Author's response to reviews

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

Authors:

Qingsen Ming (mingqingsen@163.com)
Yun Zhang (zhangyun917@163.com)
Jinyao Yi (jinyaoyi@gmail.com)
Xiang Wang (wang0916xia@gmail.com)
Xiongzhao Zhu (xiongzhaozhu@163.com)
Shuqiao Yao (shuqiaoyao@163.com)

Version: 3
Date: 28 July 2015

Author's response to reviews: see over
July 19, 2015
Mr. Carlo Rye Chua

Re: Manuscript 3132956017373473 entitled “Serotonin transporter gene polymorphism (5-HTTLPR) L allele interacts with stress to increase anxiety symptoms in Chinese adolescents: a multiwave longitudinal study”.

Dear Mr. Carlo Rye Chua,

Thank you for your letter of July 14, 2015 and we would like to thank you and the two reviewers for your carefulness and helpful comments. We have revised the manuscript carefully conforming to the journal style and responded to the reviewers’ comments point by point. The revised manuscript was corrected thoroughly by a native English speaker, and the changes we have made are highlighted with yellow marker in the revised manuscript.

Response to Reviewer 1: Dr. 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.
Response: Thank you very much for spending time and effort on our work and for your suggestions.
We might not describe clearly about the sample in the discussion. Because anxiety symptoms have not been measured in previous depression study, along with different number of follow-up measures, the subjects of previous study were not included in the current study. So there was no overlap in samples of the two studies. We just carried out the current study in a larger sample. We corrected the statement as following, which has been highlighted with yellow on page 12, line 18-19.
we performed the current study in a larger sample and controlled for the effects of depressive symptoms on fluctuation of anxiety symptoms to improve reliability of the results.

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?
Response: Researchers of the current study was less than that of the previous study. It was more difficult to carry out the assessments in a larger sample recruited from two public senior high schools. Moreover, some subjects transferred to other schools after one year’s study and could not
be assessed sequentially. Therefore, in order to ensure the data integrity, we administrated three follow-up measurements finally.

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.

**Response:** As you said, we used within-subject approach by modeling the random slope of anxiety on stress within individuals. Conventional methods used in longitudinal study were primarily between-subject designs and analyses. If we use between-subject approach, random effect of time must be considered. Because between-subject approach examines whether between-subject differences in levels of stressful life events, in conjunction with different 5-HTTLPR genotypes, to predict individual differences in levels of distress (e.g., depression and anxiety) in a single timepoint. By within-subject approach, we examined within-subject fluctuations in stress level, in conjunction with 5-HTTLPR, to predict within-subject fluctuations in the level of anxiety symptoms within subject over times. So in the within-subject model, a three-way interaction of 5-HTTLPR, stress and time could not be investigated. This approach can generate a relatively reliable estimate of each subject’s degree of stress reactivity and minimize the impact of individual difference variables. Another advantage of this approach is the flexibility in handling missing data. Unlike conventional methods, the approach can readily incorporate all participants who have been observed at least once. Results of the analysis can be interpreted as if no missing data were present under the assumption that the data are missing at random.

   We added the statements how the clustering within individuals was handled as following, which have been marked on page 9, line 6-9.

   A two-level structure of model was handled. Level 1 was 4-time follow-up measurements within each subject (fluctuation in anxiety and stress over times within each subject), and level 2 was all subjects (different 5-HTTLPR genotypes between subjects).

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?

**Response:** The baseline adjustment aimed to control for individual differences in baseline levels of anxiety symptoms and to improve reliability of effects of stress and 5-HTTLPR × stress interaction on anxiety. For example, if an individual experience high level of anxiety symptoms all the time during follow-up period, it is uncertain that whether the high level of anxiety symptoms is induced by stress or high level of baseline. If the baseline anxiety symptoms were not controlled for, the reliability of effects of stress or 5-HTTLPR × stress interaction would decrease.

5. I was not able to view a figure, only read the figure legend.

**Response:** We regret about that. We have uploaded the figure as a PDF file which should be listed
at the end of the manuscript. There might be something wrong with it. We will reupload the figure with the revised manuscript.

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?

