Reviewer's report

**Title:** A PARK2 Polymorphism (rs1784594) Associated with Delayed Neuropsychological Sequelae after Carbon Monoxide Poisoning

**Version:** 3 **Date:** 4 August 2013

**Reviewer:** Conceicao Bettencourt

**Reviewer’s report:**

Liang et al. present a case-control study, comparing the frequencies of two PARK2 polymorphisms (rs1784594 and rs1893895) between DNS and acute CO poisoning patients, to determine if those PARK2 SNPs are associated with DNS susceptibility. Their hypothesis seems interesting. However, there are several issues in the manuscript that should be improved.

**Major Compulsory Revisions:**

1) In the DNA extraction and genotyping section of Methods, an excessive and unnecessary description of certain details is given (e.g. DNA purity and concentration measurements). However, the genotyping method is not clearly described. Did the authors use the Infinium HD® 660W-Quad Assay or PCR followed by RFLP? Was the Infinium HD® 660W-Quad Assay used just in a few samples to help deciding which SNPs should be studied? Were all samples of each group pooled together or genotyped individually? The authors state “To assess transcript integrity, PCR products were digested…”, but all the analyses were performed using DNA not RNA samples. Was the digestion with Cail performed to see the amplicons integrity or to determine the genotypes for each SNP? This entire section should be improved and made clearer.

2) More details about the statistical analysis should be provided. What was the model used for logistic regression of each SNP’s main effect (i.e. what independent variables were included?). Was the major allele considered as the referent? What was the mode of inheritance used in the model (dominant, recessive, additive, co-dominant)? Did the authors really test for interactions (SNP-SNP or any other)? If so, how was the model constructed? What variables were considered? What mode of inheritance was used in the model? From tables 3 and 5, and without a detailed description of statistics in the methods, it seems that the results shown are not for any interaction, contrary to what the legends indicate. The two tables show the coefficient for an intercept, but that intercept is not the same as an interaction. The results shown in those two tables are just the results for two models of logistic regression adjusted for each of the variables shown in the tables and without any interaction term. Moreover, if DNS is the outcome (dependent variable) it cannot be said for example “Logistic analysis for the interaction between… and DNS”. Did the authors want to say “Logistic analysis for the association between... and DNS”? I believe the authors are confusing association with interaction.
3) As stressed by the authors in the Background section, “the incidence of DPHL is higher in patients over 40 years of age and increases progressively with age”. Therefore, the authors should take into account a possible effect of age, and should adjust their models by age.

4) Are there cases of individuals exceptionally presenting signs of DNS after a latent period of more than 90 days?

5) The two PARK2 SNPs analyzed by the authors are intronic SNPs. Which functional effect do the authors expect? Are these SNPs in linkage disequilibrium with coding SNPs?

6) Besides the association found between the rs1784594 and DNS, it seems that gender is also associated with DNS. Therefore, this result should be highlighted in the manuscript, since it also contributes for the identification of a vulnerable patient group (females) to permanent disability following CO toxicity. Can the authors speculate on why females are more susceptible to DNS?

7) In table 2, I assume that the numbers in parenthesis are the genotype frequencies. However, they are not correct (the sum surpasses 100%). For example, for the first SNP the frequency of the CC genotype in cases should be 17.2% and it is 49.2%. The authors should carefully correct this table.

Minor Essential Revisions:
1) The gene symbol (PARK2) should be in italics.

2) In the Background, the authors state “In a previous study, we conducted genome-wide SNP genotyping of CO poisoning patients with or without DNS using the Infinium human 660W-Quad array [13]”. However, reference number 13 does not belong to the authors nor mentions genome-wide SNP genotyping of CO poisoning patients with or without DNS. The correct reference should be provided. What were the most relevant results from the authors’ previous study? Does the Infinium human 660W-Quad array include the SNPs analyzed in the present study or other PARK2 SNPs?

3) Infinium HD® 660W-Quad Assay is an Illumina platform, not QIANGEN’s as stated in the methods.

4) The authors say in the Statistics section of Methods: “The allele frequencies between cases and controls were first tested for Hardy-Weinberg equilibrium”. Did the authors wanted to say “The genotype frequencies in cases and controls were first tested for Hardy-Weinberg equilibrium”?

5) In the Results, it is “the 1784594 TT genotype frequency was significantly lower versus CC”, and it should be “the 1784594 TT genotype frequency was significantly lower versus CC”. Also, the authors say “Finally, rs1784594 was still associated with DNS when rs1894895 genotype and hypertension were included in the logistic analysis (Table 5)”. but table 5 shows gender instead of the rs1894895 genotype. Thus, it should be “…rs1784594 was still associated with DNS when gender and hypertension were included…”.

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

**Quality of written English:** Needs some language corrections before being published

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

**Declaration of competing interests:**

I declare that I have no competing interests.