Author's response to reviews

Title: Association between Paraoxonase gene and stroke in the Han Chinese Population

Authors:

Guojun Zhang (zguoj1018@yahoo.com.cn)
Wenjin Li (wenjinliwjl@gmail.com)
Zhiqiang Li (lizqsjtu@gmail.com)
Hong Lv (lvhonghong@yahoo.com.cn)
Yonghong Ren (yhren@capitalbio.com)
Ruimin Ma (ma.ruimin@yahoo.com.cn)
Xiaohong Li (xhli@capitalbio.com)
Xixiong Kang (kangxx@VIP.sina.com)
Yongyong Shi (shiyongyong@gmail.com)
Yimin Sun (ymsun@capitalbio.com)

Version: 4 Date: 11 July 2012

Author's response to reviews: see over
Dear Editor,

We would like to thank the editor for giving us a chance to revise the paper, and also thank the reviewers for giving us constructive suggestions which would help us improve the quality of the paper. Here we submit a new version of our manuscript with the title “Association between Paraoxonase gene and stroke in the Han Chinese Population”, which has been modified according to the reviewers’ suggestions. Efforts were also made to correct the mistakes and improve the English of the manuscript. We mark all the changes in red.

Sincerely yours,
Wenjin Li, wenjinliwjl@gmail.com

Corresponding author:
Name: Yimin Sun, E-mail:ymsun@capitalbio.com;
Name: Yongyong Shi, E-mail: shiyongyong@gmail.com.

The following is a point-to-point response to the three reviewers’ comments.

Reviewer #1:

Major comments:

- Poor presentation and formatting: The manuscript should be double spaced, 12 font size and page numbered. It also contains lots of errors in spacing, punctuations, grammar and jargon language.

Answer: Thank you for the comment. We have revised it as suggested.
Do not use measurement unit consistently throughout the manuscript. For example, Odds Ratios and P-values should be 2 decimals, as they were used in Table 1.

Answer: Thank you for the comment. We have revised it in the updated tables as suggested.

Specific comments:

Subjects and methods:

Subjects

- Page 3, paragraph 3: Please provide information on how to ascertain the outcome. Also, please provide the sources of definition of outcomes (i.e., hemorrhagic stroke or ischemic stroke).

Answer: Cerebrovascular diseases include some of the most common and devastating disorders: ischemic stroke, hemorrhagic stroke, and cerebrovascular anomalies such as intracranial aneurysms and arteriovenous malformations (AVMs). A stroke, or cerebrovascular accident, is defined by this abrupt onset of a neurologic deficit that is attributable to a focal vascular cause. Cerebral ischemia is caused by a reduction in blood flow that lasts longer than several seconds. If the cessation of flow lasts for more than a few minutes, infarction or death of brain tissue happens. Transient ischemic attack (TIA) requires that all neurologic signs and symptoms resolve within 24 hours regardless of whether there is imaging evidence of new permanent brain injury; stroke has occurred if the neurologic signs and symptoms last for >24 hours. However, a newly proposed definition classifies those with new brain infarction as ischemic strokes regardless of whether symptoms persist.

Intracranial hemorrhage is caused by bleeding directly into or around the brain. In general, hemorrhagic stroke includes subarachnoid hemorrhage (SAHs) and intracerebral hemorrhage. SAHs are produced by trauma and rupture of intracranial aneurysms. Hypertension and arteriolosclerosis are the most common causes for
intracerebral hemorrhage.

- Please state who read the brain images? Were they trained technologists or physicians?

Answer: The brain images were read by a trained technologists and physicians.

- How did you measure total plasma cholesterol, tryglicerides, high density lipoprotein, and low density lipoprotein?

Answer: Total plasma cholesterol, tryglicerides, high density lipoprotein, and low density lipoprotein were measured by means of enzymatic method, Roche Diagnostics, Germany.

- Even though, common definition of hypertension and diabetes were acknowledged, please provide sources of definition of these outcomes.

