Randomised, double blind, placebo controlled study of interferon $\beta$-1a in relapsing-remitting multiple sclerosis analysed by area under disability/time curves

Clarence Liu, Lance D Blumhardt

Abstract

Objectives—The commonly employed outcome measures on disability and relapse rates in treatment trials of relapsing-remitting multiple sclerosis have well demonstrated sensitivity to treatment effects, but their clinical interpretation is problematic. An alternative method of analysis, which is more clinically meaningful and statistically appropriate to a condition with a fluctuating disease course, uses the summary measure statistic “area under the disability/time curve (AUC)”, to estimate each patient’s total in trial morbidity experience.

Methods—The AUC technique was applied in an intention to treat analysis of serial disability data derived from the expanded disability status scale (EDSS), the Scripps neurologic rating scale (SNRS), and the ambulation index (AI), collected during a double blind, randomised, placebo controlled, phase III trial of subcutaneous interferon $\beta$-1a (INF$\beta$-1a) in relapsing-remitting multiple sclerosis (PRISMS Study). The results were compared with the often quoted “conventional” end point of change in rating scores from baseline to trial completion. Analyses were also carried out on subgroups with entry EDSS stratified above and below 3.5.

Results—EDSS data analysed by AUC normalised to baseline scores disclosed that both doses of INF$\beta$-1a (22 or 44 µg) were superior to placebo (p= 0.008 and 0.013, respectively). In addition, the high dose (44 µg) was more beneficial than placebo using SNRS (p= 0.038) and AI data (p= 0.039). AUC analysis of SNRS scores also showed that for patients with baseline EDSS>3.5, the 44 µg (but not the 22 µg) dose was more advantageous than placebo (p=0.028).

Conclusions—Summary measure analysis using the AUC of serial disability/time plots, confirms and extends the results of conventional end point analysis of disability from the PRISMS Study data. AUC evaluations show that high dose INF$\beta$-1a (44 µg three times weekly) was beneficial on all of the clinical rating scale scores used in this study. This method provides a statistically powerful and clinically meaningful assessment of treatment effects on in trial disability in patients with multiple sclerosis with fluctuating and highly heterogeneous disease courses.

Keywords: multiple sclerosis; outcome measures; interferon $\beta$-1a

In the past 5 years, immunomodulatory therapies for relapsing-remitting multiple sclerosis have achieved a useful reduction in relapse rate, but less cut benefits on permanent disability, based on the outcome measures employed. These measures have been associated either with a change in rating scores from study baseline to completion, or with “confirmed progression” as defined by a certain increase in disability scores (for example, the expanded disability status scale, EDSS) at two visits 3 or 6 months apart. Interpretation of these measures is difficult, as the first end point ignores the instability and variance associated with two snapshot assessments in time, as well as the fluctuating disabilities that commonly occur in relapsing-remitting multiple sclerosis, whereas the so called confirmed progression (and its graphical depiction using Kaplan-Meier survival curves) includes an unknown number of erroneous treatment failures (that is, cases with recovery to baseline after satisfying so called progression criteria). End points which are both clinically and statistically meaningful and which incorporate the amount of disability associated with each attack into the overall disability calculations “should be preferred.”

One method of analysis that indexes all in trial morbidity changes is the area under the disability/time curve (AUC). This technique is appropriate for the mainly transient disability experienced early in the disease course in relapsing-remitting multiple sclerosis and should be more responsive to change as it utilises all the collected serial data. In the present study, we have reanalysed with AUCs the clinical rating scores obtained in a randomised, double blind, placebo controlled trial of interferon $\beta$-1a (INF$\beta$-1a: Rebif, Ares Serono) given subcutaneously in relapsing-remitting multiple sclerosis, recently published by the PRISMS (prevention of relapses and disability by interferon-beta 1a subcutaneously in multiple sclerosis) Study Group. The results are compared with a conventional method of disability analysis (2 year change in clinical rat-
were calculated by two methods to obtain the AUCSUM and the AUCCHANGE. For AUCSUM, the total area under the disability/time curve throughout the trial was determined. For AUCCHANGE, the AUCSUM was normalised to the baseline score by subtracting the area defined by the product of the initial rating score and the study period (see appendix). Between EDSS of 5.5 to 7.0, each 0.5 point increment was rescaled to 1.0 point to adjust for non-linearity of the EDSS. Three types of data analyses were performed. In the initial analyses (combined data), all 7060 datapoints from the scheduled and unscheduled visits of 560 patients were included in the AUC calculations (intention to treat analysis). The trapezium rule for determining the AUC was applied throughout, as most objectively confirmed attacks had only one additional neurological assessment and the speed and timing of relapse onset and offset would be difficult to define with certainty if other techniques were employed. Secondly, separate analyses (scheduled visit data) were carried out using solely the 6468 datapoints obtained at routine appointments. Thirdly, due to the different estimated disease course characteristics and DSS score staying times of patients with DSS<3 and DSS>4 in natural history series, analyses of subjects classified by baseline EDSS≤3.5 (n=466) or EDSS>3.5 (n=94) were attained.

