Author's response to reviews

Title: The association between neighborhood greenness and cardiovascular disease: an observational study

Authors:

Gavin Pereira (gavin.pereira@uwa.edu.au)
Sarah Foster (sarah.foster@uwa.edu.au)
Karen Martin (karen.martin@uwa.edu.au)
Hayley Christian (hayley.christian@uwa.edu.au)
Bryan Boruff (bryan.boruff@uwa.edu.au)
Matthew Knuiman (matthew.knuiman@uwa.edu.au)
Billie Giles-Corti (b.giles-corti@unimelb.edu.au)

Version: 3 Date: 30 April 2012

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

Thank you for facilitating the review of our manuscript entitled *The association between neighborhood greenness and cardiovascular disease: an observational study* for publication as an original research article in *BMC Public Health*. We appreciate the effort that the reviewers have taken to comment on our manuscript. We have revised the manuscript according to every one of these comments, which we expect will satisfy both reviewers. However, we note that we could not fully address Major Compulsory Revision 6 for the first reviewer as we no longer have participant addresses due to an ethics agreement with the state Department of Health. Our response (red text) to the reviewers comments (black italicized text) is attached below.

Please do not hesitate to contact me if you require any further information.

Faithfully,

Gavin Pereira

---

**Contact details for correspondence:**
Dr Gavin Pereira
Telethon Institute for Child Health Research
100 Roberts Road, Subiaco, Western Australia, 6008 | PO Box 855, West Perth, Western Australia, 6872
T +61 8 9489 7945 (Direct line) | 0413 630 308 (Mobile) | gpereira@ichr.uwa.edu.au (Email)

**Contact details for billing:**
Ms Jan Van der Heyden
Centre for the Built Environment and Health | School of Population Health | Mailbag M707 | University of Western Australia | 35 Stirling Highway Crawley WESTERN AUSTRALIA 6009
T +61 8 6488 7581 (Direct line) | +61 8 6488 1199 (Fax) | jan.vanderheyden@uwa.edu.au (Email)
Response to Editor

1. Please include the email addresses of all authors at the title page. The author list and email addresses must be identical in the manuscript file and on the submission system, and it must be clear which affiliation pertains to each author.

Response: Completed?

2. Remove the Acknowledgement in the title page and place it right after the Authors' Contributions section.

Response: Completed?

3. The figure legend and title should be part of the manuscript file, given after the reference list.

Response: The figure title/description is now placed immediately after the reference list instead of after the tables, as requested. However, the legend contains a color scale, which means that it cannot be represented as text and therefore cannot be moved from the figure to the manuscript.

Query: At the request of additional information from the Reviewers we have added a Supplementary Material file to this manuscript. Please provide this to the Reviewers. Please let us know the preferred citing for this material. Currently it is cited in the revised manuscript as “Supplementary Material (Additional File 1)” initially and “Supplementary Material” thereafter.

Response to Reviewer 1900853798702381 (Dr Bell)

Major Compulsory Revisions

1. There is no consensus in the literature about the size of spatial unit in which greenness ought to be measured. What was the rationale for choosing a 1600 m service area? In what ways might the size of the spatial unit have influenced the study findings or their interpretation?

Response: We agree with the reviewer that the rationale for selecting a 1600m service area was not clear in the original manuscript. We have amended the manuscript to include the following description [lines 97-107]: “Neighborhoods around participants’ homes were defined using 1600m (network distance) service areas. The rationale for selecting a 1600m service area was based on the assumption that physical activity is the most likely pathway by which neighborhood greenness is associated with cardiovascular disease risk. A 1600m service area represents how far a participant could walk from home at a moderate to vigorous intensity pace, within 15 minutes, which equates to half the recommended level of daily physical activity for adults[1]. That is, the daily recommended level of physical activity of 30 minutes would be attained for a return trip. Moreover, a 1600m service area has been shown to be a critical distance for examining the relationship between parks and walking. Sugiyama et al found that although proximity to parks was generally associated with walking, it was the
presence of parks within 1600m that was most associated with sufficient walking (>150 minutes/week)[2]. “

2. **Provide summary information (similar to that given in Table 1) for the NDVI variable across the study setting, including the mean and the range. Can you also interpret these details to assist the reader to understand the overall greenness of the study setting?**

**Response:** We agree with the reviewer and have amended the manuscript to include the median and range for levels and variability of greenness (last rows of Table 1). It is difficult to compare the greenness measures with other study settings because they do not give enough information on greenness at a comparable spatial scale. The most comprehensive description of greenness is the density/distribution of these measures. Therefore, we have also included two additional figures (Figure 4 and Figure 5) in the *Supplementary Material* to graphically illustrate the density (i.e. distribution) of the mean NDVI and the SD NDVI. We have provided additional description in the revised manuscript of service areas with the highest and lowest values of mean NDVI and SD NDVI [lines 157-168]: “The median and range for levels and variability of greenness for 1600m service areas is shown in Table 1. The highest level of greenness (mean NDVI=0.337) was observed for a service area that mostly contained bush land (typically eucalypt vegetation) and was located on the urban-rural fringe. The service area with the lowest level of greenness (mean NDVI=-0.059) was located in a region with a large proportion of commercial and industrial land use. The service area with the highest variability in greenness (SD NDVI=0.205) was located in a densely populated area with large parks and bush land, a dense network of tree-lined streets and commercial and retail destinations. The lowest variability in greenness (SD NDVI=0.048) was observed for a service area with little vegetation and large houses that aligned waterways. Complete distributions of the levels and variability in greenness are provided in the *Supplementary Material.*”

3. **Justify why the NDVI variable treated categorically rather than continuously. What was the rationale for the cut-points that were chosen?**

**Response:** The exploratory process by which we arrived at the decision to categorize NDVI was not included in the original version of the manuscript. The process was as follows:

1. We separately plotted the event risk against SD NDVI and against mean NDVI. As the event was a binary outcome, we applied a loess smoother (i.e. a nonparametric smoothing function that performs localized regression for a given value of the smoothing parameter) and then transformed these values to the log(odds) scale.

2. We looked for non-linearity in these plots for the region between the 10\(^{th}\) percentile (P10) and the 90\(^{th}\) percentile (P90) i.e. for the bulk of the data.

3. We repeated the above steps for a range of values of the smoothing parameter.

4. Visual assessment of both the figure and the density enabled us to select cut-points to ensure sufficient data in each category.

In response to your comment we have created *Supplementary Material* that describes the results of this process. We have amended the main text of the manuscript to refer to this material [lines 121-126]: “The mean and standard deviation of NDVI values were analyzed as both continuous linear variables (scaled by their interquartile range) and categorical variables (tertiles). The rationale for using tertiles is
provided in the Supplementary Material (Additional File 1). Briefly, cut-points at tertiles provided a compromise between capturing the pattern of the association and ensuring sufficient data within each category. Tertiles also provide an equal amount of data in each category and allow interpretation relative to “low”, “medium” and “high” values.

We have constructed two additional figures, placed in the Supplementary Material (Figures 4 and 5). The figures illustrate both the log(odds) of the loess smoothed risk as well as the density of the greenness variable. Both figures indicate an irregular pattern between P10 and P90. For both figures, we inferred that the pattern may be approximated parsimoniously by a piecewise constant model. Cut-points at tertiles (P33 and P67) seemed to provide a compromise between capturing the pattern of the association and ensuring sufficient data within each category. We did not use the data outside the P10-P90 range to inform our choice of cut-points because of the potential for edge-effects i.e. patterns that are an artifact of insufficient data at the lowest and highest extents of the greenness variable.

Tertiles also provide an equal amount of data in each category and allow interpretation relative to “low”, “medium” and “high” values. The effects of extreme (small and large) greenness values (i.e. outliers) can also be reduced by such categorization. Also, odds ratio estimates cannot be as readily extracted had we opted for higher order polynomial terms instead.

However we appreciate that examination of greenness measures each as a continuous linear term is the most parsimonious model, it is useful for comparison with future studies, and despite our best efforts the loess results might still be sensitive to the choice of the smoothing parameter. Therefore we have amended Table 2 to include effects for linear increases in the mean NDVI and SD NDVI. We scaled a unit increase by the interquartile range to make comparable the effects for mean and variability of greenness. Also, a 1 unit increase is uninterpretable as it is greater than the range of these values.

