Author’s response to reviews

Title: Association of NOS3 gene polymorphisms with Essential Hypertension in Sudanese patients: a case control study

Authors:

Sahar Gamil (dr_sahar_sj@hotmail.com)
Jeanette Erdmann (jeanette.erdmann@iieg.uni-luebeck.de)
Ihab Abdalrahman (drtarawa1@hotmail.com)
Abdelrahim Mohamed (abdelrahim_osman@yahoo.com)

Version: 1 Date: 25 May 2017

Author’s response to reviews:

Reviewer no.1 Tirunilai Padma

1. The sample size of the patients is 260 and the controls 144 and the frequency of the sexes in the two groups is not comparable. The frequency of males and females are respectively 43% and 57% among patients and 77% and 33% among controls. The low no. of females among controls is acceptable only to certain extent since usually aged females do not co-operate to participate in the studies and hence in many studies the frequency of female controls will be lesser than the males.

Response: This is exactly what happened, the controls were mainly males as they were more likely to participate as volunteers. To control for gender as a possible confounding factor on the genotype distribution we did logistic regression (with age, sex and smoking as covariates) in all comparisons and also we looked at the distribution of the genotypes between males and females and no effect was seen.

2. Were the control subjects verified in detail regarding the presence of EHT, family incidence and other associated conditions like diabetes etc.

Response: yes, some of the controls had family history of hypertension but we didn’t consider that in the analysis. None of the controls had diabetes.

3. Is the age of the subjects studied represents the age at onset/diagnosis of EHT or the age reported at the time of recording the cases?
Response: the age studied represent age reported at the time of recording the cases.

4. The mean age of patients among patients (59.68) is significantly higher than the mean age of the controls (36.12). This could have happened because of more no. of control subjects at the lower age group as compared to patients. A table of distribution of age groups with 3-5 yrs of class intervals among controls and patients should be worked out. It will show the frequencies of age of controls and patients at the lower age groups recorded by the authors.

Response: A table of frequencies has been done as suggested by the reviewer and based on the table, we decided to include only participants with age more than thirty or less than 65 as mentioned in the lines 114-116 in the methods part.

5. The younger control subjects must be carrying susceptible genotypes of causative genes and may express the EHT at a later age. Therefore frequency of the age groups of controls should match with the frequency of similar age groups among EHT patients.

Response: After removing participants with age less than thirty and more than 65, the frequencies were pretty much similar between the two groups in each age group.

6. The age range in patients is 28 to 87 and in controls is 19 to 70 years. So the controls below the age of 28/30 years should be removed from the data & entire data should be analysed again.

Response: we deleted participants with age less than thirty as suggested, also we removed age more than 65 to make the two groups closer.

7. The EHT and cardiac diseases are known to have early onset in certain populations like India (around 30yrs) as compared to Western populations. Similarly is the onset of EHT is at younger age among the Sudanese.

8. Are the EHT patients associated only with one or more than one condition like diabetes, MI, stroke, renal failure or hypercholesteremia.

Response: some of the patients had only one disease alongside EH and some had more than one.
9. The data collected should be grouped as 1) patients only with hypertension 2) patients with EHT and diabetes 3) patients with EHT and MI etc. and comparison of distribution of polymorphisms should be made between the groups (if the no. of cases in each group are sufficient).

Response: We didn’t categorize the patients as suggested as the numbers were small especially for diseases other than D.M. but we looked for the distribution of the genotypes according to the presence of these diseases as mentioned in the result section (lines 244-248). And no association was found.

Is EHT expressed in patients first and later on followed by the occurrence of other associated conditions. Renal failure might have occurred due to prolonged and malignant hypertension.

Response: Renal failure cases were managed carefully as we recognize that kidney problem could lead to hypertension (those are considered secondary hypertension and they were excluded from the study) and also prolonged hypertension can been complicated by renal failure (those were included).

Comparison of the data of patients only with EHT and those with EHT and Diabetes (with 80 cases) can be made since the association of a gene variant may be with diabetes rather than EHT among patients with the two conditions.

Response: this was done as we looked at the genotype distribution in patients with and without diabetes (lines244-248) in results section and no association was noticed.

10. Under results, only the mean levels of systolic pressure is mentioned and not the diastolic pressure. It should be added

Response: we added the diastolic blood pressure levels.

11. The title of table -1 can be:  Primers, restriction enzymes, and fragment lengths of the major 614 and minor alleles of 615 rs1799983, rs2070744, and VNTR polymorphisms of NOS3 gene

Response: we didn’t understand what does 614 and 615 represent??
12. The details given in tables 4, 5 and 6 can be combined into one table since the information categorized are similar in the 3 tables.

Response: we combined the three tables into one (Table 4)

13. Since the information in tables 7, 8 and 9 are statistically insignificant. So they can be removed and the details can be given in text under results.

Response: we omitted the tables and mentioned their results in the text (lines 244-248) results section.

14. The title of table -10 can be: Pairwise linkage disequilibrium between the markers rs1799983, VNTR, and rs2070744

Response: title was changed as suggested

15. Haplotype analysis can be done after modify the data as suggested.

Response: data were analyzed as suggested.

