Reviewer’s report

Title: Neurocognitive function in bipolar disorder: a comparison between bipolar I and II disorder and matched controls

Version: 1 Date: 25 February 2013

Reviewer: Ivan Torres

Reviewer’s report:

The main purpose of this article is to evaluate differences in cognitive functioning between patients diagnosed with Bipolar Disorder I and Bipolar Disorder II. This is an important topic as findings can assist with delineating potential clinical and physiological differences between these proposed subtypes of the illness.

The strengths of this study include the detailed and comprehensive diagnostic assessment procedures, the use of validated neuropsychological tests overseen by a neuropsychologist, and the use of a population-based healthy control group.

Despite these strengths, several concerns with the study are raised and enumerated below:

Major Compulsory Revisions:

1. The review of existing work in this area is incomplete, as there are several studies comparing neuropsychological functioning in bipolar I and II that are missing (e.g. Chang et al., 2011; Hsiao et al., 2009). In particular, there is a meta-analytic study that quantified cognitive deficits between bipolar I and II (Bora et al., 2011) that was omitted, and which seems highly relevant to the topic of discussion.

2. Related to point 1 above, there is not sufficient development of what this particular study attempts to add to the existing literature in light of existing work. What is special or unique about the present study that will provide a better or more valid comparison of cognitive deficits between these two subgroups of BD? How will this study resolve the unclear or conflicting cognitive findings in this area of research? The novel contribution from this study needs to be better developed in the introduction. Alternatively, should this study be viewed as a replication effort with no other novel contribution per se?

3. The patient sample is described as euthymic, but in the methods section it is stated that “euthymia was determined by the physician’s overall diagnostic judgment.” This is peculiar, especially since objective mood ratings (MADRS, YMRS) were obtained on the day of testing. It would be better to use these ratings (e.g. scoring below a certain point on each scale) as the basis for classifying patients as euthymic on the day of testing. Defining euthymia in this way would capitalize on one of the methodological strengths of this study-namely the fact that mood ratings were obtained on the day of testing.
4. Were there any comorbid conditions such as learning disorder/ADHD or anxiety disorders in the patient sample? More importantly, were there differences between the two patient groups on these comorbidities, and if so, how could this have influenced findings?

5. Although the recruitment procedures for the control group are generally well described in the “control group” section, it is not clear how the final number of 86 controls was achieved. If 7 demographics matched persons were selected for each enrolled patient, this would create a very large pool of potential controls for inclusion (7 x 110 total patients = 770 controls). How were these 770 potential controls reduced to the final sample of 86?

6. For the various neuropsychological measures it was not clear whether raw scores were used or demographics corrected scores (e.g. z-scores, t-scores, standard scores, etc.). If the latter, where did these corrected scores come from (e.g. test manuals)?

7. There are a large number of comparisons in Table 3, and thus significant correlations could be occurring just by chance. There should be some type of correction for multiple comparisons for the data in this table.

8. The conclusion that treatment with antipsychotic drugs could associate with or lead to cognitive deficits should be stated more tentatively based on the following points. First, the magnitude of the correlations between cognitive variables and antipsychotic use are rather small (most significant correlations are under .30). Secondly, these presented correlations do not control or partial out other potential influences that could be driving this relationship (such as diagnostic patient group, history of psychosis, residual symptoms, treatment with other medications).

Minor essential revisions:

1. For the Claeson-Dahl Learning and Memory test, how many learning trials were involved? Moreover, as described this test sounds more like a typical test of episodic memory than a test of working memory.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

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.