4. Discuss the rationale for the choice of control variables. A rationale was given for air pollutants (lines 112 – 118) but not for the other variables. Total minutes of walking per week may be over-controlling given that physical activity was given as the hypothesized mechanism for a greenness-CVD relationship. Could physical activity be examined as an outcome?

Response: The control variables are well established risk/protective factors for cardiovascular morbidity - see manuscript reference 4. We have included this reference in the revised manuscript [lines 127-133]: “Adjustments were made using representations of a range of well-established factors [4] obtained from the Health and Wellbeing Survey: sociodemographics (age, sex, possession of a healthcare card, education, household income), biological factors (non-gestational diabetes, BMI, hypertension, high cholesterol), behavioral factors (daily serves of fruit and vegetables, risky drinking in the last month (>6 standard drinks for men, >4 standard drinks for women), and smoking (never versus ever smoked)), and a proxy for air quality.” We could not control for other correlates of CVD (e.g. familial) as this information was not collected by the state Department of Health’s Health and Wellbeing Survey. We have added this as a study limitation [lines 240-242]: “Adjustment for a wide range of risk factors, such as nutrition, was an advantage of this study. However, adjustment was not made for known risk factors for cardiovascular disease such as heritability, which were not available from the health survey.”

We provided a full description for the rationale for inclusion of air pollutants because although readers will be aware of, and accept, the aforementioned risk factors (e.g. diet, BMI, hypertension), it is unlikely
that they will be aware of the effects of air pollutants or, for example, the Scientific Statement of the American Heart Association – see manuscript reference 15.

We debated the inclusion of physical activity as a candidate adjustment factor for the original version of the manuscript and decided to include it as a way of adjusting for physical activity conducted outside the neighborhood (i.e. 1600m service area). Unfortunately, the only information on physical activity was minutes of total walking, irrespective of location. We re-ran the analyses without adjusting for this variable and observed little difference in the parameter estimates and their precision. We agree with you that this variable potentially lies on the causal pathway and is the main hypothesized mechanism for the association. We have amended the manuscript to exclude physical activity as an adjustment variable from models in Table 1. Importantly, not having information on physical activity specific to the neighborhood (i.e., 1600m service area) limits its use for investigation as an outcome.

5. To the paragraph in lines 112 – 118 outlining the rationale for including air pollutants, add the description of the variable(s) and data source(s) used in the current study to measure “environmental factors”.

Response: Thank you for raising our attention to this omission. We have included this text after the information on air pollution [lines 141-146]: “Adjustment was made for the total length of main roads within a 400m service area as a proxy for exposure to ambient air pollution. Main roads were defined as all roads traversed by ≥6000 vehicles/day. The metropolitan road network was obtained from Landgate (the state government source of land information and geographic data) at four time points (2005, 2006, 2008, 2009). The year of road network data was matched to participant’s year of interview for the Health and Wellbeing Survey.”

6. What was done to measure and account for possible spatial autocorrelation in the statistical models? What were the results?

Response: We agree that it is ideal to account for residual spatial autocorrelation in the outcome not indirectly accounted for by adjustment for socioeconomic factors (e.g. education and household income), main roads, and the greenness variables. Unfortunately, it is not possible for us to account for such spatial autocorrelation due to the ethics agreement surrounding our provision of this data. Specifically, we first obtained geographic coordinates of participants’ home locations (without any survey or hospital data) from the state Department of Health. We then calculated the greenness measures and the length of main roads for each participant. This information was sent back to the Department of Health, who replaced the geographic coordinates of the home addresses with a de-identified person number. Thus, we do not have addresses (or geocodes of participant’s addresses) to identify where e participants lives to enable calculating a measure of spatial autocorrelation. Although we have described this as a limitation of the study in the revised version of the manuscript, it is important to consider some of the drivers for such spatial autocorrelation in the cardiovascular outcome:

- Residual spatial autocorrelation due to socioeconomic status: This is likely to have been minimized by adjustment for the participants’ socioeconomic status (education, household income, and possession of a government healthcare card) which tends to cluster at larger spatial scales (e.g. census collection district, suburb and postcode).
• Spatial autocorrelation due to location of health services: This is likely to have been partially explained by socioeconomic status. Although cases with the disease might be more likely to live closer to such health services, those with the disease are more likely to be of lower relative socioeconomic position. Moreover, many of the hospitals in Perth are surrounded by the greenest areas. This means that these areas would also have high levels of variability in greenness (because the hospital itself would be picked up as having much lower NDVI values). Therefore, the result might have been an attenuation of the true effects (should an effect exist).

