Reviewer’s report

Title: Polymorphisms in xenobiotic metabolizing genes (EPHX1, NQO1 and PON1) in B-cell lymphoma according to residential and occupational exposure: a case control study

Version: 1 Date: 10 November 2012

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.

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).

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).

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.

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.

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.

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.

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.

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.

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

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.”

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?

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?

3. The number of decimal places / significant figures retained for p-values is not consistent throughout the paper.

**Level of interest:** An article of limited interest

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

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.