**Statistics**

For both conventional and AUC data, an analysis of variance (ANOVA) model was employed with factors for treatment and centre. One degree of freedom contrasts from the ANOVA model were used to compare the treatment groups in a pairwise fashion. As there was a strong effect of baseline entry scores, all further analyses of AUCSUM were carried out with baseline disability (EDSS, SNRS, or AI) as a covariant. ANOVAs on the ranks (Kruskal-Wallis tests) were performed to determine consistency of results and validity of the parametric (ANOVA) conclusions. As our data were not normally distributed (Shapiro-Wilk W test), p values on treatment comparisons were obtained from ANOVAs on the ranks. Mean estimates with 95% confidence intervals (95% CIs were also calculated to compare treatment effects with placebo.

**Results**

**Total cohort analysis with combined data**

"2 year EDSS difference" demonstrated significant benefits favouring the 22 µg dose (p=0.026) and a tendency in favour of 44 µg (p=0.052) over placebo. AUCSUM calculations similarly showed that 22 µg IFNβ-1a was beneficial compared with placebo (p=0.046). Although the size of the treatment effect was similar, significance was not reached with the 44 µg dose (p=0.064).

The median AUCCHANGE was +0.06, +0.05, and +0.48 EDSS-year for 44 µg, 22 µg IFNβ-1a, and placebo, respectively (table 1 A). Both doses conferred significant advantages.
Table 2 Low (EDSS (<3.5, n=466) and high (EDSS>3.5, n=94) entry disability subgroups: Disability changes as measured by (A) EDSS, (B) SNRS and (C) AI, analysed using 2 year disability/time curves.

<table>
<thead>
<tr>
<th>Entry EDSS ≤3.5 subgroup</th>
<th>AUC&lt;sub&gt;CCHANGE&lt;/sub&gt; (in EDSS-years)</th>
<th>Entry EDSS&gt;3.5 subgroup</th>
<th>AUC&lt;sub&gt;CCHANGE&lt;/sub&gt; (in EDSS-years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two year difference</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>IFNβ-1a 44 μg (3x/week)</td>
<td>0 (1.0) +0.21 (1.13)</td>
<td>+0.04 (1.79)</td>
<td>+0.23 (1.58)</td>
</tr>
<tr>
<td>IFNβ-1a 22 μg (3x/week)</td>
<td>0 (1.0) +0.12 (1.27)</td>
<td>+0.05 (1.79)</td>
<td>+0.20 (1.79)</td>
</tr>
<tr>
<td>Placebo</td>
<td>+0.5 (1.25) +0.37 (1.18)</td>
<td>+0.38 (1.66)</td>
<td>+0.64 (1.84)</td>
</tr>
<tr>
<td>Two year difference</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>IFNβ-1a 44 μg (3x/week)</td>
<td>0 (7) -1.34 (7.66)</td>
<td>0 (12.68)</td>
<td>-1.11 (11.54)</td>
</tr>
<tr>
<td>IFNβ-1a 22 μg (3x/week)</td>
<td>0 (9) -1.57 (9.15)</td>
<td>-0.32 (9.51)</td>
<td>-1.15 (11.49)</td>
</tr>
<tr>
<td>Placebo</td>
<td>-4 (12) -2.49 (8.48)</td>
<td>-1.46 (9.86)</td>
<td>-2.34 (11.22)</td>
</tr>
<tr>
<td>Two year difference</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>IFNβ-1a 44 μg (3x/week)</td>
<td>0 (1) +0.19 (0.93)</td>
<td>0 (0.86)</td>
<td>+0.16 (1.25)</td>
</tr>
<tr>
<td>IFNβ-1a 22 μg (3x/week)</td>
<td>0 (1) +0.34 (1.13)</td>
<td>+0.10 (0.85)</td>
<td>+0.41 (1.38)</td>
</tr>
<tr>
<td>Placebo</td>
<td>0 (1) +0.29 (1.03)</td>
<td>0 (1.27)</td>
<td>+0.48 (1.44)</td>
</tr>
</tbody>
</table>

