Author's response to reviews

Title: Central Nervous System Antiretroviral Efficacy in HIV infection: A Qualitative and Quantitative Review and implications for future research

Authors:

Lucette A Cysique (lcysique@unsw.edu.au)
Edward K Waters (ewaters@nchecr.unsw.edu.au)
Bruce J Brew (b.brew@unsw.edu.au)

Version: 2 Date: 31 October 2011

Author's response to reviews: see over
Responses to reviewers

MS: 1565843603593469 - Central Nervous System Antiretroviral Efficacy in HIV infection: A Qualitative and Quantitative Review and implications for future research

Reviewer: Supriya Mahajan

Reviewer's report:
REVIEW: BMC Neurology
This is a very significant study because inhibition of HIV replication in the CNS is critical in treating patients who have HIV-associated neurocognitive disorders and treatment with ARV drugs is inadequate due to poor penetration of these drugs into the CNS allowing continued HIV replication. CNS targeted therapy is important for reducing CSF viral load as viral suppression in CSF results in greater neurocognitive improvement in patients with various stages of HIV-associated neurocognitive disorders.

The authors have used novel review strategies and appropriate statistical methods to evaluate the efficacy of neuroHAART. This strategy is important given the limited amount of studies in post HAART era that systematically evaluate the spectrum of HIV associated neurological disorders than range from asymptomatic neurocognitive impairment to mild and moderate neurocognitive impairment and finally HIV-associated dementia.

Appropriate scoring methods are used by the authors, that enabled them to identify qualitatively similar studies and analyse them further to provide some quantitative conclusions inspite of the methodological limitations, common to clinical Neuro-AIDS research that include clinical heterogeneity of the samples studied, unclear inclusion/exclusion criteria, lack of stringent neuropsychological evaluation.

The blind review strategy used by the authors allowed applying the same stringency for evaluation of all studies thereby allowing only studies with a higher quality, hence greater comparability to be assessed in the quantitative phase. This reviewer concurs with the conclusion of the authors that this study suggests that a large randomized longitudinal trial that includes the entire spectrum of HAND is needed to determine the efficacy of neuroHAART.

We thank the reviewer for their positive comments
Reviewer: Bernardino Roca
Reviewer’s report:
All comments are “Major Compulsory Revisions” to be made. Abuse of initials makes the article hard to read and understand.

We believe that the reviewers meant that “abbreviations” were often used in our manuscript. As per editorial guidelines, all abbreviations were clearly described on their first use and a list of abbreviation has been provided.

Most content of the last paragraph of Introduction should be included in Method.

The last paragraph of the Introduction was placed in this section of the manuscript to clearly delineate the aims of the study. Because this is a review involving some statistical decisions, we believe that it was important to inform the reader at this stage of the manuscript.

The authors call the article a review. I think it is an original, assessing the methodology of articles on “neuroHAART”

The title of our manuscript includes the terminology review to outline that a literature review work had taken place. We believe that this is the correct terminology.

Conclusions in page 16 are not real conclusions of the article, but ongoing discussion/comments/recommendations the authors make. Similar objections can be made regarding the Conclusion in the abstract. In my view the real conclusion is: Methodological problems in published studies impede reaching any conclusion on efficacy of “neuroHAART”.

1. The reviewer makes a general comment regarding page 16 and the conclusion. Because of the general nature of the reviewer’s comment it is difficult to actually address it by changing these specific parts of our manuscript, but in an attempt to address this general comment, our discussion now includes a limitations’ paragraph where meta-analytical limitations are discussed.

Regarding the reviewer’s comment on the abstract’s conclusion, we feel that without providing some recommendations on how the admitted methodological limitations of the body of literature could be overcome, the study would only promote the continuation of a status quo that had resulted in much heat and little light. By focusing on overcoming the methodological issues that have resulted in the current situation, we hope that the current work may play a role in moving the science from a position of “we can’t reach any conclusion on neuroHAART efficacy” to a situation where we can.

2. The qualitative and quantitative phases of our review process were explicitly intended to be a way to assess the amount of confidence that should be placed in a body of literature with admitted methodological issues. We believe we devised a very
strict set of inclusion criteria for our review that should give confidence in the results of studies that we analysed. Because these studies meet very strict criteria, we believe that the results of these studies should be considered when examining the issue of neuroHAART efficacy

3. We indicated in the Introduction that: "recommendations", based on the review findings, would be provided for future trial addressing the issue of “neuroHAART”. This was clearly stated as one the aims of the current work, because the links between unanswered questions regarding neuroHAART efficacy and future experimental work are poorly developed in the current HIV literature.

After reading the article the proposed title “Central Nervous System Antiretroviral Efficacy in HIV infection: A Qualitative and Quantitative Review and implications for future research” seems inappropriate to me. I would propose this other one: “Assessment of methodology of published articles on efficacy of ‘neuroHAART’”.

We believe that it is important that titles of manuscripts inform readers of the methods employed, the general subject matter of the article, and indication of the results presented. The reviewer’s proposed title does not inform readers of the methods used or the results given, and does not adequately reflect the fact that the scope of the review included, but was not restricted to, considering the methodological limitations of the current literature.

We believe that it is important that the title carries this information.
Reviewer: Scott Letendre

Reviewer's report:
The submitted manuscript is an impressive work of scholarship by highly experienced and accomplished investigators and addresses a very important and timely topic, the importance of antiretroviral distribution into the CNS in people with HIV. While others have summarized previously published studies, none have taken the careful and systematic approach described here. Since the investigators use qualitative assessments to rate the reviewed manuscripts, the exact methods will be debated by reviewers and readers but the authors’ decisions in this regard are reasonable, defensible, and perhaps even overly conservative. Ideally, a manuscript of this importance should be published on an expedited timeline but there are a few issues that should be addressed first.

