Reviewer’s report

Title: Effects of Multicomponent Exercise on Cognitive Function in Older Adults with Amnestic Mild Cognitive Impairment: A Randomized Controlled Trial

Version: 2 Date: 28 June 2012

Reviewer: Carl W Cotman

Reviewer’s report:

Major compulsory Revisions.

In the manuscript “Effects of multicomponent exercise on cognitive function in older adults with amnestic mild cognitive impairment: A randomized controlled trial,” 25 older adults (65+yrs) engaged in a multicomponent exercise program, while 25 age matched controls underwent educational/cognitive training over a 12 month period. In this presumably repeated measures study, these 25 older adults were enrolled in “biweekly” exercise training sessions, with some of those sessions including some aspect of cognitive stimulation while completing the exercise training sessions. These patients were evaluated with a series of language and memory based neuropsychological assessments, MRIs and physiological testing before and after completion of the exercise training period. The control group underwent similar testing before and after the educational sessions.

General comment: Quite a lot of work went into this study and this is an important area. As it stands however in essence the study shows that either intervention shows an improvement over baseline and that a 3 time education intervention is just about as effective as an exercise regimen. As it stands the paper is unacceptable. The methods has places where it is confusing as to exactly how many times the patients engaged in the organized exercise sessions (though its described well in the abstract but not the body of the manuscript??), and the analytical steps were poorly described and possibly incorrect. Similarly, the results section has a minimum level of sufficient explanation, however several key aspects are missing in order for the reader to really understand and interpret the results of the study.

Specific comments/suggestions:

1. Please insert page numbers and line numbers into this manuscript! Difficult to review a paper without being able to site specific lines and pages.

2. Intervention section:

1. The manuscript states" the multicomponent program lasted 12 months and involved biweekly 90 minute sessions.......” and then goes on to say, “One of the four classes during the intervention period included approximately 20-30 minutes of consecutive outdoor walking.....” The term bi-weekly implies that 2 days per week (as stated in the abstract) the exercise group engaged in the structured
program, resulting in approximately >100 exercise sessions for this group. How can “one of four classes during the intervention period....” be accurate then? I would suggest creating a figure similar to that used to describe the inclusion/exclusion steps for describing exactly what the exercise group did. Also need to reword this entire section and be extremely clear at each step who did what, when, for how long, and in what order. Currently, it is too confusing to really understand how many times the patient engaged in these activities.

2. The description of which activities were performed is well written and should be kept in the revision of this paragraph.

3. Analysis section: Here, there were a couple of nice elements of this section including the use of the Kolmogorov-Smirnov tests to confirm the normalcy of the distribution of the continuous data. However, there are several aspects of this section that are confusing and need to be clarified.

1. It states that a Pearson's Chi square (or correlation) was used to analyze the “categorical” data, however up to this point no categorical data was been put forth as being an aspect of the study that would be analyzed. What are these mysterious categorical variables? The only one reported in Table 1 would have to be gender, because all of the other variables are made up of continuous data. But, is a Chi square test necessary to tell us the percentage of males and females in each group or to tell us that this study took care to both age and gender match the populations? Please clarify and expand on what categorical variables are included in the Chi square analysis and be clear as to the question being asked by this analysis. Both are elusive in this current form.

2. In table 1, we find out for the first time what the baseline characteristics of the population are, and this is not pointed out until the Results section. Yet, this is a description of the population and some of these elements (age, gender, education level) should be included in the description of the population right along side the CDR mean values given in the Participants section of the methods (again its hard to cite where this should go because no page numbers or line numbers were provided in the manuscript). These are baseline characteristics that aren't really a part of the 'analysis' but merely descriptions of the population that was recruited. Move these elements to their proper place in the methods section. An appropriate reference in the analysis section to these characteristics and how they are described elsewhere should also be made.

