Five-Year Clinical Trial Follow-up

Carmen Tur, MD; Xavier Montalban, PhD; Mar Tintore, PhD; Carlos Nos, MD; Jordi Rio, PhD; Francesc Xavier Aymerich, MSc; Luis Brieva, PhD; Nieves Tellez, PhD; Hector Perhal, MD; Manuel Comabella, MD; Ingrid Galan, MD; David Calle, BSc; Jaume Sastre-Garriga, MD; Alex Rovira, PhD

Objectives: To investigate, during the 5-year period without treatment after termination of a 2-year clinical trial of interferon beta-1b for the treatment of primary progressive multiple sclerosis, differences in the evolution of clinical variables and magnetic resonance imaging results between trial arms and to investigate correlations between in-trial changes in Multiple Sclerosis Functional Composite (MSFC) score and magnetic resonance imaging variables and Expanded Disability Status Scale (EDSS) score evolution.

Design: Five-year clinical trial follow-up.

Setting: Clinical Neuroimmunology Unit, Multiple Sclerosis Centre of Catalonia, Autonomous University of Barcelona, Spain.

Patients: Seventy-three patients received interferon beta-1b or placebo during the trial.

Main Outcome Measures: After 5 years without treatment, the EDSS and MSFC measures were scored for 63 and 59 patients, respectively. Neuropsychological and magnetic resonance imaging assessments were performed for 59 and 50 patients, respectively.

Results: After 5 years without treatment, the interferon beta-1b group had better 9-Hole Peg Test (P = .02) and Word List Generation Test (P < .001) scores, and their magnetization transfer ratio measures in the normal-appearing white matter were significantly higher (P = .02, P = .009, and P = .03 for the mean, peak location, and peak height magnetic transfer ratios, respectively). During the entire study period (from trial baseline to assessment at 5 years without treatment), the placebo group showed a greater decrease in brain parenchymal fraction (P = .004). The in-trial increase of lesions correlated with the worsening of the EDSS score during the 5-year period without treatment (P = .004).

Conclusions: Modest but beneficial effects of interferon beta-1b on clinical variables and brain atrophy development were observed 5 years after trial termination. Moreover, in-trial lesion activity correlated with EDSS progression after trial termination. Therefore, we provide evidence to consider immunomodulation as a sensible approach to treat primary progressive multiple sclerosis.

Arch Neurol. 2011;68(11):1421-1427

Beneficial effects of interferon beta on relapse-related outcome measures have been shown for patients with secondary progressive multiple sclerosis (MS) with superimposed relapses, whereas positive results regarding disability progression outcomes have been less frequently reported. In primary progressive MS (PPMS), few randomized controlled clinical trials have been published. Trials with interferon beta-1a and glatiramer acetate failed to show positive effects on the Expanded Disability Status Scale (EDSS). A trial with rituximab, a monoclonal antibody that selectively depletes B cells, failed to show an effect on confirmed disease progression, although a subgroup analysis showed a treatment effect in younger patients with gadolinium-enhanced lesions. Finally, a single-center trial with interferon beta-1b in PPMS also showed no significant treatment effect on EDSS, although a positive effect on Multiple Sclerosis Functional Composite (MSFC) scores and magnetic resonance imaging (MRI) variables was observed.

The response to interferon beta can be highly heterogeneous, reflecting the presence of multiple complex processes responsible for treatment effects. Differences in the underlying condition may be one of the explanations for the dissimilar effects of interferon beta in relapsing vs progressive MS. However, another reason for such dissimilarities could be the too-short time frames of the clinical trials performed so far. The need for progression-
based-only primary outcome measures in progressive MS might imply the need for longer study periods to observe clinically meaningful treatment effects. Moreover, the poor sensitivity of the EDSS to measure clinical changes contributed to the failure in detecting disease progression. Instead, the MSFC would be a more sensitive tool to measure clinical worsening, especially when combined with the EDSS.

We have assessed clinically, neuropsychologically, and via MRI those patients who participated in the 2-year double-blind phase of a trial of interferon beta-1b 5 years after trial termination, with the general objective of exploring the long-term effects of their treatment. More specifically, we aimed to investigate whether those clinical and MRI differences between treatment groups observed during the trial were still present after 5 years without treatment; differences in clinical and MRI evolution between treatment groups from trial baseline to the 5-year-without-treatment follow-up; and whether the intratrial changes in MSFC and MRI variables, considering the whole cohort, correlated with progression as measured by EDSS score after trial termination.

