Reviewer’s report

Title: Exposure to Hazardous Air Pollutants and Risk of Incident Breast Cancer in the Nurses' Health Study II

Version: 0 Date: 11 Oct 2017

Reviewer: Julia Brody

Reviewer’s report:

This study evaluates breast cancer risk associated with modeled Hazardous Air Pollutants among the nationwide Nurses Health Study II cohort, applying the methods used in the California Teachers Study.

I appreciate the value of applying the methods used by the California Teacher's Study (CTS). However, I think it would be better to also think through ways in which those methods could be adapted or improved in this dataset.

Please replace the term "replicate" throughout. It is conceptually important in environmental health to understand that human studies cannot be exactly repeated. In this study, specifically, we know that breast cancer is influenced by exposures during windows of susceptibility in the life cycle and that the disease includes subtypes that are more/less common by menopausal status, so we expect that analysis in the mostly premenopausal Nurses II is not the same as the mostly older CTS.

The differences between the two studies in parity and age distribution are substantial, and I recommend noting this early in the article, so that readers can keep this in mind throughout. Otherwise, readers who are not already familiar with CTS do not encounter this important "design difference" until the discussion.

Are you able to consider the effects of occupational exposures to the same chemicals? Some of the HAPs are likely exposures in hospitals, and other carcinogens and endocrine disruptors are, too. For example, when did the ethylene oxide standards go into effect? Are the nurses likely exposed to this sterilant? Please discuss nurses' exposures to these and other mammary carcinogens and EDCs, and how this relates to the study design.

P 5, line 36 "Women were excluded from the current study if they had been diagnosed with cancer … during follow-up." Obviously, women with breast cancer weren't excluded, so maybe you mean other cancers?

p. 6. In this study, women are aged 25-42 at the beginning of the exposure period (residential histories start in 1989). Past breast cancer research suggests that exposure to carcinogens early in life is most important, before the first full-term pregnancy when the breast becomes mature. Does NHSII have adequate power to analyze risk limited to exposure before the first pregnancy?
The use of only 2002 HAPs data is explained, but this remains a limitation. How much does this matter - do locations change very much in rank order across time periods dating back to earlier years?

How does the distribution of absolute exposure levels in NHSII compare with CTS (Garcia Figure 1)? Off the top of my head, I might think more CTS women are exposed to substantial air pollution than would be found nationwide.

Diesel exhaust contains many mammary gland carcinogens, and this should be mentioned in addition to classifying diesel as estrogen-disrupting. For example, search for "diesel" in the Mammary Carcinogens Database: http://sciercview.silentspring.org/mamm_search.cfm.

Potential confounders. In addition to my comment above about chemical exposures in nursing, it seems important to consider the possible role of shiftwork. How might that confound or modify this analysis?

Do you have data on exposure to environmental tobacco smoke? What about models limited to smokers? It seems like these could be important to further evaluate in a study of air pollutants.

The women are aged 25-42 at recruitment in 1989. Follow-up extends to 2011, when they would be aged 47-64. Please clarify the definition of the average age 44.4, so people understand what "throughout follow-up" means. Similarly, % premenopausal. CTS describes its study population at baseline (1995). It appears CTS women are older. 63% of CTS cases and 49% of non-cases are menopausal. 23% of CTS cases and 26% of non-cases are nulliparous. These differences between the NHSII and CTS study populations are relevant to hypotheses about effects of pollutant exposures.

You may be interested to know that there is elevated breast cancer risk on Cape Cod in locations affected by 2,4-dinitrotoluene from explosives at a military base. https://www.ncbi.nlm.nih.gov/pubmed/12003750?dopt=Abstract

If indeed diesel increases breast cancer risk, even with relatively modest HR, this would be of major public health importance, because diesel exposure is so widespread, breast cancer is so common, and diesel is a modifiable exposure. In general, the use of the word "only" seems unhelpful.

Did you consider analyzing mixtures other than diesel?
The footnote in the supplemental tables S2 and S3 seems to be copied from Table 2, but the supplemental tables show two different models, one not fully adjusted, so this is confusing.

In the methods, you say "each variable (or set of variables) was added to models adjusted for age and calendar period to assess confounding, but we don't learn how that led to the inclusion of the large set of variables in the multivariable models.

**Level of interest**
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons
CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal