

## Reviewer's report

**Title:** A Multi-city Time-series Analysis of Air Pollution and Emergency Department Visits for Cardiac and Respiratory Conditions

**Version:** 1 **Date:** 24 January 2009

**Reviewer:** Jennifer Peel

### Reviewer's report:

This is a relatively well-written and interesting manuscript. However, there are several details that need to be clarified throughout the manuscript. Additionally, there is an over-emphasis on the statistical significance of the findings and not enough synthesis of the numerous results presented. The statements about the power of the study are difficult to evaluate because the power of a time-series study is really determined by the average daily number of visits and the length of the time-series, and it is difficult to find this information in the manuscript (it appears that the length varies by city).

### Major Compulsory Revisions

- **Methods:** Many details are missing from the Methods section. For example, how many pollution monitors are available for each city? What is the geographic size of each city (a map may help here)? What temporal resolution is available for the pollutants (hourly, daily, etc.)? How do you handle missing data? What years are included in the study?
- **Methods, ED visits:** Were all ED visits included? Or just residents of a geographic area? What years were included? How did you handle repeat visits (e.g., repeat visits by a single person within the same day or week)?
- **Methods:** Do correlations among the pollutants.
- **Methods:** The AIC and other tests do not evaluate confounding. Why were these used?
- **Methods:** Sensitivity analyses should be performed for all associations, not just those that are statistically significant (so the authors can get a better sense of the effect of the sensitivity analyses).
- **Methods:** In the various lag models, did you match the lag for meteorologic variables with the lag used for pollution? This is generally preferred.
- **Methods:** Do you ever present information about how many knots for time are in each model (the final model)? The number of knots in the sensitivity analyses seem to be very high compared to other published studies (e.g., 4 knots per year or 12 knots per year is typical).
- **Why did you choose to use the daily average for pollutants?** Many previous studies have examined 1-hour max or 8-hour max, particularly for gaseous pollutants.

- Methods: Is the method used for combining city-specific results typical? How does this compare to what NMMAPS or APHENA or similar multi-city studies have used?
- Method: Why did you choose to examine the individual lags 0-2? A moving average or distributed lag model would have been easier to interpret. As it reads now, there are numerous statistical tests and the statistically significant results are emphasized. This method has been shown to be biased. It seems that you could have narrowed down the amount of tests by using previous literature. For example, many studies have shown that cardiovascular events often have a short lag period. Some respiratory outcomes (e.g., asthma) have been observed to have a longer lag structure with pollution. One approach is to pick a primary lag structure, and then do secondary analyses to evaluate that choice.
- Results: It would be helpful to be able to see the sub-daily results. Could these be placed in an online supplement?
- Discussion: What have other studies observed with regard to seasonal (warm vs. cold) analyses?
- Table 1: Provide more descriptive statistics (it may need to be in more than one table) – e.g., standard deviation, IQR. What were the percent of days with data for the gaseous pollutants?
- It seems that some cities should not have been included for certain pollutants (e.g., Ottawa with 11% and 31% of PM10 and PM2.5 data).
- Table 2: provide the average daily number of visits and the standard deviation (more percentiles would be helpful, too).

#### Minor Essential Revisions

- Abstract: Add in some information about the time period of the study
- Throughout the manuscript (starting in the abstract), you need to clarify the comparison group when making statements that an estimate is x-fold larger (larger than what?). It is not always clear.
- Background, page 4: The percent of ED visits that are hospitalized would vary by condition, correct? Is it not much higher for CVD visits?
- Methods: Provide reference for EPA methods.
- Background: More and more relevant references should be included in the first sentence (no references are currently provided for hospital admissions): e.g., large multi-city studies (NMMAPS, APHENA), review articles, etc.
- Methods: How was the unit increase in pollution used for results chosen? Is it the mean? Typically studies use the IQR.
- Results: Can you provide more details about the missing data? E.g., are there any patterns to the missing data (e.g., certain seasons)? Is ozone measured year-round?
- Results, page 12: provide 95% confidence intervals rather than p-values.

- Discussion: There is an updated publication for the Atlanta ED study (Tolbert et al., JESEE, 2007).
- Discussion: The statement about the false positive results is misleading. Performing numerous statistical tests does not change the probability of false positive results (the alpha level is set), but you may see more of the false positive results (i.e., if you do 100 tests you would expect to see 5 statistically significant results and if you do 1000 tests you would expect 50, but the probability is the same).
- Discussion: The negative results are not discussed much. The results for some of the smaller outcomes are largely negative (not necessarily statistically significant for the smaller outcomes). Are there other reasons for this (other than chance)? Perhaps the control of temporal trends? Inverse correlations with other pollutants? Are these consistent across seasons?

#### Discretionary Revisions

- The term “weather” is not very technically specific. “Meteorologic factors” is generally preferred by atmospheric scientists.
- Discussion, first sentence: the term “significance” should be reserved for statistical significance. Consider using another term, such as “important.”
- Abbreviation NAPS is probably not needed.

**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