Reviewer's report

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

Version: 1 Date: 19 November 2013

Reviewer: Sjerp Dr de Vries

Reviewer's report:

Major Compulsory Revisions

1. Two multivariate Cox proportional hazards regression models are presented. The reader has to look very carefully to find the differences in the set of predictors. More importantly, why precisely and only these two models? Why is use of city parks only included in the model for non-fatal CVD, and not in the one for hard CVD? Why not run a model for non-fatal CVD with use excluded, since it is likely to be an important mediator of accessibility of green space? The list of covariates was selected a priori. But the reasons for including these covariates sometimes is quite unclear. E.g., why include self-rated health as a covariate? In other studies this covariate itself has been shown to be related to accessibility of green space (and therefore also might be a mediator).

2. Results are reported in an unclear way. For example, it is said that the proportion of aged 65 years or older was significantly higher in the 3rd tertile. Higher than which other group? It does not appear to be consistent with the percentages shown in Table 2. Moreover, in Table 2 only specific contrasts are tested (1st tertile against other tertiles). Why only these contrasts? Also in this table the n differs for gender and age (n = 5000) and education (n = 5224). It seems strange that level of education is known, whereas gender and age are not. Please explain. Smoking status is related to distance to green space, but the authors fail to mention how. Park use is used in the analysis, but has not been properly introduced beforehand in the Methods section (only mentioned in the Statistical analysis sub-section). How is it defined and measured?

3. The Discussion section as a whole is not very informative/helpful. Some examples:

Discussion on accessibility to green space and health is very minimal. The final sentence states that the discrepancy between this study and some others, that did find a relationship between accessibility and e.g. morbidity might be due to different study populations, study designs, population sizes and the contribution of other risk factors. (Or different accessibility measures, I might add.) This is not very helpful.

In the Discussion section the authors state that the frequency of green space use significantly declined with increasing of distance from green space for both men and women. I did not see data on frequencies in the Results section (only users
versus non-users). The time spent in city parks per week is available, but is not used in the Results section.

The Discussion sub-section on Green space, morbidity and mortality is not a discussion, but a review of some other studies. For example, it is not discussed why the findings were different for men and women (Only for men third tertile higher HR than first tertile for hard CVD (fatal and non-fatal). Only for women second and third tertile and being non-user higher HR than first tertile and user for non-fatal CVD). It is also not discussed why there were no relationships at baseline, whereas there were relationships in the cohort analyses. It also remains unclear why, once park use is known, accessibility should still matter (perhaps higher frequency of park use, or more time spent in parks?).

Why do the authors believe that the low response rate has not seriously biased results? (Actually I think it is quite high). And how precisely does the study inform public health policies aimed at promoting healthy lifestyles in urban setting? Are there specific recommendations that already can be made, given the exploratory nature of the study?

Minor Essential Revisions

1. The definition of green space is said to include city parks larger than 1 ha. Does it also include other types of green space within or outside the city (groves, gardens, nature reserves and agricultural areas are mentioned in the sentence before the definition). And/or what is considered a city park?

2. Participants provided their residential address during the self-report and the official address was also provided by the National population register office. Why is it important to have both and which one was used if they were not the same?

3. The authors state that because of the scoring of each participant cognitive tests varies, test scores were standardized. I expect that this should be: Because the scoring of each cognitive test varies.

4. “Smoking habits were assessed according to current smoking status. The respondents were classified to three groups: smokers, former smokers and never smokers.” Current does not apply.

5. The authors also evaluated the European guideline indicator of green space larger than 0.5 hectare within 300 m of the residence. What qualified as green space in this second environmental characteristic?

6. Section 2.3.5 on Green space exposure assessment. This section is unclear. On the one hand it is said that based on the home address the distance to the nearest park is estimated (and subsequently classified in tertiles). However, at the end of the section they say that they used a binary variable to estimate whether the responder’s address fell in to an estimated buffer surrounding the boundaries of a city park to address proximity to the nearest city park. Why ‘estimated buffer’? Why buffers at all, if the distance to the nearest park is
known?

7. How is “use of city parks” defined? Is it time spent in the city parks per week? Have the respondents used the same definition of city parks as the authors? Note that later on it becomes clear that is a dichotomous variable (users versus non-users)

8. Why are only the contrasts of the first tertile with the other two tertiles tested, and not that between the second and the third tertile? Contrasts appear to be post hoc contrasts. Does this test that has been used reflect this post hoc nature?

Discretionary Revisions

These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential.

1. “Respondents were categorized into two groups according to their physical activity in leisure time: active (10 and more hours/week), and inactive (< 10 hours/week).” Why 10 hours as cut-off point? This seems very high, given standard guidelines for adults (being at least moderately intensive active for at least 30 minutes on at least 5 days a week)

2. ICD: explain acronym the first time that it is used.

3. All data were age-adjusted to the total Kaunas population. Explain why this was done.

**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:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests