Reviewer's report

Title: A systematic examination of brain volumetric abnormalities in recent-onset schizophrenia using voxel-based, surface-based and region-of-interest-based morphometric analyses

Version: 1
Date: 25 February 2015
Reviewer: Ulysses Torres

Reviewer's report:

In this well-written manuscript, John et al. report on the use of voxel-based morphometry, Free-Surfer-based whole brain analysis and ROI-based analyses to assess whole brain and regional morphometric abnormalities in schizophrenia. The authors studied a sample of 90 subjects, controlling the analyses for multiple confounding factors (disease duration, exposure to antipsychotics, age, gender, etc) that could potentially affect the results and that are not traditionally taken into account in most of studies in this field. After controlling for all these variables, they found significant whole-brain reductions in schizophrenic individuals in comparison to healthy controls, although such changes were not statistically significant at a regional level after correction for multiple comparisons. These findings are interesting and somewhat provocative when one takes in mind the extensive literature of structural brain changes in schizophrenia, and especially when taking into account the high methodological standards that the authors adopted in the present study. I have just a few issues to address:

Major compulsory revisions:

1. The authors employ throughout text the expressions “volumetric abnormalities” and “morphometric abnormalities”. Such abnormalities, as reported in the literature, do not only encompass volumetric reductions. Extensive volumetric increases are also reported in schizophrenia, usually (but not exclusively) related to antipsychotics. Although this sample of patients had limited exposure to antipsychotics, did the authors perform analyses searching for volumetric increases? If yes, what were the results? If not, what are the reasons justifying not even performing them? The same applies for white-matter changes. Why such analyses were not performed?

2. Why severity of disease (as measured by PANSS, for example) was not included as a nuisance in the analyses?

3. Regarding antipsychotics usage, why typicality was not entered as a confounding factor?

4. Correction for multiple comparisons using FDR is a quite conservative approach that should not make invalid the interesting results obtained when using uncorrected p<0.001 (Supplementary Figures and Tables). It is lacking in
the text a discussion about the uncorrected results, their meaning, and comparison with the appropriate background literature on the theme.

5. Is it valid the approach of replacing missing values of antipsychotics usage/doses by the mean value in statistical analyses? What are the substrates in the literature for making this? (Supplementary Table ST-2)

6. Could the authors make some comments /explanations for the different results/ different anatomical regions of reductions between Free-Surfer and VBM analyses? (Supplementary Figures and Tables).

7. Considering that some variables such as “antipsychotics doses” and “disease duration” are uniquely related to patients (not to controls), is it valid to entering them as co-variates in an analysis control > schiz, as performed in Supplementary Figures S1,S2, etc? By viewing the outputs, I suppose that for these variables a value of 0 was entered for the control subjects. Is it usual, as some authors use, instead, only negative/positive correlations for these variables?

8. Why the authors employed an extent threshold of 20 voxels in some VBM analyses?

9. The authors argue in the Discussion that their results could not be attributed to inadequate sample sizes, as some studies with very small samples yielded extensive findings of morphometric abnormalities. However, as the authors themselves say, such studies have not adopted a rigid methodology (absence of FDR correction and of correction for clinical confounding factors, etc), as that adopted in the present study. Perhaps if these small studies had employed similar methodologies, they also had not found positive results. At the same time, the recent mega-analyses in schizophrenia also fail in accounting for confounding factors. It is possible, theoretically, that mega-analyses using such rigid methodology could reach positive results, conversely to the negative results of such small-sized analyses herein reported (some of them with 13 patients x 10 subjects, Supplementary Figure S7). So, I raise this issue again: these results really cannot be attributed to the sample size?

Minor essential revision

10. I did not found in the text a definition for ROS and HCS. Such terms are cited in the main text, but are defined only on the figure legends.

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

**Quality of written English:** Acceptable

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