Reviewer's report

Title: Serotonin transporter gene polymorphism (5-HTTLPR) L allele interacts with stress to increase anxiety symptoms in Chinese adolescents: a multiwave longitudinal study

Version: 2  Date: 24 June 2015

Reviewer: Kelly Benke

Reviewer's report:

This article provides an opportunity to evaluate a fairly well studied GxE relationship that is inconsistent in the literature. The advantages are that the study is conducted in a Chinese population during adolescence. Measures of anxiety and stress were taken longitudinally. The statistical model the authors consider does not explicitly evaluate time trends or genotype x time x stress interactions, but rather considers a multilevel model for individual fluctuations of stress and anxiety. A deeper understanding of why this model was chosen is in order.

Major Compulsory Revisions

1. It is apparent in the discussion that this study represents an extension of a previously published study, including more samples. This should be mentioned in introduction, and it should be clear up front how many previously collected individuals and how many new samples were included here.

2. Further, in the previously published study, the CESD and ALEQ were measured from 3-24 months of follow up. In this study, only follow up measures for 3, 6 and 9 months were considered. What was the reason for this?

3. Many people are familiar with mixed effects models in the context of trajectory changes over time. In these models, an explicit fixed and (oftentimes) random effect for time is provided in the model statement, and the interaction of time with genotype is interpreted. Here, you model the random slope of anxiety on stress within individuals. Please explain why you chose this approach rather than modeling a fixed or random effect of time in the model, interpreting a three way interaction. Some text to explain to readers how the clustering within individuals is handled, given these are repeated measures, rather than a subscript of t in the equation, would be an important extension to the methods section.

4. Is baseline adjustment for anxiety influential to the 5HT by stress fixed effect term? In a traditional trajectory analysis, adjusting for baseline can introduce a spurious association by inducing a regression to the mean effect. Is that a potential concern here?

5. I was not able to view a figure, only read the figure legend.
6. What is the correlation between anxiety and depression in this study? How is the 5HT by stress interaction term affected by the control for depression?

7. Your group means are stratified by gender, however, in the model, it was a covariate. Did you consider models within gender?

8. It is not clear if every subject completed every measure, or if there is missing information. If the latter, is that missing information related to any of the other variables in the study and was a multiple imputation approach considered?

9. How is the anxiety scale distributed? Is it by eye normally distributed? Some mention of this should be provided in the methods or results.

10. In the discussion, you mention that a change in the direction of effect of an allele could result because of differences in allele frequencies that tag an underlying haplotype between ethnic groups - please connect the dots for readers about how this translates to an opposite direction of effect.

Minor Revisions

11. P-values to the second or third digit should be provided in a separate column in Table 1 and 2 rather than the use of subscripts

12. Table 1 means should be further stratified by 5HT genotype groups

13. English in line 10 needs to be clear

14. How many people were excluded before arriving at the final analytic sample? It seems that the KSADS does not need to be discussed in detail with the other measurements since it is not featured in the association model, but rather serves to diagnose for exclusions.

15. CES-D often has a threshold effect. Did the authors consider a cutoff rather than a continuous measure for this adjustment variable?

16. Sentence 14-16 redundant please revise.

17. The mention of La and Lg alleles felt out of the blue and not properly introduced.

Discretionary Revisions

18. It would be nice if authors provide the variance components in their articles. Please consider doing so here.

19. Table 2 is the first representation of the parameter for the genotype by stress interaction term. It may help to anticipate this by providing the model equations as a single equation, grouping the fixed and random effects terms. This would allow the explicit writing of the interaction term that is then represented in Table 2.
**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

I declare that I have no competing interests