Reviewer's report

Title: Spatial and temporal interpolation of air pollutants in New York City: exposure assignment for use in a birth outcomes study

Version: 1 Date: 25 April 2013

Reviewer: Jo Kay Ghosh

Reviewer's report:


This manuscript describes the exposure modeling methodology for a birth cohort in New York City, based on vital statistics data. The analysis evaluates the performance of the exposure models for PM2.5 and NO2, and provides comparisons by buffer size. The analysis and topic area are highly relevant to air pollution health effects studies. The manuscript would be greatly improved, however, with some clarifications in the Methods section (detailed below) as well as a more focused Background section that clearly states the objectives of the work.

-------------

- Major Compulsory Revisions

1. Methods, analysis, paragraph 3 – description of how you averaged the PM2.5 data for every day vs every third day is confusing. For NO2, were there only 2 monitors in total – please specify in methods. Please provide examples of how you adjusted based on the relationship between one monitor to the average of all monitors.

2. Please discuss why the spatial model for NO2 was not as predictive of measured levels as the model for PM2.5. In the Discussion paragraph noting the correlations between PM2.5 and NO2, a discussion of the differences between these pollutants (their sources, photochemistry, etc) would improve the paper.

3. Methods, paragraph 4 – why were two-week windows used for the temporal adjustment, rather than the actual date span of the pregnancy period? This would alleviate the issue with the final two-week window extending beyond the birth date.

- Minor Essential Revisions

1. Background, First paragraph – too wordy. Try to trim it down to improve readability.

2. Background, End of paragraph 2 – the statement about ambient monitors networks being sparse in urban areas is correct, but the argument to justify the need for better exposure assessment methods is weak. Make an argument for why using exposure models that capture fine spatial variations in air pollutants is
important in epidemiologic studies, e.g. reduced exposure misclassification. Alternatively, emphasize that one of the aims of this paper is to compare the temporal adjustment method against the nearest monitor approach (I believe this comparison is a strength of the paper, although the results are largely tucked away in the Supplement).

3. Background, Paragraph 3 – “While daily and hourly data ... does not generally EXIST...”. Also, Citation #12 does not appear to support this statement.

4. Methods, paragraph 1 – specify that you are using an LUR method

5. Methods, paragraph 1 – how many monitors were used for the temporal adjustment, and were all available data points averaged? Please provide details on your approach.

6. Methods, Data, paragraph 1 – please list all the variables included in your model

7. Methods, Data, paragraph 2 – for the daily air pollution averages, how many observations were excluded? Any evidence of bias?

8. Methods, Data, paragraph 3 – describe distribution of gestational ages in the final dataset. Also, provide details on how you defined the begin/end dates of each trimester. For those pregnancies that had incomplete 2nd and 3rd trimesters, did those subjects have 2nd/3rd trimester averages computed?

9. Results, spatial estimates, paragraph 1 and 2 – add SD to your description of annual PM and NO2 levels

10. Results, last paragraph – use “seasonal” rather than “temporal”

11. Discussion – first paragraph – while this work would certainly be used for a birth outcomes study, a more appropriate description here (and in the title perhaps) would be “prenatal air pollution exposures”

12. Discussion – comment on limitations due to lack of data on residential mobility and time-activity patterns. Have there been any studies detailing the extent of mobility in this population, and how it might impact exposure measures? Similarly, comment on the potential for bias based on time-activity patterns expected in this population.

13. Figure 4 – vertical axis needs labels

- Discretionary Revisions

1. Methods, paragraph 1—“we could not use a single source...”. Please rephrase this to emphasize how combining these data sources is effective in capturing spatial and temporal variations in air pollution.

2. Methods, data, paragraph 3 – by “unique locations”, I assume you mean a residential parcel, e.g. apartment building, house, etc. You may want to clarify this in the text.

**Level of interest:** An article of importance in its field

**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 I have no competing interests.