Reviewer's report

Title: Prevalence of birth defects and risk-factor analysis from a population-based study in Inner Mongolia

Version: 3 Date: 28 March 2012

Reviewer: Analee Etheredge

Reviewer's report:

Major Compulsory Revisions

1. Methods, paragraph 4. Since birth defects are considered to be a rare event, it is statistically appropriate to use Poisson regression for data analyses instead of logistic regression, which assumes that the outcome is not rare. Additionally, Poisson is standard in the birth defects literature. It is strongly suggested that the authors modify their analyses.

2. Discussion. I had hoped the authors would talk about some of the seemingly contradictory prevalence estimates. Specifically, I am interested to know what they think about the very high prevalence in third (or later) pregnancies but also a very high prevalence in women under 20 years of age. Given what we know about reproductive habits in very young women and the age cutoffs selected by the authors, it doesn’t seem likely that women are on their third pregnancy before they hit 20. What do the authors make of this? This seems like one of the most interesting results in the study, but it is not discussed at all.

Minor Essential Revisions

3. Generally, when talking about the occurrence of birth defects, it is more appropriate to report prevalence rather than incidence. Since incidence refers to new events in a population at risk (e.g., number of birth defects divided by the number of conceptions), and there is not adequate information to calculate incidence (number of birth defect cases lost due to early pregnancy loss), prevalence is the more appropriate terminology.

4. Methods, paragraph 3. Refers to hematomas as a birth defect included in this study, and Table 2 specifies “hemartoma.” Please clarify if it is hamartoma (or even hamartoses) that is meant.

5. Equinovarus in Table 2 should be: Talipes equinovarus

6. The first paragraph in the Results section is a little confusing. It is unclear from the text what the final denominator is. For the overall prevalence, it appears that the denominator is 62,529, as suggested in Table 1 (although the text says 62,443), and the numerator is 976, as suggested in the text. Table 1 has a total of 961 cases—are there 15 cases missing nationality data?

7. The categories in Table 4 do not appear to be complete or accurately labeled. Please include in the table or the text (or both) what variables were included in the multivariate models used to calculate the adjusted estimates.
Discretionary Revisions

8. In the Background section, paragraph 1, the authors present a reference from Western Australia, but do not clarify why this region was chosen to highlight. Is there an aspect of the referenced study that is similar in some way to the current study?

9. Background, paragraph 1, sentence 5. Please provide reference for the statement “…the proportion of congenital abnormalities varies greatly among countries and regions.”

10. Background, paragraph 1, sentence 6 and Discussion, paragraph 1, sentence 7. Contextually, Taskinen, 1990 seems an unlikely primary reference.

11. Background, paragraph 1, sentence 8. Please provide reference for the statement “…less than half of birth defects can be attributed to a single factor.”

12. Methods. Please provide more detail on ascertainment and surveillance.

13. Methods, paragraph 1. Please define stillbirths and abortions. Do stillbirths include spontaneous fetal deaths < 20 weeks gestational age? Are abortions inclusive of those spontaneous and induced?

14. Methods, paragraph 1. Please clarify whether or not the denominator includes stillbirths and abortions, in addition to live births. It is most common for prevalence estimates of birth defects to include live births only in the denominator.

15. Methods, paragraph 1. Please clarify how cases were counted. If a child had multiple malformations, were they counted separately for each defect?


17. Methods. Were cases classified as syndromic vs. non-syndromic?

18. Methods. How were chromosomal anomalies classified?

19. Methods. Given that there is much debate about the shared etiology of anencephaly, spina bifida, and encephalocele, it might be helpful to separate the neural tube defect classification into the three respective phenotypes.

20. Overall, the results, while very interesting, are hard to read. It is suggested that all of the results go into tables. As a reader, I value the ability to look at the count data so that I can see how the authors arrived at an answer or even to do my own calculations in the margin. Also, including count data makes a publication more tractable for inclusion in meta-analyses or even for comparison across studies.

21. Figure 1 is not effective for conveying these important results. Actual prevalence values need to be reported, preferably in a table.

22. The opening sentence of the Discussion is transitory-- perhaps the first sentence is missing?

23. How were your comparison countries (Discussion, page 8, paragraph 1) selected? Any particular significance?
24. Discussion, paragraph 2, second sentence, the authors reference 2 studies but mention one.

25. In addition to agricultural exposures, is there any chance that the rural locations might provide higher exposures to other environmental teratogens compared to the urban locations? For example, are there mining activities or groundwater contamination issues in any of the rural locations? We know that there are some of these exposures in Inner Mongolia and neighboring Shanxi, but it would be most interesting to know if the rural sites included in this study share these characteristics.

26. Discussion, paragraph 2, sentence 4. The sentence “After adjustment for other factors…” seems out of place here and does not clearly convey that a comparison is being made between the current study and the one referenced, nor does it describe the purpose for the comparison.

27. Discussion, paragraph 3, last sentence. This sentence suggests that the authors have reason to believe that smoking is a risk factor for birth defects. Is there a teratologic mechanism or epidemiologic data that can be used here to support the authors’ hypothesis?

28. Discussion, paragraph 4, first sentence. The comparison made between risk of birth defects of any kind (this study) and the specific birth defects evaluated in other studies does not seem appropriate. Since the current study did not present data evaluating the variables of interest on each birth defect, it is not necessary to include this reference. Also, reference 17 appears to refer to a study on paternal age?

29. Discussion, paragraph 5. This is a difficult paragraph. First sentence: The references for this general statement about consanguinity seem a little out of place. Maybe reference studies that evaluate broader populations? Second sentence: it is clear that consanguinity, family history and race/ethnicity share a common thread that relate to genetics, however, the second part of the sentence “…however, familial inheritance and ethnicity were related to birth defects” really does not belong here.

30. Discussion. If family history and race/ethnicity remain important after the Poisson analyses are conducted, the authors really need to provide more discussion on these results. They are of great interest epidemiologically, and perhaps etiologically. Are there future plans for genetic component to this study?

31. Discussion, paragraph 7. This paragraph is not terribly informative. I would much rather read about what the authors are excited about studying next, or their ideas on how this study can be used to advance the field.

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

**Quality of written English:** Not suitable for publication unless extensively edited

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

I declare that I have no competing interests