

## 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

<b>TITLE (PROVISIONAL)</b>	The association of exercise and sedentary behaviors with incident end stage renal disease in the Southern Community Cohort Study
<b>AUTHORS</b>	Pike, Mindy; Taylor, Jacob; Kabagambe, Edmond; Stewart, Thomas; Robinson-Cohen, Cassianne; Morse, Jennifer; Akwo, Elvis; Abdel-Kader, Khaled; Siew, Edward; Blot, William; Ikizler, T. Alp; Lipworth, Loren

## VERSION 1 – REVIEW

<b>REVIEWER</b>	Ryoma Michishita Laboratory of Exercise Physiology, Faculty of Health and Sports Science, Fukuoka University, Fukuoka, Japan
<b>REVIEW RETURNED</b>	19-Apr-2019

<b>GENERAL COMMENTS</b>	<p>This study is to investing the association of physical activity and sedentary behavior with the incidence of ESRD. This study has a large number of study subjects, and the association of physical activity and sedentary behavior with the incidence of ESRD is clearly shown. However, there are several questions in this study.</p> <ol style="list-style-type: none"><li>1. Reviewer considered that physical activity and sedentary to be the most important outcomes in this study. I would like you to clarify whether physical activity or sedentary behavior is more important from the results of this study. In addition, the authors should explain in more detail the description of physical activity and sedentary behavior measurements.</li><li>2. Do the authors do a sub-analysis on the intensity of physical activity? Which intensity of physical activity impacts the incidence of ESRD? (low or moderate or vigorous intensity?). Similarly, what kind of sedentary behaviors impacts the incidence of ESRD? (Car or work or watching TV, home computer?).</li><li>3. In this study, the incidence of ESRD is determined from the value of eGFR. Reviewer considered that the proteinuria in addition to eGFR value to be an important factor in the assessment of kidney function. Does the author also evaluate the association of physical activity or sedentary behavior with proteinuria? In relation to the above, the definition of ESRD is ambiguous in this study. The author should explain in more detail the description of ESRD in this study. Additionally, the authors also oversimplify the method of eGFR calculation.</li><li>4. This study is generally poor information on the subject's background. All the subjects of this study were healthy subjects?</li></ol>
-------------------------	---

	<p>Was the taking medications, dialysis treatment and other complications such as cardiovascular disease and stroke included in subjects of this study? Authors needs to describe the presence or absence of taking medications and other complications.</p> <p>5. Do you need a flow chart of the subjects as Fig 1?</p>
--	---

<b>REVIEWER</b>	Professor Tazeen Jafar Duke-NUS Medical School
<b>REVIEW RETURNED</b>	16-Jun-2019

<b>GENERAL COMMENTS</b>	<p>Pike et al report a case-control study embedded within a cohort of low-income community clinic patients. In general, the study is well written, and the conclusions are supported by the results. I have the following concerns that can be addressed:</p> <p>Main concerns:</p> <ol style="list-style-type: none"> <li>1. The case-control design is “retrospective” –albeit the cohort study within which it was embedded was prospective.</li> <li>2. The authors examined interactions between sedentary time or physical activity and baseline kidney function on ESRD. Please indicate clearly how the interaction term was constructed. Was nonlinear term included in the interaction?</li> <li>3. The study population was sampled from individuals who donated blood and about 46% of the whole cohort (n=86000) donated blood samples. This raises concerns re generalizability of the findings to the whole cohort which must be stated.</li> <li>4. Please provide additional information for a comparison between the selected and unselected population in terms of the risk profiles.</li> </ol> <p>Abstract:</p> <ul style="list-style-type: none"> <li>• In the abstract please clarify why the population is at high risk for ESKD</li> </ul> <p>Methods:</p> <ul style="list-style-type: none"> <li>• Page 5, 3rd paragraph: please specify your type of probability sampling method used for subcohort (e.g. simple random or stratified random). Is 13% the sampling fraction? How did you choose this fraction?</li> </ul> <p>Also, citation No. 15 is an unpublished paper. I would suggest briefly reporting the results from this paper as supplement materials</p> <ul style="list-style-type: none"> <li>• Page 6, 1st paragraph: define “overweight or obese”</li> <li>• Page 6, 1st paragraph: Data on most covariates were collected via self reports, such as height, weight, hypertension, diabetes, and high cholesterol. Self-reported data likely bias the results. The study population had low SES and were at high risk of CKD or ESRD; thus, it is possible that chronic conditions such as hypertension or diabetes could be underdiagnosed and underreported due to limited access to health care. Did you evaluate the accuracy of these self-reported data? If not state this as a limitation.</li> <li>• Data analysis methods: <ul style="list-style-type: none"> <li><input type="checkbox"/> Please specify which variance estimate method was used in the Cox regression?. Did you test the proportional hazards assumption for the covariates?</li> <li><input type="checkbox"/> In the COX model, the authors added continuous variables as restricted cubic splines with four knots including total sedentary time, physical activity, age, eGFR, and BMI. But it was not</li> </ul> </li> </ul>
-------------------------	---

	reported which continuous variable has a significant nonlinear relationship with the outcome. Please provide P value for tests of the nonlinear relationship. If the p-value is insignificant, then the continuous variable should be modeled as a linear term.
--	---

### VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

This study is to investigating the association of physical activity and sedentary behavior with the incidence of ESRD. This study has a large number of study subjects, and the association of physical activity and sedentary behavior with the incidence of ESRD is clearly shown. However, there are several questions in this study.

1. Reviewer considered that physical activity and sedentary to be the most important outcomes in this study. I would like you to clarify whether physical activity or sedentary behavior is more important from the results of this study. In addition, the authors should explain in more detail the description of physical activity and sedentary behavior measurements.

Reply:

We would like to thank the reviewer for bringing up this point. The two primary exposures in our study were total sitting and total physical activity, which were modeled simultaneously as independent exposures. In our interpretation of our results, we prefer not to make statements as to which is a more “important” exposure. As presented, in our study population, our results show that among individuals with preserved kidney function, higher physical activity is associated with lower risk of ESRD. We also observed that high levels of sitting time were associated with increased kidney disease risk among those with advanced kidney disease, but due to attenuation of this finding after excluding the first two years of follow up, we postulate that the observed positive association is likely due to reverse causation.

Regarding the measurement of physical activity and sedentary behavior, we thank the reviewer for pointing this out and have now provided a more detailed description of the methods. The expanded paragraph on measurement of physical activity and sedentary behaviors on page 6-7 of the Methods section now reads:

“Usual sedentary and active behaviors were assessed using a validated physical activity questionnaire (PAQ) developed specifically for the SCCS [18]. For sedentary behaviors, participants were asked questions about the amount of time per day typically spent sitting in a car or bus, at work, viewing television or movies, and other activities that involve sitting such as sitting at meals, talking on the phone, reading, playing games, or sewing. For physical activity, participants were asked about time typically spent performing light, moderate, and strenuous activities at home and at work, as well as time spent doing moderate and vigorous exercise/sports. Time spent doing work and home activities was assessed separately for week and weekend days, and exercise and sports participation was assessed for a typical week. Examples of light work were given to participants and included standing at work, shopping, cooking, and child or elderly care. Moderate work examples included shop work, cleaning house, gardening, mowing lawn, and home repair. Examples of strenuous work included loading or unloading trucks, construction, farming, or other hard labor. Moderate sports included activities such as bowling, dancing, and golfing, while vigorous sports included jogging, aerobics, tennis, swimming, and weight lifting. For all questions, participants provided open-ended duration responses (hours and minutes). The reliability and validity of the SCCS physical activity questionnaire was evaluated in 118 randomly selected SCCS participants via use of accelerometers [18].”

