Author's response to reviews

Title: Cognitive performance in relapsing remitting multiple sclerosis patients during 2 years treatment with intramuscular interferon-beta-1a: an observational study in daily practice using a brief computerized cognitive battery

Authors:

Peter J Jongen (ms4ri@kpnmail.nl)
Chris Edgar (chris.edgar@unitedbiosource.com)
Evert Sanders (ResearchNeurologie@amphia.nl)
Christian Sindic (Christian.Sindic@uclouvain.be)
Sophie Goffette (docteurgoffette@yahoo.com)
Michel Dupuis (mi.dupuis@clinique-saint-pierre.be)
Philippe Jacquere (p.jacquere@clinique-saint-pierre.be)
Daniel Guillaume (daniel.guillaume@cnrf.be)
Regine Reznik (regine.reznik@cnrf.be)
Keith Wesnes (Keith.Wesnes@unitedbiosource.com)

Version: 3 Date: 21 March 2011

Author's response to reviews: see over
Responses to reviewer comments:

The authors would like to thank the reviewers for their feedback and feel the comments have supported important revisions to the detail and focus of the manuscript. Detailed responses to each point raised are provided below.

Reviewer 1

1. Abstract - Background: "we investigated the effectiveness...in postponing cognitive decline" is not entirely correct because there was no control group. The authors can only write "we investigated cognitive function during a 2 year treatment with ....." Same for text in the Background on page 4 line 19.

Response: The corresponding passages have been changed and text has been edited to remove references to effectiveness of treatment.

2. In the discussion the authors should mention something on the natural history of cognitive decline in RRMS over a 2 year period. What decline do we expect on these psychomteric tests over a 2 year period?

Response: The following text has been added:

‘In a 3-year controlled study the MS patient group deteriorated on the majority of neuropsychological measurements, and 21% of the MS patients met criteria for significant deterioration over 3 years [Bernardin et al. 1993] In another 3-year controlled study 35% of the MS patients who were cognitively intact had slightly deteriorated, whereas 77% of the patients in the impaired group showed significant deterioration [Kujala et al. 1997]. So, once cognitive dysfunction occurs, it is unlikely to remit and it often progresses [Amato et al. 2006]. In a study in early onset MS patients initial deficits in verbal memory and abstract reasoning remained more or less stable on a retest after 4 years, however, linguistic disturbances had than appeared [Amato et al. 1995]’

In addition, it would be appropriate to discuss the underlying mechanism of cognitive decline in MS and the mechanism by which INF beta might stabilize cognitive functions.

Response: The following text has been added:

‘Cortical involvement related to MS may arise from local demyelinating lesions, meningeal inflammation, neuronal injury, and Wallerian or transsynaptic degeneration [Calabrese et al. 2009]. MRI data on the relation between cognition and regional white
matter lesions, central atrophy, and whole brain atrophy suggest that white matter changes also contribute to cognitive impairment. In RRMS axonal damage is related to inflammation, even in the early phase. In fact, inflammatory processes are considered crucial in the development of cognitive symptoms. As the DMDs are known to reduce immune-mediated inflammation in a clinically significant degree, it may be conceived that a beneficent effect of DMD treatment on cognition is mediated by reducing inflammation.'

Reviewer 2

ABSTRACT

In the Abstract and throughout the manuscript, the term “information processing” should be further qualified – all cognition involves the processing of information. The term “Power of attention” etc, should be defined. Are these psychometric tests? If so they should be capitalized as other proper nouns. As there is no control group, the authors should not offer conclusions about the effects of treatment on cognition.

Response: Use of information processing clarified to indicate nature of cognitive task(s). Composite cognition measures (Power of Attention, etc) capitalised. Conclusions regarding effectiveness of treatment removed.

INTRODUCTION

More robust correlation is found between cognitive function and whole brain and regional gray matter atrophy, rather than lesion volume. The literature review should reflect this. The authors should acknowledge that other DMTs also reduce MRI lesion load and relapse rate. The term “real life” is vague and depending on what the authors are trying to convey, the point may be irrelevant.

Response: The text following text has been added:

‘More robust correlations have been found between cognitive function and whole brain atrophy [Benedict et al. 2004] [Sanchez et al. 2008] and regional gray matter atrophy [Calabrese et al. 2011].’

As there is no control group, the study was not designed to investigate “the effectiveness of IM INF#-1a in postponing cognitive decline in RRMS patients treated in daily practice.”
Response: In line with this comment and those from the other reviewers we have changed the relevant passages, removed the discussion of treatment effectiveness and instead focussed on preliminary validation of the CDR battery.

The authors note that “three of the MACFIMS tests (D-KEFS sorting, 10/36 and especially BVMTR) are dependent on complex motor responding.” This is false – the motor responding is simple – for example, on DKEFS subjects only move cards around.

Response: The word ‘complex’ has been deleted.

At the bottom of page 5, DSST is not defined. It is later but not referenced. Why did the authors use this test, which to my knowledge has not been validated in MS?

Response: One issue we have tried to explore is that cognitive tests vary in their specificity for cognitive and other domains of function. DSST can perhaps be considered a more ‘global’ measure in MS as opposed to a measure of attention, due to motor, working memory and other requirements. The authors feel this is an important issue to highlight when considering test selection and evaluating the profile of impairment.

At the point of introducing the CDR, a new paragraph should be designated.

Response: New paragraph has been added

Please explain how information is “recalled” on this computer administered test – I believe this is probably a recognition memory task. Please reference data that shows the alternate forms of CDR are equivalent in difficulty.

Response: Text has been added to state that verbal responses were recorded by the administrator.

At least 30 alternate forms are available for all tasks within the CDR system test battery, though more may be available dependent on the specific task and language. The use of computerization allows for the creation of multiple forms of tasks using the same criteria and rule sets, to create multiple equivalents. For several of the tasks the use of random or pseudo-random selection of targets or manipulation of inter-stimulus interval is used to create multiple forms of tasks. The presentation of targets and probes for tasks is also randomized across visits for each participant (Table 1).
<table>
<thead>
<tr>
<th>Domain</th>
<th>Test</th>
<th>Alternate Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attention, Concentration and Vigilance</strong></td>
<td>Simple Reaction Time</td>
<td>The application randomly determines the duration of the interval between two stimuli being displayed based on a minimum and maximum inter-stimulus interval.</td>
</tr>
<tr>
<td></td>
<td>Choice Reaction Time</td>
<td>The application randomly determines the duration of the interval between two stimuli being displayed based on a minimum and maximum inter-stimulus interval. The stimuli ‘YES’ and ‘NO’ are presented on a pseudo-random basis, ensuring half of all stimuli presented are ‘YES’ and half ‘NO’.</td>
</tr>
<tr>
<td></td>
<td>Digit Vigilance</td>
<td>A target digit between 1 and 9 is selected on a pseudo-random basis. Files are used which contain largely pseudo-random strings of probe digits, within which target digits are equally interspersed. The use of these strings is randomized across task administrations for each participant.</td>
</tr>
<tr>
<td><strong>Verbal and Visuospatial Working Memory, and Executive Control</strong></td>
<td>Spatial Working Memory</td>
<td>Files are used which contain sets of target and probe pictures generated according to the same rule set. The use of these sets is randomized across task administrations for each participant.</td>
</tr>
<tr>
<td></td>
<td>Numeric Working Memory</td>
<td>Files are used which contain sets of target and probe digits generated according to the same rule set. The use of these sets is randomized across task administrations for each participant.</td>
</tr>
<tr>
<td><strong>Verbal and Visual Episodic/Declarative Memory</strong></td>
<td>Word Presentation</td>
<td>Multiple word lists (total number is language dependent) are available, each having been generated according to criteria specifying lists with the same number of words of a given length, frequency and imagery. The use of these sets is randomized across task administrations for each participant.</td>
</tr>
<tr>
<td></td>
<td>Immediate Word Recall</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delayed Word Recall</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Word Recognition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Picture Presentation</td>
<td>Pictures are paired with similar pictures to give a presented target group and a distractor group, randomly interspersed with the targets during recognition. The use of these pairs is randomized across task administrations for each participant.</td>
</tr>
<tr>
<td></td>
<td>Picture Recognition</td>
<td></td>
</tr>
</tbody>
</table>
METHODS

The DSST is not widely used. The authors may be confused about the difference between SDMT and DSST. Ironically, after criticizing other tests from MACFIMS for reliance on motor function, the authors employ the DSST which requires copying symbols as fast as possible.

Response: The use of the DSST highlights an issue with assessment of cognition, which is that to a greater or lesser extent tests are not pure assessments of a single domain, requiring several aspects of cognitive function and other functions e.g. motor control in the case of DSST. This means that results must be cautiously interpreted with respect to determining the domain of function assessed, but also that some tests may serve as more global measures of function.