16. A brief comparison of the detection of the contribution of NOS3 gene polymorphisms to EHT by investigations in different populations is preferred. A couple of references are:

2. Sushma Patkar, et. al, Risk conferred by 786 T>C polymorphism of NOS3 gene to Essential Hypertension in synergy with smoking and elevated Body Mass Index International J of Current Res. Jan 2011,

Response: the first reference was added but the second was not available

Since it is an attempt to conduct the study of the Sudanese population the manuscripts can be considered for publication after modification
Reviewer 2 “

1. How where the SNPs studied selected? Was there a prior literature review, or do these SNPs have particular interest in Sudanese based on reported allele frequencies or functions? The authors should include this in the introduction or the methods. There are parts of each of these sections that can be deleted - so there should be room (see comment #10). Additionally, why was the GWAS significant level SNP in NOS3 that has been reported in at least 3 HTN GWAS studies not included? (rs3918226). The authors need to justify their SNP selection, and why they included some SNPs in this gene, but not others that have been studied and highly associated.

Response: those three SNPs were selected as they were the most studied during our literature review in 2011 (the rs3918226 is an interesting one to study too but it is relatively a newly studied one (after our literature review). We reported why those 3 were selected in lines 134-136 in the methods part.

2. It is unclear if the genotyping methods are accurate. The authors should include duplicate concordance rates, and genotyping rates for each SNP/variant.

Response: duplicate concordance rates were added in lines 145-146
Genotyping rates were added in lines 202, 203,207,208, 218 and 219

Additionally, for rs1799983, as it was digested with two different restriction enzymes - where the results from these two digestions consistent and concordant? These numbers should be reported.

Response: yes they were concordant as reported in lines 156,157

Also, how were the RFLP results determined: by just one person reading the gels, or by two people reading them, and making sure their results were concordant and employing a third person or additional method when the two people didn't agree?

Response: we mentioned that in lines 155 amd 156

Finally, how do the other allele frequencies compare to 1000 Genomes, and do the authors feel comfortable with the deviation from 1000 Genomes they observe for rs1799983? Does it make sense given the individual African populations - as opposed to the overall African population?
Or with other populations that could represent any admixture that is occurring within the Sudanese population?

Response: yes we attribute the differences noticed to geographic differences and population admixture as mentioned in lines 263, 264

3. Another issue is in the statistical methods. Currently all of the results are unadjusted and there is no correction for multiple comparison. The authors need to either adjust for multiple comparisons, or state why they are not adjusting for multiple comparisons. Given the differences between the cases and controls in Table 2, adjusted analyses seem more appropriate (logistic regression), adjusting for at least age, and gender, and possibly smoking status.

Response: multinomial and binomial logistic regressions were performed to look for the effect of age, gender and smoking status on the genotype distribution.

4. The linkage disequilibrium analysis does not make sense. P-values are not usually presented with linkage disequilibrium results, just the D’ and r-squared values. Additionally, it seems the authors are perhaps unclear on what the differences between D’ and r-squared are, based on the discussion. All of the SNPs are within moderate-high D’ - which would be expected as they are all probably in the same LD block, or at least close to one another as they are in the same gene. The r-square values however, are low, indicating that the SNPs cannot be used as proxies for one another. Thus they likely did not evolve at the same time. Likely SNPs/variants were just created/occurred on one allele version of the SNPs/variants that already existed - giving higher D’ values, and lower r-squared values, as the allele frequencies would be different and not predictive of one another. Using a program like haploview or PLINK may be useful for the LD analysis.

Response: we omitted the p values and changed the discussion as suggested.

5. The difference in age between the cases and controls is a little alarming. Was family history of EH collected within the controls? It is very possible that with controls that are younger than the cases, they could develop EH later in life, especially if they have a family history of EH. This should be discussed as a limitation within the discussion section.

Response: we omitted participants with age less than thirty and more than 65 as mentioned in lines 114-116
6. The authors to discuss the small sample size. A formal power calculation would be nice to determine if there is enough power to observe associations with the SNPs/variants they are testing.

Response: this has been discussed in lines 292-294 in discussion also in the conclusion in lines 333-334. Calculation of power was not feasible.

7. What does "a positive home blood pressure monitoring" mean in line 123?

Response: this has been omitted to remove any confusion. But HBPM is a way of diagnosing hypertension by making different measurements in home of blood pressures and taking the average, it is recommended by the NICE guidelines for diagnosing EH.

8. What is meant by "present as a single copy in the haploid human genome." On line 96? Human have a diploid genome - meaning we have two copies of every gene.

Response: this has also been removed to get out of the confusion but as some genes have multiple copies in the haploid genome, NOS3 has only one copy in the haploid.

9. What is meant by "trans linkage" on line 354? If this is meaning that there is LD between two SNPs even though there is another variant between them - this is not a correct use of the word trans. In genetics 'trans' usually refers to very far effects from a gene - in relational to eQTLs and expression. And usually the trans relation describes loci on different chromosomes or VERY far apart on the same chromosome.

Response: we omitted the phrase.

10. There are paragraphs in the introduction and methods that could be removed or shortened as they contain information that should be known by the audience or that could be summarized much more succinctly or referenced elsewhere:

a. Paragraph on the NOS3 gene lines 93-100

b. DNA sample preparation could be shortened: lines 143-151

c. Paragraph on PCR and Taqman could be shortened or referenced elsewhere if available: lines 164-177
Paragraphs on gel electrophoresis and cloning could be shortened: lines 178-198.

Response: those paragraphs were shortened as suggested