2. Do the authors do a sub-analysis on the intensity of physical activity? Which intensity of physical activity impacts the incidence of ESRD? (low or moderate or vigorous intensity?). Similarly, what kind of sedentary behaviors impacts the incidence of ESRD? (Car or work or watching TV, home computer?).

Reply:

In our study, physical activity was transformed from hours/day into summary measures of energy expenditure, defined as standard metabolic equivalent (MET)-hours/day. By using this summary measure, multiples of this value (METS) are used to index the intensity of the specific activities. Thus, through the use of MET-hours/day, we are capturing intensity of physical activity as well as amount.

To clarify this further, we have now provided a more detailed description on page 8 of the Methods section:

“Sedentary time was calculated as hours/day based on the sum of all individual sedentary behaviors. Total physical activity was calculated as the sum of light, moderate and strenuous household/occupational work as well as moderate and vigorous sports; values were transformed from hours/day into summary measures of energy expenditure, defined as metabolic equivalent (MET)-hours/day. MET values for specific activities and intensities were based on the Compendium of Physical Activities [19]. MET-hours reflect the weighted average of the intensity (MET) and duration (hours) of activity behaviors. Two MET-hours/day is roughly equivalent to participating in 1 hour of a light activity, 0.5 hours of a moderate activity such as walking, or 0.25 hours of a vigorous activity such as jogging [18]. For example, one MET-hour is roughly equivalent to the energy expenditure associated with walking very briskly (4 METS) for 15 minutes (0.25 hours).”

Sedentary behavior was analyzed as total sitting, or the sum of all individual sedentary behaviors in hours/day. In response to your question, individual sedentary behaviors were also examined in relation to ESRD. We modeled the data using a multivariable Cox model in which sitting time was separated into sitting in the car/bus, sitting at work, watching TV/movies, and other sitting, all in hours/day. When compared to the model with total sitting hours, the likelihood ratio test for non-nested models was non-significant ( $p=0.9835$ ); thus, we cannot conclude that the expanded model, with sitting hours separated by type, fits differently.

We have now included a brief description of the additional analysis on page 9 of the Methods section:

“To examine if the relationship with ESRD differed for different types of sitting, we also modeled the individual sedentary behaviors, sitting in the car/bus, sitting at work, watching TV/movies, and other sitting. The multivariable Cox model included sitting hours for each category modeled as restricted cubic splines and mutually adjusted. Non-nested likelihood ratio tests were used to compare this model to the Cox model including total sitting hours.”

We have also included the results from the analysis on page 12 of the Results section and page 15 of the Discussion section:

“In analyses examining the individual types of sitting, the non-nested likelihood ratio test indicated that the model with sitting hours by type did not significantly differ from the model with total sitting hours ( $p=0.98$ )...”

“Additionally, we observed that the model separating sitting time by type did not fit better than the model with total sitting time.”

3. In this study, the incidence of ESRD is determined from the value of eGFR. Reviewer considered that the proteinuria in addition to eGFR value to be an important factor in the assessment of kidney function. Does the author also evaluate the association of physical activity or sedentary behavior with proteinuria?

In relation to the above, the definition of ESRD is ambiguous in this study. The author should explain in more detail the description of ESRD in this study. Additionally, the authors also oversimplify the method of eGFR calculation.

Reply:

We apologize for any confusion. Incident ESRD in this study was not determined from the value of eGFR. Rather, incident ESRD was identified by linking the SCCS cohort with the US Renal Data System. ESRD cases in this population-based registry are certified by a physician diagnosis and filed using a medical evidence report form (to the Medicare ESRD program), or when chronic dialysis or kidney transplant occurs, irrespective of the glomerular filtration rate. The USRDS is a national registry and therefore, ascertainment of ESRD cases is virtually complete.

To clarify the ascertainment of ESRD in this study, we have added the following sentences in the Methods section on page 5:

“ESRD cases in this registry are certified by a physician diagnosis and filed using a medical evidence report form (to the Medicare ESRD program), or when chronic dialysis or kidney transplant occurs, irrespective of the glomerular filtration rate. The USRDS is a national registry and therefore, ascertainment of ESRD cases is virtually complete [1].”