Response: Depression is strongly related to anxiety in children and adolescents. High depression symptoms are always accompanied by high anxiety symptoms. In our study, the Pearson’s correlation coefficients between baseline depressive symptoms and anxiety symptoms in 4 measurements ranged from 0.36 to 0.58 (all p<0.001). Similar to adjusting baseline anxiety, controlling for the effects of baseline depressive symptoms on fluctuation of anxiety symptoms during follow-up assessments can also improve reliability of effects of stress and 5-HTTLPR × stress interaction on anxiety. We added the statement of correlation between anxiety and depression as following, which has been marked on page 10, line 18-19.

The Pearson’s correlation coefficients between CES-D scores and MASC-C scores in 4 measurements ranged from 0.36 to 0.58 (all p<0.001).

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

Response: Thank you for your valuable advice. We have done HLM analyses within each gender but have not found gender differences in 5-HTTLPR × stress interaction. So HLM model was performed in a whole sample and gender was treated as a covariate. We added the following statement, which has been marked on page 11, line 12-14.

Preliminary analyses indicated no gender differences in models with 5-HTTLPR × stress interaction, and thus fixed-effects component of the model analyses are presented for the sample as a whole.

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?

Response: Missing value in this study was infrequent. In each measurement, just two or three subjects were missed of assessing. In addition, we ensured that every subject completed at least 3 measurements finally. Given that the within-subject approach of HLM model allows missing values, it was acceptable that we did not perform missing data imputation.

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.

Response: Scores of anxiety scale all accorded with normal distribution, which were tested by one sample Kolmogorov-Smirnov tests. We added the statement as following, which has been highlighted with yellow on page 10, line 17-18.

One sample Kolmogorov-Smirnov tests showed that MASC-C scores in 4 measurements all accorded with normal distribution (all p>0.05).

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.

**Response:** Previously, we speculated that the S allele containing haplotype may harbour a disease-underlying variant in the Caucasian populations whereas the L allele containing haplotype may harbour the variant in the Asian populations. But when we downloaded data from 1000 genomes and performed the linkage disequilibrium (LD) analyses, no LD patterns around 5-HTTLPR were found. So we revised the discussion about the ethnic differences of 5-HTTLPR effect as following, which has been marked on page 13, line 12-22.

There were essential uncertainties in ethnic differences of 5-HTTLPR allele frequency, 5-HTT availabilities, 5-HTT uptake and the central serotonergic activity, which might be responsible for the inconsistency. Frequency of the L allele is much lower than that of S allele in Chinese population [32] and Japan population [33] but higher in Caucasians or African-Americans [4, 34]. Meta-analyses have demonstrated the different effects of 5-HTTLPR on selective serotonin reuptake inhibitors (SSRIs) efficacy between Caucasians and Asians [35, 36]. Moreover, functional magnetic resonance imaging (fMRI) studies have shown a link between S allele and higher amygdala activation in response to emotional stimuli in Caucasians [37-39], whereas in Asians L allele was associated with amygdala hyperactivation [40]. Therefore, these fascinating questions of different 5-HTTLPR effects between Caucasian and Asian populations require further investigation.

**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

**Response:** Thank you for your suggestion. We provided accurate p-values in a separate column in all tables.

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

**Response:** Your suggestion was useful. We provided scores of all measurements stratified by 5-HTTLPR genotypes in Table 2, which has been marked on page 22.

13. English in line 10 needs to be clear

**Response:** Thank you for your carefulness, but you did not mentioned the page. We guessed it would be on page 4, line 10 and corrected this sentence as following, which has been highlighted with yellow on page 4, line 10-11.

5-HTTLPR allele frequency varies considerably according to ethnic background

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.

**Response:** Thank you for your valuable advice. We added the information of excluded people as following, which has been highlighted with yellow on page 5, line 20-22.

Of the 692 adolescents administered the clinical interviews, 41 were excluded for meeting the criteria for lifetime major depressive disorder (9, 1.3%), generalized anxiety disorder (11, 1.6%), compulsive disorder (16, 2.3%) or specific phobia (5, 0.7%).

We also deleted the following detailed statements of K-SADS from the revised manuscript as your
suggestion:

Schedule for Affective disorder and Schizophrenia for School-Age Children

The K-SADS is a semi-structured clinical interview tool based on DSM-IV criteria [23] that assesses depressive disorders and schizophrenia in children [19]. The K-SADS has been shown to yield reliable diagnoses of depressive disorders and is frequently used in the field of clinical child psychology [19, 24]. The procedures for training programs of clinical interview using K-SADS were described in detailed previously [18].

