Reviewer’s report

**Title:** Highway proximity associated with cardiovascular disease risk: the influence of individual-level confounders and exposure misclassification.

**Version:** 1  **Date:** 20 February 2013

**Reviewer:** Ryan Allen

**Reviewer’s report:**

**Summary:**
This cross-sectional study of traffic-related air pollution and inflammation explores the sensitivity of health effect estimates based on different exposure assessment methods and adjustment covariates. Studies of chronic exposure and inflammation have had mixed results, so as a purely epidemiologic study these results make a contribution to the literature. The discussion of geo-locating errors also has value. The paper is generally well written, although I think the results section could be made more concise.

I have a few general comments/concerns, as well as several more specific comments.

**General comments:**
1. Some of the main conclusions from the sensitivity analysis portions (geo-locating, confounding) of the paper may not represent a big advance over our current knowledge:

   - Much of the paper is dedicated to the issue of confounding. A main conclusion from the paper is that it’s important to adjust for individual-level confounders (e.g., Abstract “Our analyses emphasize the importance of controlling for individual-level confounders.”) I would argue that this is already known. For example, in a previous study of chronic exposure and inflammation Hoffman et al. (ref #7) adjusted for age, area of residence, smoking behavior, ETS, BMI, waist circumference, physical activity, alcohol consumption, LDL, and HDL. Is inadequate control for individual level confounders a major problem in this area of research or, put another way, does this study go significantly beyond what previous studies have done to address potential confounding bias at the individual level?

   - In addition, when introducing the importance of confounding, the authors state that “In the absence of pollutant exposure measures, proximity could represent traffic noise or gradients of socioeconomic status (SES) near heavy traffic, raising the need to carefully address potential confounders [11].” Reference 11 is a review on the CV effects of noise. Yet, the paper does not address noise at all.

   - A recommendation is made that “future research will need to go beyond using proximity and, instead, assign individual exposures to residents.” But the paper...
names only 3 previous studies that look at inflammation using a road proximity exposure assessment approach, while at least as many studies have assessed inflammation vs. measured MODELED exposures. So while this paper provides some valuable numbers on the sensitivity of effect estimates to different geo-locating approaches, I think that the limitations of simple proximity metrics are well recognized and many studies are already going “beyond using proximity”.

2. The paper would benefit from additional discussion of how generalizable these results and methods are to other settings and to larger study populations.

For epidemiologists who might read the paper, it would be useful to discuss how easily these geo-referencing techniques could be applied in other settings (e.g. are high resolution aerial photos available and inexpensive in most locations? Is this approach easily implemented in a study of several thousand individuals?) In addition, how generalizable are the results, especially since one recruitment area was dropped because “highway geometries and street canyons complicated assignment of simple proximity values.”

3. The paper should discuss the advantages of road proximity for exposure assessment.

I agree that proximity metrics have limitations, but they also have advantages. Perhaps the most important advantage – and a major reason that proximity metrics are relatively commonly used – is the ease of communicating results to a non-scientific audience and/or policy-makers. To say “risk is elevated by X% within Y meters of busy roads” is much easier to communicate and interpret than “an X ug/m3 contrast in pollutant is associated with a Y% increase in risk.” I think this should be considered when discussing the advantages/disadvantages of road proximity metrics in epi. studies.

4. Given the major focus on road proximity as a surrogate exposure metric, the paper should provide some rationale for why these specific distance categories were selected:

Hoffman et al. (cited reference #7) used distances of <50m, 50-100, and 100-200. Williams et al. (cited reference #6) used a binary cutoff of 150m, while also exploring ranges of 0-50, 50-100, 100-200, 200-300, and 300-500. Other cardiovascular studies have also used close/far binary indicators (e.g. Gan et al., Epidemiology, 2010 defined exposed as <50 m from a road or <150 m from a highway and unexposed as anything beyond those distances). Why/how were the distances in this study chosen? How sensitive are the results in this type of analysis to alternative cutoff distances?

Specific comments:

- Abstract should specify that this was a cross-sectional analysis, and when presenting results in the abstract confidence intervals should also be provided.
- The term “exposure adjusted” is used throughout the paper. Please introduce
the term when describing the models at the end of the Methods section.

-“Only a few health studies have reported pollution exposure as well as distance.” I think I understand what is meant here, but I suggest rewriting this sentence to more clearly communicate that few studies have estimated exposure using both proximity metrics and measurements/models of concentration.

-Although much of the paper is devoted to specific issues around exposure assessment, the epidemiologic analysis and results themselves have value. Thus, I suggest adding a bit more discussion on how these results compare to previous research in this area. The magnitude of the effects (~40-50% increases in CRP near roads) seem quite high. How do these compare to previous findings?

-What floor did people live on? Someone living right next to a highway but 25 floors up will have different exposure than someone living at street level.

-Tables 2 and 3: I suggest limiting these analyses to participants for whom all variables and geo-locating methods are available. That way the tables compare the different models/methods across the exact same group of participants.

**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:**

No competing interests.