Unfortunately, data on proteinuria were not available for the study population. We have added this as a limitation on page 16 in the Discussion section, as follows:

“Finally, baseline data on proteinuria were not available.”

Regarding eGFR calculation, eGFR was calculated using the creatinine-based CKD-EPI equation, which is described in reference 20 (see Levey et al *Annals of Internal Medicine* 2009). We have clarified that the CKD-EPI equation is creatinine-based, and added the following sentences regarding the measurement of creatinine in the Methods section on page 5:

“Baseline serum levels of creatinine were measured using the Jaffe (Rate) method on a Beckman Coulter DXC 600 clinical chemistry analyzer. The creatinine assays were calibrated, and daily quality checks performed at three levels before sample testing.”

4. This study is generally poor information on the subject's background. All the subjects of this study were healthy subjects? Was the taking medications, dialysis treatment and other complications such as cardiovascular disease and stroke included in subjects of this study? Authors needs to describe the presence or absence of taking medications and other complications.

Reply:

We thank the reviewer for this comment and have expanded and clarified the description of the cohort participants. Participants in the SCCS were primarily recruited at participating community health centers, which provide primary healthcare for under-insured populations. All subjects were given standardized questionnaires at enrollment. These questionnaires obtained information on demographic, medical, and lifestyle variables, which included information on whether participants were taking medications or had been diagnosed with comorbidities. Additionally, for the current analyses, participants with prevalent ESRD at the time of cohort enrollment were excluded from the study, and history of diagnosis of diabetes, hypertension, and hypercholesterolemia were adjusted for in the Cox models. Previous myocardial infarction and stroke were not identified as covariates a priori, but in response to your question, we included MI and stroke to the models and no difference in results was seen.

We have now expanded the paragraph on the study population in the Methods section on page 4 and 5 to clarify the cohort description:

“The SCCS is a prospective cohort study that recruited ~86,000 primarily low-income black and white adults, aged 40-79 years, in the southeastern US (2002-2009) [12]. Participants eligible for enrollment spoke English and had not been treated for cancer in the 12 months before enrollment. The majority (86%) were recruited at participating community health centers (CHC), which provide primary healthcare for under-insured populations. A detailed description of SCCS methods has been published (<http://www.southerncommunitystudy.org>) [13]. All participants provided written informed consent, and the study was approved by the Institutional Review Boards of Vanderbilt University Medical Center and Meharry Medical College. We used the STROBE cohort checklist when writing our report [14].”

We have also added to the paragraph on data collection in the Methods section on page 6:

“Standardized computer-assisted personal interviews were administered at enrollment to obtain data on demographic, medical, and lifestyle variables [13]. Sections included demographic characteristics (education, income, residence), tobacco use, personal and family medical history, medication use, emotional well-being, occupation, physical activity, and diet. Body mass index (BMI) was calculated from self-reported height and weight. History of hypertension, diabetes, and hypercholesterolemia as well as stroke and cardiovascular disease were self-reported by asking whether a doctor had ever diagnosed the participant with the condition. Self-reported height and weight were compared with clinic recorded measurements for over 20% of participants. In a series of validation studies, biomarkers, repeat interviews, or medical records were used to assess the reliability of variables such as smoking status and self-reported diseases including diabetes [13].”

As indicated on page 5, participants with prevalent ESRD were excluded from the analysis.

“Participants with an ESRD diagnosis prior to SCCS enrollment (prevalent cases) were excluded from the analysis.”

5. Do you need a flow chart of the subjects as Fig 1?

Reply:

As suggested, we have now added a flow chart of the participants as Figure 1 in the manuscript titled “Figure 1. Study selection of the SCCS case-cohort.”

