Reviewer's report

Title: Association of traffic-related hazardous air pollutants and cervical dysplasia in an urban multiethnic population: a cross-sectional study

Version: 1 Date: 1 April 2014

Reviewer: Paolo Giorgi Rossi

Reviewer's report:

Major compulsory

The results are interesting and not obvious. Unfortunately, it is impossible to understand if the study design is sound or not:

1) the inclusion criteria are not reported, the HPV status is not report in tables, but only mentioned in the text and it is not possible to understand it is referred to cases or to both cases and controls (in the first case the study is very weak, in the second it can be sound).

2) cervical intraepithelial neoplasia usually grows very slowly, furthermore the clinical significance of CIN is more relevant when the lesion is older, i.e. the probability of regression is high for new lesions and lower for long lasting lesions; this means that the lag time is relevant: women exposed for a long time should have a higher risk not only for a dose-response effect, but also because there is a substantial lag time.

3) The text has several small errors in reporting, sometimes they make impossible to understand what has been observed (cancer and neoplasia are used I suspect for the same disease; the OR for continuous variables are reported without the unit increase, there is no mention of the distribution and if a transformation was useful)

Minor essential

abstract

1. define the inclusion criteria of the women, please explain if all the women are HPV positive. this is the most important point of the study design: in the presence of a necessary cause, the risk factors should be clearly divided in those influencing the probability of having the necessary condition (i.e. infection) and those influencing the probability of developing the disease once the necessary cause is present.

2. The two analyses, continuous variables and for quartiles, are testing the same concept with two different formulation of the alternative hypothesis. The authors should decide which is the most appropriate and then make their conclusions according to that analysis. the alternative analysis should be presented only to corroborate the results or to confirm the hypothesis of linear trend (if the authors decided to rely on the analysis on continuous variables). My suggestion is to present only the analysis on continuous variables and to dedicate some words to
explain the unit increase (without this explanation the OR is meaningless!).

methods
3. please define inclusion criteria for the trial. How is it possible that all the women (or only the cases???) were HPV 16 or 18 positive? Was this an inclusion criterion? If it is so, why there are 69 subjects with unknown HPV status? If it is not an inclusion criterion the positivity for HPV16 or 18, how is it possible that you did not find any other high risk HPV? HPV 16 or 18 are responsible of less than 50% of the infections in the general population and of about 2/3 in high-grade CIN.

4. Please report briefly the treatment and diagnostic management.

5. please define the endpoint (for this study the case) and the non cases: which CIN are included? CIN1, 2 and 3? The inclusion of CIN1 is critical since they are not pre-cancerous lesions. If they are included a sensitivity analysis should be added excluding CIN1.

6. In my opinion the principal analyses are those stratified for HPV status (i.e. excluding the HPV unknown).

results
7. The difference in exposure between the most and the less exposed group is not so high: is it an effect of the distribution (i.e. the centile includes also very high outliers) or of the average of census tract?

8. please better describe the cases.

9. page 9 second par: please report the unit increase for the continuous variables.

10. the analysis with categorical variables should be presented before to show the prerequisite of the continuous analysis, i.e. a substantially linear effect. After that the analyses that should be considered for the test of hypothesis should be only the ones on continuous variables: the authors should decide which are the most appropriate analyses and not leave the decision to the readers.

11. Please consider to present the analyses restricted to HPV known and just report the analyses including the HPV-unknown as a secondary analysis.

12. pag 9 par 2, line9: cervical cancer or neoplasia?

13. It could be interesting to see the OR for smoking, age and ethnicity, may be in a supplementary table, just to confirm existing data and to give more reliability to the study.

discussion
14. Again the authors should decide which is the most appropriate analysis and discuss this one, the other analysis has the meaning of sensitivity analysis (to test the robustness of the results) or to show that testing a linear relationship is a reasonable assumption.

15. discussion pag 10 first par: This is the first place where the authors explain what is the unit for the linear trend OR!
16. clearly report which kind of lesions (this is missing in the whole paper, but it should be reported also in the discussion)

17. Among the etiological hypotheses there are some including the immunological response.

18. limits: please discuss better the problem in assess temporality: the main limit is that the authors do not have the history of exposure.

19. the use of CIN2+ as surrogate outcome for cervical cancer is highly validated but should be discussed; in particular, the validation is for incidence but dealing with prevalence more assumptions are needed.

Conclusions

20. the conclusions are confusing and refer to things that are not mentioned in the introduction or in the discussion (urbanization in Latin America...)

**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'