**Reviewer's report**

**Title:** The feasibility of implementing recovery, psychosocial and pharmacological interventions for psychosis: comparison study

**Version:** 5  
**Date:** 17 December 2014

**Reviewer:** Scott C Roesch

**Reviewer's report:**

Major Compulsory Revisions

1. My primary concern is the level of detail provided for the major analyses (i.e., the chi-square tests of independence). I am not sure I completely understand what was done, other than a number of chi-square tests. Let me provide an example of what I mean. Here is verbatim text from the Results section:

   There was a significant difference in total number of blocks (#2 (5)=33.02, p<0.000) and enablers (#2 (6)=161.05, p<0.000)

   How do you get 6 degrees of freedom? What variables are being crossed in this contingency table? Also, using p<0.000 is not the best way to represent this. You obviously have a small p-value so report it as p<0.001.

2. For all analyses the authors correctly adjust using a Bonferroni correction. I strongly believe that the authors should also add an index of effect size for each test. Many of these effects in Table 2 from the pairwise comparisons look small to me. Effect sizes would greatly help with interpretation.

3. The profile presented in the figure also caused me some confusion. In the figure the authors use all 4 categories from the SAFE (Yes, Partial, No, Unable to Rate), but in the text state that “the four response categories…were restricted to the two categories Yes versus ‘non-Yes’. Why? Again, the addition of detail would help in this regard.

4. It is clear from Table 2 that the individual SAFE items were used in the pairwise comparisons. It is less clear what was used in the “omnibus” analyses. Were total scores used for the two sub-scales from SAFE? Some justification is needed for this.

   Doesn’t the SAFE manual recommend using the individual items and not using a total score? The Results section reads like the initial tests used total scores and then
the follow-up pairwise comparisons used individual items.

In short, I found the description of the analyses to be too general to make an educated assessment of how important these findings are. I would also encourage the authors to think about additional analyses that could be done to make the manuscript more attractive. For example, comparisons between specific interventions in the 3 major categories should be considered. The sample sizes are reasonable for many (e.g., n=25 for CBT; n=24 for Acute treatment).

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

No to all