Author's response to reviews

Title: Accessibility and use of urban green spaces, and cardiovascular health: findings from a HAPIEE cohort study

Authors:

Abdonas Tamosiunas (abdonas.tamosiunas@lsmuni.lt)
Regina Grazuleviciene (r.grazuleviciene@gmf.vdu.lt)
Dalia Luksiene (dalia.luksiene@lsmuni.lt)
Audrius Dedele (a.dedele@gmf.vdu.lt)
Regina Reklaitiene (regina.reklaitiene@med.kmu.lt)
Migle Baceviciene (migle75@gmail.com)
Jone Vencloviene (j.vencloviene@gmf.vdu.lt)
Gailute Bernotiene (gailuteb@yahoo.com)
Ricardas Radisauskas (ricardas.radisauskas@lsmuni.lt)
Vilija Malinauskiene (vilimali@itc.kmu.lt)
Egle Milinaviciene (egle.milinaviciene@kaunoklinikos.lt)
Martin Bobak (m.bobak@ucl.ac.uk)
Anne Peasey (a.peasey@ucl.ac.uk)
Mark J Nieuwenhuijsen (mnieuwenhuijsen@creal.cat)

Version: 2
Date: 10 January 2014

Author's response to reviews: see over
Dear Reviewers,

We would like to thank you for your valuable comments regarding our manuscript “Urban green space accessibility, use and cardiovascular health: findings from a HAPIEE cohort study ”, submitted to BMC Environmental Health (number of the manuscript MS:1495204494107488). We hope that the quality of our manuscript was improved as suggested. All modifications in the text of the manuscript and tables are in „yellow”.

Answers and comments to the reviewers:

Reviewer: Paul Villeneuve

Reviewer’s report:
MAJOR COMPULSORY 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.

Following your proposal, we have modified the abstract. We have added the information about the length of follow-up in the study and information about the types of analyses in the abstract's methods section.

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.

Following your proposal, we have modified the abstract. We have indicated that use of green space or city parks was self reported measure. We have also included in the abstract's results section the distances from home to parks. We have also modified the title of the manuscript.

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.

We have added the sentence in the abstract of the manuscript: “During the follow-up (mean duration 4.41 years) there were 83 deaths from CVD and 364 non-fatal cases of CVD among persons free from CHD and stroke at the baseline survey.”

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.
We have included in the abstract’s results section the distances from home to parks.

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.

Following your proposal, we have added the sentence in the method’s section of the manuscript: “The proportion of participants lost to follow-up was 0.86% (N=44). People who were lost to follow-up were censored at their last date of contact.“

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?

Any case of non-fatal CVD was identified before death.

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?

Our study is the first large epidemiological study in Central and Eastern Europe investigating the relation between the exposure to green space and CVD. Therefore we suggested that it was important to show the prevalence of CVD and their risk factors in relation to the distance to the green space.

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)

We added the sentence in the strength and limitations section of the manuscript: “Response rate at the baseline survey was not very high. This is a common problem in most epidemiological studies and our study non-responders were more likely to be male, younger, lower educated and less healthy than responders [42].”

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

Following your proposal, we have modified the Table 4 and Supplementary Table 3.

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.

We corrected titles of the manuscript and tables.

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.

Following your proposal, we have added the footnote to Table 3: “All data were age-adjusted to the Kaunas population census of 2006 (details in section Methods 2.4 „Statistical analysis“). Distance to green space: 1st tertile - < 347.8 m (high); 2nd tertile – 347.81 – 629.6 m (moderate); and 3rd tertile - > 629.61 m (low).”

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?

We have calculated adjusted OR and 95% CI and added footnote for AH cut points.

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.

We have corrected Table 5 according to your suggestions: deleted p values and added footnote: “Non-fatal CVD - all incident cases of non-fatal acute myocardial infarction, unstable angina pectoris and stroke. Hard CVD - all fatal and non-fatal CVD cases. Distance to green space: 1st tertile - < 347.8 m (high); 2nd tertile – 347.81 – 629.6 m (moderate); and 3rd tertile - > 629.61 m (low).”

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.

Thank you for this recommendation but saving the space of the manuscript we decided the results for hard CVD and non-fatal CVD to show in one table.

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.

We calculated adjusted OR and 95% CI, deleted χ² value in Supplementary Table 3, and added footnote: “CHD – coronary heart disease, CI – confidence interval. Distance to green space: 1st tertile - < 347.8 m (high); 2nd tertile – 347.81 – 629.6 m (moderate); and 3rd tertile - > 629.61 m (low). *adjusted by: age, education, smoking, arterial hypertension, physical
activity, total cholesterol level, fasting glucose level, body mass index, and sex (last one - only in men and women group)."

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

The English language was improved.

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).

We included into multivariate Cox proportional hazards regression models only covariates which have statistically significant HR and 95% CI in univariate models.

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?

Sorry, we made a mistake in describing results from the Table 2. Following your proposal, we have modified the results and number of cases in Table 2. Smoking status is related only with green space (park) use (Table 4) but not with distance to green space. Use of green space is introduced in the methods section „2.3.2. Variables obtained using the questionnaire“. Use of green space was classified as „yes=1“ and „no = 2“.

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?

Thank you for your valuable remarks and recommendations to the Discussion section. Following your proposal, we modified this section of the manuscript.

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?

We modified the description of green space (city parks) in methods section (2.1. Study area).

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?

We have residential addresses provided by study participants (self-reported) and official addresses provided by the National population registry office. Some addresses were different. Finally we used addresses provided by study participants. Manuscript section “2.2. Study cohort” was modified to: “Participants provided their residential address during the self-report.”

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.

Thank you for your remark: we corrected accordingly in the manuscript text.

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.

Thank you for your remark: we corrected accordingly in the manuscript text.

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
Our definition of “green space” included city parks larger than 1 hectare. We excluded the sentence “We also evaluated the European guideline indicator of green space larger than 0.5 hectare within 300 m of the residence” from the methods section because we did not show such results in this manuscript.

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?

Sorry, but we did not use binary variables to estimate distance from participant home to green space.

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)

Yes, responders used the same definition of city parks.

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?

Thank you for your remark. We additionally tested contrast between the second and the third tertile.

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)

Leisure time physical activity (LTPA) was assessed as all kind physically demanding activities, such as housework, gardening, and maintenance of the house, also walking, running, engagement in sports, games or hiking in a typical week on summer and winter seasons. Before the analysis, the averages of actively spent hours in summer and winter were calculated and all the responders according to LTPA were classified into two groups where 10 hours/week was set as the cut point (the limit of the lowest quartile of LTPA).

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

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

Our study sample was randomly selected from total Kaunas population therefore we used total Kaunas population age structure for age-adjustment.

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

The English language was improved.