Reviewer: 2

Pike et al report a case-control study embedded within a cohort of low-income community clinic patients. In general, the study is well written, and the conclusions are supported by the results. I have the following concerns that can be addressed:

Main concerns:

1. The case-control design is “retrospective” –albeit the cohort study within which it was embedded was prospective.

Reply:

We apologize for any confusion. This study was not a case-control design, but a case-cohort design, which is a prospective design. Detailed methods and more information can be found in references 15 and 16 (see Therneau et al Lifetime Data Analysis 1999 and Sharp et al PLoS ONE 2014), which have been added to the manuscript in the Methods section. The case-cohort study design is an efficient option in assessing time-to-event data in cohorts where ascertainment of information for all members of the cohort may be cost-prohibitive or problematic for other reasons. The case-cohort strategy is to generate an analysis subset by combining a stratified or probability sample of the larger cohort with all the participants who experience the outcome event of interest, in this case ESRD.

2. The authors examined interactions between sedentary time or physical activity and baseline kidney function on ESRD. Please indicate clearly how the interaction term was constructed. Was nonlinear term included in the interaction?

Reply:

We thank the reviewer for this comment. The nonlinear terms for physical activity, sedentary time, and eGFR were used to construct the interaction terms. The first interaction term used physical activity modeled as restricted cubic splines with four knots and eGFR modeled as restricted cubic splines with four knots. The second interaction term used sedentary time modeled as restricted cubic splines with four knots and eGFR modeled as restricted cubic splines with four knots.

To clarify the construction of the interaction terms, we have now added to the sentence on the interaction term on page 9 of the Methods section:

“To examine interactions between sedentary time or physical activity and baseline kidney function on ESRD risk, multiplicative interaction terms between the nonlinear, continuous predictors of sedentary time/physical activity and nonlinear, continuous eGFR were added to the model.”

3. The study population was sampled from individuals who donated blood and about 46% of the whole cohort (n=86000) donated blood samples. This raises concerns re generalizability of the findings to the whole cohort which must be stated.

Reply:

A total of 37,277 SCCS participants donated blood samples. Baseline characteristics of participants in this sample are comparable to the whole SCCS cohort. For instance, the median age was 51 years and 50 years for the whole SCCS cohort and the subset with blood samples, respectively.

Approximately 22% of those who donated blood and the whole SCCS cohort had diabetes, 45% were current smokers, 32% had an education level <12<sup>th</sup> grade, and 39% had an income greater than \$15,000 per year (see Buchowski et al J Phys Act Health 2012).

We have now added generalizability as a limitation to the Discussion section on page 16:

“Although the probability sample is comparable to the whole cohort, the findings might not be generalizable to all SCCS participants.”

4. Please provide additional information for a comparison between the selected and unselected population in terms of the risk profiles.

Reply:

Consistent with the case-cohort design, rather than choosing individual controls matched to the cases in a 2 to 1 or higher ratio, we have selected a weighted probability sample of cohort participants with available blood specimens to act as controls. The sampling produced a greater than 5-fold number of “controls” (n=4,238 subcohort members) than cases of ESRD. This probability sample constitutes 13% of all SCCS participants with a stored baseline blood sample.

The probability sample of SCCS participants with available blood specimens is comparable to the entire SCCS cohort who donated blood samples, with respect to baseline demographic and other characteristics, including racial distribution, low income and education level and high prevalence of cardiovascular risk factors. In particular, the weighted subcohort included 70.8% blacks and 29.2% whites, which was similar to 67.3% blacks and 28.6% whites in the SCCS target population. In the overall SCCS as well as the subcohort, about 32% had an education level <12<sup>th</sup> grade, 61-62% had an annual income <\$15,000, and 56% and 22% had hypertension and diabetes, respectively.

