PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Neighbourhood walkability, daily steps, and utilitarian walking in Canadian adults
AUTHORS	Hajna, Samantha; Ross, Nancy; Joseph, Lawrence; Harper, Sam; Dasgupta, Kaberi

VERSION 1 - REVIEW

REVIEWER	Gavin McCormack University of Calgary, Canada
REVIEW RETURNED	23-Jul-2015

GENERAL COMMENTS	GENERAL COMMENT Overall, the paper reads well and the findings are interesting. Including both self-report and objective measures of physical activity is a positive aspect of the study as is the inclusion of a large Canadian sample. The inclusion of a sensitivity analysis based on the different approaches to operationalizing the neighbourhood buffers is interesting and a positive aspect of the study. However, rather than reporting the results of the sensitivity analysis as a supplement, I think the method and findings should be integrated more into the manuscript, discussed in more detail, and compared with the results of other studies comparing the effect of neighbourhood scale on walking (e.g., Learnihan et al., 2011 in Geographical Research). The authors make good use of an existing dataset, but this also means that there are some major limitations with the study (i.e., exclusion of important confounders; a weak measure of self-reported utilitarian walking). Despite some concerns in regard to the methodology, the results are presented clearly and the discussion, for the most part, supports the results in their present form. A main concern is whether the findings of this study are sufficiently novel and whether it is providing evidence that is not already known in relation to the built environment and physical activity (even within the Canadian context). ABSTRACT: 1. The counter-intuitive finding regarding the relationship between neighbourhood walkability and accelerometer-measured steps should be presented in the abstract as an important finding. 2. The buffer size and type used to define the neighbourhood should be mentioned in the abstract. 3. Participants: this should highlight that this is the sample with complete data. The total participants (5605) who completed the CHMS should also be mentioned.

	INTRODUCTION: 1. Line 6, page 5: Is “self-report” related to the measurement of the built environment or to the measurement of utilitarian walking? 2. Line 13, page 5: Indicate whether the associations between biosensors and walkability in the cited evidence are based on self-reported or objectively-assessed measures of the built environment. 3. Line 32, page 5: The authors claim that combining measures of self-reported utilitarian walking and objective measures of total physical activity provides an optimal approach. Including both self-reported and objective measures for assessing physical activity is important, but is it the “optimal” approach. Combing Global Positioning Systems (GPS) data (with accelerometer data) would also provide some objective data about travel to destinations. 4. Lines 44 to 56, page 5: The outcome of focus in this paragraph is “step/counts” and it seems as though the authors are assuming here (based on their prior discussion of walking) that step/counts are for the most part capturing walking behaviour. There are North American studies that have compared accelerometer-measured physical activity between high and low walkable neighbourhoods (e.g., Saelens et al., 2003 in American Journal of Public Health; King et al., 2011 in Social Science and Medicine; Sallis et al., 2009 in Social Science and Medicine), although many use minutes as the outcome (total, moderate, vigorous). There are also studies from Australia that have examined differences in pedometer-counted steps between objectively different neighbourhood types (e.g., Giles-Corti et al., 2008 in Preventive Medicine). A study from Sweden also examined differences in accelerometer-measured physical activity and GIS-determined walkability (Sundquist et al., 2011, in Social Science and Medicine. Furthermore, these studies also include self-reported walking (and some including neighbourhood-specific walking) in addition to objectively-measured physical activity. Some of these studies also adjust for residential self-selection – an important confounder that is missing from the current study. Why have the authors decided to include only accelerometer-measured steps? Why would that also not include other accelerometer-measured physical activity outcomes? The authors should undertake a more thorough review of the evidence and provide more detail about the studies included in their introduction. For example, details about the similarity or differences in the way in which neighbourhood walkability has been operationalized would be particularly useful. Not all measures of “GIS-determined” walkability include the same built characteristics. This also gives a point of reference for the reader when they are presented with the operational definition of walkability used in the current study. METHOD 1. Study population, page 6: The sample size for the Canadian Health Measures Survey needs to be mentioned. An estimate of the response rate should also be included. Of the 5605 participants who participated in the survey (mentioned on line 15, page 10) report how many were invited (and eligible) to participate. 2. Line 8, page 7: Provide a rationale for adopting a 500-meter buffer to define participants’ neighbourhoods. Does this have policy or urban planning relevance? Furthermore, it is unclear whether the buffer is based on the street/pedestrian network or straight-line distance. If it is not line-based, then a rationale for not using this

	approach is needed. 3. Line 11, page 7: By “anonymous spatial coordinates”, are the authors referring to the coordinates for the geographical location of participant’s home? 4. Line 12, page 7: Provide a rationale for including land use mix, street connectivity, and population density in the measure of walkability. These are not mentioned as important contributors to walkability in the introduction section but should be discussed. Also, why are these built characteristics only included in the measure of walkability. The availability of transit stops is another potential correlate of transportation walking (which the authors allude to as being important in the discussion). “Total steps” is an outcome measure in this study, of which only some of these measured steps will reflect transportation walking. Why then would the authors not include built environment characteristics that might be important for other types of walking (e.g., sidewalks/pathways, parks, traffic calming devices etc.)? 5. Line 8 to 22: Mention the year in which the GIS databases were updated. 6. Line 29, page 7: Walkscore is used as a measure of “walkability” but this also was not mentioned in the introduction section. Why would the authors resort to using Walkscore when they have access to GIS data and can more accurately estimate availability of specific destinations within a participant’s neighbourhood? They link Walkscore to the participant’s neighbourhood via postal code, yet they have the participant’s geocoded complete household address. Another major concern is that in their data the authors have participants from urban and rural geographical locations. How well does Walkscore (based on postal code) reflect the walkability of the participant’s neighbourhood, particularly in rural areas? 7. Line 13, page 8: Provide a reference to support the definition of valid hours for the accelerometer data used in this study. 8. Line 6, page 8: The study would be strengthened (and comparable with many more other studies) if the authors include other accelerometer-measured outcomes, such as low, moderate, vigorous-intensity, and total minutes (in addition to steps). 9. Line 13, page 8: Did the four or more days of accelerometer data include data from weekends and weekdays? 10. Line 25, page 8: Given that self-reported utilitarian walking is one of the main study outcomes, information regarding the reliability and validity of this item is required. 11. Line 48, page 8: The authors have included season in as a covariate, which is justified given the relationships found between season and physical activity in other Canadian-based studies. However, there is the potential for a date miss-match in the collection of the self-report utilitarian walking (i.e., reporting for a typical week in the past 3-months) and season, which was determined based on the date the survey was administered. For instance, there is the potential for a participant to report a winter pattern of transportation walking, even though they completed the survey in spring. This issue should be mentioned as a study limitation or addressed, if possible. 12. Line 34, page 8: A major limitation of this study is the lack of statistical adjustment for residential self-selection and neighbourhood exposure. Other cross-sectional studies that are similar to the current study have included measures of residential self-selection. Numerous literature reviews on built environment and physical activity published over the past two decades have recommended the need to control for residential self-selection through either study design or statistical approaches. This is a major

	limitation. 13. Line 20, page 13: What do the authors mean when they say: "This is a meaningful outcome as it roughly dichotomizes participants into utilitarian walkers and non-walkers" when referring to using a cut-point of ≥1 hour/week. Participants achieving less than this cut-point could still participate in frequent utilitarian walking (e.g., walking to and from the bus stop 10-minutes/day). Non-walker, would refer to some who reports 0 (or very low) minutes of utilitarian walking. This sentence should be deleted. 14. Line 25, page 9: This statement should include references to studies with which the results of this current study will be compared. 15. Provide a rationale for choosing the cut-point for annual household income (i.e., ≥$40,000). RESULTS 1. Table 1: For some of the categorical variables the authors do not mention the excluded category while for others it is presented in parenthesis. The excluded categories should be presented in the table. 2. Table 3: Please include the r-squared estimates for each linear regression model. Please also include pseudo r-squared estimates for the each logistical regression model. DISCUSSION 1. Line 4, page 16: The authors provide an explanation of their counter-intuitive finding, whereby more walkability neighbourhoods were associated with lower daily steps. I wonder however, whether the least walkable Canadian suburban neighbourhoods would be likely to have high accessibility to public transit (as the authors suggest). This might be the exception rather than the norm, based on the several published Canadian studies examining associations between the built environment and physical activity. Accelerometer-measured steps capture both utilitarian and recreational walking (and other physical activity). Recreational walking is more common than transportation walking in Canada. It would be more likely that the low walkable, suburban neighbourhoods in Canada, for the most part, have features that facilitate or support non-utilitarian walking, rather than having better access to public transit (e.g., bus stops). The higher steps reported in the least walkable neighbourhoods in this study might reflect a more supportive environment for recreational walking or other forms of physical activity. 2. Line 8, page 18: Indicate that the two potential confounders (car ownership and residential self-sections) could not be account for because they were not measured in the CHMS. Another limitation is that the self-reported measure of utilitarian walking was not context (or neighbourhood) specific. Level of exposure to the neighbourhood environment was not taken into account in the analysis.

