Author’s response to reviews

Title: Road proximity, air pollution, noise, green space and neurologic disease incidence: A population-based cohort study

Authors:

Weiran Yuchi (weiran.yuchi@ubc.ca)
Hind Sbihi (hind.sbihi@ubc.ca)
Hugh Davies (hugh.davies@ubc.ca)
Lillian Tamburic (lillian.tamburic@ubc.ca)
Michael Brauer (michael.brauer@ubc.ca)

Version: 1 Date: 15 Oct 2019

Author’s response to reviews:

(see attached response document)

We thank the reviewers for their thoughtful, constructive comments. Our responses to each comment are provided below, and where applicable the locations of corresponding changes to the manuscript are provided in parentheses.

Reviewer #1:

(1) This an interesting paper on the association between neurologic disease incidence and exposure to road proximity, air pollution, noise and greenness. The authors investigated non-Alzheimer's dementia, Alzheimer's disease, Parkinson's disease, and multiple sclerosis. Although I find the paper important, I think that the results are in some way overstated. The only outcome consistently associated with proximity measures was NAD, while the only outcome associated with exposure to different air pollutants (PM2.5 and NO2) was PD. MS resulted associated to major road proximity and to PM2.5 (once greenness was taken into account).

We thank the reviewer for this feedback. We have revised the text in several places to be more suspect regarding results (Line 23-30, 262-266, 276-285). While it appears that the reviewer’s conclusions may have been based mainly on statistical significance, we observed elevated effect estimates with confidence intervals that were not symmetric around 1 for all road proximity models (except for Highway <150m for multiple sclerosis). We based our conclusions on consistency and detailed examination of effect estimates and confidence intervals, not merely based on statistical significance as per common practice.
(2) To rephrase the main results in the abstract, result section, and discussion.

See response to comment 1 above. For example, the abstract now states:

“Road proximity was associated with all outcomes (e.g. non-Alzheimer’s dementia hazard ratio: 1.14, [95% confidence interval: 1.07-1.20], for living &lt;50m from a major road or &lt;150m from a highway). Air pollutants were associated with incidence of Parkinson’s disease and non-Alzheimer’s dementia (e.g. Parkinson’s disease hazard ratios of 1.09 [1.02-1.16], 1.03 [0.97-1.08], 1.12 [1.05-1.20] per interquartile increase in fine particulate matter, Black Carbon, and nitrogen dioxide) but not Alzheimer’s disease or multiple sclerosis. Noise was not associated with any outcomes while associations with greenness suggested protective effects for Parkinson’s disease and non-Alzheimer’s dementia.” (Line 23-30)

In the Results we conclude:

“Associations between air pollutants, AD and MS were generally null with wide confidence intervals. There was no evidence of primary effects of noise, while greenness was associated with increased ORs for both AD and MS (Figure S3 in Additional file 1). In sensitivity analysis, no increasing patterns of effect estimates were observed between air pollution estimates from CANUE and any of the outcomes.” (Line 262-266)

In the Discussion we write:

“We observed increasing patterns of associations between road proximity across subcategories (e.g. Major road &lt;50m, Major road &lt;50m or Highway &lt;150m etc.) with hazard and odds of developing NAD, PD, AD and MS. PM2.5, black carbon and NO2 were associated with increased incidence of NAD and PD. While noise did not affect any of the associations with air pollution, there were indications that greenness was protective for the development of NAD, PD and AD. Overall, we saw indications of associations between air pollution with NAD and PD, but in general not with AD or MS. We did observe increased OR for MS in association with PM2.5, but not for any of the other air pollutants. We found associations for air pollution from locally developed LUR models, while sensitivity analysis using air pollution exposure estimated in national exposure models showed no associations.” (Line 276-285)

(3) To avoid the replication of results in tables and figures and to simplify the report. I found that 5 tables and 4 figures are too many for a paper. I suggest keeping in the tables the results for proximity measures and to keep the figures with air pollution exposure, noise and greenness.

We thank the reviewer for this feedback. We combined the previous Table 1 and 2 into one table (current Table 1). We combined the previous Table 3 and 4 into one table (current Table 2). The previous Table 5 was renamed to Table 3.
We only kept one figure in the revision, Figure 1 (previous Figure 3). We moved the previous Figures 1, 2 and 4 to the Additional File (current Figure S1, S2 and S3).

(4) In Table 1 and Table 3 specify the comorbid conditions and keep Table 1 in one page.

Change made.

(5) I suggest having a Table 2 with descriptive statistics with proximity measures for all the outcomes (NAD, PD, AD, MS)

The study design for the NAD and PD analyses was a cohort study, while study design for the AD and MS analyses was a nested case-control study. For this reason and to emphasize these differences, we separate the descriptive statistics into two tables, one for NAD/ PD and one for AD/ MS (also see response to comments of Reviewer 2).