METHODS

PATIENTS

All patients who had completed the double-blind phase of the initial trial (67 of 73) were invited to participate in this study. Patients included in the feeder study had a PPMS course, allowing for a single attack before study entry. All patients gave informed written consent before the study.

ASSESSMENTS

Clinical

Patients were assessed 5 years after the last visit of the clinical trial at the Clinical Neuroimmunology Unit of the Multiple Sclerosis Centre of Catalonia at the Autonomous University of Barcelona, Spain. We obtained MSFC and z subscores (9-Hole Peg Test [9-HPT], Paced Auditory Serial Addition Test [PASAT], and timed 25-ft walk test), ambulation index score, EDSS scores, and safety variables, such as those obtained from the Ashworth Scale, the Beck Depression Inventory, the Fatigue Severity Scale, and the Sickness Impact Profile, given that interferon beta has been found in some studies to increase spasticity, fatigue, and depressive symptoms. The MSFC and z subscores were calculated according to published guidelines, using our baseline sample (before treatment) as a reference. The z subscores (ie, standardized scores) gave us a measure that related every individual's performance to the average performance in the population. The z subscores indicated the number of standard deviation units from which a patient's score was less than (if negative values) or greater than (if positive values) the average performance. For those tests measured in seconds (ie, the 9-HPT and the timed 25-ft walk test), the final z subscore was obtained after changing its sign because higher scores in seconds would indicate worse clinical performance. On the basis of the approximately normal distribution in our baseline sample of the original numbers before transformation to z scores and based on their change over the trial period, z subscores were considered to be reliable measures.

The EDSS was administered by telephone to those patients too disabled to be examined in person. Three months later, all patients were assessed again with the EDSS to obtain the confirmed EDSS score, which was the lowest score of the 2 assessments (hereafter referred to as the EDSS score). The EDSS scores of those patients who died after trial completion and those who did not participate in the follow-up study were not included in the analysis.

Neuropsychological

The Brief Repeatable Battery of Neuropsychological Tests was administered by a neuropsychologist who was unaware of previous patient allocation and not in communication with the treating or examining neurologists. It provided measurements for attention and speed of information processing (PASAT and Symbol Digit Modalities Test), verbal learning and late recall (Selective Reminding Test), visuospatial learning and late recall (10136 Spatial Recall Test), and semantic fluency (Word List Generation Test [WLGT]). Cognitive assessments never were performed during relapses or treatment with corticosteroids.

MRI and Analysis

All patients were invited to undergo an MRI examination (Magnetom Vision Plus 1.5T; Siemens AG, Erlangen, Germany). The following sequences were obtained: transverse, T2-weighted, fast-acquisition, dual-echo sequence (repetition time, 3000 milliseconds; echo time, 14-85 milliseconds; echo train length, 5; and acquisitions, 1); transverse, T1-weighted, spin-echo sequence (repetition time, 667 milliseconds; echo time, 14 milliseconds; and acquisitions, 2); and 2-dimensional, gradient-echo, pulse sequence (repetition time, 805 milliseconds; echo time, 12 milliseconds; flip angle, 30°; and acquisitions, 1), repeated with an additional off-resonance preparation pulse to saturate bound protons. The variables of this saturation pulse are described elsewhere. The difference between the two 2-dimensional gradient-echo sequences, with and without the saturation pulse, was used to obtain the magnetization transfer images.

The T2 lesion masks initially were analyzed and marked on proton density–weighted hard copies cross-referenced to the T2-weighted images using Displimage (D.L. Plummer, University College London, England). From these masks, T1-weighted lesion masks were obtained as the hypointense region in T1-weighted images located within the T2-weighted lesion mask. T1-weighted and T2-weighted lesion volumes (T1LV and T2LV, respectively) were calculated from these lesion masks. Brain atrophy was evaluated by measuring the brain parenchymal fraction (BPF) according to a previously described algorithm. Total intracranial volume was calculated as normal-appearing white matter (NAWM) + normal-appearing gray matter (NAGM) + cerebrospinal fluid (CSF) + lesion mask. The BPF was calculated as follows: (total intracranial volume - NAGM + NAGM + lesion mask)/total intracranial volume. From magnetization transfer images, we calculated the magnetization transfer ratio (MTR) as follows: MTR = 100 (Mo - Ms)/Mo, in which Mo and Ms are the signal intensities for normal-appearing white matter (NAWM) + normal-appearing gray matter (NAGM) + cerebrospinal fluid (CSF) + lesion mask. The BPF was calculated as follows: (total intracranial volume - NAGM + NAGM + lesion mask)/total intracranial volume.
sue on a proton density–weighted image. Further details of the MRI protocol have been reported elsewhere.