REVIEWER	Andrew Kaczynski University of South Carolina, USA
REVIEW RETURNED	02-Aug-2015

GENERAL COMMENTS	Thank you for the opportunity to review this paper for BMJ Open. This study used a large and interesting data set to somewhat novel take on the relationship between neighborhood walkability and steps/day and self-reported utilitarian walking. I think the paper can

	be amended to be publishable, but there are currently numerous minor areas that should be addressed (collectively, they amount to a major revision, but I don't believe any of them are especially onerous on their own). I hope my comments are helpful to the authors and I wish them well with their future work in this area. The introduction is very brief, but generally provides the necessary context for the study. However, in the last paragraph, it seems as though China is missing from the list of countries mentioned for past studies in this area (see reference #14 … unless #14 was not supposed to be included in the previous sentence, because 9-14 is indeed six studies, not four, so please clarify). More importantly, the authors mention that comparable studies of objective walkability and step counts have not been conducted in North America. In a literal sense, this MIGHT be true, but there are plenty of studies that have used accelerometers if not step counters (pedometers) in relation to GIS metrics of walkability. Moreover, even if there weren't many U.S. or Canadian studies on walkability and steps, why is this an important gap to fill? (i.e., describe what is different about North America that necessitates specific studies here). In general, I think the rationale for the study needs to be improved to convince the reader that the ensuing methods are worth continuing forward into. For the description of the CHMS sampling strategy, there is not much detail about actual participant sampling or recruitment, only that the CHMS collects data on a (supposedly) representative sample of Canadians and that there were 15 data collection sites. Can you provide a bit more information about how PEOPLE in the study are selected and recruited, not just the sites where they attend data collection? (and, if I understand correctly, I'm a little amazed that the entirety of Canada is sampled down to only 15 sites from which those respondents are garnered – if I am mistaken, please clarify). I am familiar with the measure of walkability the authors are employing (three constructs summed using z-scores). However, I think they could provide some more rationale as to why this index and its components are appropriate for determining walkability (and not other variables instead or in addition to). Also, they could at least provide some citations for the composite measure. Couldn't the authors get a Walk Score based on the anonymous spatial coordinates instead of having to resort to postal codes? If not, I think some discussion of the size of postal codes and the appropriateness of linking individual physical activity data to larger postal codes is warranted. Also, when you say Walk Score is a validated measure, please provide a citation (e.g., Duncan, etc.). Some readers might wonder why step counts were used as the outcome variable when accelerometers were used for data collection and more precise measures of activity time and intensity were presumably available. I think some discussion of the motivation and justification for this decision would be helpful. Is there any reliability or validity information for the utilitarian walking question? Estimating the typical week over the past three months might be challenging. As well, the question would seem to exclude walking for other utilitarian purposes (e.g., to visit someone, to go to a restaurant); I believe such questions are usually phrased something like "to get to and from places". This raises further

	questions about validity, so I would suggest the authors add any psychometric information if possible. In the covariates section, the categories for perceived health seem to be out of order. Was the steps/day variable fairly normal (with respect to its distribution) and therefore appropriate for linear rather than logistic regression? (often, PA variables are not very normal) When the authors reference walkability quartiles from previous studies (page 9, line 25), can they provide some citations for studies that have used this approach? (e.g., Frank et al., 2005, etc.). The descriptive statistics describe a sample of 5605 individuals whittled down to 2949 participants with complete data (which is a big drop, but tolerable). However, earlier in the paper, it says something about there being 9,351 neighborhoods for which GIS-based walkability was calculated. This was confusing. Why is the font tiny on pages 13 and 14? In the results section about utilitarian walking (page 13), please state that these results are for the highest walkability and Walk Score quartiles in relation to the LOWEST walkability quartile (like you do in the previous section on daily step counts).

VERSION 1 – AUTHOR RESPONSE

REVIEWER ONE

GENERAL COMMENTS

Comment 1: Rather than reporting the results of the sensitivity analysis as a supplement, I think the method and findings should be integrated more into the manuscript, discussed in more detail, and compared with the results of other studies comparing the effect of neighbourhood scale on walking (e.g., Learnihan et al., 2011 in Geographical Research).

Thank you for your suggestion. We now address the importance of neighbourhood scale on walking in our revised manuscript. We also have included some relevant references, including the one that you suggested by Learnihan et al (2011). We now state:

“The choice of the geographic scale at which neighbourhoods are defined may influence the estimated associations between neighbourhood walkability and walking.1,2 To address this issue, we conducted sensitivity analyses to assess if varying the sizes and shapes of neighborhood buffers used in the calculation of the GIS-derived walkability index (i.e., 1000-m polygonal buffers; 500 and 1000-m line-based buffers) would meaningfully alter the regression results.”

We agree that our sensitivity analyses based on neighbourhood scale are interesting. Upon much thought, however, we have decided to not proceed with integrating the results into the body of the manuscript. There are two reasons for this:

1. We would like to keep our focus on assessing the relationship between walkability, utilitarian walking and accelerometer-assessed daily steps. Incorporating the results of the sensitivity would draw the focus away from this to a more methodological issue.

2. BMJ Open has a limit of five figures and/or tables in this manuscript. In order to incorporate these analyses into the manuscript, we would have to expand the size of our existing tables (Tables 3 and 4). We feel that this added information would make the tables too 'busy' and would distract from the main messages that we would like to make.

Comment 2: A main concern is whether the findings of this study are sufficiently novel and whether it is providing evidence that is not already known in relation to the built environment and physical activity (even within the Canadian context).

Upon rereading the introduction, we agree that the novelty of our study was not sufficiently highlighted. We have edited the introduction to explain what our study adds over and above what has already been done both abroad and within the Canadian context. We do this in three steps:

First, we first demonstrate that the majority of studies that have been conducted on the association between researcher-assessed walkability and researcher-assessed measures of physical activity have been done outside of North America:

"Based on a recent systematic review and meta-analyses of studies that have been done using biosensor-assessed daily steps, we know that, in Europe and in Asia, adults who live in high compared to a low walkable neighbourhoods, accumulate 766 more steps/day (95% credible interval). This accounts for approximately 8% of recommended daily steps."

Second, we give our rationale for why the relationship between walkability and physical activity may be different in the Canadian context:

"We do not know, however, what this association is like in Canada. Since adults living in Europe and in Asia are more physically active than North American adults 3 and they might also have very different opinions regarding the importance of walking-friendly neighbourhoods, the influence of walkability on total physical activity may be different in the Canadian context."

Lastly, we show that all of the studies that have been conducted in Canadian adults have been restricted to small geographic areas with limited variability in walkability and relatively small sample size or restricted to self-reported measures of physical activity:

"Canadian studies to date have been restricted to a single city with limited variability in neighbourhood walkability and/or relatively small sample sizes.4-6 A large study (n=151,318) was published on the association of neighbourhood walkability as captured by the publicly-available Walk Score® with total, utilitarian and leisure-time walking, but the measures of physical activity were based on self-report. Studies with large variability in neighbourhood walkability, a population-based sample of adults, and researcher-assessed exposures and outcomes are needed to elucidate the role of neighbourhood walkability on total physical activity in the general Canadian adult population."

The novelty of our study is that it is the first to: ours is the first:

"...to estimate the association of researcher-assessed neighbourhood walkability (measured using GIS and the Walk Score®) with both self-reported utilitarian walking and biosensor-assessed total walking in a large sample of Canadian adults.