Answer: Hypertension was defined as systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg and concomitant use of antihypertensive medications according to the WHO/ISH guidelines [1]. Diabetes was diagnosed by using American Diabetes Association criteria [2].

Genotyping
- Please provide sources of manufacturer’s protocol.

Answer: All the SNPs were genotyped according to the iPLEX Gold Application Guide from Sequenom (San Diego, CA) website. (http://www.sequenom.com/sites/genetic-analysis/applications/snp-genotyping)

Statistical analysis
- In the adjusted logistic regression model, which covariates were included in the model?

Answer: Thank you for the comment. In our study, age, HDL and the incidence of hypertension were included in the adjusted logistic regression model.
- Can PLINK be used to run logistic regression model? If not, please state clearly which software you did use to perform this analysis.

**Answer:** Thank you for the comment. Yes, PLINK can be used to run logistic regression model. And we used the command “plink --bfile mydata --covar mycovar.txt --logistic” to perform this analysis.

- Chi-square and t-test justification should be moved from Notes in Table 1 to this sub-section.

**Answer:** Thank you for the comment. We have removed them in the revised Table 1 as suggested.

**Results**

- Single site association: results from adjusted logistic regression model showed significant association, however, the authors did not provide Odds Ratios and their 95% CIs. Please correct in this paragraph and throughout the manuscript. Only P-values do not show any information about the magnitude and direction of association!

**Answer:** Thank you for the comment. We have added Odds Ratios and their 95% CIs to the P-values as suggested.

**Discussion:**

- Page 6, paragraph 2: Please consider to replace jargon language “What’s more…” (Line 6 from top)

**Answer:** Thank you for the comment and we have changed “What’s more…” to “Moreover” as suggested.

- Page 6, paragraph 2: Please consider to replace “…the speculation…” by “hypothesis” or “assumption”

**Answer:** Thank you for the comment and we have replaced “…the speculation…” by
assumption as suggested.

- When you said about the role of ethnic differences, please expand it further. Answer: Thank you for the comment. And we have expanded it in the discussion part of the paper in page 12 paragraph 2.

- Page 6, paragraph 4: Please consider to replace “… was revealed” by “… observed” or “…found”

Answer: Thank you for the comment and we have replaced “… was revealed” by “… observed” as suggested.

- What are the strengths of this study?

Answer: Thank you for the comment. A major strength of this study is this study was conducted in a relatively large Chinese sample set with careful assessment of two stroke subtypes. Other strength was that we selected common variants in all three members of the PON gene family for genetic analysis. And we have added it in the discussion part of the paper in page 13 paragraph 5.

- Limitations: when you said about the possibility of “survival bias”, please expand more and find a convincing reason to show to audience, do not just leave the paragraph with unavailable data reason!

Answer: Thank you for the comment. Survival bias is the phenomenon by which individuals are excluded from analysis of a trait because of mortality related to the expression of that trait. In case-control studies, variants increasing risk for disease onset as well as risk of disease-related mortality could be difficult to detect. This could possibly result in the underestimation of a variant's effect on disease risk, especially when we study diseases or events with a relatively high mortality rate. However, the effect size erosion because of survival bias is <20% in ischemic stroke [3]. Considering the age of samples studied here, survival bias should not have big influence on our results.
Conclusion:

- Please remove the “… association” before “study” as it is redundant.

Answer: Thank you for the comment and we have removed the “… association” before “study” as suggested.

Reviewer #2:

1. In epidemiological terms, the study is not a cohort. You should call the study a case-control.

Answer: Thanks for the comment. Yes, this study is a case-control study and we mentioned it in the background part of the paper in page 4 paragraph 2.

2. Was the genotyping 100% successful? This would be quite unusual. What was the number of successful calls per SNP and outcome?

Answer: Thank you for reminding us the missing description on the study. The total successful rate of genotyping is 98.6%, which has been added in the paragraph of page 8 in the present paper.

