Reviewer's report

Title: Influence of family size and birth order on risk of cancer: a population-based study

Version: 2  Date: 8 August 2010

Reviewer: Jacqueline Clavel

Reviewer's report:

General comments

This paper reports an analysis of the influence of birth order and sibship size on the risk of adult cancer. It is based on the large and unique Swedish family cancer database. The question is well defined and interesting. The paper tends to be too fast on many points and should be improved by a more precise and rigorous presentation.

Minor Essential Revisions

There are some misleading phrasings:

1- Introduction:

a. the sentence line 2 suggests the relationship between birth order and birth weight is well established. Is it really so? If yes, please explain and develop this statement, else rephrase it so that local observations do not appear as generally admitted results.

b. Similarly, observing a negative relationship (as reported between testicular cancer and high birth order) should not be described as detecting a protective effect unless the relationship is causal. Therefore, the sentence should be either developed, in case evidences for a causal association are available, or moderated.

c. Again, family size and birth order influence the early environment and lifestyle, but they are not actual environmental factors.

d. Lastly, some information given at the top of page 4 should rather take place in the Methods section.

2- Methods

a. Clinical information on tumors is available for most cases. However, only CIM7 is used. What information is available and for how many cases and for which cancers?

b. The reason why individuals with a family history of cancer are excluded should be explained more clearly. If the motivation by the risk of confounding, it would have been more interesting, in my opinion, to analyze the data with and without these individuals and to elicit this confounding, since this database offers this rare opportunity. If it is a question of selection bias due to the fact that parental cancer could have censored the size of the family, it should be explained (may
be the authors have a possibility to evaluate such a bias, which could be interesting? May be the motivation is elsewhere.

c. In an attempt to alleviate the issue of multiple comparisons, the authors used 1% confidence intervals. However, given the number of tests, it is far to address the issue and finally, does not really weaken it either. I think that 5% confidence intervals are sufficient.

d. Please, give the power of the study for different cancers before and after 50 years old.

3- Results:

a. Page 6: The comments on the methods used in table 2 should take place in the Method section and in footnotes, rather than in the text of the Results.

b. page 7 lines 4: “Smaller family size had a protective effect on”. Something like “smaller family size was negatively associated with” would be more adequate, given that causality is not established. Also Page 6: “the effect of family size” would be replaced by something like “the association between family size…”.

c. Page 7 lines 6-8: “as most… non-significant results” is not exact. It means that the results would have been significant if the numbers had been sufficient. This is only an assumption, the facts are that numbers are small, power is too low, and results are not significant.

d. “Risk” is sometimes used instead of “Relative Risk”.

e. Legends of the tables:
   i. poisson should begin with a capital letter
   ii. the legend of table 3 on family size indicates “first-born child” as reference. Is it really the case?

Major Compulsory Revisions

4- Results

a. Is there any possibility to distinguish subtypes given the available clinical information? In particular, could chronic lymphoid leukemia, chronic myeloid leukemia and acute leukemia be separated? papillary and medullary thyroid cancers?

b. What are the relationships between the variables of interest, birth order and size of family, and between them and socioeconomic status? Are these relationships different before and after 50 years old?

c. Is obesity available in the database? If yes, the relationship between endometrial cancer and birth order.

d. Figures: I would have preferred tables rather than figures, all the more that figure 1 puts together figures on lung cancer with birth order and figures on stomach cancer and family size for stomach cancer, without confidence interval.

e. Tables 2 and 3 are the most important and should be included in the paper rather than additional. The figures should be put in a table. Table 1 could be described in the text and given as an additional file.
5- Discussion
a. The discussion section does not discuss the validity of the results, and often takes the observations as if they were causal. The discussion of the relationships between endometrial cancer, socioeconomic status and obesity is too far from the facts and the observations and should be improved.

In conclusion, the paper could be good because the data are very interesting but it has not been written thoroughly enough and should be improved

**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