Reviewer's report

Title: Objectively measured urban green space accessibility, use and cardiovascular health: findings from a cohort study

Version: 1 Date: 25 October 2013

Reviewer: Paul Villeneuve

Reviewer's report:

MAJOR COMPULSORY Revisions

Objectively measured urban green space accessibility, use and cardiovascular health: findings from a cohort study

This study addresses an important research area, namely, the impact that access to green space has on health. An emphasis is placed on cardiovascular outcomes. Unlike most other studies in this area that have used a cross-sectional study design, this study had the advantage of including a longitudinal component. This offers a distinct advantage. This paper addresses an important area, but really needs major revisions.

Some comments

1. Abstract: the authors indicate that this is a cohort study, but provide no information about the length of follow-up in the abstract. This information is needed. The methods section does not provide any information about the types of analyses that were performed, they need to specify the types of regression models used (logistics, Cox model, etc), and discuss what variables were available as confounders and effect modifiers.

2. Abstract: it is confusing because in the title the authors give the impression that the study uses “Objectively measured green space accessibility” yet in the results section of the abstract they refer to that the prevalence of several risk factors was LOWER among park users. The only way you could determine whether someone was a park user or not (short of following them around) is to ask them. Asking people whether they are park users or not is NOT AN OBJECTIVE MEASURE OF GREEN SPACE. This is a self-reported measure which is subject to important biases. For me, objective measures of green space include: distance from home to parks, or use of remote sensing satellite images to assign green space.

3. Abstract. In my view, the longitudinal analyses are far more informative than cross-sectional analyses. I think the authors should explicitly state the number of incidence CVDs that were identified during follow-up in the abstract.

4. Some of the statements are too vague in the Abstract. For example, “Men living further away from parks had a higher risk of non-fatal and fatal CVD compared to living nearby”. What distances define “nearby” and “further away”. The reader is unable to interpret this hazard ratio without this being defined.
5. The methods section should describe how individuals were censored during follow-up. Presumably those who developed CVD were censored at the time they were diagnosed, while those who died of CVD were censored at the time of death. In section 2.3.4 the authors need to reassure the reader that most cases of CVD were identified during follow-up. Were any subjects lost to follow-up. Among the 83 CVD deaths how many of these were identified before death? Were any of these deaths included in the 364 non-fatal cases? Were there differences in the distribution of types of CVD between those identified with mortality data, and those identified as incident (non-fatal) cases? For example, were there more myocardial infarctions among CVD fatal cases, than non-fatal?

6. Two types of analyses were performed those based on prevalence, as well as those based on incidence. The authors never take the time to explain why both analyses were completed. They should provide a rationale. I can understand the need to look at relationships between green space and obesity and physical activity (or other risk factors for CVD), but what value is there in looking at the relationship with CVD prevalence when incidence data are available?

7. The response rates in the survey were fairly modest. The authors need to better discuss the impact of non-response. Simply stating the fact that “we believe that non-response has not seriously biased the results” is inadequate. Non-response bias would likely be larger for the cross-sectional analyses than for the longitudinal analyses. Can the authors say anything about how the characteristics of the responders differed from the non-responders? Was the level of participation related to either access to green space, or risk factors for CVD (socio-demographic status, smoking, obesity, etc, etc, etc)

8. The tables should be more explicit in describing whether prevalent or incident health outcome measures were used.

9. Some of the titles are misleading. For example, stating the “Distribution of the Lithuanian urban population aged 45-72 according to distance to green space” assumes that participants who participated in the survey are representative of the whole Lithuanian population. How can the authors make this claim? They had a 61% response rate. The study was done in Kaunas.

10. Table 3. The authors indicate that age-standardization was applied. Table should be able to be interpreted as stand-alone pieces. They should footnote how the age-standardization was done. Further, the reader has no clue as to what the tertiles represent. The authors should tell the reader what distances these tertiles correspond to.

11. Table 4: The authors need to indicate that these data are self-reported data. I would strongly recommend that they calculate adjusted odds ratios and their 95% confidence intervals for these factors, and adjust these comparisons for risk factors such as age, sex, etc. The unadjusted comparisons provide little value. For AH they indicate it is measured in mmHG, then somewhat inexplicably describe it as a YES/NO variable. Was a cutpoint applied?

12. Table 5: Again, the authors need to describe the distances that the tertiles correspond to. Presenting a confidence interval AND a p-value is redundant. I know whether the finding is significant by having the CI.
13. It appears that some exposures were only available for incident CVD and others only available for fatal CVD. I would recommend putting these different outcomes in separate tables.

14. Supplementary table 3. The title should indicate whether these represent INCIDENT or PREVALENT outcomes. Again, the tertiles do not describe the differences. Regression analyses where risk estimates are calculated and adjusted for relevant confounders is the preferred way of presenting these associations. Knowing that the prevalence varies across tertiles groups has no value to me without taking the time to adjust for relevant risk factors like smoking, age, BMI, etc. The number of events (e.g., stroke, diabetes, etc) should be made clear. Again, the authors don’t need to present the chi-square AND p-value. They are redundant against each other.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.