**STUDY END POINTS AND STATISTICAL ANALYSIS**

**Analysis at 5-Year-Without-Treatment Follow-up**

The EDSS score and the z scores of the MSFC and its subtests, radiologic variables, and neuropsychological test scores were compared between treatment groups by the Wilcoxon rank sum test. Demographic variables were compared, if quantitative, by means of a 1-way analysis of variance (the data analyzed met the assumption of analysis of variance; otherwise, the Kruskal-Wallis [for independent groups] would have been used) and, if dichotomous or qualitative, using the $\chi^2$ test.

**Changes From the End of the Trial and Trial Baseline to the 5-Year-Without-Treatment Follow-up**

Changes in EDSS score, the z scores of the MSFC and its subtests, and MRI variables were calculated. According to EDSS score change from trial baseline, patients were classified as responders and nonresponders. Nonresponders were defined according to 3 criteria: criterion A: at least 1–EDSS point increase if the baseline EDSS score was 3 or lower or 0.5-point increase if the baseline EDSS score was 3.5 or higher (1 step change); criterion B: 2.5–EDSS point increase if the baseline EDSS score was 0, 2.0-point increase if it was 1.0 to 4.0, 1.5-point increase if it was 4.5 to 5.0, and 1.0-point increase if it was 5.5 or higher (2 step changes); and criterion C: 3.5–EDSS point increase if the baseline EDSS score was 0, 3.0-point increase if it was 1.0 to 3.0, 2.5-point increase if it was 3.5 to 4.0, 2-point increase if it was 4.5 to 5.0, and 1.5-point increase if it was 5.5 or higher (3 step changes).

The MSFC changes from trial baseline were calculated from month 6 of the trial to avoid interference with learning effects. Wilcoxon tests (between-groups comparison) and Wilcoxon rank sum tests (intragroup comparison) were used to compare the change in MSFC and EDSS scores, neuropsychological test scores, safety, and MRI variables over time between treatment groups. Two-sided statistical tests were used.

**Correlation Between In-Trial Changes and Progression as Measured by EDSS Score After Trial Termination**

Correlations between in-trial changes in MSFC score, T1LV, T2LV, BPF, and total number of ALs and changes in EDSS scores from trial termination to 3-year-without-treatment follow-up were performed, considering the whole sample of patients independently of their treatment group. Spearman correlation coefficients were calculated.

All statistical analyses were performed using SAS statistical software, version 8.2 (SAS Institute, Inc, Cary, North Carolina). The level of statistical significance was set at $P < .05$. All assumptions for each test (when necessary) were checked for violations.

**RESULTS**

**PATIENT DISPOSITION AND DEMOGRAPHIC CHARACTERISTICS**

Sixty-seven of the 73 initially randomized patients had completed the clinical trial.

**STUDY END POINTS**

**Assessment at 5-Year-Without-Treatment Follow-up**

Patients from the interferon beta-1b group had better scores on the 9-HPT ($P = .02$) and the WLGT ($P < .001$). No significant differences were observed in any other clinical or neuropsychological variable (Table 1). Regarding the MRI variables, no significant differences were observed in T1LV, T2LV, the proportion of patients with ALs, or the mean number of ALs per patient. Patients in the interferon beta-1b group showed significantly higher MTR measurements (Table 1).

**Changes From Trial Termination to 5-Year-Without-Treatment Follow-up**

Both groups scored worse on the EDSS ($P = .002$ for the interferon beta-1b group and $P < .001$ for the placebo group) and MSFC ($P < .001$, for both groups) measures. Such worsening, as well as the proportion of nonresponders, was not significantly different between groups (Table 2).

**Changes From Trial Baseline to 5-Year-Without-Treatment Follow-up**

Both groups scored worse on the EDSS ($P < .001$) and MSFC ($P < .001$ for the interferon beta-1b group and $P = .001$ for the placebo group) measures except on the PASAT z subtest. However, neither this worsening nor the proportion of nonresponders was significantly different between groups (Table 2).