• Spatial autocorrelation due to an unknown factor: We believe this is the most likely reason for spatial autocorrelation for our study. However, as stated above, we are unable to calculate a measure of spatial autocorrelation as in accordance with our ethics agreement for our state Health Department we no longer have information on participant’s addresses.

The following text was included in the revised manuscript [lines 253-259]: “Finally, spatial autocorrelation in cardiovascular outcomes was not directly modeled in the analyses. However, such autocorrelation would be limited to that not already accounted by adjustment for socioeconomic factors (e.g. education and household income) which also cluster spatially and are strong predictors for cardiovascular disease. Perhaps more importantly, it is impossible to completely rule out the chance of spatial autocorrelation in the cardiovascular outcomes due to an unknown factor.”

7. Lines 138 – 143. Please clarify here, and in Table 3, whether the association with mean greenness controlled for variability in greenness and vice versa. The footnote in the table referring to mean greenness under “environmental factors” is confusing.

Response: We have combined the results of Tables 2 and 3 into a single table in the revised manuscript (Table 2). We have now included a column for the unadjusted models (Model A) and excluded physical activity from all models (see response to comment 4). The last column now states adjustment was made “All greenness”. The footnote now explicitly states that “the model included terms for both mean and SD of NDVI”. Therefore, the ambiguity has now been removed and it is clear that the fully adjusted model includes sociodemographics, biological factors, behavioral factors, air quality, mean NDVI and SD NDVI. We have also used “full adjustment” in the revised manuscript’s Results section. The meaning was defined at the first use of this term [lines 169-172]: ‘After “full adjustment”, which included terms for both mean and variability in greenness, the odds ratio for...’

8. Present additional results for the association of a) mean greenness (without controlling for variability) with the study outcomes; and for variability (without controlling for mean) with the study outcomes. (If these are among the models already presented in Tables 2 and 3, they should be clearly labeled accordingly.)

Response: This is now presented as Model E (fully adjusted except not including the other greenness variable) and Model F (fully adjusted model including terms for both mean NDVI and SD NDVI) in Table 2 of the revised version of the manuscript.

9. In the discussion, can you speculate further and provide any explanation(s) for the result presented in line 141 – 143?

Response: We were conscious of explicitly not interpreting any of the results using the criterion of p<0.05, or equivalently a dichotomous interpretation as to whether the 95% CI for the odds ratio
intersects 1, as the results become prone to misinterpretation. Our findings have been interpreted in accordance with the Uniform Requirements for Manuscripts Submitted to Biomedical Journals.

In this situation, the sizes of the effect estimates were similar and their intervals overlapped considerably. Therefore, the effect estimates are compatible and might indicate that moderate and high levels of greenness variability could have been combined into one category for this particular comparison i.e. a threshold effect is reached.

We have now made explicit the overlap in the intervals in the revised version of the manuscript [line 173-178]: “Stronger associations were observed for variability in greenness and effect sizes were similar for adults with moderate and high levels of variability in greenness, with a considerable overlap in the confidence intervals. The odds ratio for coronary heart disease or stroke (self-report) was 0.76 (95% CI: 0.62, 0.94) for adults in areas of moderate variability in greenness and was 0.84 (95% CI: 0.68, 1.03) for those in areas with high variability in greenness.”

10. The possibility of confounding by SES is important and could be elaborated. Also, provide citations or other justification for the claim that SES would be expected to be more highly correlated with levels of greenness than with variability. (There are a lot of non-green elements that might lead to high variability in high SES neighborhoods e.g., swimming pools, tennis courts, big homes).