ABSTRACT

Comment 3: The counter-intuitive finding regarding the relationship between neighbourhood walkability and accelerometer-measured steps should be presented in the abstract as an important finding.

Thank you for your suggestion. We have rearranged the presentation of our results and conclusions in the abstract to better highlight this counter-intuitive finding. The results section now reads:

“No important relationship was observed between neighbourhood walkability and daily steps. Participants who reported more utilitarian walking, however, accumulated more steps (<1 hours/week: 6613 steps/day, 95% CI 6251, 6975; 1 to 5 hours/week: 6768 steps/day, 95% CI 6420, 7117; ≥6 hours/week: 7391 steps/day, 95% CI 6972, 7811). There was a positive graded association between walkability and odds of walking ≥1 hour/week for utilitarian purposes (e.g., Q4 versus Q1 of GIS-derived walkability: OR=1.66, 95% CI: 1.31, 2.11; Q3 versus Q1: OR=1.41, 95% CI: 1.14, 1.76; Q2 versus Q1: OR=1.13, 95% CI: 0.91, 1.39), independent of age, sex, body mass index, married/common-law status, annual household income, having children in the household, immigrant status, mood disorder, perceived health, ever smoker, and season.”

The conclusion section now reads:

“Contrary to expectations, living in more walkable Canadian neighbourhoods was not associated with more total walking. Utilitarian walking and daily steps were, however, correlated and walkability demonstrated a positive graded relationship with utilitarian walking.”

Comment 4: The buffer size and type used to define the neighbourhood should be mentioned in the abstract.

Thank you for your comment. We now incorporate the buffer size and type into the abstract. We now state that:

“Contrary to expectations, living in more walkable Canadian neighbourhoods was not associated with more total walking. Utilitarian walking and daily steps were, however, correlated and walkability demonstrated a positive graded relationship with utilitarian walking.”

Comment 5: Participants - this should highlight that this is the sample with complete data. The total participants (5604) who completed the CHMS should also be mentioned.

Thank you for your comment. As you suggested, we now provide more details regarding the samples sizes at various stages of the analyses including 1) the number of individuals who participated in the CHMS, 2) how many participants attended each of the assessments (household interviews and in-clinic assessments), and 3) how many participants had complete data on exposures, outcomes, and covariates.

“5605 individuals participated in the survey. 3727 adults (≥18 years) completed a computer-assisted interview and attended a mobile clinic assessment. Analyses were based on those who had complete exposure, outcome, and covariate data (n= 2949).”

INTRODUCTION

Comment 6: Line 6, page 5: Is “self-report” related to the measurement of the built environment or to the measurement of utilitarian walking?

Thank you for drawing our attention to this. This was in reference to the measurement of utilitarian walking. Upon rereading our manuscript, we felt that including the term 'self-report' was redundant since utilitarian walking can only be assessed feasibly with self-report. This sentence has been edited our due to all of the other suggested changes, but we have carefully reviewed the rest of our manuscript to make sure that our descriptions of self-report and researcher-assessed measures are well aligned with the exposures and outcomes that we are speaking about.

Comment 7: Line 13, page 5: Indicate whether the associations between biosensors and walkability in the cited evidence are based on self-reported or objectively-assessed measures of the built environment.

As suggested, we now indicate that the association between biosensors and walkability in the cited evidence is based on objectively-assessed measures of the built environment.

"Numerous studies have been conducted on the association between biosensor-assessed physical activity (e.g., minutes spent in moderate-to-vigorous intensity physical activity) and Geographic Information System (GIS)-derived measures of street connectivity, land use mix and population and/or population/residential density.7-10"

Comment 8: Line 32, page 5: The authors claim that combining measures of self-reported utilitarian walking and objective measures of total physical activity provides an optimal approach. Including both self-reported and objective measures for assessing physical activity is important, but is it the "optimal" approach? Combing Global Positioning Systems (GPS) data (with accelerometer data) would also provide some objective data about travel to destinations.

Yes, we completely agree. In fact, we are conducting a study in which we are investigating the link between neighbourhood walkability and neighbourhood-specific physical activity captured using integrated GPS-accelerometer technology. The approach that we suggested in this study (i.e., assessing physical activity using both self-report and biosensors) is advantageous, but not necessarily optimal. Other more optimal, albeit more labor-intensive options, could be used. We have replaced the statement:

"Combining self-reported utilitarian walking with objective measures of total physical activity provides the optimal approach."

with

"Given the potential for biases associated with self-reported measures of physical activity,11,12 combining self-reported utilitarian walking with objective measures of total physical activity is advantageous. It allows researchers to isolate the policy-amenable subset of total physical activity (utilitarian walking) while also providing estimates of the potential for walkable environments to influence total physical activity."

Comment 9: Lines 44 to 56, page 5: The outcome of focus in this paragraph is "step/counts" and it seems as though the authors are assuming here (based on their prior discussion of walking) that step/counts are for the most part capturing walking behaviour. Why have the authors decided to include only accelerometer-measured steps? Why would that also not include other accelerometer-measured physical activity outcomes?

Yes, we are indeed assuming, based on previous work, that daily steps are a good measure of total walking.13,14 We recognize, however, that by focusing exclusively on steps/day, we were not giving

credit to the important work that has been done using other objectively-assessed metrics of physical activity (e.g., MVPA). As suggested, we have edited the introduction to acknowledge the important work that has been done using other accelerometer-assessed measures of physical activity (e.g., Oakes, 2007; Ding, 2014; Saelens et al., 2003). Please see our response to Comment 18 for our reasons for not using other accelerometer-measured physical activity outcomes in our study.

Comment 10: The authors should undertake a more thorough review of the evidence and provide more detail about the studies included in their introduction. For example, details about the similarity or differences in the way in which neighbourhood walkability has been operationalized would be particularly useful. Not all measures of "GIS-determined" walkability include the same built characteristics. This also gives a point of reference for the reader when they are presented with the operational definition of walkability used in the current study.

Thank you for your suggestion. We have expanded the introduction to include a more thorough review of the literature. We have also taken your suggestion of including a discussion about different ways of measuring walkability (i.e., perceived/participant-reported versus researcher-assessed). We trust this is helpful in positioning our choice of exposures (i.e., GIS-derived street connectivity, land use mix and population density) in the larger context of the other measures that exist.

METHOD

Comment 11: Study population, page 6: The sample size for the Canadian Health Measures Survey needs to be mentioned. An estimate of the response rate should also be included. Of the 5604 participants who participated in the survey (mentioned on line 15, page 10) report how many were invited (and eligible) to participate.

Thank you for you comment. We now include the sample size for the Canadian Health Measures Survey in the methods section (including the response rate and the number of people who were included in our final analyses). We felt, however, that including this information again in the results section would be redundant so we left the results section as is. If the editor wishes for us to repeat this information in the results section we would be happy to do so.

Comment 12: Line 8, page 7: Provide a rationale for adopting a 500-meter buffer to define participants' neighbourhoods. Does this have policy or urban planning relevance? Furthermore, it is unclear whether the buffer is based on the street/pedestrian network or straight-line distance. If it is not line-based, then a rationale for not using this approach is needed.

Thank you for you suggestion. Our reason for selecting a 500-meter buffer was that it would approximate a five to ten minute walk from one's neighbourhood. Areas farther away were hypothesized to be less important for walking. This is supported by a study by Learnihan and colleagues in 2011 (Geographical Research, 49(2):183-191) in which the authors demonstrated that the more accurate the neighbourhood scale (e.g., the suburb versus a 15-minute buffer around the neighbourhood), the stronger the association with utilitarian walking. Our choice of a 500-m buffer was also consistent with the buffer size used in previous North American studies (e.g., Adams et al., Int J Health Geog 2014 reporting on the results of the large International Physical Activity and the Environment Network study).

The buffer that we used was based on the street-network (using only pedestrian-accessible streets). Transit routes that were not deemed to be walkable (e.g., highways) were excluded from the network when creating the neighbourhood buffers. We have clarified these two points in the methods. We now

state that:

“Buffers were defined based on street networks. Streets that were not pedestrian-accessible (e.g., highways) were excluded from the creation of the neighbourhood buffers. Five hundred meter buffers were chosen as the scale of analysis as they approximated a five to ten minute walk from the home and would capture the environment to which participants are most exposed.”

