Author's response to reviews

Title: Polymorphisms in xenobiotic metabolizing genes (EPHX1, NQO1 and PON1) in lymphoma susceptibility: a case control study

Authors:

Pablo Conesa-Zamora (pablo.conesa@carm.es)
Javier Ruiz-Cosano (xavierumxxiii@hotmail.com)
Daniel Torres-Moreno (daniel-torres@orange.es)
Ignacio Español (daniel-torres@orange.es)
María D Gutiérrez-Meca (mdgumemae@yahoo.es)
Elena Pérez-Ceballos (elena.perez@carm.es)
Rocío González-Conejero (rocio.gonzalez@carm.es)
Javier Corral (javier.corral@carm.es)
Vicente Vicente (vicente.vicente@carm.es)
Miguel Pérez-Guillermo (miguel.perez-guillermo@carm.es)

Version: 2 Date: 20 December 2012

Author's response to reviews:

Reviewer: Olga Gra
Reviewer's report:

Discretionary Revisions

-It is desirable that the authors have pointed out in the discussion of the obtained data the cytogenetic location of genes EPHX1, NQO1 and PON1, because the absence of information about the location of these genes on different chromosomes generates a number of questions regarding the linkage analysis of studied polymorphic variants.

Author response (AR): Information on cytogenetic location of EPHX1, NQO1 and PON1 is now included in the discussion

-It is also desirable to analyze gene-gene interactions to assess the combined effect of genotype and allele frequencies for the development of haematopoietic malignancies.

AR: Following the reviewer comment we have now included linkage disequilibrium and haplotype analyses. Therefore, the following sentences have been added in the Methods:

"Hardy-Weinberg, linkage disequilibrium (LD) and haplotype analyses were performed using the online program Shesis Page (http://analysis2.bio-x.cn/myAnalysis.php) [28]"

Results:

"In order to ascertain the possible combined effect of the genotypes a LD and
haplotype analyses were performed. D´ and r2 were far from 1 value and thus demonstrated that the polymorphisms studied were not in linkage disequilibrium (Figure 1). Haplotype analysis showed that the only significant combination was T-C-G for EPHX1, NQO1 and PON1 polymorphism, respectively rendering an OR of 1.73 (95% CI: 1.22-2.45; p=0.00192) (Table 4).”

and Discussion sections:

“As demonstrated by the haplotype analysis, the combination of the PON1, EPHX1 and NQO1 polymorphisms did not increase the lymphoma risk associated with the PON1 polymorphism alone (OR: 1.7 vs. 1.5)”

Reviewer: Kathryn Barry
Reviewer's report:

This case-control study aimed to evaluate interactions between single nucleotide polymorphisms (SNPs) in three xenobiotic metabolizing genes (EPHX1, NQO1 and PON1) and exposure to smoking and chemical pollutants (both environmental and occupational) with respect to B-cell lymphoma. This is an important topic that has been relatively little studied, in part because of limited statistical power to assess interactions in many studies. The present study was also limited by small numbers. Even with examination of broad groupings of exposures, there were still small cell counts for many exposure/genotype combinations, including some cells with a count of zero (Table 2). Participants were classified as having residential exposure to environmental pollutants based on residence in a heavy industrial area in Spain for 10 or more years. Additionally, participants were classified as having occupational exposure in general if they regularly used or were exposed to any of a myriad of chemical compounds in certain occupations (including chemical industry, construction, agriculture and metallurgy). However, the study was unable to evaluate exposure to individual chemicals. My principal concern with the paper is that the authors did not acknowledge the limitations of the study. In addition, there are some problems with interpretation that need to be addressed before the paper can be considered for publication. I have also described some additional concerns below.

Authors’ response (AR): We acknowledge this comment. Certainly, this is one of the most important limitations of our study, and accordingly, it has been clearly stated in the revised manuscript as we indicate latter.

Major Compulsory Revisions
1. Why was a recessive genetic model selected? Sometimes there were very small cell counts for different genotype/exposure combinations, especially for the
homozygous variant group (Table 2). It seems that this problem could have been reduced by choosing a different genetic model (for example, the dominant model, such that heterozygous and homozygous variant genotypes would be combined).