The MS54-QoL may be mislabeled. Is this the MSQOL-54? Authors reference #22 but in the bibliography 22 refers to PASAT.

Response: Abbreviation has been revised and reference corrected (Vickery et al 1995)

It is not clear how the various indices from the CDR are calculated for example “power of attention” or whatever that may be.

Response: The calculation is included in the word document supplied with the manuscript ‘additional file 2’.

The analysis does not model treatment effects as there is no control group. The ANOVA is not a mixed model, as there is no between-groups effect. There is no reason to include age as a covariate. The rationale for selecting the identified pairings with t tests is not explained. The best approach is to examine the effects of the model and then examine simple effects is the general effect is significant. The general effect should include all time points.

Response: ANOVA model by time-point re-run excluding age as covariate (includes all time-points).

It is not valid to arbitrarily select cases from another database which is not even peer-reviewed and then use the data as a control group.

Response: The data in the database are aggregate data derived from healthy subjects enrolled in industry sponsored clinical trials run in accordance with GCP and using FDA CFR 21 Part 11
compliant systems. Although there is not a specific publication focussed on the database alone, the data have been included in a number of prior publications e.g.


I believe there is a typo, a Bonferroni correction would not be $p=0.05$.

Response: $P=0.05$ is the level prior to correction in all cases. Since different numbers of tests are conducted in each case, 0.05 is divided by a different value. So the specific alpha level is reported separately for each table/analysis approach.

**RESULTS**

The test-retest correlations are of interest, but all time points should be included. This would be a better study if it were focused exclusively on the reliability of CDR in MS.

Response: All time-points included in correlation. Focus shifted to preliminary validation of CDR battery.

*Cronbach’s alpha is not a measure of “stability.” Its use in this context is not clear.*

Response: Cronbach’s alpha now included to assess internal reliability of composite score i.e. inter-relatedness of measures comprising composite.
The entire ANOVA model is never presented.

Response: We have tried to clarify description of the revised ANOVA model in the ‘statistical analysis’ section.

DISCUSSION

There is no basis for concluding that medication had any influence on cognition in this study.

Response: Discussion and conclusions have been revised to remove comment on effectiveness of treatment and focus on preliminary CDR battery validation.

Reviewer 3

- The main limitation of the study is the lack of a control group (untreated patients). In this condition is not possible to evaluate the real effect of IFN on cognitive functions. For instance, the stability observed in the study may actually embody a worsening when an improvement due to practice effect may be expected. Indeed, in this study the practice effects are only partially controlled by repeated administration in the first four months. This limitation should be at least commented in the Discussion

Response: Discussion and conclusions have been revised to remove comment on effectiveness of treatment and focus on preliminary CDR battery validation.

- The neuropsychological test battery used in this study is not widely applied in MS population and may be not fully tailored for the profile of cognitive dysfunction in MS. The Authors provide some psychometric properties of this battery in a sort of validation of the battery. However, the validation of the battery is beyond the objectives of the study and cannot be performed in this relatively small sample size. Moreover, a validation study should include as reference criterion the presence of cognitive impairment assessed through neuropsychological tests validated for MS (for instance the Rao’s Brief Repeatable Battery). The correlation with PASAT scores alone is not sufficient.

Response: In line with this and other comments discussion and conclusions focus on preliminary validation/utility and the need to conduct further studies in larger samples and with comparison to established batteries e.g. BRNB and/or MACFIMS.

- There is no information on the cognitive functioning at the individual level.

Moreover, there is no definition of overall cognitive impairment (failure on how many domains?).
It has been documented that cognitive functioning in MS can remain stable and tends to progress over time particularly in subjects with a cognitive impairment at baseline (see for instance Amato MP et al Arch Neurol 1995, Amato MP et al Arch Neurol 2001). This information should be retrieved. In particular, it should be of interest to know the proportion of impaired patients at baseline and the evolution of cognitive functioning over time in this subgroup of subjects. Furthermore, baseline cognitive functioning should be presented before the follow-up.

Response:

A corresponding text has been added at the start of the Discussion (See above).

- Changes over time have been assessed including only age as a covariate. The analysis should be performed including other demographic and clinical variables, such as gender, education, disease duration and disability levels on the EDSS. At this regards, there is no information on depression and fatigue, that are widely recognised as possible confounders of cognitive functioning in MS. Also this variables should be included as covariates in the model.

Response: Limitations of the present sample have been acknowledged in the discussion including the importance of following up this preliminary validation in a larger sample, more extensively characterized.