Comment 13: Line 11, page 7: By “anonymous spatial coordinates”, are the authors referring to the coordinates for the geographical location of participant’s home?

We apologize for not making this clear. The “anonymous spatial coordinates” are referring to the coordinates for the geographical location of participant’s home. We have clarified this in the text. We now state that:

“Home neighborhoods were approximated using 500-meter polygonal buffers around latitude-longitude coordinates that corresponded to the centroid of the participants’ home postal codes.”

Please note that we removed our reference to “9,351 anonymous spatial coordinates”. As pointed out by Reviewer Two, this would be confusing to readers given that initial sample size of the CHMS was only 5,604. The reason for the additional coordinates was that when we requested access to this data, Statistics Canada was not allowing access to any spatial information from the CHMS as they thought that this could be a breach of confidentiality. The only way that they would allow us to get access to this data was if they gave us a separate blinded xy coordinate file with 9351 observations (only 5,604 of these corresponded to the addresses of the CHMS participants). We then had to derive our built environment variables and send the full file back to Statistics Canada. At that point Statistics Canada merged our built environment file with the individual-level data collected by the CHMS, removed all of the spatial data, and sent the file to the McGill-Concordia Research Data Centre so that there would be no way that the location of any of the CHMS participants could be traced.

Comment 14: Line 12, page 7: Provide a rationale for including land use mix, street connectivity, and population density in the measure of walkability. These are not mentioned as important contributors to walkability in the introduction section but should be discussed. Also, why are these built characteristics only included in the measure of walkability. The availability of transit stops is another potential correlate of transportation walking (which the authors allude to as being important in the discussion). “Total steps” is an outcome measure in this study, of which only some of these measured steps will reflect transportation walking. Why then would the authors not include built environment characteristics that might be important for other types of walking (e.g., sidewalks/pathways, parks, traffic calming devices etc.)?

Thank you for your suggestion. Upon rereading our manuscript, we agree that we did not provide a strong enough rationale for why we were specifically including land use mix, street connectivity, and population density in the measure of walkability. We have edited our introduction to better introduce these three measures. A brief summary is provided below:

As we now describe in our opening paragraph, neighbourhoods were historically designed to have greater land use mix, higher street connectivity, and higher population density (e.g., South London in the 1850s). With the advent of the automobile, people began to rely heavily on motorized transport. With this, came the development of sprawling suburbs, characterized by low land use mix, higher street connectivity, and higher population density and concurrent drops in physical activity. The literature has speculated that these three large-scale neighbourhood features are the main culprits behind reduced levels total walking. Our aim was to assess the link between these large-scale features and total walking in Canada. While we chose to focus exclusively on these large-scale

features of neighborhood designs, we completely agree that there are many other characteristics of the built environment that may influence an individual's propensity to engage in physical activity (e.g., aesthetics, neighbourhood safety, transit stops). In the revised manuscript, we have added some lines that acknowledge that there are other features that may be important drivers of higher levels of physical activity. We state:

"…we focused on associations of walking with large-scale features of neighbourhood designs. We acknowledge there are other potentially important features of the built environment (e.g., aesthetics, neighbourhood safety, transit stops) that may be importantly associated with both utilitarian and total walking."

Comment 15: Line 8 to 22: Mention the year in which the GIS databases were updated.

Thank you for drawing our attention to this. The year was buried in the references that we included with our supplementary files. As you have suggested, we have now added the year (i.e., 2009) of the GIS databases into the body of the manuscript. After reading the methods section again, we felt that it would also be better to include the operational definitions of our main exposures in the body of the manuscript rather than referring the readers back to the Supplemental Files. We have therefore removed Supplementary File 1 and incorporated all of this information into the main text.

Comment 16a: Line 29, page 7: Walk Score is used as a measure of "walkability" but this also was not mentioned in the introduction section. Why would the authors resort to using Walk Score when they have access to GIS data and can more accurately estimate availability of specific destinations within a participant's neighbourhood?

Thank you for drawing our attention to the fact that we did not introduce the Walk Score as a measure of "walkability" in the introduction. In the revised version of the manuscript we now mention the Walk Score in both the abstract and the introduction of the manuscript. We have also inserted a couple of lines into the methods section to explain how the Walk Score is linked to the concept of walkability. We state:

"The Walk Score® is a validated measure of the walkability of a geographic location based on its proximity to 13 walkable destinations.15-17 The score ranges from 0 (car-dependent) to 100 (walker's paradise) and is calculated based on an algorithm that assigns equal weights to each walkable destination.18 It is relevant to the construct of walkability as it reflects the diversity and density of neighbourhoods. A higher Walk Score® is indicative of a greater diversity of services and also higher population density which creates a higher demand for such services.19"

We agree that measures other than the Walk Score could have been used to capture proximity to destinations (e.g., distance to stores). In fact, we did use one such measure (i.e., land use mix). We acknowledge that land use mix does not capture exactly the same construct as the Walk Score, but they are closely related. The more land uses that are in a neighbourhood, the more likely it will be that the participants' homes will be located close to a walkable destination (i.e., a higher Walk Score). This was confirmed in our analyses where we found a moderate correlation between the Walk Score and land use mix (R=0.67, 95% CI 0.65 to 0.69).

Given that we did use GIS-derived land use mix as a primary exposure of interest, we would not say that we used the Walk Score at the expense of another variable, as you suggest. Instead we used the Walk Score as a supplemental measures that is being increasingly used for research purposes and that may be more easily understood by the general public compared to more technical GIS-based measures (i.e., as evidenced by its regular integration into real estate advertisements in the USA to promote walkability as a sellable feature of a property).

Comment 16b: They link Walk Score to the participant's neighbourhood via postal code, yet they have the participant's geocoded complete household address.

We did not have the household addresses of participants – only the xy coordinates for the centroid of their postal code area. We apologize if this was not clear. The methods have been edited to make this clearer. We now state:

"Home neighborhoods were approximated using 500-meter polygonal buffers around latitude-longitude coordinates that corresponded to the centroid of participants' home postal codes."

At the time that we requested data from Statistics Canada, address information was not made available to researchers. A special arrangement had to be made between McGill, Statistics Canada and the McGill-Concordia Research Data Centre in order for us to get access to this data. Even then, the only spatial data that Statistics Canada would release were the postal code centroids. Please see our Response to Comment 13 for more details regarding this arrangement.

Comment 16c: Another major concern is that in their data the authors have participants from urban and rural geographical locations. How well does Walk Score (based on postal code) reflect the walkability of the participant's neighbourhood, particularly in rural areas?

This is an interesting point. The Walk Score is a measure based on a location's proximity to 13 walkable destinations, where a higher value is indicative of a more walkable location (Range: 0-100). It has been validated in a range of settings,15,16,20 but further validation in rural settings is needed as pointed out by Dustin Duncan in 2013.17 We agree that in rural locations where land use data may not be available (e.g., in Nunavut in Northern Canada), the validity of the Walk Score would be low since the number of destination could not be quantified. However, the 'rural' locations that were included in our study were no more than 100 kilometers away from large urban centres and the majority are perhaps better termed 'peri-urban' with landscapes more similar to urban environments. From our experience with land use files, coverage in Canada is good surrounding the urban areas that were home to the CHMS assessment sites. Assuming that the land use files that the Walk Score algorithm used are comparable in coverage, we expected that the Walk Score was a good measure of walkability (based proximity to destinations) even in these peri-urban areas. We tested our hypothesis in two ways:

1. Only a small subset of our study sample lived in peri-rural neighbourhoods (n=424). Of these, 76.7% (n=325) had Walk Score of 0 (the locations that would be most likely to suffer from measurement error since a value of 0 could indicate that there were no land use data available in these locations). We took a random sample of these addresses and checked (using Google maps) that they were indeed in locations that did not have any walkable destinations nearby. Indeed, this was the case. Because of this, we concluded that the Walk Score was able to estimate the walkability even in these peri-rural locations.