15. CES-D often has a threshold effect. Did the authors consider a cutoff rather than a continuous measure for this adjustment variable?
Response: Although CES-D has a threshold effect, it is not intended as a clinical diagnostic tool after all. It was designed mainly to measure current level of depressive symptomatology in general populations. Furthermore, we have excluded individuals with MDD by structural clinical interviews. A high CES-D score of someone might be due to high level of stressful life events rather than episodes of depressive disorder. So in this study, CES-D scores were just considered as continuous measure and were treated as controlled variable.

16. Sentence 14-16 redundant please revise.
Response: Thank you for your suggestion, but you did not mentioned the page. We guessed it would be on page 9, line 14-16 and revised this sentence as following, which has been highlighted with yellow on page 9, line 16-17.

To control for individual differences in age, gender, and baseline anxiety symptoms, these variables were included in this model.

17. The mention of La and Lg alleles felt out of the blue and not properly introduced.
Response: We agreed with your suggestion and added the introduction of rs25531 as following, which has been highlighted with yellow on page 15, line 12-15.

Finally, a single nucleotide polymorphism (SNP) rs25531 within L allele has been described [45]. Due to this SNP, the L allele can be further categorized into LA and LG allele, with LG allele functionally equivalent to S allele [16].

Discretionary Revisions
18. It would be nice if authors provide the variance components in their articles. Please consider doing so here.
Response: Thank you for your suggestion. We added the following statements marked on page 11, line 7-12 of revised manuscript and provided the variance components of model in Table 3 on page 23.

To select a covariance structure for our analyses, we fitted the models utilizing each structure and chose the best fit based on Akaike information criterion (AIC and AICC) and Schwarz Bayesian criterion (BIC). The best fit was a heterogeneous autoregressive structure (ARH[1]). With respect to random effects, the ARH[1] parameter (p<0.001), random slope (p<0.01) and random intercept (p<0.001) were all retained in the model. Table 3 showed the estimates of covariance parameter for the final model.
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.

**Response:** Your suggestion was excellent. We provided the single equation as following, which has been highlighted with yellow on page 10, line 4-7.

\[
\text{ANXIETY}_i = \gamma_{00} + \gamma_{10}(\text{STRESS})_i + \gamma_{05}(5\text{-HTTLPR})_i + \gamma_{01}(\text{timepoint 1 MASC-C})_i + \gamma_{02}(\text{timepoint 1 CES-D})_i + \gamma_{03}(\text{Age})_i + \gamma_{04}(\text{Gender})_i + \gamma_{11}(\text{STRESS})_i(5\text{-HTTLPR})_i + [u_{0i} + u_{1i}(\text{STRESS})_i + e_{ii}]
\]

---

**Response to Reviewer 2: Dr. Kristin Nicodemus**

Reviewer's report:

1. Page 5, lines 18-9: individuals with various psychiatric disorders were excluded, but not those with anxiety disorders. Why?

**Response:** Thank you for your carefulness. We have excluded the participants with anxiety disorders, but have not reported that in the text due to negligence. We added relevant information as following, which has been highlighted with yellow on page 5, line 19 and page 6, line 10-15.

   Participants with neurological diseases, and/or past or current episodes of anxiety disorders, major depression disorder, manic disorder, bipolar disorder, schizoaffective disorder or schizophrenia were excluded.

   Neurological physical examination and the clinical interviews were conducted one-on-one with each participant outside of class time. Interviews consisted of two parts. The Anxiety Disorders Interview Schedule for Children for DSM-IV was used to assess anxiety disorders [19]. Affective disorder, schizoaffective disorder and schizophrenia were diagnosed using the Schedule for Affective disorder and Schizophrenia for School-Age Children [20].

2. Table 1: What explanation might be given for the decrease in MASC-C and ALEQ scores across the timepoints?

**Response:** Decrease in MASC-C and ALEQ during the follow-up period might due to the beginning time of our study. We made the possible explanation as following, which has been marked on page 12, line 5-10.

   Stress and anxiety symptom levels were highest at the initial assessment and decreased during the follow-up period. A possible explanation for this phenomenon is that the students had just entered senior high school when the first assessment was carried out. This major transition may have involved increased competition and academic pressure, elevating stress and anxiety symptoms. The decreases in stress and anxiety levels may reflect acclimation to senior high school.

