Reviewer's report

**Title:** Ethnicity, medical history and incidence of Hodgkin's disease in Canadian population

**Version:** 1  **Date:** 18 October 2008

**Reviewer:** Marie-Claude Rousseau

**Reviewer's report:**

Summary: Ethnicity has been suggested as an important factor in the epidemiology of Hodgkin's disease (HD). The authors investigated, using data from a population-based case-control study conducted in six provinces of Canada, the association between ethnicity, medical history, and incidence of HD adjusting for pesticide exposure. A questionnaire was sent by mail to 316 men diagnosed with HD and to 1506 controls eliciting information on ethnicity and personal medical history. Information regarding exposure to pesticides was collected via a telephone interview. The study found that compared to North Americans descendents, the risk of HD was significantly greater among the Eastern European and Western European descendents.

General impression: This is an interesting analysis based on sound data collected as part of a population-based case-control study conducted in six Canadian provinces. The manuscript needs to be improved before publication, and specific suggestions for improvement follow.

**MAJOR COMPULSORY REVISIONS**

Clarification of the research question and presentation of the background:

1- The research question is not clearly stated, either in the abstract or the background section. The authors do provide a literature review of studies that considered ethnicity as a potential etiological factor for HD, along with a rationale. However, the background does not address medical history except to mention Epstein-Barr virus (EBV) infection which is not included in their study. The rationale for studying extensive medical history is not presented. The authors also mention studies that focused on the association between pesticides and HD. It is not clear whether they consider exposure to pesticides as a main exposure or as a potential confounder. Clarifying the aim of the study will help structure the background and provide a stronger rationale for this manuscript.

Appropriateness and description of the methods:

2- The methods seem appropriate and were previously published elsewhere. However, readers would benefit from a clearer description of the methods specific to this article. On the 5th line of p. 6 (Methods), the statement “...three matched controls for each NHL case.” is confusing. I suppose that there were 3 matched controls for each case in that study, which included Soft Tissue Sarcoma, Hodgkin’s Disease, Non-Hodgkin’s Lymphoma, and Multiple Myeloma.
cases, and that only the HD case series is used in the current analyses? Also, it is not clear whether all or only a subset of the controls were used. Please clarify. The methodology should cover the current analysis, and not the whole study.

3- The use of questionnaires and interviews in population-based case-control studies is appropriate. On the one hand, the authors provide general categories of the data collected (demographic information, personal medical history, cigarette smoking history, lifetime occupational history and specific occupational exposures of interest). On the other hand, a detailed list of variables relating to personal medical history is presented for the first time only in Table 3. As mentioned previously, the authors need to justify this long list of medical conditions.

4- The definition of ethnicity is interesting and clever, making use of the reported ethnicity of the four grandparents. The description is clear, except for one aspect. How are subjects classified if they have only 2 grandparents of the same ethnicity or if all 4 grandparents are from different ethnic groups? Among the 7 ethnicity categories, the “Other” category has a high prevalence in some provinces (38% in the Prairies). Could the authors provide a little more information, even if just descriptive, about what are the relative proportions of the different ethnicities included in the group? This would surely vary greatly by province.

5- The paragraph “Pesticide” on p. 8 does not seem necessary. However, the authors could clarify the definition of “Pesticide Exposure”. Was exposure to pesticides defined according to answers to the broad questions (major classes), to the chemical groups, to the individual substances, or through a combination of all of those? Was it coded as ever/never exposed? Which classes or groups of pesticides were considered? The text on p. 7 says “the listed pesticides”, but there was no list provided.

6- Statistical analysis section (p. 9): The authors need to clarify why conditional logistic regression was used if the controls were frequency-matched, not individually-matched, to cases?

Presentation of results

7- There is no mention of the number of eligible subjects and hence no mention of participation rate. In the description of the distribution of ethnicity by province on p. 10, end of 1st paragraph, it is not clear why these numbers were chosen and why they are presented in this order. This leaves the reader without a clear message.

8- For the reader’s interest, I suggest that Table 1 be modified to include some sociodemographic and other characteristics (e.g., average age, education, smoking, pesticide exposure) stratified by province. Otherwise, very little information is given to describe the study population.

9- Table 2 should also be revised to differentiate between cases and controls in each region. More information should be provided on the “other” category for
ethnicity. Maybe the authors could provide the most common ethnicities under this category for each province as a footnote to the table?

10- In Tables 4 and 5, the authors should include the numbers of cases and controls. It is not clear why Table 4 presents the parameter estimate and SE, while the other tables don’t. This information is redundant with the OR and 95% CI.

Discussion and conclusions:
11- Although interesting, the first paragraph of the discussion is unrelated to the results presented in this manuscript, and should be removed, unless the objective of the paper is also to look at incidence according to age.

12- The authors mentioned that there are several ways of defining ethnicity in epidemiological research, and make a case for the importance of a good definition. How does their definition compare to that used in other studies?

13- The authors need to discuss the strength of the epidemiological evidence in relation to the association between ethnicity and HD. They should better distinguish between studies that supported and those that refuted the association. Is their study in agreement with most previous studies? How does it differ or resemble previous investigations?

14- What do the authors mean by the sentence: “Further, because of financial constraints, women were unable to participate in this study.” on p. 15?

15- The conclusion section is a repetition of the results, in telegraphic style. What do these results mean? What are the implications of these findings and future research directions?

MINOR ESSENTIAL REVISIONS
16- Could the authors verify that references 2 and 3 really pertain to the importance of ethnicity in the etiology of HD?

17- p. 5, 2nd paragraph, 3rd line: “Exposure to pesticides has been associated with an increased incidence of Hodgkin’s disease in many studies, but not in all.” References are not provided for studies which did not suggest an association. In addition, it would be interesting to mention what studies in Canada have shown on this topic.

18- p. 8, 2nd paragraph, last line: No references as to which pesticides were shown to be teratogenic, mutagenic or carcinogenic.

DISCRETIONARY REVISIONS
There were instances where the writing was not clear, or where there were typographical mistakes. Here are some of them:

19- p. 4: same sentence repeated at the end of 1st paragraph (sentence before last) and beginning of 2nd paragraph.
Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

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.