2. In a study was published this month in Preventative Medicine (Thielman et al., 2015 Aug;77:174-80), the authors examined the association between the Walk Score and self-reported total walking, walking for transport and leisure-time walking using data from the Canadian Community Health Survey. In their paper, they expressed the same concern regarding the validity of the Walk Score in rural locations. To address this they removed all of the participants that were living in the rural locations from their analyses. We did the same in our analyses and found that the results were the same regardless of whether or not we included or excluded participants with rural addresses from our analyses.

In summary, given that we had 1) little reason to suspect that the Walk Score did not do a good job in capturing walkability in rural areas included in this study and 2) that including or excluding these data had no important effect on our results, we decided to retain all rural participants in our analyses. This allowed us to retain the maximal sample size and to not unnecessarily limit the generalizability of our results.

Comment 17: Line 13, page 8: Provide a reference to support the definition of valid hours for the accelerometer data used in this study.

Thank you for drawing our attention to this section. This was an oversight on our part. The criteria for defining a valid day was not applicable to the step count data. Steps per day equaled the total steps accumulated divided by the number of days for which valid step counts were recorded. The majority of participants (87%) had step data for 7 days. Less than 5% had step data based for fewer than 4 days. Thus, even if there is bias associated with fewer wear days, the effect on our results is negligible. We have noted this, however, in our limitation section as a possibility:

"…walking was assessed for up to seven days, a snapshot that may not be representative of habitual walking levels."

Comment 18: Line 6, page 8: The study would be strengthened (and comparable with many more other studies) if the authors include other accelerometer-measured outcomes, such as low, moderate, vigorous-intensity, and total minutes (in addition to steps).

Thank you for this comment. We agree that other accelerometer-assessed outcomes would have been interesting to include. We, however, decided a priori to restrict our analyses to walking as captured by steps taken per day for three reasons. We have updated our introduction as follows to highlight these three reasons:

"Daily steps are an outcome of particular interest for the study of neighbourhood walkability for several reasons. First, daily steps are a good estimate of total habitual physical activity, something that is not captured by other biosensor-assessed metrics, such as MVPA. Second, walking is the most common and preferred form of physical activity among adults. Understanding if neighbourhood walkability is associated with daily step among adults would suggest that neighbourhood-level interventions may have the potential to impart a benefit to large segments of the population. Third, daily steps are highly interpretable.21 by both the scientific and lay communities. This is unlike other biosensor-assessed metrics (e.g., accelerometer counts) that may also be good a capturing total levels of physical activity, but are of less value when trying to explain the association of walkability with physical activity in an easily interpretable and relevant way. For example, saying that living in a high compared to a low walkable neighbourhood is associated with x more steps/day is more interpretable and readily understood by the public than saying that it is associated with x more activity counts."

Comment 19: Line 13, page 8: Did the four or more days of accelerometer data include data from weekends and weekdays?

Please see our response to your Comment 17. (The criteria for defining a valid day was not applicable to the step count data. This was an oversight on our part.) Most people had data from both the weekend and the weekdays given that the majority of participants (87%) had step data for 7 days. Less than 5% had step data based for fewer than 4 days. Thus, even if there is bias associated with fewer wear days and not having data from the weekend or weekday, the effect on our results is negligible. We note in the limitations that steps were only assessed for a short period of time:

"…walking was assessed for up to seven days, a snapshot that may not be representative of habitual walking levels. If steps were measured over a longer period of time an influence of walkability on daily steps may have emerged."

Comment 20: Line 25, page 8: Given that self-reported utilitarian walking is one of the main study outcomes, information regarding the reliability and validity of this item is required.

This measure of utilitarian walking has been used in other previous Canadian health surveys including the Aboriginal Health Survey22 and the National Population Health Survey23 (we now note this in our methods section). To our knowledge, no validation studies have been conducted on this measure of utilitarian walking. Although we cannot make definite claims about validity due to a lack of studies, we tested the construct validity of this measure by comparing it with daily steps. As expected, we demonstrated that higher self-reported utilitarian walking was associated with more daily steps (<1 hours/week: 6613 steps/day, 95% CI 6251, 6975; 1 to 5 hours/week: 6768 steps/day, 95% CI 6420, 7117; ≥6 hours/week: 7391 steps/day, 95% CI 6972, 7811).

Comment 21: Line 48, page 8: The authors have included season in as a covariate, which is justified given the relationships found between season and physical activity in other Canadian-based studies. However, there is the potential for a date miss-match in the collection of the self-report utilitarian walking (i.e., reporting for a typical week in the past 3-months) and season, which was determined based on the date the survey was administered. For instance, there is the potential for a participant to report a winter pattern of transportation walking, even though they completed the survey in spring. This issue should be mentioned as a study limitation or addressed, if possible.

We see the reviewer's point but to counter this issue we grouped seasons into two categories: spring/summer and fall/winter (given similar step counts in both the paired categories). This serves to lessen the degree of mismatch present (i.e., rather than four possible mismatches there would be only two possible mismatches).

Further, even if there were mismatch for every single person near this borderline, this would be for a very small percentage of people (i.e., <5%) and could not have an important effect on the results. Specifically, a mismatch could be expected to occur two to four weeks every six months (i.e., utilitarian walking could be reported as occurring in the winter in the first two to four weeks of spring, and utilitarian walking could be reported as occurring in the summer in first two to four weeks of fall). Based on the visit dates of participants included in our study (n=2949), 2.3%, 2.9% and 4.7% of participants would have a walking-season mismatch assuming a two, three and four-week mismatch period at the beginning of each fall/winter and spring/summer periods.

We thank you for your comment. We have inserted a line into the limitations section to address the possibility of this mismatch.

Comment 22: Line 34, page 8: A major limitation of this study is the lack of statistical adjustment for residential self-selection and neighbourhood exposure. Other cross-sectional studies that are similar to the current study have included measures of residential self-selection. Numerous literature reviews on built environment and physical activity published over the past two decades have recommended the need to control for residential self-selection through either study design or statistical approaches. This is a major limitation.

Yes, we agree this lack of information is a limitation of this study. A recent study (Cao, 2010) demonstrated that not accounting for residential self-selection may over estimate the association of neighbourhood type with utilitarian walking by 64%. If this applies to our cohort then our estimated effect of 66% increased odds of walking ≥1 hour/week for utilitarian purposes (95% CI: 1.31, 2.11)

when living in the most walkable neighbourhoods, may be an overestimate of the true association.

Unfortunately we could not be address residential self-selection in our study as data on participants' reasons for moving to their neighbourhoods were not queried in the Canadian Health Measures Survey. We note this in the discussion as an important limitation:

"Third, two potential confounders (i.e., car ownership and residential self-selection24-28) could not be accounted for in our analyses."

Comment 23: Line 20, page 13: What do the authors mean when they say: "This is a meaningful outcome as it roughly dichotomizes participants into utilitarian walkers and non-walkers" when referring to using a cut-point of ≥1 hour/week. Participants achieving less than this cut-point could still participate in frequent utilitarian walking (e.g., walking to and from the bus stop 10-minutes/day). Non-walker, would refer to some who reports 0 (or very low) minutes of utilitarian walking. This sentence should be deleted.

We completely agree. We have deleted this sentence.

Comment 24: Line 25, page 9: This statement should include references to studies with which the results of this current study will be compared.

Thank you for your suggestion. Upon reading this statement again we have decided to change the line from:

"Associations were estimated across quartiles of walkability to maximize comparability with previous studies"

to:

"Associations were estimated across quartiles of walkability."

While it is true that we decided to use quartiles to keep in line with what other researchers have done, (e.g., Frank, 2005; Leslie, 2007) and to increase the interpretability of our effect estimates (as opposed to keeping the variable in its continuous form), the studies to which we compare our results define high versus low walkability using other category cutoffs (e.g. lowest and highest deciles). To avoid confusion regarding exactly which studies we are comparing our results to, we have decided to remove this clause.

Comment 25: Provide a rationale for choosing the cut-point for annual household income (i.e., ≥$40,000).

$36,000 has also been suggested as being the minimum household income to qualify a household as middle class for a family of four.29 The closest category that this corresponded to in our dataset was $40,000. We chose $40,000 as a cut-point as it would roughly dichotomize people into either low versus middle-income households. We now state this in our revised manuscript:

"A cut-off of ≥$40,000 for total annual household income was selected as it corresponded to the minimum income required to qualify a household with four members as middle class."