Response: The mentioned non-green elements (e.g. tennis courts, swimming pools and big homes) would be mostly captured by our socioeconomic-related adjustment variables e.g. household income and education. It is difficult to speculate as to what elements of the environment these variables do not capture. Perhaps it does not matter. Even adjusting for these two strong correlates of socioeconomic status had little change on the effect estimates i.e. there was at most a 10 percentage point change in the effect estimates from Model A to Model B. However, we admit that residual confounding by SES should be mentioned as a limitation. We also appreciate that data (or cited studies) is needed to back the claim that SES would be expected to be more highly correlated with levels than variability in greenness. We have removed this speculative statement from the revised version of the manuscript.

We have also elaborated on SES in the Discussion section [lines 243-246]: “The possibility of residual confounding by individual level socio-economic status cannot be dismissed. However, adjustment was also made for multiple correlates of socioeconomic position, including education, which has been shown to be strongly predictive. Adjustment was also made for biological and behavioral factors that also exhibit gradients in socioeconomic status.”

11. Include the sample sizes in Tables 2 – 4. Also, clarify in text which analysis was performed on the full sample (n = 11,406) versus the subsample who granted permission for data linkage (74%). (Table 1 suggests the full sample was used for all analysis.) Did the subsample differ from those in the full sample on any of the key study variables?

Response: The full sample mentioned in the main text and in the title of Table 2 consists of 11,404 adults, which is 2 fewer than originally reported. These two individuals did not have values for the outcome variables and have been subsequently excluded.
We have now clearly stated in the revised manuscript that the 11,404 are those individuals that gave permission for data linkage [lines 80-82]: “There were 15,502 adult residents in Perth who completed the survey between 2003 and 2009. Of those participants, 11,404 (74%) participants granted permission for data linkage.”

Of the 11,404 adults, many had missing values for at least one of the adjustment variables. We have provided the sample sizes in Table 2. The difference between the sample sizes for self-reported cardiovascular disease (11,374) and hospital admissions (11,198) in Model A can be explained by the presence of missing values for the particular outcome variable. The total sample size of 11,404 exceeds both of these values because the 11,198 is not completely a subset of the 11,374. In general, the effect estimates and their confidence intervals changed minimally with successive adjustment, indicating our results after full adjustment (Model F) are generalizable to the full study population.

However, we acknowledge that we did not show that the results are externally generalizable to the 26% of adults who were interviewed (i.e. we had their responses to the health survey) but did not grant permission for data linkage (i.e. we could not calculate their neighborhood greenness). The manuscript has been revised to include a brief description of the comparison between these groups [lines 191-198] “Assessment of external generalizability. Of the 15,502 adult residents in Perth who completed the Health and Wellbeing Survey between 2003 and 2009, there were 4,098 (26%) adults who did not grant permission to obtain their hospital records or calculate their neighborhood greenness variables. A comparison between the age, sex, education, household income and BMI distributions for the study population and the excluded (non-linkable) population is provided in the Supplementary Material. Although the study population differed from the non-linkable population in terms of these variables, the extent of this difference was marginal.” More detail has been provided in the Supplementary Material.

Minor Essential Revisions

1. It would be helpful if the variables described in lines 107 – 111 were directly mapped in text to the labels “socio-demographic factors, biological factors, behavioral factors and environmental factors” given in lines 119 – 120 before these labels were introduced.

Response: We have now introduced these labels at the suggested position in the revised manuscript [line: 127-133] “Statistical analysis. Adjustments were made using representations of a range of well-established factors [4] obtained from the Health and Wellbeing Survey: sociodemographics (age, sex, possession of a healthcare card, education, household income), biological factors (non-gestational diabetes, BMI, hypertension, high cholesterol), behavioral factors (daily serves of fruit and vegetables, risky drinking in the last month (>6 standard drinks for men, >4 standard drinks for women), and smoking (never versus ever smoked)), and a proxy for air quality.”

2. Can you also help the reader to understand how to interpret the findings for mean NDVI controlling for its variability? And for NDVI variability, controlling for mean NDVI? The figures are helpful for considering variability on its own—but it does not appear that variability was modeled as independent variable without mean greenness.