AR: The reviewer is right, the dominant model might increase the statistical power but it failed to find significant associations between rs662 and B-cell lymphoma. This result suggests that only the homozygous genotype might have consequences, therefore sustaining a recessive model. When considering allelic distribution and the co-dominant model this significant association is also observed. This information has been included in the revised manuscript. Small cell counts are only present in specific sub-analysis, (according to exposure). This limitation and the requirement of a cautious interpretation of these results is now stated in the manuscript.

2. The authors do not mention any limitations of their paper in the Discussion section. Specifically, it is important to mention the small cell counts, and more attention needs to be paid to the fact that the authors were unable to evaluate individual chemical exposures. There was some mention of this in relation to residential exposure, but not occupational exposure. The authors should be clear that this limitation could introduce noise into the analysis because the pattern of association for the various SNPs with B-cell lymphoma by exposure could differ for different chemicals (based on different pathways of activation and detoxification).

AR: We agree with the reviewer’s opinion. We have now included in the discussion the following critical comment concerning the limitation of the study:

“However, these results need to be taken with caution given important limitations of the study. First, the limited sample size did not allow us to test a gene-environment interaction. Therefore, in order to take into account the possible environmental influence we stratified both cases and controls according to exposure. The observation of higher frequency of PON1 G allele in cases than in controls has the aim of drawing attention on possible genotype selection in the exposed population that has to be confirmed in subsequent independent studies with larger sample size in order to ascertain this effect. Another limitation of the study is that the heterogeneous nature of the industries and pollutants in the study area could introduce a certain degree of noise when analyzing polymorphisms in some, but not all, genes involved in the metabolism and detoxification of the chemicals involved. Taken together, our results could serve as a starting point for future studies in which pollutant activity and the genotype influence could be more defined.”

3. There are also several problems with the interpretation. The authors state that they observed an association between the PON1 SNP and B-cell lymphoma in males, but not females. However, it appears the p-value for females was also statistically significant (although barely, rounding to 0.05) based on the fact the CI for females did not cross 1.0. In addition, the OR point estimates for males and females were relatively similar to each other (4.1 and 3.2) and the associated CIs were wide and had substantial overlap. The authors should modify their interpretation to say that they did not observe much evidence for a
differential effect by gender. Also related to interpretation, the authors conclude that their study is the first to demonstrate a relationship between the PON1 polymorphism and BCL risk where the association was dependent on proximity to chemical industries. This is an overly strong conclusion given that there was a zero cell count for one of the genotype groups with residential exposure, which likely influenced the findings.

AR: Concerning the effect on gender, we did not mention in the previous version that the Bonferroni correction in the estimation of significance had been taken into account, thus considering p < 0.016 as significant. This has now been added in material and methods. With this proviso, a dissimilar association seems to be observed between genders despite the fact that males and females were equally represented in the study population. The reviewer comment on the conclusion is pertinent and the strength of this conclusion in the abstract and at the end of the manuscript has now been reduced by saying: “This is the first study demonstrating a relationship between the rs662 PON1 polymorphism and BCL risk whereas the observed association dependent on the proximity to chemical industries to be confirmed in future studies”

4. Although it appears the major aim of the paper was to evaluate interactions, no p-values for interaction were presented. The authors should add this information to Table 2 and also describe the associated methods in the Methods section. In fact, the authors do not mention anywhere in the Statistical Analysis section how they handled the exposure information in the analysis.

AR: As now stated in the discussion (see response to point 2) we could not evaluate gene x environmental interactions. In fact, there was no aim of identifying interactions in the manuscript but just an intention to compare the genotype distribution in non-exposed and exposed subjects. Considering possible gene-gene interactions we have now included LD and haplotype analyses which are now presented as figure 1 and table 4, respectively.

5. The discussion needs to be more balanced to include null studies as well as those that showed an association. For example, the authors state that there is some support for a role of EPHX1 rs1051740 in the genetic susceptibility to cancer, but they do not mention studies that did not demonstrate an association.

AR: We agree with the reviewer’s comment. Accordingly, in the revised manuscript we have included references finding no association for the three SNPs with different neoplasias.Refs: 30, 31, 34, 35, 38, 39.