RESULTS

Comment 26: Table 1: For some of the categorical variables the authors do not mention the excluded category while for others it is presented in parenthesis. The excluded categories should be presented in the table.

The reason that we did not list the excluded categories for some of the dichotomous variables was that for the alternate category for these variables is relatively intuitive. (e.g., “being a woman” versus “being a man”). However, for the variable “good/very good/excellent perceived health” the alternate categories are not that obvious. We have inserted the alternate response options, but if the editor feels that any of these categories are unnecessary, we would be happy to remove these.

Comment 27: Table 3: Please include the r-squared estimates for each linear regression model. Please also include pseudo r-squared estimates for the each logistical regression model.

Thank you for your comment. We know include r-squared estimates for each linear regression model and pseudo r-squared estimates for each logistic regression model.

DISCUSSION

Comment 28: Line 4, page 16: The authors provide an explanation of their counter-intuitive finding, whereby more walkability neighbourhoods were associated with lower daily steps. I wonder however, whether the least walkable Canadian suburban neighbourhoods would be likely to have high accessibility to public transit (as the authors suggest). This might be the exception rather than the norm, based on the several published Canadian studies examining associations between the built environment and physical activity. Accelerometer-measured steps capture both utilitarian and recreational walking (and other physical activity). Recreational walking is more common than transportation walking in Canada. It would be more likely that the low walkable, suburban neighbourhoods in Canada, for the most part, have features that facilitate or support non-utilitarian walking, rather than having better access to public transit (e.g., bus stops). The higher steps reported in the least walkable neighbourhoods in this study might reflect a more supportive environment for recreational walking or other forms of physical activity.

We completely agree and thank you for your comment. We have incorporated this into our discussion. We now state:

“This is counterintuitive as it suggests that more walkable neighborhoods are associated with lower daily steps. This may be a result of Quartile 1 being representative of suburban neighborhoods characterized by good access to public transit. It has been demonstrated that even in very low walkable neighborhoods, if there are transit stops, residents will walk to board express buses and trains.30 It may also be a result of the lowest walkable neighbourhoods having desirable aesthetic or other features that encourage leisure walking.”

Comment 29: Line 8, page 18: Indicate that the two potential confounders (car ownership and residential self-section) could not be account for because they were not measured in the CHMS. Another limitation is that the self-reported measure of utilitarian walking was not context (or neighbourhood) specific. Level of exposure to the neighbourhood environment was not taken into account in the analysis.

Yes, we completely agree and we would go a bit further to say that our accelerometer-based measure of steps/day was also not context specific (We are actually investigating this using integrated GPS-accelerometer technology in a current study in Montreal). This limitation is what we were trying to express in the discussion when we stated that:

"…we do not know the degree to which participants were exposed to their neighborhoods. Neighborhoods can only influence walking if people spend time in their neighborhoods. Although studies on the association between neighborhoods and location-based physical activity are emerging more studies are needed."

Perhaps this was not stated as clearly as it could have been. We have edited the discussion to express this concept more clearly. We now state:

"…our measures of walking were not context-specific. We do not know how much of the reported utilitarian walking and the accumulated number of steps occurred in the home neighbourhood. Studies on the association between neighborhood walkability and neighbourhood-based physical activity are emerging, but these do require a high respondent burden and are not generally feasible for national-scale studies like that presented here."

REVIEWER TWO

GENERAL COMMENTS

Comment 1: The introduction is very brief, but generally provides the necessary context for the study. However, in the last paragraph, it seems as though China is missing from the list of countries mentioned for past studies in this area (see reference #14 … unless #14 was not supposed to be included in the previous sentence, because 9-14 is indeed six studies, not four, so please clarify).

Thank you for your comment. We have expanded our introduction to more thoroughly discuss the previous work that has been done in this area and to provide a stronger rationale for our study.

Also thank you for pointing out the inconsistency in our references. In our original manuscript we stated that:

"Four previous studies have compared the total daily step counts of adults living in low and high walkable neighborhoods based on indices derived using Geographic Information Systems (GIS).31-36 These studies were conducted in Belgium, the Czech Republic, and Japan."

Indeed, there were only four studies. The Scottish and Chinese studies should not have been included in the reference list. Further, there may have been an apparent contradiction with only three countries being named, but this was because two studies were conducted in Belgium. These sentences have been deleted in the revised manuscript as a result of the other edits that we have made.

Comment 2: The authors mention that comparable studies of objective walkability and step counts have not been conducted in North America. In a literal sense, this MIGHT be true, but there are plenty of studies that have used accelerometers if not step counters (pedometers) in relation to GIS metrics of walkability. Moreover, even if there weren't many U.S. or Canadian studies on walkability and steps, why is this an important gap to fill? (i.e., describe what is different about North America that necessitates specific studies here). In general, I think the rationale for the study needs to be improved to convince the reader that the ensuing methods are worth continuing forward into.

Thank you for drawing our attention to this. We recognize that by focusing on a very specific subset of studies (i.e., those that focused exclusively on biosensor-assessed steps/day and GIS-derived measures of street connectivity, land use mix and population density), we were not giving credit to the other important work that has been done using other objectively-assessed metrics of physical activity

(e.g., MVPA). We have edited our introduction to acknowledge the important work that has been done using other accelerometer-assessed measures of physical activity in relation to GIS metrics of walkability. We also provide a more detailed explanation for why conducting this study was important and what it adds to the literature over and above what has already been done (both abroad and in the Canadian context).

Comment 3: For the description of the CHMS sampling strategy, there is not much detail about actual participant sampling or recruitment, only that the CHMS collects data on a (supposedly) representative sample of Canadians and that there were 15 data collection sites. Can you provide a bit more information about how PEOPLE in the study are selected and recruited, not just the sites where they attend data collection? (and, if I understand correctly, I'm a little amazed that the entirety of Canada is sampled down to only 15 sites from which those respondents are garnered – if I am mistaken, please clarify).

Thank you for your comment. We have incorporated more information into our manuscript about the sampling and recruitment strategy of the CHMS.

To answer your question regarding the number of sites, yes, there were only 15 sites from which respondents were garnered: The CHMS employed a multistage sampling strategy collecting data from 15 sites from five regions across Canada, including British Columbia (including the Yukon), the Prairies (Alberta, Manitoba, Saskatchewan and the Northwest Territories), Ontario, Québec and the Atlantic provinces (Newfoundland and Labrador, Prince Edward Island, Nova Scotia and New Brunswick). The number of data collection sites within each region was proportional to the size of the population, with two sites in British Columbia, two sites in the Prairies, six sites in Ontario, four sites in Québec and one site in the Atlantic provinces. Using the 2006 Census, all households within the 15 data collection regions were stratified into one of five age groups (6 to 11, 12 to 19, 20 to 39, 40 to 59, 60 to 79) using the respondents' age at the time of the census. This ensured that the dwellings in each stratum had a high probability of having at least one occupant in the desired age range. From these strata, a simple random sample of households was selected. Of the 8,772 households that were contacted and asked to provide information on current household composition, 6,106 complied (69.6%). Using the household composition lists obtained from these households, 7,483 individuals were requested to participate in the CHMS. Of these, 6,604 completed the household questionnaire (88.3%) and of those that completed the household questionnaire, 5,604 attended the Medical Examination Centre (MEC) (84.9%) – a high response rate likely attributable to the CHMS being a government survey and the practices implemented by Statistics Canada to maximize participation.37

Comment 4: I am familiar with the measure of walkability the authors are employing (three constructs summed using z-scores). However, I think they could provide some more rationale as to why this index and its components are appropriate for determining walkability (and not other variables instead or in addition to). Also, they could at least provide some citations for the composite measure.

Thank you for your comment. We agree that we did not, in our original submission, provide a strong rationale for focusing exclusively on land use mix, street connectivity, and population density. In our revised manuscript we address our choice of measures directly (Please see our response to Comment 14 of Reviewer One). We also provide citations for other studies in which a composite index based on these three measures has been used.