Both groups increased their T1LV and T2LV over time ($P < .001$ for both groups) and worsened regarding their BPF ($P < .001$). Some improvement was observed in the NAWM mean MTR ($P = .001$) and the NAWM peak location MTR ($P < .001$) in the interferon beta-1b arm (Table 2).
Table 1. Demographic, Clinical, Cognitive, and MRI Variables at 5-Year-Without-Treatment Follow-up

<table>
<thead>
<tr>
<th>Variablea</th>
<th>Interferon Beta-1bα</th>
<th>Placeboα</th>
<th>P Valuec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>57 (40 to 71)</td>
<td>55 (39 to 71)</td>
<td>.34</td>
</tr>
<tr>
<td>Male, No. (%)</td>
<td>11 (37)</td>
<td>18 (62)</td>
<td>.05d</td>
</tr>
<tr>
<td>Disease duration, y</td>
<td>16 (8 to 39)</td>
<td>16 (8 to 39)</td>
<td>.73</td>
</tr>
<tr>
<td>EDSS scoree</td>
<td>6.5 (4.0 to 9.5)</td>
<td>6.5 (4.0 to 9.5)</td>
<td>.84</td>
</tr>
<tr>
<td>9-HPT z score</td>
<td>0.32 (−3.39 to 2.19)</td>
<td>−0.30 (−3.29 to 1.93)</td>
<td>.02</td>
</tr>
<tr>
<td>PASAT z score</td>
<td>0.48 (−2.19 to 1.72)</td>
<td>0.55 (−2.19 to 1.72)</td>
<td>.66</td>
</tr>
<tr>
<td>MSFC z score</td>
<td>−0.53 (−3.07 to 1.29)</td>
<td>−0.49 (−3.07 to 1.18)</td>
<td>.46</td>
</tr>
<tr>
<td>Ambulation Index score</td>
<td>0.78 (−0.59 to 13.70)</td>
<td>−0.13 (−0.54 to 13.70)</td>
<td>.62</td>
</tr>
<tr>
<td>BDI score</td>
<td>8 (0 to 34)</td>
<td>8.5 (1 to 31)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>SIP score</td>
<td>20.23 (9.59 to 36.70)</td>
<td>28.99 (0 to 43.48)</td>
<td>.19</td>
</tr>
<tr>
<td>Fatigue Severity Scale score, mean (SD)</td>
<td>5.47 (1.07)</td>
<td>4.97 (1.31)</td>
<td>.20</td>
</tr>
</tbody>
</table>

*Brief Repeatable Battery of Neuropsychological Tests scoresa

- Spatial Recall Test: 18 (10 to 27) vs. 17 (8 to 28) (P = .73)
- Spatial Recall Test: delayed recall: 6 (0 to 10) vs. 5 (1 to 10) (P = .50)
- SDMT: 24 (13 to 37) vs. 22 (11 to 38) (P = .66)
- WLG: 19.0 (9.0 to 28.0) vs. 11.5 (6.0 to 22.0) (<.001)
- Selective Reminding Test: 87 (19 to 145) vs. 80 (4 to 145) (P = .22)

*MRI variablesa

- T1LV, mm3: 6071.09 (411.99 to 29276.90) vs. 10448.46 (1161.58 to 35599.50) (P = .22)
- T2LV, mm3: 14502.53 (1553.50 to 74947.40) vs. 25311.48 (4594.8 to 82772.20) (P = .16)
- Patients with ALs, No. (%): 9 (30) vs. 8 (28) (P = .76d)
- No. of ALs: 0 (0 to 12) vs. 0 (0 to 10) (P = .83)
- BPF: 72.86 (58.13 to 79.13) vs. 70.71 (58.87 to 75.75) (P = .07)
- NAWM mean MTR: 39.54 (37.35 to 40.82) vs. 38.64 (35.67 to 40.69) (P = .02)
- NAWM PL MTR: 40 (38 to 42) vs. 39 (36 to 41) (P = .09)
- NAWM PH MTR: 0.09 (0.07 to 0.10) vs. 0.08 (0.06 to 0.10) (P = .03)

Abbreviations: ALs, active lesions (defined as new and enlarging lesions observed in the current scan [at 5-year-without-treatment follow-up] when compared with the previous scan performed at the end of the clinical trial); BDI, Beck Depression Inventory; BPF, brain parenchymal fraction; EDSS, Expanded Disability Status Scale; MRI, magnetic resonance imaging; MSFC, Multiple Sclerosis Functional Composite; MTR, magnetization transfer ratio; NAWM, normal-appearing white matter; 9-HPT, 9-Hole Peg Test; PASAT, Paced Auditory Serial Addition Test; PH, peak height; PL, peak location (mean PL and PH are MTR histogram variables); SDMT, Symbol Digit Modalities Test; SIP, Sickness Impact Profile; T1LV, T1-weighted lesion volume; T2LV, T2-weighted lesion volume; WLG, Word List Generation Test.

a All values shown as median (range), except where indicated.

For MRI variables, n=25 for each group; for the other variables, n=30 for the interferon beta-1b group and n=29 for the placebo group.

c The Wilcoxon test was used for all comparisons except for sex and patients.

d Test.

e Sample sizes are 32 for the interferon beta-1b group and 31 for the placebo group.

f Measures long-term storage plus consistent long-term storage plus total delay.

the placebo group worsened on the delayed version of the Spatial Recall Test (P = .03), the Symbol Digit Modalities Test (P < .001), and the WLG (P < .001) (Table 3). Changes in the delayed version of the Selective Reminding Test were significantly different between the 2 groups (P = .04) in favor of the interferon beta-1b group. No other significant differences were seen between groups (Table 3).

Correlation Between In-Trial MRI and Clinical Changes and EDSS Changes After Trial Termination

The total number of ALs during the trial significantly correlated with EDSS changes from trial baseline to 5-year-without-treatment follow-up (correlation coefficient=0.36, P = .004). No other significant correlations were obtained (eTable 1; http://www.archneurol.com).

At 5-year-without-treatment follow-up, the interferon beta-1b group scored higher on the Ashworth scale than the placebo group. However, no differences were observed in changes from trial baseline to 5-year-without-treatment follow-up in any of the safety variables (eTable 2).

Our results suggest that treatment with interferon beta-1b could have a modest long-term effect on clinical and MRI variables in patients with PPMS. The first piece of evidence supporting this statement is the fact that after a period of 5 years without treatment, patients from the interferon beta-1b group had better scores on the 9-HPT and on some neuropsychological tests. This finding is in keeping with that of patients taking interferon beta-1a who obtained better MSFCz scores than the placebo group in all time points from month 6 to month 24 during the double-blind phase of this trial.8 Of interest, the IMPACT (Intramuscular Interferon Beta-1a in Secondary Progressive MS) trial9 and the trial of interferon beta-1a in PPMS10 also showed a positive effect of interferon beta on the 9-HPT, and a trial with low-dose methotrexate in progressive MS showed less clinical progression as measured by the 9-HPT in those patients who received treatment.11 It was then suggested that measures of upper limb function might be more sensitive to change than measures of lower limb function because patients with PPMS
Table 2. Clinical and MRI Changes Over Time