(6) I suggest a figure or a Table with air pollution, noise and green exposure for all the outcomes.

The current Table 3 (previous Table 5) contains all such information.

(7) Discussion. Page 15. Lines 308-309. The authors should discuss all the results, including the "harmful effects" of greenness

We thank the reviewer for this feedback. We added: “Greenness was associated with higher incidence of both AD and MS. While this may indicate a harmful effect of greenness proximity on these outcomes, this finding may also result from residual confounding due to unmeasured and/or unidentified spatially-varying risk factors for AD and MS. Overall, we did not find associations with AD or MS for any exposures besides road proximity.” (Line 287-291)

(8) Section results, page 12, lines 230, 239, 240 use "non cases" instead of "controls"

Change made.

(9) Format the references according to the journal guidelines

Changes made.

Reviewer #2:

(1) This is a very well conducted cohort/nested case controls study of multiple environmental factors (proximity to roadways, noise, air pollution and greenness) and several important neurological outcomes in the elderly. This study is based on a very strong resource i.e. the Canadian health care system that allows identifying diseased individuals from a health care database that covers almost everyone and also has access to extensive data to
model the exposures of interest well for this population. The comparison of the satellite derived air pollution and land use regression results is interesting and another strength, as suggests that local contrasts in pollution levels (micro- rather than national scale derived estimates of exposure) and sources are the most important contributors to the outcomes in the Vancouver environment. Overall this is a strong study and the manuscript is well written.

We thank the reviewer for this positive feedback and for highlighting several of the unique aspects of this analysis.

(2) None of these diseases are treated in hospitals until late in disease and some are not even very treatable (AD and NAD) and treatment information may not be a great tool to identify cases early; but according to the validation studies conducted in Canada the procedures used in this study to identify cases electronically from the medical records have good sensitivity and excellent specificity.

Despite high sensitivity (78-84%) and specificity (99-100%) in the validation studies, we acknowledge that we are still unable to identify all incident cases. Therefore, we added this as a limitation in the Discussion section (Line 345-347):

“Fourth, the issue of under-representation exists. While cases were ascertained based on criteria with very high specificity and relatively high (78–84%) sensitivity compared to physician diagnosis, not all incident cases were captured in our study.”

(3) However, would this also include age at onset? Especially for such insidious diseases as AD and PD the prodromal stage is very long, and they may come to the attention of the medical system (and especially hospitals) relatively late depending on patient characteristics and attitudes. For example, the median age of onset of PD of 72 years is older than the median age otherwise reported in population-based studies of about 68 years. Hence, the environmental measures may not reflect exposures prior to disease onset unless individuals have a very low frequency of moving. Also, during disease progression these disorders are disabling, and the patients may choose to move to accommodate their diseases - thus the exposure measures could include exposures after changes in residence to accommodate worsening of the disorder rather than exposures that are contributing to the disease onset. I would like the authors to consider and discuss this issue and possibly provide some additional information on how much the exposure measures do indeed reflect long term exposures prior to disease onset.

We restricted our study population to those older than 45 years and the median age (68 years) was close to the median age of PD onset (72 years). As a result, we did capture a substantial proportion of cases. However, we agree that we only characterized exposures for a short period of time (1994-1999). We have added this as a limitation in Discussion (Line 347-350):

“Fifth, the exposure period was relatively short (1994-1998), in contrast to the longer period during which PD or AD may develop. Given that exposures before 1994 were not available, this
limitation may lead to non-differential exposure misclassification and bias towards the null.” (Line 347-350)

We further agree with the reviewer that it is possible and indeed likely that people moved during disease progression. We therefore added it as a limitation in Discussion (Line 350-355). While only a small proportion of people (15%) moved during the relatively short exposure period (1994-1999) [1], and while we did account for moving during the exposure period, we have no information on moving (or exposure) prior to the period before 1994. There might be, for example, another 15% of people who moved before 1994, but we do not have information to verify this possibility. We have added this possibility to the Discussion:

“While we did account for changes in address during the exposure period, during the period of disease progression, people may choose to move in order to accommodate the disabling conditions of their diseases. Therefore, the exposures that we assessed may only include those exposures occurring after changes in residence to accommodate worsening of disease, but not include exposures that contribute to the onset of the disease.” (Line 350-355)

(4) It may be less confusing to avoid referring to cases and controls in the results section when the analyses are for the cohort i.e. when referring to NAD and PD results. On the other hand please avoid saying "in the cohorts of AD and MS" since these were analyzed with a nested case control approach

We revised the text in which cases/non-cases were used for NAD/PD results. Cases/controls were used for the AD/MS results (nested case-control analysis).