We have now added to the description of the cohort in the Methods section on pages 5 and 6:

“This sample constitutes 13% of SCCS participants who donated blood, and is comparable with respect to baseline sociodemographic characteristics including racial distribution, low income, and high prevalence of CKD risk factors [17]. The weighted subcohort included 70.8% blacks and 29.2% whites, and the SCCS population included 67.3% blacks and 28.6% whites. In the subcohort and overall SCCS population, about 32% had an education level below 12<sup>th</sup> grade, the majority had an annual income of <\$15,000, and the prevalence of hypertension and diabetes was similar at 56% and 22%, respectively.”

Abstract:

- In the abstract please clarify why the population is at high risk for ESKD

Reply:

We agree that the previous wording was not clear. The SCCS is a cohort of mostly black, low-income individuals with a large prevalence of CKD risk factors including hypertension, diabetes, and obesity, and this is why the population is at higher risk for ESRD. We have now changed the first sentence in the conclusion of the Abstract:

“In a population with a high prevalence of CKD risk factors such as hypertension and diabetes, ...”.

Methods:

• Page 5, 3rd paragraph: please specify your type of probability sampling method used for subcohort (e.g. simple random or stratified random). Is 13% the sampling fraction? How did you choose this fraction?

Also, citation No. 15 is an unpublished paper. I would suggest briefly reporting the results from this paper as supplement materials

Reply:

The subcohort was constructed from several previous nested case-control studies performed within the SCCS for which creatinine had already been measured. We calculated the sampling probability into the subcohort from the entire SCCS population with donated blood. Using the sampling probabilities, we constructed trimmed inverse sampling probability weights. We selected a weighted probability sample from the cohort participants with available blood specimens as outlined above. This weighted probability sample constitutes 13% of all SCCS participants who donated blood sample.

The manuscript titled “Baseline kidney function and racial disparities in end-stage renal disease risk in the Southern Community Cohort Study” by Bock et al. is currently in press and will be published shortly. We will update the citation accordingly once published. Until then, we have confidentially provided for the reviewer Table 1 from the manuscript in this response.

Table 1. Baseline Characteristics of ESRD cases, Weighted Subcohort, and Overall SCCS Population Who Donated Blood at Enrollment, 2002-2009

Characteristic	ESRD Cases N=737	Subcohort N=4,238	SCCS (with stored blood) N=37,277
Age at enrollment, median (25 <sup>th</sup> , 75 <sup>th</sup> percentile), years	53 (47, 59)	50 (45, 58)	50 (45, 57)
eGFR, median (25 <sup>th</sup> , 75 <sup>th</sup> percentile), ml/min/1.73m <sup>2</sup>	63.3 (36.0, 98.2)	103.2 (86.0, 117.9)	
eGFR, %, ml/min/1.73m <sup>2</sup>			
≤30	19.9	1.0	
31-60	26.0	4.8	
61-90	23.5	24.6	
>90	30.6	69.6	
Female, %	52.6	58.8	59.5
Race-Sex categories, %			
Black women	43.6	40.0	40.0
Black men	43.4	30.8	30.1
White women	9.0	18.8	19.5
White men	4.0	10.4	10.3
Marital status, %			
Married	32.8	31.1	32.4
Separated	31.6	35.1	35.4
Widowed	12.2	10.7	9.6
Single	23.4	23.1	22.6
Education < 12 <sup>th</sup> grade, %	40.7	32.8	32.4
Income < \$15,000, %	66.6	62.5	60.7

BMI, median (25 <sup>th</sup> , 75 <sup>th</sup> percentile), kg/m <sup>2</sup>	31.2 (26.3, 37.8)	29.2 (24.8, 34.3)	29.3 (25.1, 34.9)
BMI categories, %			
Underweight (<18.5)	0.5	1.6	1.1
Normal (18.5-24.9)	17.8	24.2	23.4
Overweight (25-29.9)	27.7	28.7	29.1
Obese (30+)	54.0	45.6	46.5
Smoking status, %			
Current	34.5	47.1	44.6
Former	24.0	20.3	21.3
Never	41.5	32.5	34.1
Hypertension, %	85.9	54.9	56.2
Diabetes, %	68.7	22.4	22.2
Stroke/TIA, %	12.6	6.9	6.6
MI/Bypass, %	14.5	7.0	7.1

Abbreviations: BMI, body mass index; eGFR, estimated glomerular filtration rate; MI, myocardial infarction; SCCS, Southern Community Cohort Study; TIA, transient ischemic attack

- Page 6, 1st paragraph: define “overweight or obese”

Reply:

In the manuscript we have now added “(BMI≥25 kg/m<sup>2</sup>)” as the definition of overweight or obese on page 9 of the Results section. We have also added this definition to Table 1 and Table 2.