3. The analysis is not clear. Please explain what you did and do not only cite software. More specifically:

A. Why did you adjust for covariates? Due to Mendelian randomization the genetic factors cannot possibly be confounders (please refer to the epidemiological definition of a confounder!). Please drop all adjusted analyses, or if you opt to keep them, provide an epidemiologically convincing explanation for your analytic choices.

Answer: Thanks for the comments. In this case-control study, we wanted to examine the influence of the genetic factors of stroke without being influenced by other factors. However, previous studies have revealed that diabetes, age, sex, hypertension, HDL and LDL do influence the development of strokes. For example, HDL is an established risk factor for stroke and several retrospective studies showed statistically significantly lower plasma values of HDL and LDL in patients with ischemic stroke.
[4-6]. There is also a correlation between blood pressure, age and stroke [7-9].
Moreover, in our study, there is also significant difference in age, HDL and the incidence of hypertension between cases and controls. Therefore, we conducted analyses with adjustment and found that the association of one SNP (rs705381) with stroke remained significant after the adjustment. On the other hand, we also find that multiple logistic regression was conducted with adjustment for factors such as age, sex, history of hypertension, diabetes as covariates in some other genetic researches [10, 11].

B. Is the contrast always referring to the frequency of the minor allele? Please provide a Table with the genotyping results in cases and controls for all outcomes.
Answer: Yes, the contrast always refers to the frequency of the minor allele. We have added the genotyping results in the revised Table 2, 3 & 4.

C. For haplotype analyses, explain how you reconstructed haplotypes (which method?). In the statistical reconstruction of haplotypes we obtain for each person a vector of probabilities corresponding to each possible haplotype. How did you proceed from this stage onwards? Explain whether you used the probabilities for reconstructed haplotypes in the statistical analysis, or whether you rounded them up in some way. Also, how did you handle unsuccessful genotyping calls?
Answer: Thanks for the comment. In our study, the haplotype analyses were carried out by using PLINK software, which is based on multimarker predictors using the standard Expectation Maximization (EM) algorithm and performs simple tests based on the distribution of probabilistically-inferred set of haplotypes for each individual. And in the haplotypes reconstructing step of PLINK, sample with unsuccessful genotyping calls will be excluded from the analysis.

D. For haplotype analyses, describe how you performed the association analysis, i.e., exactly which statistical analysis the SHEsis (?) software does. Because all haplotypes are (should be!) considered jointly, I was expecting to see an omnibus test for
including vs not including the haplotype information in a logistic regression; This could be a joint hypothesis Wald test, or a likelihood ratio test of nested models. At any rate, I was expecting a P-value. The ORs (presumably from coefficients of a logistic regression) refer to an arbitrarily coded reference category and are not useful unless we see all of them for all analyzed haplotypes, and have their covariance matrix as well. Please amend the description of the methods, and provide complete reporting of the analyses.

**Answer:** Thank you for the comment. Haplotype was reanalyzed by using PLINK software (http://pngu.mgh.harvard.edu/~purcell/plink/haplo.shtml) as shown in the revised Table 5. Omnibus test was added with or without adjustment for covariates. The case/control omnibus test is a H-1 degree of freedom test, if there are H haplotypes. We have also revised the methods part accordingly.

E. In Tables 2, 3 & 4, what is the "T-statistic OR"? Is this an unadjusted OR? If yes keep this column and drop the adjusted analyses as per comment 2A.

**Answer:** Yes, the "T-statistic OR" is an unadjusted OR. According to the reviewer’s comment and our reply to comment 3A, we opted to keep the adjusted analyses in the revised Tables 2, 3 & 4.

F. In Tables 2, 3 & 4 list the numerators and denominators for the comparisons in each row. For example in Table 2, additive model you should list the number of minor alleles over the total allele pool (- the unit is the allele -- the denominator should be 2*500=1000 alleles -- or whatever the rate of successful calls was). For the other models the unit is a person, so the denominators should be 500. Do not list the minor allele frequency, it is not useful. People who do meta-analyses would need that information to include your study.

**Answer:** Thanks for the comment. We have revised the tables according to this suggestion in the updated Table 2, 3 & 4.

