Author's response to reviews

Title: Cognitive performance in relapsing remitting multiple sclerosis: a longitudinal study in daily practice using a brief computerized cognitive battery

Authors:

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

Version: 8 Date: 16 May 2011

Author's response to reviews: see over
Responses to reviewer comments: Richard Benedict, 11 April 2011

The authors are confident that these final edits unambiguously present the results without suggestion of this as a treatment effects study, and that the manuscript now provides an open, fair and unbiased presentation of the data. We thank Dr Benedict for providing the critical feedback crucial to this revised description of the data.

We have addressed the points as follows.

Abstract

Explain that Power of Attention etc are subtests from the CDR and what they measure.
We have clarified the ‘domain’ assessed by each measure as well as the derivation in the Methods section. In addition, detailed task descriptions and composite score derivations appear in the additional files 1 and 2.

Explain who are the controls, this is very important. Otherwise the d values are meaningless.
We have clarified in the Abstract, Results and Discussion sections that these normative data are derived from healthy volunteers enrolled in prior clinical trials and not a control group from the present study. In addition, some of the potential caveats of such a comparison are now noted in the discussion.

Stable is not defined, do you mean that the mean values did not change, good test-retest reliability?
The term ‘stable’ has been removed to avoid any ambiguity.

Introduction

“Cognitive dysfunction has a significant impact on social functioning and health-related quality of life (HR-QoL).” Requires reference, actually several studies have shown that self-report QOL measures are more related to depression than NP measures.
As the focus of the study is now fully on the validation of the CDR System, this sentence has been deleted.

The authors are stubbornly holding on to the idea that this is a treatment effects study as they state “The present study investigated cognition during 2 years treatment with IM INFβ-1a in RRMS patients, treated in daily practice.”
This sentence has been deleted. All references suggesting this was a study to assess treatment of cognition have been deleted. The Title has also been changed accordingly.

This sentence is still problematic: “Three of the MACFIMS tests (D-KEFS sorting, 10/36 and especially BVMTR) are dependent on motor responding, in a patient group where this may be impaired.” First the 10/36 is not in the MACFIMS. Second the DKEFS requires no more motor function than does using a mouse or keyboard. BVMTR has a copy trial to control for manual coordination effects.
This section has been deleted. However, we have highlighted that the CDR battery is designed to minimize motor requirements and does not require use of mouse or keyboard by the participant.
The authors seem to be promoting (selling) the CDR and the first author is employed by the company. In the Competing Interests section these relationships need to be explicated. 

In the Competing Interest section the relation of the first author to the company has been described and also that of the last author Dr. Wesnes.

The issue of alternate forms requires elucidation. The CDR randomly generates alternate forms using a computer algorithm as is common with computerized batteries. However, that does not mean that the forms are equivalent in difficulty – this is an untested assumption inherent in the program. It could also hamper test-retest reliability, which should be acknowledged later.

This point is addressed in the Discussion.

Methods and Results

I believe the authors wish to state that oral responses are written down by the examiner. Yes, this is correct and is now described in the Methods section.

In addition, there is no evidence that visual recall is assessed. Do the subjects describe figures and the examiner writes down these responses also.

Visual recall is not assessed, only recognition. We hope this is clear both in the task descriptions (Methods and additional files) as well as being explicitly noted in the Discussion as an aspect of function not covered by the battery.

Please explain how the subtest scores are calculated.

Calculation of subtest scores has been further described in an updated ‘additional file 1’. These have previously been reported elsewhere e.g. in the cited Wesnes KA, Ward T, McGinty A, Petrini O (2000) The memory enhancing effects of a ginkgo biloba/panax ginseng combination in healthy middle-aged volunteers. Psychopharmacology (Berl) 152 (4):353-361.

What does (training) mean on the time point measures?

The purpose of this training has been added to the Methods section. The potential benefits of such training in overcoming initial and marked learning/practice effects, have been addressed in the Discussion.

All of the Pearson r reliability coefficients should be reported, that is all time points. See for example Benedict RHB, Duquin JA, Jurgensen S, Rudick R, Feitcher J, Munschauer F, Panzara MA, Weinstock-Guttman B (2008). Repeated Assessment of Neuropsychological Deficits in Multiple Sclerosis using the Symbol Digit Modalities Test and the MS Neuropsychological Screening Questionnaire. Multiple Sclerosis, 14:940-946. These are now reported in the results.

I remain very concerned about the selection of controls from an industry developed database. Much more explanation is needed here. I suspect that this is not a control group per se, but rather a standardization sample which is routinely employed by the users of the CDR. This is tantamount to a study on the WAIS in MS and comparing patient performance to the manual norms. This is not necessarily a fatal flaw, but the authors should be up front about this and fully describe these control data.
As stated above, we have clarified in the Abstract, Results and Discussion sections that these normative data are derived from healthy volunteers enrolled in prior clinical trials and are not a control group from the present study. In addition, some of the potential caveats of such a comparison are now noted in the Discussion. This is an approach used in several prior publications including other areas of neurology (see below) and has the benefit of allowing for comparison across such analyses. References: Wesnes KA, McKeith I, Edgar C, Emre M, Lane R (2005) Benefits of rivastigmine on attention in dementia associated with Parkinson disease. Neurology 65 (10):1654-1656 and Wesnes KA, Edgar C, Dean AD, Wroe SJ (2009) The cognitive and psychomotor effects of remacemide and carbamazepine in newly diagnosed epilepsy. Epilepsy Behav 14 (3):522-528.

For all validity coefficients, the time point needs to be clear. I suggest that the authors stick to correlating the baseline values and perhaps the termination values. For the correlations with EDSS and HR-QoL both baseline and month 24 correlations are provided. For the correlations between cognitive outcomes baseline alone has been reported.

Discussion

The Discussion should be shortened and the authors should stick much more to their data. The discussion has been shortened and relates more to the data.

The authors state: In this study, clinically active RRMS patients who started IM INFb-1a treatment in daily practice showed stable cognition, both in terms of change over time and proportion of patients with cognitive impairment. This is a veiled intimation that the drug was effective, and it is repeated in the conclusions. ALL such references should be removed, as there is no control group. We have removed all references suggesting that the drug treatment may have been effective.

The CDR System is the most widely used computerized assessment of cognition in clinical trials. Citation needed or remove the sentence. This sentence has been removed.

Again in the Discussion there is exaggeration of the merits of CDR, suggesting bias among the authors. The CDR does not cover the same domains as the MACFIMS [e.g. learning over multiple trials, visual recall memory, abstract reasoning, verbal fluency]. We have clarified in the Discussion that there are aspects of cognition which are not assessed by the CDR battery. So although it is more concise, this is at the cost of some specific measures. Whilst we have removed the suggestion that the batteries cover all of the same domains, conceptually we do not agree that these specific measures constitute domains. Rather they are measures or aspects of cognition within domains and there is likely to be considerable overlap in the cognitive processes utilized.

Later, the authors’ descriptions of the BRB and MACFIMS memory tests are not correct. There are several errors here and I will not list them – I suggest the authors consult the manuals or more detailed descriptions of these tests. These descriptions have been removed.
The conclusions regarding sensitivity and reliability of CDR may hold, depending on what is found with the recommended analyses above. I am especially interested in the reliability data from the other time points not yet mentioned.

Do the conclusions regarding the sensitivity and reliability of the CDR hold?

The analyses show that the conclusions hold in part i.e. for four of the composite scores. However, for two of the scores there was lower and more variable test retest.

Kind regards,

On behalf of the authors,

Peter Joseph Jongen, neurologist