Comment 5: Couldn't the authors get a Walk Score based on the anonymous spatial coordinates instead of having to resort to postal codes? If not, I think some discussion of the size of postal codes and the appropriateness of linking individual physical activity data to larger postal codes is warranted. Also, when you say Walk Score is a validated measure, please provide a citation (e.g., Duncan, etc.).

Thank you for you for raising our attention us to the fact that we did not provide a citation to support our statement that the Walk Score is a validated measure. We do so now in the revised manuscript.

We had access to longitude and latitude data but these were derived based the centroid of the postal codes (using a postal code conversion file). We did not have the xy coordinates that corresponded to the street addresses of the participants (due to confidentiality policies by Statistics Canada) For more information on this, please see our response to Reviewer One (Comment 13).

We do agree with you there may be some measurement bias when using postal code centroids as opposed to street address longitude and latitude. However, in the Canadian context, postal codes are accurate proxies for home addresses: 87.9 and 96.5% of Canadian postal codes fall within 200 and 500 meters of the true home address, respectively.38 Given that neighbourhood walkability is not expected to vary dramatically within a 200 to 500 meter limit, use of postal codes in our study was appropriate. We now insert the reference regarding the validity of using postal codes into the body of our manuscript:

“In the Canadian context, postal codes are accurate proxies for home addresses with 87.9 and 96.5% of postal codes falling within 200 and 500 meters of the street address, respectively.”

Comment 6: Some readers might wonder why step counts were used as the outcome variable when accelerometers were used for data collection and more precise measures of activity time and intensity were presumably available. I think some discussion of the motivation and justification for this decision would be helpful.

We agree that some readers will wonder this. Thank you for suggestion of providing justification for our decision to focus solely on step counts. There are several reasons why we chose to use daily steps as opposed to other accelerometer-assessed metrics. We have inserted a discussion about this into our introduction. We now state that:

“Daily steps are an outcome of particular interest for the study of neighbourhood walkability for several reasons. First, daily steps are a good estimate of total habitual physical activity, something that is not captured by other biosensor-assessed metrics, such as MVPA. Second, walking is the most common and preferred form of physical activity among adults. Understanding if neighbourhood walkability is associated with daily step among adults, would suggest that neighbourhood-level interventions may have the potential to impart a benefit to large segments of the population. Third, daily steps are highly interpretable.21 by both the scientific and lay community. This is unlike other biosensor-assessed metrics (e.g., accelerometer counts) that may also be good a capturing total levels of physical activity, but are of less value when trying to explain the association of walkability with physical activity in an easily interpretable and relevant way. For example, saying that living in a high compared to a low walkable neighbourhood is associated with x more steps/day is more interpretable and readily understood by the public than saying that it is associated with x more activity counts.”

For this paper we have decided to restrict our analyses solely to the association of neighbourhood walkability with walking as captured by steps taken per day. While this is not something that we are currently planning to do, we agree that using other accelerometer-assessed metrics of physical activity would make an excellent additional paper.

Comment 7: Is there any reliability or validity information for the utilitarian walking question? Estimating the typical week over the past three months might be challenging. As well, the question would seem to exclude walking for other utilitarian purposes (e.g., to visit someone, to go to a restaurant); I believe such questions are usually phrased something like “to get to and from places”.

This raises further questions about validity, so I would suggest the authors add any psychometric information if possible.

Thank you for your comment. This measure of utilitarian walking has been used in other previous Canadian health surveys including the Aboriginal Health Survey and the National Population Health Survey (we now note this in our methods section). To our knowledge, no validation studies have been conducted on this measure of utilitarian walking. Although we cannot make definite claims about validity due to a lack of studies, we tested the construct validity of this measure by comparing it to daily steps. As expected, we demonstrated that higher self-reported utilitarian walking was associated with more daily steps (<1 hours/week: 6613 steps/day, 95% CI 6251, 6975; 1 to 5 hours/week: 6768 steps/day, 95% CI 6420, 7117; ≥6 hours/week: 7391 steps/day, 95% CI 6972, 7811), suggesting that this measure indeed captured higher levels of walking.

This was the only question available in Cycle 1 of the CHMS that we could use to estimate utilitarian walking. The International Physical Activity Questionnaire has more specific questions regarding utilitarian walking and has been extensively validated in adult populations, but it was not incorporated into the questionnaire until Cycle 3 (2012-2013). On its own, we would agree that our measure of utilitarian walking might have excluded walking for other utilitarian purposes (e.g., while doing chores) and could carry more measurement bias than alternatively phrased self-reported measures. However, the potential for additional measurement bias may have been minimized in the CHMS by the fact that, prior to answering this question, participants were prompted with the following statement:

"Next, some questions about the amount of time spent in the past 3 months on physical activity at work, while doing daily chores around the house, or doing errands, but not leisure time activity."
CHMS Household Questionnaire, Cycle 1 2007-2009

This may encouraged participants to remember other types of utilitarian walking that are captured by the "work, school, or errand" categories (e.g., those around the house).

Comment 8: In the covariates section, the categories for perceived health seem to be out of order.

Yes, they are. Thank you for catching this. We have made the necessary change. The sentence now reads:

"…and perceived health (poor, fair, good, very good, excellent) were assessed as part of the computer-assisted interview."

Comment 9: Was the steps/day variable fairly (with respect to its distribution) and therefore appropriate for linear rather than logistic regression? (often, PA variables are not very normal).

Thank you for your comment. The normality assumption in linear regression applies to the residuals of the models not to the distribution of the y variable. It is possible for the y variable to have a skewed distribution but for the residuals to be normal.

We did check the normality of the residuals and found them to be approximately normal. Usually we would provide reviewers with a plot of the residuals, however, release of this data is not permitted by Statistics Canada. Although the residuals were nearly normal, they were not perfectly so. To ensure that our choice of linear regression was appropriate, we conducted a few additional analyses. These are described below. (Please note: We chose to report the results of the linear regression models in our manuscript as we favored the interpretation of these effect estimates and the results were the same across all of the test methods.)

1. We conducted sensitivity analyses to see if removing the 5% of participants (n=135/2949) who had outlying steps/day (i.e., more than 15,000 steps/day) altered our findings. Deleting outliers improved the normality of the residuals, but the results of the regression models were the same when all of the data were retained.

2. We ran logistic regression models between steps and walkability (for odds of achieving ≥5,000 steps across quartiles of walkability). These results were the same as when keeping steps/day in their continuous form.

3. We conducted robust linear regression (mm-estimation). This model does not make assumptions regarding normality. The results of these analyses were consistent with those of our regular linear regression analyses.

Comment 10: When the authors reference walkability quartiles from previous studies (page 9, line 25), can they provide some citations for studies that have used this approach? (e.g., Frank et al., 2005, etc.).

Please see our response to Comment 24 (Reviewer 1). We have changed the line:

“Associations were estimated across quartiles of walkability, to maximize comparability with previous studies.”

to

“Associations were estimated across quartiles of walkability.”

As we explain in our response to Comment 24, the references are no longer needed.

Comment 11: The descriptive statistics describe a sample of 5,605 individuals whittled down to 2,949 participants with complete data (which is a big drop, but tolerable). However, earlier in the paper, it says something about there being 9,351 neighborhoods for which GIS-based walkability was calculated. This was confusing.

Thank you for drawing our attention to this. To avoid confusion, we have removed our reference to “9,351 anonymous spatial coordinates”.

The reason for the 3,747 additional coordinates (9,351 - 5,604) was that at the time that we requested this data, Statistics Canada would not allow researchers to have access to spatial information as they thought that this could be a breach of confidentiality. The only way that they would allow us to get access to this data was to give us a separate xy coordinate file with 9351 observations (only 5,604 of these corresponded to the addresses of the CHMS participants). We then had to derive our built environment variables and send the full file back to Statistics Canada. Statistics Canada then merged our built environment file with the individual-level data collected by the CHMS, removed all of the spatial data, and sent the file to the McGill-Concordia Research Data Centre so that there would be no way that the location of any of the CHMS participants could be traced.

Comment 12: Why is the font tiny on pages 13 and 14?

The proof that we received from BMJ had 12-point Times New Romans font on pages 13 and 14, so this must have occurred as part of the editorial-reviewer uploading process. Unless you are referring to the font in the tables? In this case we did use 11 point for some of the subheadings. We will updated the tables to include only 12-point font, but if this is another technical issue, please advise.