3. It says that the baseline characteristics (that we later find out are found in table 1) between the two groups were examined using t-tests. It should be clear that these are independent samples t-tests assuming equal variance that are looking at between group comparisons only. It is evident that no within group examinations occurred, even though they should have and this will be described in the comments below regarding the results section. All reported T, df, and CI should be reported for each t-test, not just the p value which is a useless statistic without the other information being provided.

4. The text then goes on to say that an ANOVA was used to evaluate cognitive functions before and after the 12mos intervention. Was this a repeated measures ANOVA?? Because it should've been!! It doesn't appear so because the text
goes on to describe that a repeated measures ANCOVA was used to look at group effects, but it doesn’t mention what the IV, DV or how many levels were used. Expand each of these sentences and include a very detailed description of which IV, DV, levels and covariates were used. All analyses conducted should include repeated measures ANOVAs. For example, the first analysis should have been a repeated measures mixed design ANOVA with MMSE total scores as the dependent variable, and “time” (e.g. baseline vs. 12mos) as the repeated measure and “group” (exercise vs. control) as the between subjects factor. Thus, the reader knows that one level of the analysis includes looking at between group effects (e.g. exercise vs. control), and the other level of the analysis is looking at within group effects (e.g. what happened after 1yr), all with an emphasis on the MMSE scores. This process should be repeated for all neuropsych tests and appropriately described. The covariates described in the ANCOVA are acceptable and should be maintained in this revision of the analysis.

NOTE: Note this may be why there weren’t significant between group effects--the analysis could’ve been setup better. This could leave someone to erroneously conclude that exercise is no more or less effective than cognitive stimulation when it might just depend on how one compares the two interventions.

5. What the “intention to treat” analysis is and how it was used is a mystery, and needs to be expanded or left out. What is this for? Why did you do this? How did it help the study?

4. Results Section: Need to include all F, df, CI, and/or t values for every reported result. P values are useless stats without this other information to tell us the shape of the distribution (even if it is “normal”!)

1. No analysis was done that utilized the blood pressure, diagnosis, cholesterol, glucose or TMIG indexes. Why???? Did this not change over 1yr? Really? Even for the exercise group? It’s not worth reporting the change over time? Then why do you report them in a table if you didn’t include them in any analysis? At a minimum these variables need to be evaluated with respect to the cognitive scores to see if any predictions can be made from biomarkers indicative of body health as they pertain to cognitive health. Then this should be done for the 12mos follow up and then a repeated measures evaluation needs to be done to see if the change in one type of biomarker differs over the 1yr from another biomarker. Finally this should all be done between the groups (see the above described Mixed Design Repeated Measures ANOVA for the structure of these analyses).

2. Table 2 is really difficult to read because of the parentheses and a lack of formatting. Please clean up this table so it is easy to see which p value goes with which test, and include the other missing statistical information.

3. All of the sudden there is a “subgroup analyses,” and it is unclear what questions is being asked or why. Please clarify and expand this section so the reader can follow the analyses/results. How are we defining “old” and “young?” Isn’t everyone over the age of 65yrs? This paragraph is just poorly written.
4. Where is the MRI data?? Why was this shown to have occurred but it is also not analyzed or included in the results section. Was this used as a diagnostic measure? Please clarify.

5. The introduction and discussion sections are decently written, but at the beginning of the discussion section it talks about the “intention to treat analysis” which at this point comes out of nowhere because the reader doesn't have any reference as to what this is and why we should care about it. Clarify?

6. The conclusion suggesting a composite approach, and the claim that this study supports that idea, is completely unfounded. There were no combination therapy approaches (e.g. an exercise + education group) included in the study, only between group approaches (e.g. either exercise or education). And the analytical approach is questionable as to how this data was evaluated.

In general there are positive elements to this manuscript, as previously stated. However, multiple key elements in the methods and results section need to be redone. Moreover the initial design is really not balanced for amount of contact/socialization. And it seems odd there is so much change relative to baseline.

**Level of interest:** An article of importance in its field

**Quality of written English:** Not suitable for publication unless extensively edited

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.