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: 30 November 2012

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.

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 performed 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.

Further minor points.
1. Please, add the necessary functional information on the SNPs that were tested.
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.
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?
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.

Quality of written English: Acceptable

Declaration of competing interests:

I declare that I have no competing interests.