Total cohort analyses with scheduled visit data

Utilising data from scheduled visits only, AUC<sub>CCHANGE</sub> analysis on EDSS scores disclosed a trend favouring 22 μg IFNβ-1a over placebo (p=0.058). AUC<sub>CCHANGE</sub> calculations showed that on EDSS scores, there were significant benefits of both doses over placebo (p=0.024 and 0.014 for 44 μg and 22 μg, respectively). For SNRS and AI data, treatment with IFNβ-1a 44 μg was also superior to placebo.

Subgroup analyses

EDSS

For patients with entry EDSS ≤3.5, treatment comparisons showed 22 μg to be better than placebo, for both “2 year EDSS difference” (p=0.016) and AUC<sub>CCHANGE</sub> analysis (p=0.043). Using AUC<sub>CCHANGE</sub>, there were significant effects in favour of both treatment doses compared with placebo (p=0.036 and 0.016 for 44 μg and 22 μg IFNβ-1a), with estimated benefits of −0.4 and −0.5 EDSS-year, respectively (95% CI −0.8 to 0; −0.8 to −0.1) (table 2 A, fig 1 A-B).

For patients with baseline EDSS>3.5, neither conventional “2 year EDSS difference”, nor AUC analyses, showed significant differences between treatment arms, although over placebo (p=0.013 and 0.008 for 44 μg and 22 μg IFNβ-1a, respectively), with an estimated treatment effect of −0.5 EDSS-year for each dose (95% CI −0.9 to −0.1 and −0.8 to −0.1, respectively) (fig 1 A-B).

SNRS

There were no significant group differences in efficacy based on either “2 year SNRS difference”, or AUC<sub>CCHANGE</sub> analysis. With AUC<sub>CCHANGE</sub>, the 44 μg treatment arm improved by a median of +0.17 SNRS-year compared to deteriorations of −0.25 and −1.68 SNRS-year for the 22 μg and placebo, respectively (95% CI −0.8 to 0; −0.8 to −0.1) (table 2 A, fig 1 C-D).

There were no significant group differences for either the “2 year AI difference”, or AI AUC<sub>CCHANGE</sub>. On AUC<sub>CCHANGE</sub>, the 44 μg dose (median of 0 AI-year) was significantly superior to placebo (median worsening of +0.11 AI-year) (p=0.039) with a mean benefit of −0.4 AI-year (95% CI −0.7 to −0.1) (table 1 C, fig 1 F).

Figure 1 AUC<sub>CCHANGE</sub> data obtained on total visits with intention to treat analysis for the entire cohort (n=560). AUC<sub>CCHANGE</sub> EDSS: (A) basic statistics and (B) treatment comparisons in EDSS-years. AUC<sub>CCHANGE</sub> SNRS: (C) basic statistics and (D) treatment comparisons in SNRS-years. AUC<sub>CCHANGE</sub> AI: (E) basic statistics and (F) treatment comparisons in AI-years. (Basic statistics: histogram=group mean, horizontal bar=median, vertical line=interquartile range. Treatment comparisons: histogram=estimated mean, vertical line=95% CI, p value from ANOVA on ranks).
AUC\textsuperscript{CHANGE} showed an estimated mean benefit for the 44 µg dose over placebo of −1.0 EDSS-year (95% CI −1.9 to 0) (fig 2 A-B).