3. Table 2: The coefficient for 5-HTTLPR x stress is actually negative, but it is not interpreted in the text. The authors should explain what this means in light of the interaction.
Response: Because we coded a 5-HTTLPR genotype as -1, 0, 1 for LL, SL, and SS genotype, respectively. (If we coded LL, SL and SS genotype as 1, 0, and -1 respectively, the coefficient of interaction would be positive.) The negative coefficient of 5-HTTLPR × stress interaction in our study means that with the increase of stress, LL carriers may experience more anxiety symptoms than SS carriers. We added the mixed model equation with explicit writing of the interaction on page 10, line 4-7 and explanation of 5-HTTLPR × stress interaction on page 11, line 18-20.

\[
\text{ANXIETY}_{it} = \gamma_{00} + \gamma_{10}(\text{STRESS})_{it} + \gamma_{05}(5\text{-HTTLPR})_{i} + \gamma_{07}(\text{timepoint 1 MASC-C})_{i} + \gamma_{08}(\text{timepoint 1 CES-D})_{i} + \gamma_{09}(\text{Age})_{i} + \gamma_{10}(\text{Gender})_{i} + \gamma_{11}(\text{STRESS})_{it}(5\text{-HTTLPR})_{i} + [u_{0i} + u_{1i}(\text{STRESS})_{i} + e_{it}]
\]

Specifically, 5-HTTLPR acted as a moderating role in relationship between stress and anxiety symptoms. Compared with SS carriers, individuals with LL genotype exhibited higher levels of anxiety symptoms in relation to SLEs.

4. Figure 1: I don’t know if I am missing something here, but it appears that the intercepts for the different genotypes are different, but their slopes are not, which would indicate no interaction.

Response: In Figure 1, it’s not hard to notice that anxiety scores in three genotype groups would not vary widely without effect of stress. So the intercepts of different genotypes are nearly the same, which confirm no main effect of 5-HTTLPR (B=0.32, \(p>0.05\)). The differences of three slopes seem to be modest, but were demonstrated to be statistically significant (B=–0.08, \(p<0.01\)). Therefore, the interaction was significant indeed.

5. Page 12, lines 14-16: The authors report that they found the opposite effect on depression (in girls only) than the results reported in this manuscript. They controlled for depression in the models presented, but it is unclear if they also controlled for anxiety in the depression analyses reported in [18]?

Response: Anxiety symptoms have not been measured in the previous depression study. So we have not controlled for anxiety in previous study. This was our thoughtlessness. We will pay more attention to this issue in future study.

6. The authors claim that differences might be found due to differential linkage disequilibrium patterns between populations previously studied (e.g., African American, European-ancestry samples). It would be useful to add information on LD patterns from 1000 Genomes instead of speculating on the differences between populations.

Response: Your suggestion is excellent. We downloaded data from 1000 genomes according to your suggestion and performed the linkage disequilibrium (LD) analyses. But no LD patterns around 5-HTTLPR were found either in Caucasians or Asians. So we revised the discussion about the ethnic differences of 5-HTTLPR effect as following, which has been marked on page 13, line 12-22.

There were essential uncertainties in ethnic differences of 5-HTTLPR allele frequency, 5-HTT availabilities, 5-HTT uptake and the central serotonergic activity, which might be responsible for the inconsistency. Frequency of the L allele is much lower than that of S allele in Chinese population [32] and Japan population [33] but higher in Caucasians or African-Americans [4, 34]. Meta-analyses have demonstrated the different effects of
5-HTTLPR on selective serotonin reuptake inhibitors (SSRIs) efficacy between Caucasians and Asians[35, 36]. Moreover, functional magnetic resonance imaging (fMRI) studies have shown a link between S allele and higher amygdala activation in response to emotional stimuli in Caucasians [37-39], whereas in Asians L allele was associated with amygdala hyperactivation [40]. Therefore, these fascinating questions of different 5-HTTLPR effects between Caucasian and Asian populations require further investigation.

Major Essential Revisions
7. There are several typos – please proofread.
Response: Thank you for your carefulness. We have made some corrections with a careful proof reading.

Thank you again for your attention and consideration. We shall look forward to receiving good news from you.

Sincerely yours,

Shuqiao Yao, M.D. & Ph.D.
Professor of Psychiatry and Psychology