- Page 6, 1st paragraph: Data on most covariates were collected via self reports, such as height, weight, hypertension, diabetes, and high cholesterol. Self-reported data likely bias the results. The study population had low SES and were at high risk of CKD or ESRD; thus, it is possible that chronic conditions such as hypertension or diabetes could be underdiagnosed and underreported due to limited access to health care. Did you evaluate the accuracy of these self-reported data? If not state this as a limitation.

Reply:

Many of the questions on the SCCS questionnaire were adapted from questionnaires used and validated in other settings and a series of validation studies have been carried out to assess the reliability of the questionnaires within the SCCS. The reliability and validity of the SCCS physical activity questionnaire was evaluated in 118 randomly selected SCCS participants via use of accelerometers and the validity was comparable for blacks and whites (see Buchowski et al J Phys Act Health 2012). Self-reported height and weight were compared with clinic recorded measurements for over 20% of participants. Biomarkers, repeat interviews, and medical records were used to assess the reliability of variables such as smoking status and self-reported diseases including diabetes (see Signorello et al J Health Care Poor Underserved 2010).

To clarify the validity of the self-reported data, we have now added the following sentences to the Methods section on page 6 and 7:

“Self-reported height and weight were compared with clinic recorded measurements for over 20% of participants. In a series of validation studies, biomarkers, repeat interviews, or medical records were used to assess the reliability of variables such as smoking status and self-reported diseases including diabetes [13].”

“The reliability and validity of the SCCS physical activity questionnaire was evaluated in 118 randomly selected SCCS participants via use of accelerometers [18].”

We have also clarified the sentence about self-reported data in the Discussion section on page 16:

“Moreover, the physical activity, sedentary behaviors, and covariates were self-reported by participants rather than objectively measured.”

Data analysis methods:

- Please specify which variance estimate method was used in the Cox regression? Did you test the proportional hazards assumption for the covariates?

Reply:

The standard variance estimate method was used: model-based variance estimate from the maximum partial likelihood estimator for the Cox proportional hazards model. The proportional hazards assumption was verified using graphical checks.

- In the COX model, the authors added continuous variables as restricted cubic splines with four knots including total sedentary time, physical activity, age, eGFR, and BMI. But it was not reported which continuous variable has a significant nonlinear relationship with the outcome. Please provide P value for tests of the nonlinear relationship. If the p-value is insignificant, then the continuous variable should be modeled as a linear term.

Reply:

We thank the reviewer for this comment. Total sedentary time, physical activity, age, eGFR, and BMI had significant non-linear relationships with the outcome with p-values less than 0.01.

FORMATTING AMENDMENTS (if any)

Required amendments will be listed here; please include these changes in your revised version:

- Please combine your Figures 1A and 1B; 2A and 2B into one to have a single file figure and make sure that they have a resolution of at least 300 dpi and at least 90mm x 90mm of width. Figures in DOCUMENT, EXCEL and POWER POINT format are not acceptable. Note: If you can't convert your figure into one, kindly renumber the figure legends into Figure 1, Figure 2, etc.

Reply:

We have now combined the files for Figure 1A and 1B and Figure 2A and 2B into two single files, “Figure 2” and “Figure 3”.

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Professor Tazeen Jafar Singapore
<b>REVIEW RETURNED</b>	27-Jul-2019
<b>GENERAL COMMENTS</b>	The reviewer completed the checklist but made no further comments.