We thank the reviewer for their positive comments

1. Abstract: Shouldn’t future studies also carefully control for comorbid conditions?
   We agree with the reviewer. And this was added in the abstract.

2. Aren’t analyses that meet a pre-defined alpha value (e.g., 0.05) for their primary hypotheses inherently adequately powered? Stated differently, shouldn’t the standard for adequate power be applied only to those projects that did not find a statistically significant effect for their primary hypotheses?
   In order to consider all studies with the same power criterion, we chose to focus on *a priori* power only. This is because there is a difference between nominal (desired) and achieved false positive and false negative (type I and II error) rates, and hence a difference between nominal and achieved power. The false positive rate, is given by the significance level. There are many scenarios where underpowered studies assume significance at 0.05, but their achieved precision is much lower than this due to low sample size or repeated testing between multiple covariates, meaning they are much more likely to find spurious effects than the nominal rate of 0.05 would indicate. Something similar happens with nominal precision, which is calculated as one minus the false negative rate multiplied by 100. If for some reason an inadequate sample size is collected, or multiple testing occurs, the false negative rate can be much higher than the nominal false negative rate specified, resulting in power much less than the nominal power. Thus, just as a significant result is not always an accurate indicator of a real effect, it is not always an indicator of correctly rejecting the null. These issues are explored in depth in Ioannidis (2005) in PLoS Med.

3. Under the section “Study selection and data extraction”, was this truly a Boolean “and” search or were combinations of these keywords used. In other words, would a more apt description of the search approach be the use of the Boolean “or” term (or at least a combination)?
   A clarification of the searching methodology is has been included in our revised version.
4. An error in Table 1: The correct finding in the Smurzynski et al manuscript is that higher CPE was associated with better NP performance when ART was composed of more than 3 drugs (as opposed to at least 3 drugs as stated in the table). This was corrected in Table 1.

5. An error in Figure A1: The correct sample size for the Smurzynski et al manuscript is 2,636 (not 3,046). The reviewer is correct regarding the total number of participants included in the study. However, to correctly estimate the effect size in regards to the longitudinal analyses that were performed to test a neuroHAART effect, we included the outcome of relevance which was the number of visits rather than the number of participants.

6. Please clearly state in the Methods that you used 15 quality criteria that fell into 5 categories (Blinded: Design, Outcomes, Subjects, Controls, and Unblinded Outcomes). Also, state that analyses had to meet at least 13 of 15 criteria to qualify (if this is correct). Readers will be able to conclude this once they refer to Table A2 but including this information in the Methods will ease the understanding of the readers and will not require that they refer to the Appendix.

   a. Please consider summarizing in the Results how many studies met different ranges of criteria (e.g., < 5, 5-8, 9-12, 13-15)

We have modified our manuscript to include the reviewer’s first suggestions. Regarding the reviewer’s second point, we note that our analyses were specified as follows in the Methods section: “A score less than or equal to 80% meant that a study presented at least three or more significant methodological limitations.” Regarding reviewer’s last point, we had already included an interpretative comment on this aspect in the first and second paragraph of the discussion. Moreover we have included the final scoring of each study in Figure A2 of the appendix. This presentation of our data avoided labelling single studies as “bad studies” and focused on the overall methodological and quantitative review results. Still, the overall scoring is provided in appendix and each limitation has been clearly defined in Table A1, this hopefully adding further transparency to our methods.

7. Some publications included multiple elements. For example, Marra et al included an assessment of CSF viral loads (n = 79) and NP performance (n = 26). DeLuca et al included both cross-sectional and longitudinal analyses with somewhat different findings. Should you group manuscripts with more than one major analysis into separate component assessments?

   The effect size and quantitative analyses were actually done separately on those aspects within each study concerned. This is further explained in the Appendix for the effect size computations. Figure 1, provides a summary of all the study designs used and therefore separate power computations were also made separately in this instance as well.

However, when the qualitative analysis of the papers was considered, each study was defined “as a whole” as this would have otherwise provided redundant scores for one single study. Nevertheless, to avoid penalizing studies which did not include certain quality criteria (such as study including no neuropsychological testing), the item was
counted as “not applicable” in the qualitative review and an overall percentage score was computed. This was explained in the Appendix and is now briefly included in the Methods section.

8. I suggest revising the following sentence to read “clinical trials” (plural) since not single clinical trial suffices to answer all aspects of other important clinical questions and this same standard should be applied for the treatment (and perhaps the prevention) of HAND. “Nonetheless, the most definitive answer to the issue of the potential superior efficacy of neuroHAART remains a randomised controlled clinical trial.” This applies to the sentence in the Conclusions that begins, “Therefore, a large randomised trial is needed…”. This may seem a subtle point but the field has suffered from clinicians placing too many expectations on a single publication even though the likelihood that a single analysis, even a randomized clinical trial, will definitively answer this complex question is small.

We agree with the reviewers and have included his suggestions.

9. Regarding the recommendation to use the CPE score as an estimate of antiretroviral effectiveness in the CNS, I would also recommend planning secondary analyses with the version of the CPE that is current at the time enrollment in future clinical trials ends. The CPE will continue to evolve as new drugs are approved and new evidence on the effectiveness of these drugs in the CNS is generated so it is likely that different versions will be available when a clinical trial is designed and when the data are analysed.

Indeed, our manuscript included the notion that the CPE rank score is an evolving score. We have also added the reviewer’s suggestion.

10. A final comment: One must admire the authors objectivity in excluding one of their own analyses for not having a quality score > 80% and two others that had estimated power below 80%.

We thank the reviewer for their positive comments.