Response: We did model variability as an independent variable without mean greenness and have modified the revised version of the manuscript to make this clearer (see response to Comments 7 and 8 above). We agree with the reviewer that readers would benefit from further explanation on the
interpretation of effects for mean NDVI (or SD NDVI) controlling for SD NDVI (or mean NDVI). We thought that a text-based description of the statistical interpretation might confuse readers further. You mentioned that the figures helped interpretation for variability on its own. We created another figure in the revised version of the manuscript to illustrate four simulated service areas (A,B,C and D) each representing one of the four combinations of Low or High mean NDVI with Low or High SD NDVI (Figure 1).

In the figure, the pairs of areas with the same SD NDVI are A-B, C-D. The pairs of areas with the same mean NDVI are A-C, B-D. The results are now readily interpretable using this figure. For example, the model for SD NDVI controlling for mean NDVI is analogous to comparing A to C, and B to D in relation to the cardiovascular outcome.

3. Are you able to sort what the mix or variability of greenness means in the data you have on hand—or at least speculate further for the reader about why “the mix of greenness is more relevant than the extent of greenness”? For instance, do the areas with higher variability include parks (greenness) that are accessible to cars (with parking lots)? Cycling or jogging paths along green settings? Bricks and mortar facilities for physical activity with outside landscaped recreation areas?

Response: We agree with the reviewer that readers need further information as to the specific attributes of highly variable neighborhoods that might promote better health. We have now more explicitly described some of these attributes, albeit speculative [lines 216-228]: “Neighborhood attributes that may contribute to a high variability in greenness might include prevalence of tree lined streets/cycleways/footpaths, presence of parks with parking, or green reserves with good road connectivity. The coexistence of both aesthetically pleasing natural vegetation to entice people out of their homes and destinations within walking distance would also contribute to variability in neighborhood greenness. A review of the environmental influences on walking concluded that aesthetic neighborhood attributes (which included ratings of natural features) were found to be associated with walking [7]. Furthermore, a cross-sectional study conducted in Seattle, US, reported that the most frequently walked non-green destinations were grocery stores, restaurants, libraries, coffee shops, and post offices[19]. Further studies are needed to identify the specific attributes of neighborhoods with a high degree of variability in greenness, such as the co-location of green and non-green areas (e.g. tree-lined paths, parks surrounded by well-connected streets).”

Unfortunately, the resolution of the satellite imagery does not allow us to ascribe greenness to features less than the pixel size (Landsat Thematic Mapper 30m x 30m resolution). This resolution is not suitable for a systematic analytic approach to identifying what the mix (or variability) of greenness means for the study area.

4. It is odd that the highest NDVI tertile is not significant across all models in Table 2 while the moderate category shows at least a weak association. Again, providing an explanation for the choice of NDVI cut-points would be helpful as well as presenting the findings with mean NDVI is modeled without NDVI variability and vice versa. The discussion should also include an explanation for the unexpected finding about why the highest category of NDVI might not be significant while the moderate category is significant.

Response: Inference cannot be dichotomized as “significant” vs “non-significant” (see response to comment 9 above). Using similar terminology the high category would be borderline significant. The
reason for the difference between significance for moderate and high categories is most likely due to the drop in the sample size due to addition of more adjustment variables available only for a subset of the sample population. Fortunately, the effect estimates are relatively robust (i.e. drop in sample size does not seem to bias results) and the confidence interval widths only change by a small amount (the drop in sample size does not seem to have too much influence on the variance of the effect estimates); see response to comment 11 above. The modifications made to the revised version of the manuscript in response to comments 9 and 11 come part way at addressing the reviewer’s concerns.

Minor Issues not for Publication

1. **Consider rewording the sentence in lines 39 – 40.**

   **Response:** We reworded the sentence in the Abstract section of the revised manuscript [lines 43-46]: “The odds of hospitalization for heart disease or stroke was 37% (95% CI: 8%, 57%) lower among adults in neighborhoods with highly variable greenness (highest tertile) compared to those in predominantly green, or predominantly non-green neighborhoods (lowest tertile).”

2. **Consider rewording (perhaps dividing) the sentence in lines 64 – 68.**

   **Response:** We have divided the long sentence into two shorter sentences in the revised version of the manuscript [lines 67-71]: “Various studies have examined the association between green neighborhoods and body weight, which is an intermediate outcome on the pathway to heart disease. However, few studies have explicitly assessed its association with cardiovascular outcomes, despite the potential for greener neighborhoods to promote physical activity.”

