Reviewer’s report

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

Version: 1 Date: 12 October 2008

Reviewer: Lindsay M Morton

Reviewer’s report:

The authors report results of a population-based case-control study of Hodgkin lymphoma conducted in Canada. Main findings include differences in Hodgkin lymphoma risk by ethnicity, some medical history factors, and pesticide use.

Major compulsory revisions:

1. It is not clear why conditional logistic regression was used to analyze the data. How were the strata defined? What is the rationale?
2. The “univariate” results presented in Table 4 are irrelevant, in light of stated differences in the age distribution by cases and controls. Also, variables used in the design of the study for matching (age, province) should automatically be included in the models.
3. The abstract could be improved to more clearly state the rationale and methodology. Currently, the rationale does not explain at all why ethnicity, medical history, and pesticide exposure are included in the same analysis, which is important to justify as these are typically thought of as three distinct risk factors. Use of the word “database” implies that perhaps you had registry data, but in fact the design is a population-based case-control study. Include information about the telephone survey to collect pesticide information. The abstract should also include terminology to make it clear that conditional logistic regression is the appropriate analytic approach. The abstract also implies a large amount of missing data regarding ethnicity (27% unspecified), but in fact those participants just had more diverse ethnicities and thus weren’t grouped into your categories, right? It would be more accurate to describe them as “other” ethnicity. Finally, the conclusions state information about herbicides, but this finding was not mentioned in the results.
4. The introduction also needs to more clearly state the rationale and, as suggested above, explain why ethnicity, medical history, and pesticide exposure are included in the same analysis.
5. Methods, page 6: It seems unlikely that participants would be able to provide information regarding the specific chemicals that are in the pesticides they used. If participants were really queried regarding chemicals (rather than just deriving them from known formulations), provide justification that this is valid. Also, the method of ascertainment is unclear – the text currently implies that participants were just asked about ever/never use, but then there is mention of the number of hours of use. Please clarify.
6. Methods: More information about the ascertainment of medical history data is needed.

7. Results, page 10: Risk estimates by ethnicity should be discussed with the descriptive data on ethnic distribution in the first paragraph of results.

8. Table 4: Results are only presented for those risk factors with $p<0.2$, but the authors never provide a comprehensive list of what was evaluated. The parameter estimate and standard error are unnecessary if odds ratios and confidence intervals are presented. Sample sizes would be helpful to add, though.

9. The discussion is too lengthy and contains much information unrelated to the analyses at hand, yet completely lacks information comparing the findings of these analyses to those in the literature.

10. The text would benefit from editing to smooth the flow of sentences and ideas.

Minor essential revisions:
1. The appropriate terminology is “Hodgkin lymphoma” according to current standards.

2. Abstract, Results: Suggest rewording phrase “Based on our data, the distribution of ethnic groups in Canada constitutes…” to “In our study population, the distribution of ethnic groups was…”

3. Background, page 4: Provide references for statements regarding changes in incidence, statements regarding “Another study conducted in the United States…” and statements regarding “Findings from both parts of our study…”.

4. Background, page 4: The sentence “Research has shown that ethnicity is a significant factor in the epidemiology of Hodgkin's disease” is vague. Be more specific about the findings.

5. Methods, page 6: The authors state that frequency matching was used, but then describe caliper matching (for individuals) by age within +/- 2 years. Clarify whether individual or frequency matching was used. If frequency matching, give age groups (e.g., 5 year age groups).

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**
I declare that I have no competing interests.