We used “in the analysis of AD and MS” to replace “in the cohorts of AD and MS”.

(5) To point out in the results that cases and non-cases differed (for NAD and PD) in terms of age and comorbidities while this is not the case for AD and MS is trivial as the controls were matched to cases by age for the later but not former; please consider rewording this.

We removed the sentence related to AD and MS in the results section: “In the cohorts of AD and MS, distributions of characteristics and exposures were similar between cases and controls (Table 3 and 4).” (Original manuscript: Line 241-243)

(6) The authors point out the co-morbidity differences between cases and non-cases for NAD and PD but not AD and MS, but from table 1 and 2 it is clear that this is solely due to age i.e. the cohorts are not age matched or adjusted while the nested case control samples are. Please either adjust for age in the cohort or do not point out such comorbidity differences as currently done.

We adjusted for age and comorbidity in all models in the cohort of NAD and PD (“Household income, education, ethnicity and comorbidity were included as covariates with additional adjustment for age and sex in the Cox proportional hazard models”, original manuscript: Line 204-206). We still, however, think that the descriptive statistics related to comorbidity differences between cases and non-cases are worth mentioning. We added minor modifications
to the text to indicate the age dependence or the comorbidity differences: “For NAD, the median age (76 years) of cases was older than that of non-cases (57 years) with a corresponding much higher percentage of comorbidity in cases (49% vs 26%)” and “Similarly, PD cases were much older (median age 72 vs 58 years old) and with a higher proportion of comorbidity (44% vs 27%) compared to non-cases.” (Line 234-235, Line 242-243).

(7) Do results vary by gender (given the preponderance of different outcomes in either males (PD) or females (AD and MS) or by major races (e.g. PD is reported to be less common among Asians; Vancouver has a large Asian population - is race information only available at the neighborhood levels as well?)

We conducted stratified analyses by sex, ethnicity and age for NAD and PD, while we conducted stratified analyses by ethnicity for AD and MS as age and sex were matched. Results are presented in Appendix Table 1, 2 and 3. We added text related to stratified analyses both in the methods and results section:

In the methods section: “Next, we evaluated relationships with noise and NDVI individuals and in joint models with both (road proximity and air pollution) relationships. Further, in the cohort of NAD and PD, sex, ethnicity and age were stratified in analyses. For AD and MS, ethnicity was stratified.” (Line 218-220)

In the results section: “Males who lived near a highway were at a higher risk of developing NAD and PD compared to females. Individuals living in neighborhoods where &gt; 10% of the population was Chinese had higher incidence of NAD when they lived near major road or a highway compared to those in neighborhoods with &lt; 10% Chinese residents. Both road proximity and air pollution had greater effects on incidence of NAD and PD among people aged under 65. Individuals living in areas with &gt; 10% visible minorities who lived near major road had much greater risk of developing AD and MS than individuals living in areas with &lt; 10% visible minorities (Table S3-S5 in Additional File 1).” (Line 267-274)

Race/ethnicity information was only available at neighborhood-level as stated: “Neighborhood-level covariates included household income, education (as indicators of socio-economic status) and ethnicity from the 2001 Statistics Canada Census.” (Original manuscript: Line 189-190).

We had erroneously left out the strata definitions for ethnicity and have now added them to the manuscript: “Ethnicities included Chinese, South Asian and Visible Minorities with strata defined as living in neighborhoods with &gt; 10% of the population of the given ethnic status.” (Line 198-200)

(8) Please remind the reader which covariates in table 1 are measured at the neighborhood level and not individual

We added this information for current Table 1 and 2.
(9) What is the influence of living in high rise buildings (with air filtration etc) - is there a way to explore this in sensitivity analyses excluding certain neighborhoods with a high proportion of high rises?

In our study, the postal code data alone do not allow us to identify which type of building (e.g. high-rise, low-rise, single house) participants lived in. Therefore, we are unable to conduct a sensitivity analysis on this issue but have indicated it as a limitation (Line 355-359):

“Last, as the postal code data alone do not allow us to identify types of residences (e.g. high-rise, single house), we were unable evaluate the potential influence of vertical gradients in distance or pollution. A study in Hong Kong reported that including vertical gradient information did not lead to meaningful differences or changes in estimates of effect for the association between air pollution with mortality [2].”

(10) Why not use a cut-off for noise rather than a linear model, such as 65 dB and ≥ 65 dB and nighttime noise as &lt;55 dB and ≥55 dB since the effect of noise on the outcomes may not necessarily be linear. Also, how common is the use of noise protecting windows in Vancouver near major roadways?