SNRS
For patients with entry EDSS ≤3.5, neither “2 year SNRS difference”, AUC\textsuperscript{SUM}, nor AUC\textsuperscript{CHANGE} showed any significant distinctions between treatment arms (table 2 B, fig 2 C-D).

For patients with a baseline EDSS >3.5, there was a significant effect in favour of the 44 µg dose over placebo using conventional “2 year SNRS difference” (p=0.016). AUC\textsuperscript{SUM} did not show differences between treatment arms. AUC\textsuperscript{CHANGE} confirmed significant effects of 44 µg IFNβ-1a versus placebo (p=0.028) with a mean benefit of +9.0 SNRS-years (95% CI +2.9 to +15.1) (fig 2 C-D).

AUC\textsuperscript{SUM} and AUC\textsuperscript{CHANGE}. AUC\textsuperscript{SUM} summates the total disability data serially over the study period and data independence is assured (the independence is lost in analyses involving any change in disability scores as each data point is influenced by its preceding values). However, it would be expected to be insensitive to small changes in short trials if there were a large disability range in the cohort at baseline. On the other hand, AUC\textsuperscript{CHANGE} obtained by normalising the AUC\textsuperscript{SUM} value to the baseline rating is sensitive to in trial changes and has been previously utilised in a neurorehabilitation pilot study.26 However, neurological improvements and deteriorations may cancel and the technique is susceptible to unstable scores at trial entry, a problem partly resolved by ensuring a stable run in period.

In the present study, comparison of the two AUC techniques showed that AUC\textsuperscript{SUM} essentially confirmed the “2 year disability difference” analysis (treatment benefit over placebo for EDSS, but not for SNRS or AI data) without improved responsiveness. This can be attributed to the small in study changes relative to the wide range of disability (EDSS 0–5) at baseline. By contrast, the increased sensitivity of AUC\textsuperscript{CHANGE} for detecting positive therapeutic effects not only showed savings in terms of EDSS-years for both IFNβ-1a doses compared with placebo, but also significant beneficial effects in favour of the high dose (44 µg three times weekly), using SNRS and AI data.

AUC analysis has several advantages over conventional techniques for this type of study. The method can be applied to any clinical ran-
These patients have a tendency to deteriorate EDSS >3.5 (when gait dysfunction ensues) are relapsing-remitting multiple sclerosis and an separately in this study because there is data from unscheduled visits. were similar with or without the inclusion of out, at least in part, the effects of transient dis-

reduced mean relapse rate in the treatment performed to reduce any sampling bias due to the aspects of transient dis-

ects with the EDSS and AI data, confirmed the conventional end point of 2 year SNRS di-

cfirmed progression and Kaplan-Meier analy-

morbidity remains unaltered throughout. This issue, as well as the problem of erroneous treatment failures wrongly assigned by con-


AUC analyses also provide no information on disability trends over time: a subject with a period of improvement followed by deterioration may have the same AUC score as one with the opposite tem-

oral sequence, or another in which the morbidity remains unaltered throughout. This issue, as well as the problem of erroneous treatment failures wrongly assigned by con-


be more sensitive to the treatment e-

subjects with baseline EDSS>3.5 (n=94), may have the opposite trend followed by deterioration. This may occur over short periods despite the slow accumulation of irreversible disability. AUC analyses might be regarded as having baseline (including the area under the curve between each pair of con-

secutive scores given by the trapezium rule. If disability scores (y0, y1, y2, ..., yn) are plotted versus their times of assessment (t0, t1, t2, ..., tn) totalling n+1 measurements y, at times t (t0, t1, ..., tn), then

AUC=1/2 \sum_{i=0}^{n-1} (t_{i+1}-t_i)(y_i-y_{i+1})

AUC^CHANGEP = AUC^CHANGEP - (y_n - y_t)

1 IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. 1. Clin-


4 Johnson KP, Brooks BR, Cohen JA, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-

remitting multiple sclerosis: results of a phase III multicenter, double-blind, placebo-controlled trial. Neuro-

logy 1995;45:1268–76.

5 Fazekas F, Deisenhammer F, Strasser-Fuchs S, et al. Randomised placebo-controlled trial of monthly intra-

Liu, Blumhardt