<table>
<thead>
<tr>
<th>Variable</th>
<th>Interferon Beta-1b&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Placebo&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Comparison of Interferon Beta-1b vs Placebo&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Interferon Beta-1b&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Placebo&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Comparison of Interferon Beta-1b vs Placebo&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>Change&lt;sup&gt;b&lt;/sup&gt;</td>
<td>p Value&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Change&lt;sup&gt;b&lt;/sup&gt;</td>
<td>p Value&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Change&lt;sup&gt;b&lt;/sup&gt;</td>
<td>p Value&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>MSFC z score</td>
<td>−0.36 (−2.22 to 0.48)</td>
<td>&lt;.001</td>
<td>−0.49 (−2.70 to 0.56)</td>
<td>&lt;.001</td>
<td>−0.41 (−2.30 to 1.28)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>9-HPT z score</td>
<td>−0.48 (−2.98 to 0.52)</td>
<td>&lt;.001</td>
<td>−0.67 (−2.15 to 0.46)</td>
<td>&lt;.001</td>
<td>−0.69 (.001)</td>
<td>−2.77 (0.31)</td>
</tr>
<tr>
<td>PASAT z score</td>
<td>−0.13 (.04)</td>
<td>.005</td>
<td>−0.13 (−4.17 to 0.39)</td>
<td>&lt;.001</td>
<td>−0.61 (−2.61 to 2.33)</td>
<td>.001</td>
</tr>
<tr>
<td>AI z score</td>
<td>0.36 (.04)</td>
<td>.84</td>
<td>0.14 (.04)</td>
<td>&lt;.001</td>
<td>0.76 (.007)</td>
<td>.07</td>
</tr>
<tr>
<td>EDSS score</td>
<td>0.8 (.02)</td>
<td>&lt;.001</td>
<td>0.5 (&lt;−2.0 to 2.5)</td>
<td>&lt;.001</td>
<td>1.5 (.001)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Nonresponders, %&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criterion A</td>
<td>59.26</td>
<td>...</td>
<td>53.85 (&lt;−2.0 to 2.5)</td>
<td>...</td>
<td>70.97 (38)</td>
<td>...</td>
</tr>
<tr>
<td>Criterion B</td>
<td>45.16</td>
<td>...</td>
<td>35.48 (&lt;−2.0 to 2.5)</td>
<td>...</td>
<td>58.06 (28)</td>
<td>...</td>
</tr>
<tr>
<td>Criterion C</td>
<td>16.13</td>
<td>...</td>
<td>&gt;.99 (&lt;−2.0 to 2.5)</td>
<td>...</td>
<td>25.81 (38)</td>
<td>...</td>
</tr>
<tr>
<td>MRI variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1LV</td>
<td>726.70</td>
<td>&lt;.001</td>
<td>952.72 (&lt;−2.0 to 2.5)</td>
<td>&lt;.001</td>
<td>832.55 (.001)</td>
<td>1856.80 (&lt;.001)</td>
</tr>
<tr>
<td>T2LV</td>
<td>466.30</td>
<td>.94</td>
<td>183.1 (.001)</td>
<td>.94</td>
<td>2265.9 (.001)</td>
<td>2986.90 (&lt;.001)</td>
</tr>
<tr>
<td>BPF</td>
<td>−1.33</td>
<td>&lt;.001</td>
<td>−2.36 (&lt;−4.79 to 2.00)</td>
<td>&lt;.001</td>
<td>−1.78 (&lt;−6.99 to 1.29)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM mean</td>
<td>0.61</td>
<td>&lt;.001</td>
<td>0.08 (&lt;−4.97 to 2.00)</td>
<td>&lt;.001</td>
<td>0.16 (&lt;−6.87 to 2.37)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM MTR</td>
<td>−1.35</td>
<td>&lt;.001</td>
<td>−1.30 (&lt;−8.21 to 2.23)</td>
<td>&lt;.001</td>
<td>−1.76 (&lt;−2.77 to 3.570)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM PL MTR</td>
<td>−0.001</td>
<td>&lt;.001</td>
<td>0 (&lt;−2.62 to 0.55)</td>
<td>&lt;.001</td>
<td>0.06 (&lt;−1.17 to 1.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM PH MTR</td>
<td>−0.001</td>
<td>&lt;.001</td>
<td>−0.003 (&lt;−3.0 to 3)</td>
<td>&lt;.001</td>
<td>−0.003 (&lt;−2.77 to 3.570)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>MRI variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1LV</td>
<td>726.70</td>
<td>&lt;.001</td>
<td>952.72 (&lt;−2.0 to 2.5)</td>
<td>&lt;.001</td>
<td>832.55 (.001)</td>
<td>1856.80 (&lt;.001)</td>
</tr>
<tr>
<td>T2LV</td>
<td>466.30</td>
<td>.94</td>
<td>183.1 (.001)</td>
<td>.94</td>
<td>2265.9 (.001)</td>
<td>2986.90 (&lt;.001)</td>
</tr>
<tr>
<td>BPF</td>
<td>−1.33</td>
<td>&lt;.001</td>
<td>−2.36 (&lt;−4.97 to 2.00)</td>
<td>&lt;.001</td>
<td>−1.78 (&lt;−6.99 to 1.29)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM mean</td>
<td>0.61</td>
<td>&lt;.001</td>
<td>0.08 (&lt;−4.97 to 2.00)</td>
<td>&lt;.001</td>
<td>0.16 (&lt;−6.87 to 2.37)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM MTR</td>
<td>−1.35</td>
<td>&lt;.001</td>
<td>−1.30 (&lt;−8.21 to 2.23)</td>
<td>&lt;.001</td>
<td>−1.76 (&lt;−2.77 to 3.570)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM PL MTR</td>
<td>−0.001</td>
<td>&lt;.001</td>
<td>0 (&lt;−2.62 to 0.55)</td>
<td>&lt;.001</td>
<td>0.06 (&lt;−1.17 to 1.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>NAWM PH MTR</td>
<td>−0.001</td>
<td>&lt;.001</td>
<td>−0.003 (&lt;−3.0 to 3)</td>
<td>&lt;.001</td>
<td>−0.003 (&lt;−2.77 to 3.570)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>MRI variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Al, Ambulation Index; BPF, brain parenchymal fraction; EDSS, Expanded Disability Status Scale; ellipses, not applicable; MRI, magnetic resonance imaging; MSFC, Multiple Sclerosis Functional Composite; MTR, magnetization transfer ratio; NAWM, normal-appearing white matter; 9-HPT, 9-Hole Peg Test; PASAT, Paced Auditory Serial Addition Test; PH, peak height; PL, peak location (mean PL and PH are MTR histogram variables); T1LV, T1-weighted lesion volume; T2LV, T2-weighted lesion volume.