Previous literature has used noise as a continuous variable [3-5] so we elected to do the same for consistency with prior analyses. Further, categorizing noise values into only two levels is likely to lose information. Similar logic would also apply to all the exposures (besides road proximity) and including cut-offs for all would greatly expand the number of comparisons. Were we to have observed strong effects of noise, such sensitivity analyses may be warranted. Overall, we did not observe strong effects of noise.

We do not have any data with respect to the use of noise protecting windows.

(11) In Table 5, it seems that the interquartile range used to generate the estimates is taken from cases of each disease rather than from the non-cases or controls? This is different from the usual recommended use of the control population exposure distribution as the reference; what was the reason for using case exposures? Also, why not use one IQR for the whole population (or all non-cases) instead - this would allow comparing the results more validly across diseases; using one IQR will probably change little but still it would allow for more valid comparisons.

We thank reviewer for this feedback. In fact, we did use one IQR for the whole population and have clarified this in the Table caption (current Table 3). We have fixed the typos in current Table 3 (previous Table 5).

(12) In Table 5, please also list which confounding variable were included in the models, this information can also go into a footnote.

Change made (previous Table 5 was renamed to Table 3)
Please explain whether you controlled for comorbidities such as diabetes and CVD or stroke in all models including the NAD model. These diseases might be mediators on the pathway to NAD i.e. would also be caused by air pollution or noise exposures and contribute to NAD, hence it might not be appropriate to adjust for them. Do results change if these covariates are not included in the NAD models?

Diabetes, coronary heart disease and stroke were included as comorbidities and adjusted in all models including sensitivity analysis using the CANUE data. We revised the text to indicate which covariates were included in sensitivity analyses: “In sensitivity analysis, we assessed the relationships between air pollution and the outcomes using alternative national models provided by CANUE, adjusting by the same covariates listed above for each outcome” (Line 220-222). We also added information with respect to why these conditions were adjusted in all models: “In addition to age, sex and socioeconomic status, these comorbidities are accepted risk factors for neurodegenerative pathology [6-9].” (Line 188-189)

The increase in AD risk with greenness seems rather strange and needs an explanation. Furthermore, not only is greenness positively associated with AD and MS but also seems to confound the air pollution estimates for these diseases. Would it be possible to check whether the distribution of cases for these two outcomes with a relatively small N is different across the city compared with cases of NAD and PD?

We added information for possible reasons (residual confounding) of the ‘harmful’ effects of greenness on AD and MS in Discussion section (Line 287-291); also see response to Reviewer 1. We also checked the distributions of greenness by income, education and ethnicity for each outcome. We found no obvious differences between cases and non-cases or cases and controls. This further indicates that there might be other factors that influence greenness and were not accounted in our study (Appendix Table 4 and 5).

We do not have access to postal code data for mapping as data linkage requirements do not allow us to map any of the outcome data (postal code data were used only for linking to exposure data and then stripped from analytical files based on policy of Population Data BC).

I do not understand the conclusion that AD was linked to air pollution while MS was not when all effect estimates for AD and air pollutants where less than one or very close to one while for MS there is a large and even statistically significant effect for PM2.5 (when controlling for greenness [OR= 1.43 (1.09-1.97)]).

We thank reviewer for this comment as this was a mistake in our Conclusions section. We therefore revised the text to: “Although results were not entirely consistent, air pollution was linked with NAD and PD, but not AD or MS.” (Line 362-363).

Our conclusions remain the same as we wrote in the Results section: “Associations between air pollutants, AD and MS were generally null with wide confidence intervals” (Line 262-263). In abstract: “Air pollutants were associated with incidence of Parkinson’s disease and non-Alzheimer’s dementia … but not Alzheimer’s disease or multiple sclerosis.” (Line 25-28)
Further, we also revised the text in the Discussion section: “While noise did not affect any of the associations with air pollution, there were indications that greenness was protective for the development of NAD, PD and AD. Overall, we saw indications of associations between air pollution with NAD and PD, but in general not with AD or MS. We did observe increased OR for MS in association with PM2.5, but not for any of the other air pollutants.” (Line 279-283).

(16) Would the authors have any explanation for the fact road proximity is a better - i.e. more consistent - predictor of the outcomes than most of the air pollutants?

Predictors (Major road, highway etc.) of road proximity are binary categorical variables, while air pollutants (e.g. PM2.5, black carbon etc.) are continuous variables. Additionally, measured or unmeasured factors associated with road proximity and air pollution are also different which may contribute to different results patterns.

To avoid confusion, we removed the sentence: “Overall, results related to air pollution were less consistent than those for road proximity”. See response to comment 2 for Reviewer 1.

(17) Page 30 table 3 has a typo in the last column and row, visible minority is 21.74 not 11.74; there is another mistake for NAD and highways<150m in table 2 page 28; it should read under cases 12.9 and not 22.9%. It seems that the lengthy tables need some additional scrutiny for typos.

Done.

References