We would be happy to send the document in another format.

Comment 13: In the results section about utilitarian walking (page 13), please state that these results are for the highest walkability and Walk Score quartiles in relation to the LOWEST walkability quartile (like you do in the previous section on daily step counts).

Thank you for drawing our attention to this inconsistency. We now make sure that in the “Utilitarian Walking” section we note that the results for the highest quartiles are in relation to the lowest quartiles.

References

1. Learnihan V, Van Niel KP, Giles-Corti B, Knuiman M. Effect of Scale on the Links between Walking and Urban Design. Geographical Research. 2011;49(2):183-191.
2. Oliver LN, Schuurman N, Hall AW. Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. Int J Health Geogr. 2007;6:41.
3. World Health Organization. Global Health Observatory: Prevalence of Insufficient Physical Activity. 2015; http://www.who.int/gho/ncd/risk_factors/physical_activity_text/en/index.html. Accessed February 17, 2015.
4. McCormack GR, Shiell A, Doyle-Baker PK, Friedenreich CM, Sandalack BA. Subpopulation differences in the association between neighborhood urban form and neighborhood-based physical activity. Health Place. 2014;28:109-115.
5. Riley DL, Mark AE, Kristjansson E, Sawada MC, Reid RD. Neighbourhood walkability and physical activity among family members of people with heart disease who participated in a randomized controlled trial of a behavioural risk reduction intervention. Health Place. 2013;21:148-155.
6. Prince SA, Kristjansson EA, Russell K, et al. Relationships between neighborhoods, physical activity, and obesity: a multilevel analysis of a large Canadian city. Obesity (Silver Spring). 2012;20(10):2093-2100.
7. Ding D, Sallis JF, Norman GJ, et al. Neighborhood environment and physical activity among older adults: do the relationships differ by driving status? J Aging Phys Act. 2014;22(3):421-431.
8. Saelens BE, Sallis JF, Black JB, Chen D. Neighborhood-based differences in physical activity: an environment scale evaluation. Am J Public Health. 2003;93(9):1552-1558.
9. Oakes JM, Forsyth A, Schmitz KH. The effects of neighborhood density and street connectivity on walking behavior: the Twin Cities walking study. Epidemiol Perspect Innov. 2007;4:16.
10. De Meester F, Van Dyck D, De Bourdeaudhuij I, Deforche B, Sallis JF, Cardon G. Active living neighborhoods: is neighborhood walkability a key element for Belgian adolescents? BMC Public Health. 2012;12:7.
11. Schmidt MD, Cleland VJ, Thomson RJ, Dwyer T, Venn AJ. A comparison of subjective and objective measures of physical activity and fitness in identifying associations with cardiometabolic risk factors. Ann Epidemiol. 2008;18(5):378-386.
12. Lim S, Wyker B, Bartley K, Eisenhower D. Measurement error of self-reported physical activity levels in New York City: assessment and correction. Am J Epidemiol. 2015;181(9):648-655.
13. Tudor-Locke C, Hatano Y, Pangrazi RP, Kang M. Revisiting "how many steps are enough?". Med Sci Sports Exerc. 2008;40(7 Suppl):S537-543.
14. Tudor-Locke CE, Bell RC, Myers AM, Harris SB, Lauzon N, Rodger NW. Pedometer-determined ambulatory activity in individuals with type 2 diabetes. Diabetes Res Clin Pract. 2002;55(3):191-199.
15. Carr LJ, Dunsiger SI, Marcus BH. Validation of Walk Score for estimating access to walkable amenities. Br J Sports Med. 2011;45(14):1144-1148.
16. Carr LJ, Dunsiger SI, Marcus BH. Walk score as a global estimate of neighborhood walkability. Am J Prev Med. 2010;39(5):460-463.
17. Duncan DT. What's your Walk Score(R)?: Web-based neighborhood walkability assessment for

health promotion and disease prevention. Am J Prev Med. 2013;45(2):244-245.
18. Carr LJ, Dunsiger SI, Marcus BH. Walk Score™ as a global estimate of neighborhood walkability. American journal of preventive medicine. 2010;39(5):460-463.
19. Leslie E, Coffee N, Frank L, Owen N, Bauman A, Hugo G. Walkability of local communities: using geographic information systems to objectively assess relevant environmental attributes. Health Place. 2007;13(1):111-122.
20. Duncan DT, Aldstadt J, Whalen J, Melly SJ, Gortmaker SL. Validation of walk score for estimating neighborhood walkability: an analysis of four US metropolitan areas. Int J Environ Res Public Health. 2011;8(11):4160-4179.
21. Tudor-Locke C, Bassett DR, Jr. How many steps/day are enough? Preliminary pedometer indices for public health. Sports Med. 2004;34(1):1-8.
22. Government of Canada. Aboriginal Peoples Survey, 2006: An overview of the health of the Métis population. In: Statistics Canada, ed2006.
23. Government of Canada. National Population Health Survey: Household Component - Logintudinal Cycle 9 (2010-2011) - Questionnaire. In: Statistics Canada., ed. Ottawa, 2009.
24. Ding D, Sallis JF, Norman GJ, et al. Neighborhood environment and physical activity among older adults: do the relationships differ by driving status? J Aging Phys Act. 2014;22(3):421-431.
25. Zander A, Rissel C, Rogers K, Bauman A. Active travel to work in NSW: trends over time and the effect of social advantage. Health Promot J Austr. 2014;25(3):167-173.
26. Frank LD, Saelens BE, Powell KE, Chapman JE. Stepping towards causation: do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Soc Sci Med. 2007;65(9):1898-1914.
27. McCormack GR, Shiell A. In search of causality: a systematic review of the relationship between the built environment and physical activity among adults. Int J Behav Nutr Phys Act. 2011;8:125.
28. Boone-Heinonen J, Gordon-Larsen P, Guilkey DK, Jacobs DR, Jr., Popkin BM. Environment and Physical Activity Dynamics: The Role of Residential Self-selection. Psychology of Sport and Exercise. 2011;12(1):54-60.
29. McMahon T. Who belongs to Canada's middle class? 2014; http://www.macleans.ca/economy/who-belongs-to-canadas-middle-class/. Accessed August 19, 2015.
30. Wasfi RA, Ross NA, El-Geneidy AM. Achieving recommended daily physical activity levels through commuting by public transportation: unpacking individual and contextual influences. Health Place. 2013;23:18-25.
31. Van Dyck D, Deforche B, Cardon G, De Bourdeaudhuij I. Neighbourhood walkability and its particular importance for adults with a preference for passive transport. Health Place. 2009;15(2):496-504.
32. Van Dyck D, Cardon G, Deforche B, De Bourdeaudhuij I. Urban-rural differences in physical activity in Belgian adults and the importance of psychosocial factors. J Urban Health. 2011;88(1):154-167.
33. Dygryn J, Mitas J, Stelzer J. The influence of built environment on walkability using geographic information system Journal of Human Kinetics 2010;24:93-99.
34. Kondo K, Lee JS, Kawakubo K, et al. Association between daily physical activity and neighborhood environments. Environ Health Prev Med. 2009;14(3):196-206.
35. Robertson LB, Ward Thompson C, Aspinall P, Millington C, McAdam C, Mutrie N. The influence of the local neighbourhood environment on walking levels during the Walking for Wellbeing in the West pedometer-based community intervention. J Environ Public Health. 2012;2012:974786.
36. Zhang Y, Liu ND, Liu X. Relationship Between Built Environment, Physical Activity, Adiposity and Health in Adults Aged 46-80 in Shanghai, China. J Phys Act Health. 2014.
37. Statistics Canada. Canadian Health Measures Survey (CHMS) Data User Guide: Cycle 1. Ottawa, ON: Statistics Canada; 2011.
38. Bow CJ, Waters NM, Faris PD, et al. Accuracy of city postal code coordinates as a proxy for location of residence. Int J Health Geogr. 2004;3(1):5.

VERSION 2 – REVIEW

REVIEWER	Gavin McCormack University of Calgary, Canada
REVIEW RETURNED	21-Sep-2015

GENERAL COMMENTS	The reviewer completed the checklist but made no further comments.