Reviewer’s report

Title: NQO1 gene rs1800566 variant is not associated with the risk for multiple sclerosis

Version: 1 Date: 26 February 2014

Reviewer: Giulio Disanto

Reviewer’s report:

Agundez et al. aimed to investigate the potential association between a genetic polymorphism located in the NQO1 gene (involved in oxidative stress) with risk of multiple sclerosis (MS), age of onset, gender and clinical type of MS. This gene is not included in the list of MS associations provided by large GWAS in MS. However, a previous Greek study has reported the presence of an association, therefore it seems reasonable to attempt to replicate their findings. In contrast with the Greek study, Agundez et al found no evidence of association. I believe it is important to report negative studies to provide evidence against the large number of published false positive associations. However, I have a number of concerns that should be addressed.

MAJOR COMPULSORY REVISIONS

-The quality of English across the manuscript is not great and needs to be improved.

-In the introduction, the authors say that the majority of MS genetic associations are associated with other autoimmune diseases. However, based on the latest GWAS in MS (Beecham et al. Nature Genetics), only “~22% of signals overlapped at least one other autoimmune disease signal”. Please correct

-Table 1: could you please clarify whether each finding comes from EAE or MS? For some studies, this was not clear. Having an additional column providing this information would be helpful.

-Were all individuals included in the study of Spanish Caucasian origin? How was ethnicity assessed?

-How was the association between the SNP and risk of MS assessed? Was it done using logistic regression? The author mention Chi Square, Fisher, t test, Mann-Whitney but none of these statistical tests provide OR with 95%CI. I think the authors used logistic regression but this information should be included.

-The authors provide a power calculation assuming an OR=1.5. Can the authors please provide a power calculation for OR=1.1, 1.2, 1.3 and 1.4? This is important to understand how sensitive this analysis was to detect a difference. Also, are the sample size and statistical power in this study greater than in the Greek study? If so, this validates your results and should be commented.
The authors say that the power of the study was 84.6% and 90.9% for bilateral and unilateral associations respectively. Can the authors clarify what they mean by unilateral and bilateral?

Table 2: EDSS is an ordinal variable and it does not make much sense to provide mean and SD. Provide instead median and IQR.

Table 3: It is not clear to me how the authors calculated the OR and which genotype and allele were used as reference categories. In the legend, the authors write that “Major alleles and genotypes were assumed as reference values”. This is ok, but then I do not understand the meaning of the OR and 95%CI of CC genotypes and allele C. These are supposed to be the reference categories. What were they tested against? In any case, there is no reason to perform the same test twice (e.g. if you test the effect of allele T using allele C as reference, there is no need to test the effect of allele C using allele T as reference. These two tests are equal and indeed the p value is the same). The authors should simply use logistic regression with major alleles as reference category and provide the OR, 95%CI and p values obtained for the minor allele.

In the text the authors say that “allelic and genotype frequencies were not influenced by gender”. If I look at table 3, it seems the authors compared genotypic and allelic frequencies between male cases vs male controls and female cases vs female controls. If so, this is not testing whether frequencies are influenced by gender. This means testing whether the association between this SNP and risk of MS is influenced by gender which is not exactly the same. Please correct the sentence in the text.

Please provide details about each test performed (e.g. age, EDSS and progression index stratified by genotype with the corresponding p values).

The authors say that this SNP is classified as “pathogenic”. I believe the authors need to clarify to what extent the allele T is pathogenic? Based on what evidence? Which disease? It has anyway a high frequency in the general population.

The following sentence is not clear: “The present study has some limitations. First, the size of analyzed cohorts may not be sufficient for strict conclusions about NQO1 role in MS. Second, despite the sample size is adequate to detect an OR as small as 1.5, a more modest association would not be detected (this is a usual weakness of genetic association studies). Third, because the cohort study included MS patients with different degrees of severity, it is not adequate for the investigation of the influence of NQO1 genotypes on the disability or severity of MS (the ideal study for this purpose should include genotyping of patients with a recent diagnosis of MS with similar follow-up periods)”.

The first two limitations are actually the same thing (i.e. the small sample size does not allow conclusive evidence because there is no power to detect smaller effects). The third limitation is also unclear. The problem is not that there are patients with different degrees of severity (you actually need these people
because if every patient had the same severity there would be nothing to test!).
The problem is rather that these patients have different disease durations. Also, the authors say that the ideal cohort to test this should include patients with a recent diagnosis of MS. This is also unclear because if the diagnosis was recent, then you would not be able to distinguish between cases with slow or fast progression. Please clarify.

MINOR ESSENTIAL REVISIONS

-I believe the introduction is too long and the authors provide information that in my opinion is not needed. For example, there is a long paragraph describing the role of putative environmental associations in MS. This could be either deleted or reduced to a single sentence. Please give only the information that the reader actually needs to understand the paper.

-As for the introduction, the discussion can be considerably shortened. For example there is no need to list the IDs of all non-synonymous SNPs located within NQO1.

-I suggest table 1 be included as supplementary material

-Table 3: Are percentages with their 95%CI within the brackets? Please clarify this because it is not clear.

Level of interest: An article of limited interest

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