3. **Sentence beginning line 148, “odds” are plural (i.e., should read odds were lower)**

   **Response:** Resolved

---

**Response to Reviewer 1075023021702453 (Dr Hu)**

**Major Compulsory Revisions**

1. **The subdivision of NDVI and SD of NDVI into low, moderate, high: What criteria did you use to set the break values? E.g., why do you think NDVI is high if it is greater than 0.10? If you change the break values, do you expect that your statistical analysis will have same results?**

   **Response:** Please see our response to Major Compulsory Revision 3 by the first reviewer. To summarise, we have included additional information addressing this comment in the revised version of the manuscript [lines 121-126] and we have included a full description of our analytic approach to identifying appropriate break values in the *Supplementary Material*. We selected tertiles as the break values because they seemed to capture the variation in the outcome and each tertile contained sufficient/comparable sample i.e. same number of adults, which mean that the intervals would be directly comparable.
2. The paper needs to discuss/explain more why the disease is associated with the variability of greenness than the absolute level of greenness.

Response: Most of those defined as cases (i.e., having the disease of interest) in the study were diagnosed with coronary heart disease. There is insufficient evidence to suggest that the absolute level of greenness is protective. However, we agree with the reviewer that there needs to be more explanation as to why variability of greenness might be negatively associated with cardiovascular disease. Please also see our response Comment 3 (Minor Essential Revision) by the first reviewer. We have provided some speculative suggestions in the revised version of the manuscript: “Neighborhood attributes that may contribute to a high variability in greenness might include prevalence of tree-lined streets/cycleways/footpaths, presence of parks with parking, or green reserves with good road connectivity. The coexistence of both aesthetically pleasing natural vegetation to entice people out of their homes and destinations within walking distance would also contribute to variability in neighborhood greenness. A review of the environmental influences on walking concluded that aesthetic neighborhood attributes (which included ratings of natural features) were found to be associated with walking [7]. Furthermore, a cross-sectional study conducted in Seattle, US, reported that the most frequently walked non-green destinations were grocery stores, restaurants, libraries, coffee shops, and post offices [18]. Further studies are needed to identify the specific attributes of neighborhoods with a high degree of variability in greenness, such as the co-location of green and non-green areas (e.g., tree-lined paths, parks surrounded by well-connected streets).”

Minor Essential Revisions

1. What does “service area” mean? What does “1600m” mean? Figures 1 and 2 show two 1.6 km service areas of irregular shapes. What does the 1.6 km measure? Perimeter? Radius (but they are not circle)? I thought 1600m was circular buffer distance around home, but they are actually not circle. I am very confused.

Response: We have amended the manuscript as per our response to the first reviewer, Major Compulsory Revision 1. The service areas were not defined to be circular. Instead, they were defined as the irregular polygon spanned by a 1600 meter route along the road/footpath network from the participant’s home (i.e., network distance).

2. While Table 1 shows the distributions of socio-demographic, biological, health behavior risk factors, the paper does not show the distribution of greenness, which is the focus of the research.

Response: Thank you for identifying this omission. Table 1 has now been modified to include unadjusted measures of the distribution of greenness.

3. In Table 3, what do the bold numbers mean? Significant?

Response: Tables 2 and 3 in the original manuscript have been combined into a single table (Table 2) in the revised manuscript. We have added a new footnote to this table. “Bold text indicates statistical significance at the 5% (α) level”

Discretionary Revisions

1. Methods, paragraph 1, “Cross-sectional... who ... who... metropolitan area” has grammar error. This is not a sentence.
A cross-sectional study was undertaken for 11,404 adults aged 25 and over who (i) were residents of the Perth metropolitan area, (ii) consented to data linkage and (iii) completed the Western Australian Health and Wellbeing Survey between 2003 and 2009.

2. **Discussion, parag. 2, line 2. Add “by” after “influenced”.

Response: The omission has now been corrected (lines 208-210): “The lower prevalence of heart disease may be attributable to higher levels of physical activity, such as neighborhood walking which is positively influenced by the natural and built environment.”

References