<sup>a</sup>n=25.
<sup>b</sup>Median change (range).
<sup>c</sup>Wilcoxon test.
<sup>d</sup>Signed rank test.

<sup>e</sup>All clinical changes (except changes in the EDSS score) were considered from month 6 of the trial. Sample sizes are 32 for the interferon beta-1b group and 31 for the placebo group. For the MSFC z score and subscores, negative changes imply clinical worsening. For the EDSS and AI scores, positive changes imply clinical worsening.

<sup>f</sup>Depending on the amount of change of EDSS score, we defined responders and nonresponders according to 3 different criteria. For criteria A and B, we used the Signed rank test to calculate P values. For criterion C, we used the Fisher exact test to calculate P values.

Table 3. Changes in the Brief Repeatable Battery of Neuropsychological Test Scores Over Time

<table>
<thead>
<tr>
<th>Test</th>
<th>Change&lt;sup&gt;c&lt;/sup&gt;</th>
<th>p Value&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Change&lt;sup&gt;c&lt;/sup&gt;</th>
<th>p Value&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Change&lt;sup&gt;c&lt;/sup&gt;</th>
<th>p Value&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Recall Test</td>
<td>−1 (−7 to 7)</td>
<td>.43</td>
<td>−1 (−10 to 7)</td>
<td>.15</td>
<td>.44 (−4 to −26)</td>
<td>.96</td>
</tr>
<tr>
<td>Spatial Recall Test: delayed recall</td>
<td>0 (−4 to 4)</td>
<td>.56</td>
<td>0 (−6 to 3)</td>
<td>.03</td>
<td>0 (−6 to 3)</td>
<td>.04</td>
</tr>
<tr>
<td>Spatial Recall Test: total</td>
<td>−1 (−9 to 11)</td>
<td>.61</td>
<td>−2 (−15 to 6)</td>
<td>.09</td>
<td>−1 (−9 to 11)</td>
<td>.28</td>
</tr>
<tr>
<td>SDMT</td>
<td>−4 (−13 to 13)</td>
<td>.13</td>
<td>−5 (−22 to 5)</td>
<td>&lt;.001</td>
<td>−4 (−13 to 13)</td>
<td>.20</td>
</tr>
<tr>
<td>WLG</td>
<td>−3 (−12 to 6)</td>
<td>.004</td>
<td>−5 (−22 to 6)</td>
<td>&lt;.001</td>
<td>−3 (−12 to 6)</td>
<td>.05</td>
</tr>
<tr>
<td>Selective Reminding Test</td>
<td>3 (−21 to 35)</td>
<td>.19</td>
<td>−4 (−23 to 26)</td>
<td>.96</td>
<td>3 (−21 to 35)</td>
<td>.35</td>
</tr>
</tbody>
</table>

Abbreviations: SDMT, Symbol Digit Modalities Test; WLG, Word List Generation Test.

<sup>a</sup>n=30.
<sup>b</sup>n=29.
<sup>c</sup>Median (range).
<sup>d</sup>Wilcoxon test.
<sup>e</sup>Signed rank test.

©2011 American Medical Association. All rights reserved.
achieve high levels of lower limb dysfunction at earlier stages of the disease.6

Furthermore, after 5 years without treatment, patients in the placebo group showed lower NAWM MTR measures and a trend toward a lower BPF. It is known that lower NAWM MTR measures chiefly reflect higher levels of demyelination in the NAWM (ie, outside the lesions),3 and a lower BPF probably reflects a higher degree of neuroaxonal loss26 in the entire brain. Instead, the MRI measures directly related to lesions were not significantly different between treatment arms after 5 years without treatment, although significant differences had been observed in those variables during the trial.8 This finding suggests a delayed effect of interferon beta-1b treatment on nonlesional tissue, which might be observed years after treatment termination, in contrast to a rapid but brief beneficial effect on lesion variables. The clinical underpinnings of this time shift in observed MRI beneficial effects have yet to be determined.

Changes from the end of the trial to the 5-year-without-treatment follow-up also were investigated to detect possible delayed effects of interferon beta-1b. We failed to demonstrate different behavior in the 2 treatment arms in terms of clinical evolution after the trial. However, MRI measures, such as BPF and NAWM MTR, showed borderline significant differences between groups, favoring the interferon beta-1b arm and suggesting some treatment effect on nonlesional tissue, which might be seen even after years of treatment termination.