6. The Introduction should cite some gene x environment interaction papers that have already been conducted with xenobiotic metabolic genes and environmental/occupational exposures in relation to lymphoma.

AR: References with the information indicated by reviewer are included in the revised manuscript: Refs: 17, 18, 19

Minor Essential Revisions

1. The authors state in the Background that they aimed to evaluate correlations
between SNPs in xenobiotic metabolic genes and various exposures; however, it
seems a more appropriate description would be an evaluation of interactions
between these factors.

AR: As stated in response to point 2, we could not evaluate interactions because
of the limited sample size and thus we stated that the association observed when
comparing the genotype distribution in exposed and non-exposed has to be
taken with caution and it should be validated in future studies. In order to a better
description of the aim of the study, the revised manuscript states (Page 3):
“The aim of our work was (...) to compare the genotype distribution between
subjects, with and without chemical pollutant exposure in cases and controls,
using information obtained from medical records and a detailed ad hoc
questionnaire”.

2. In the Statistical Analysis section in the Methods, the authors state that the
sample size reached in their series was “optimum according to the sample size
estimation for a gene-only study.” This statement is misleading because the
major aim of the study was not to look at main effects of the SNPs, but rather to
look at interactions between the SNPs and environmental and occupational
exposures.

AR: As stated in response to point 2, we could not evaluate interactions and, for
this reason, the gene-only estimation was performed as this is the main objective
of the work.

3. Table 3: typo (LNH should be NHL).

AR: This error has now been corrected

4. Please also check the grammar (for example, comma use) throughout. For
example, in the first paragraph of the Background, a comma should be added
after “may increase the risk of NHL” and the comma should be removed before
“remains controversial.”

AR: A new revision by an English native has been performed to detect this and
others grammatical errors

Discretionary Revisions

1. More detail is needed on study subject selection. The authors state in the
Methods that their study subjects come largely from a previous described series
of patients. How were the rest of the participants selected?

AR: In order to solve this issue the following sentence has been added: “The rest
of participants included in the study cases were those diagnosed during 2011
who accepted the participation in the study.”

2. More detail is needed on the specific chemical compounds to which subjects
were exposed. Although it appears that some of this information was presented
in a previous paper, some information along these lines is also needed here.
Also, the authors state that occupational exposure was defined as regular use or
exposure to chemical compounds. How was regular use defined?
AR: Following the reviewer’s suggestion we included in page 4 a brief description of pollutants and now define more precisely the term regular use or exposure to chemical compounds by adding the following sentence: “Regular use was defined as daily manipulation of chemical compounds in the referred industries”

3. The number of decimal places / significant figures retained for p-values is not consistent throughout the paper.
AR: The number of decimal places has now been uniformed in the manuscript

Reviewer: Pierluigi Cocco
Reviewer’s report:
Scientists in various branches of medical research frequently criticize case-control studies because of their retrospective nature, as, in their opinion, this study design would inherently convey bias. To prevent such limitations and criticism, epidemiologists have developed well defined standard procedures for case-control studies, which are particularly important when exploring associations with non genetic risk factors, whether occupational, environmental, dietary or other lifestyle related risk factors. One of the major concerns relates to the source of controls, that should be defined as an unbiased sample of the general population within which cases are recruited. Therefore, using a definitely selected group of controls such as blood donors, raises fundamental issues of validity for the parts of this study that consider non genetic risk factors.

Besides, in their analysis, the authors use an extremely broad definition of environmental and occupational exposure, with no indication of what would be the environmental emissions in the area in question, how far they might reach, how distant the study subjects resided from the source of each emission, and in what direction in respect to the prevalent wind. Not to mention the use of GIS mapping and models for the distribution of airborne emissions, which, nowadays, are standard approaches in environmental epidemiology. The same applies to the stratified analysis by occupational exposures, which cannot be defined as the universe of industrial work or the whole book of agrochemicals: these are thousands of different chemical agents, each with its own toxicological profile.

Old fashion occupational epidemiology used job titles and groups thereof (and many explorative papers keep doing the same), but this has always been meant as a preliminary scrutiny for further more in depth investigation of the true risk factor.