G. Contextualize this study given previous meta-analyses in the field. One I am aware
of is PMID: 20856122, and studies 2 of the SNPs you included.

**Answer:** Thank you the reviewer for the comment. We have contextualized our studies with two meta-analyses, PMID: 20856122 and PMID: 18511872 as shown in the result part.

H. Please drop the p-values from Table 1. There is no reason for statistical comparison between the cases and the controls.

**Answer:** Thank you for the comment. We have dropped the p-values in the revised Table 1.

I. Please explicitly mention the family of the comparisons that have been Bonferroni adjusted (i.e., adjustments for how many multiple comparisons?). Please explain which P-values have been adjusted (e.g., all, including haplotypes?; only the SNP-ones?)

**Answer:** Thank you for the comment. Bonferroni adjustment was made for then multiple comparisons. P-values were Bonferroni adjusted only for the number of single site analysis, and did not include haplotype analysis in the previous manuscript.

**MINOR REVISIONS**

J. Change "multivariable test correction" to "multiple comparisons adjustment" in the middle of the paragraph "Single site association" in the Results.

**Answer:** Thank you for the comment. We have revised it as suggested.

Reviewer #3:

1, control subjects are younger than cases, but with much higher prevalence of hypertensions. The stroke cases have 17% of hypertension is also questionable. What is the proportion of recurrent of stroke? What is the proportion of cases taking anti-hypertensive medications?

**Answer:** We are very sorry for our negligence. We mistook the incidence of hypertension for the percent of the people who do not suffered from hypertension in
cases and controls. And thus, the percentage shown as Hypertension (%) was been miscalculated. Accordingly to the reviewer’s comment, we carefully checked the data and revised it. So as revealed by the updated Table 1, control subjects are younger than cases and with lower proportion of hypertensions. The stroke cases have 83% of hypertension rather than 17%. The two-year recurrence rate is 12.3%. The proportion of cases taking anti-hypertensive medications is 47.8%.

2, all subjects recruited in the study have 76% and 74% of diabetes, that is exceptionally high. I wonder if stroke cases were from consecutive patients and control subjects were randomly recruited from the clinic?

Answer: We are very sorry for our negligence in the previous manuscript. We mistook the incidence of diabetes for the percent of the people who do not suffered from diabetes in cases and controls. And thus, the percentage shown as diabetes (%) has been corrected as being subtracted by 100%. So as revealed by the revised Table 1, all subjects recruited in the study have 26% and 24% of diabetes in stroke cases and controls rather than 74% and 76%. The stroke cases were from consecutive patients, and the controls were recruited and selected from the health examination department of the Beijing Tiantan Hospital.

3, the rs705381 is associated ischemic stroke, since authors have reread removed cerebral embolism and other rare cases, it would be interesting to know the association of SNPs with large vessel and small vessel strokes.

Answer: Yes, your opinions inspired us. But regrettfully, we do not get the complete information concerning the classification of the ischemic stroke patients. So the association of SNPs with large vessel and small vessel strokes is not available to us as shown in page 14 paragraph 1 in the discussion part of the paper. We also think it will be better and more interesting in our future research by adding this part.

4, does rs705381 associate with HDL level? Does HDL attenuate the association of rs705381 with ischemic stroke?
**Answer:** Thank you for the comment. According to this question, we tested the specific relationship between rs705381 and HDL level and found that rs705381 was not associated with HDL level ($P$ value = 0.8242). As reviewed by Table 1, there is significant difference between stroke cases and controls at the HDL level. However, HDL did not attenuate the association of rs705381 with ischemic stroke as shown by the Table 2, the association of rs705381 with ischemic stroke remained significant after adjustment for age, HDL and hypertension as covariates.

5. English should be improved.

**Answer:** Thank you for the comment. We have carefully revised the English of this paper.

6. there are some redundancy of reference, number of reference should be reduced.

**Answer:** Thank you for the comment. We have reduced some redundant reference according to the suggestion.

**Reference:**