However, because a 2-year treatment input could be too short to observe delayed effects after 5 years without treatment, the investigation of changes from trial baseline to the 5-year-without-treatment follow-up was performed. No significant differences were seen between treatment arms in worsening scores on the EDSS or MSFC. Nevertheless, the interferon beta-1b group showed better evolution than the placebo group regarding verbal delayed memory and verbal fluency, an aspect of the executive function. Along the same lines and during the same period, the placebo group developed significantly more brain atrophy than the interferon beta-1b group. However, from the present data it is not possible to determine whether the delayed beneficial effect of interferon beta-1b on brain atrophy represents the pathologic substrate of clinically observed neuropsychological effects of interferon beta-1b.

Finally, the in-trial total number of ALSs correlated with worsening EDSS score from trial termination to 5-year-without-treatment follow-up. This is relevant because one of the most remarkable effects of interferon beta-1b during the trial was the reduction of ALSs,8 which indicates that, even in PPMS, in which the role of lesion load is less clear than in relapse-onset MS, inflammatory activity could indeed play a part in the accrual of disability, as previously suggested.37,38 Moreover, it seems that it is precisely the presence of new or enlarging lesions that determines most of the clinical progression, as noted in the 5-year assessment of the PPMS Magnetic Imaging in Multiple Sclerosis cohort.37 Our findings also agreed with those of the recently published trial with rituximab in PPMS,9 in which some positive treatment effect was observed in younger patients with inflammatory lesions.

We also confirmed the safety profile of interferon beta-1b and provided the longest follow-up of a cohort of patients with PPMS from a clinical trial, to our knowledge. The interpretation of these results is reinforced by the high proportion of patients who participated in this extension study. However, some patients, probably the most disabled ones, dropped out of the study. Therefore, the remaining, less disabled group might not have been fully representative of the whole population. Although the negative results of previous trials on PPMS have been attributed to the lack of progression of the study population,7 the exact impact of dropouts on our results cannot be known with certainty.

Another limitation is that no gadolinium-enhanced imaging was performed, reducing the possibility of knowing the degree of inflammatory activity at baseline and therefore preventing us from investigating its influence on all subsequent analyses. Furthermore, several statistical tests were performed, placing us in jeopardy of committing a type 1 error at a much higher rate than 0.05, which is also a limitation to this study. Nevertheless, considering the exploratory nature of this study, we thought it was not appropriate to adjust for multiple comparisons.39

In conclusion, our results showed a modest but beneficial effect of interferon beta-1b on clinical variables and on the MRI measures reflecting nonlesional damage 5 years after trial termination. This finding, together with the fact that in-trial lesion activity was associated with progression of disease as measured by worsening EDSS scores after trial termination, provides some evidence that immunomodulation could be explored still further in the search for an effective treatment for PPMS.
editorial boards of *Multiple Sclerosis Journal*, *Journal of Neurology, The International MS Journal*, *Revista de Neurologia*, and *Therapeutic Advances in Neurological Disorders*; serves as a consultant to Bayer Schering Pharma, Biogen Idec, Merck Serono S.A., Teva Pharmaceutical Industries Ltd, sanofi-aventis, Novartis AG, Almirall S.A., and Eli Lilly and Company; has received research support for clinical trials from Genentech Inc, Genzyme Corporation, Wyeth, and the organizations for which she has served as a consultant, and receives research support from the Fundación Esclerosis Múltiple. Dr Tintoré has served on scientific advisory boards for Teva Pharmaceutical Industries Ltd, Novartis AG, and sanofi-aventis and has received funding for travel and speaker honoraria from Bayer Schering Pharma, Biogen Idec, and Novartis AG. Dr Rovira serves on the editorial board of the *American Journal of Neuroradiology* and *Neuroradiology* and has received speaker honoraria from Bayer Schering Pharma, sanofi-aventis, Bracco S.p.A., Merck Serono S.A., Teva Pharmaceutical Industries Ltd, and Biogen Idec; receives research support from Bayer Schering Pharma; and serves as a consultant for Novartis AG.

**Funding/Support:** The present study was sponsored by Schering España S.A.

**Online-Only Material:** The eTables are available at http://www.archneurol.com.

**Additional Contributions:** We thank all patients for their participation in the trial and in the extension; we extend special thanks to the nurses (María José Vicente, BN, and Rosalía Horno, BN), without whom this study never would have been performed.

**REFERENCES**