The authors engaged themselves in the very difficult task of exploring gene-environment (G x E) interactions. To do so reliably, one must refer to an E factor that, if not matching the extreme precision of the G side (the single nucleotide polymorphism), would at least restrict the range of possibilities to the group of chemicals that might be affected by the polymorphism of that specific metabolic gene. For instance, PON1 is implicated in the metabolism of organophosphate insecticides. Did the authors make any attempt during the detailed interview to collect information on what type of crop did the study subject with farm work cultivate, if not what type of crop disease did they treat or what
type of pesticide did they use? Unless these requirements had been matched in designing and conducting the study, I would strongly recommend the authors to focus their attention to the genetic side of their work, and to cut any reference to environmental and occupational factors from the abstract, methods, results and discussion. For instance, as it concerns the results and discussion, drop the presentation of results from line 16 (Interestingly…) on page 6 through the end of the page. Leave the last sentence of the results (on lymphoma subtypes). Cut the last two columns of Table 1, and the last four rows of Table 2 (any result related to occupational and environmental exposures). Cut also the lines 11-14 and from line 16 to the bottom of the page on page 8.

Authors´ response (AR): We agree with the reviewer´s comment. Therefore, the study association without taking into account the residential or occupational exposure has now been considered as the main objective and conclusions of the work. Consequently, the reference to chemical exposure has been deleted from the title and the important limitations of this study in this issue have now been highlighted throughout the manuscript. In order to provide the scientific community some clues that could be of help when studying the complexity of gene x interaction we have decided to keep the discussion of our results when comparing exposed cases and controls stressing the importance that these should be taken with caution and that they are subjected to validation in independent studies with larger sample size and a more adequate exposure assessment.

Another issue is related to the inclusion of Hodgkin lymphoma in the B-cell lymphoma definition. The authors should make clear what type of lymphoma classification was applied, and whether was any diagnostic revision preformed by an expert pathologist. If they want to maintain HL in their case series, they need to refer to lymphoma overall and not B cell lymphoma in their title and all over the text.

AR: The reviewer´s comment is correct. All our cases were diagnosed by our pathologists and, as suggested, a reference it has now been included the classification. The term B-cell lymphoma has now been deleted from the manuscript.

Further minor points.

1. Please, add the necessary functional information on the SNPs that were tested.
AR: This information is provided for each polymorphism in the discussion.

2. The description of epi methods is inadequate; ORs and 95% CI are reported in the results, but there is no indication on whether a conditional or unconditional logistic regression was used and what covariates were included in the model.
AR: The following sentences have been added in material and methods: “For the gene-only associations, logistic regression was used conditioned for gender and age” and in result section: “Age-conditioned logistic regression confirmed the association of GG genotype (vs. AG+GG) with disease (OR= 1.7; CI (95%):
1.1-2.7; p=0.02) as well in gender-conditioned logistic regression (OR= 1.7; CI (95%): 1.1-2.4; p= 0.01). Considering the co-dominant model, the protective effect of AG (OR=0.6; CI (95%): 0.4-0.9; p= 0.02) and AA (OR=0.6; CI (95%): 0.4-0.9; p= 0.02) genotypes was observed”.

3. The distribution of cases by lymphoma subtypes is quite unexpected. Do the authors have an explanation on the reason why follicular lymphoma has the highest prevalence in their case series?

AR: In order to explain this pertinent observation we have added the following explanation: “All patients from the HUMM institution were follicular lymphoma (FL) and, for this reason, most of the patients included in this study (135/215) belonged to this type of lymphoma on the basis of the availability of DNA specimens from a previous study”

4. The authors mention that “Major reasons for nonresponse were subject refusal, physician refusal and death”. It seems weird that cases were the lesser respondent. Presumably, the 47% acceptance rate presented in Table 1 was due to death. The authors should make this point clear.

AR: Following reviewer’s suggestion, major reasons for non response are now given in percentages in the result section. “Major reasons for non response were subject refusal, unable to be interviewed or death (for controls: 84.8%, 10.9%, 4.3% and for cases 39.5%, 17.5% and 43%, respectively)”