# PEER REVIEW HISTORY



## ARTICLE DETAILS

### Title (Provisional)

Association between Measures of Kidney function and Preserved Ratio Impaired Spirometry in Diabetes: NHANES 2007-2012

### Authors

Patel, Ikramulhaq; Gong, Hong-Jian; Xu, Hui; Chai, Yin-He; Qiao, Yu-Shun; Zhang, Jin-Yan; Zhang, Meng-Ting; Stehouwer, Coen DA; ZHOU, JIANBO

---

## VERSION 1 - REVIEW

---

**Reviewer** **1**

**Name** **Hamaguchi, Masahide**

**Affiliation** **Kyoto Prefectural University of Medicine, Department of Endocrinology and Metabolism**

**Date** **03-Sep-2023**

**COI** **No**

---

The authors performed a subanalysis of NHANES 2007-12 and found that Preserved Ratio Impaired Spirometry was associated with CKD.

This study appears to be essentially a cross-sectional study, but I assume that they are conducting a longitudinal study of all-cause mortality and incident CKD.

I had difficulty understanding which studies were cross-sectional and which were longitudinal.

I request that you re-state this point in a clearer manner.

Also, the COX model is presumed to be applied to longitudinal studies, but the definition of events is not clear. This should be clearly stated.

Also, is the longitudinal study a prospective study or a retrospective study? Is it a prospective study or a retrospective study?

This should also be clarified.

I assume that time studies are not mentioned in the flowchart. If so, the flowchart should clearly state the time study.

If it is a prospective study, omissions should be described in detail.

---


| | |
|---|---|
| **Reviewer** | **2** |
| **Name** | **Miyazaki, Mariko** |
| **Affiliation** | **Tohoku Daigaku Daigakuin Igakukei Kenkyuka Igakubu** |
| **Date** | **01-Nov-2023** |
| **COI** | **None** |


---

This research examined the relationship between kidney disease and abnormal respiratory function.

It will be useful not only in the respective specialized fields but also in the practice of general physicists.

I have some comments on confounders and inclusion.

1. Did patients included in the analysis who treated lung cancer, both primary and metastatic, infectious diseases such as pulmonary tuberculosis or organizing pneumonia, other organic lung diseases associated with collagen disease.

It would be good to have data on the complications of collagen disease.

2. Cause of GFR decline:

If the cause of CKD is polycystic kidney disease, liver cysts may occur, making it difficult to expand the thoracic cavity and affect respiratory function. Due to the nature of polycystic kidney disease, the renal prognosis is poor.

Although these factors doesn’t show statistically significant impact on the overall results, we would like to consider the clinical and comprehensive pathophysiology of kidney disease and respiratory function in discussion section.

Minor comments:

Table 1

It will help to understand if the bold items of the table variables are aligned to the left, and the small items of normal width are placed in the center.

“(%)” needs in the variables that are expressed in percentages, such as race, education, and income in the same way as sex.

In supplemental table 2.

I suggest to change "Normal with CKD" to `Normal spirometry with CKD''

## VERSION 1 - AUTHOR RESPONSE

please find the attached file

**Response to Reviewer 1**

**Dr. Masahide Hamaguchi, Kyoto Prefectural University of Medicine**

**Comments to the Author:**

1) **The authors performed a subanalysis of NHANES 2007-12 and found that Preserved Ratio Impaired Spirometry was associated with CKD.**
**This study appears to be essentially a cross-sectional study, but I assume that they are conducting a longitudinal study of all-cause mortality and incident CKD.**
**I had difficulty understanding which studies were cross-sectional and which were longitudinal.**
**I request that you re-state this point in a clearer manner.**
**Response:**
We apologise for the confusion. In the cross-sectional study, we examined the relationship between the measure of kidney function (eGFR and UACR) and abnormal spirometry findings (PRISm, VO). In the retrospective cohort part, we used mortality data available on the NHANES website to see whether the abnormal spirometry findings influence all-cause death risk. Further, how having PRISm/VO with CKD affects all-cause mortality risk was also assessed.
We have improved the clarity of this point by revising the abstract and relevant passages of the methods (Line 96, 163-171, 174-185). In the revised version, the abstract clearly describes the study design, setting, outcomes, etc. In addition, the new Figure 1 is clearer and more informative. Thank you for your comment, which helped us improve clarity.

2) **Also, the COX model is presumed to be applied to longitudinal studies, but the definition of events is not clear. This should be clearly stated.**
**Response:**
You are correct. COX model was used to assess all-cause death risk. We have revised the relevant section and clearly defined the outcome (all-cause mortality). Line 142-148 (Ascertainment of mortality)

3) **Also, is the longitudinal study a prospective study or a retrospective study? Is it a prospective study or a retrospective study?**
**This should also be clarified.**
**Response:**
It is retrospective. As mentioned earlier, the revised abstract and the main text (methods) now clearly state the study design and setting.

4) **I assume that time studies are not mentioned in the flowchart. If so, the flowchart should clearly state the time study.**
**If it is a prospective study, omissions should be described in detail.**

**Response:**
Thank you for pointing this out. We have revised Figure 1, clearly showing the study participant selection process and the type of studies, including the cross-sectional and cohort.

**Response to Reviewer: 2**

Dr. Mariko Miyazaki, Tohoku Daigaku Daigakuin Igakukei Kenkyuka Igakubu

Comments to the Author:

This research examined the relationship between kidney disease and abnormal respiratory function.

It will be useful not only in the respective specialized fields but also in the practice of general physicists.

I have some comments on confounders and inclusion.

1. **Did patients included in the analysis who treated lung cancer, both primary and metastatic, infectious diseases such as pulmonary tuberculosis or organizing pneumonia, other organic lung diseases associated with collagen disease.**
   **It would be good to have data on the complications of collagen disease.**
   **Response:**
   Thank you for your insightful comment. NHANES did not provide the information about the infectious diseases. However, NHANES excluded those with tuberculosis exposure from spirometry testing. Regarding cancer, analysis excluding those with lung cancer or metastases was not feasible as participant responses to these medical conditions were extremely low. For example, in NHANES 2007-2008, out of 10109 participants, responses of 9491 were missing regarding the question "What kind of cancer was it?".

2. **Cause of GFR decline:**
   **If the cause of CKD is polycystic kidney disease, liver cysts may occur, making it difficult to expand the thoracic cavity and affect respiratory function. Due to the nature of polycystic kidney disease, the renal prognosis is poor.**
   **Although these factors doesn't show statistically significant impact on the overall results, we would like to consider the clinical and comprehensive pathophysiology of kidney disease and respiratory function in discussion section.**
   **Response:**
   Thank you for your insightful comment. We agree that in cases of CKD due to ADPKD, confounding may be present as lung function may be reduced due to liver cysts seen in ADPKD. We acknowledged this limitation and discussed it: "Fifth, although we controlled for potential risk factors, residual confounding cannot be ruled out entirely due to the complex pathophysiology between kidney disease and lung function. For example, in polycystic kidney disease, cysts may develop in the liver, which may hinder the expansion of the thoracic cavity, potentially impacting lung function."

**Minor comments:**

1. **Table 1**
   **It will help to understand if the bold items of the table variables are aligned to the left, and the small items of normal width are placed in the center.**
   **Response:**
   Duly revised

2. **"(%)" needs in the variables that are expressed in percentages, such as race, education, and income in the same way as sex.**
   **Response:**
   Duly revised

3. **In supplemental table 2.**
   **I suggest to change "Normal with CKD" to `Normal spirometry with CKD"**
   **Response:**
   Duly revised. Thank you for your thorough review of our paper. We have revised the manuscript as per your suggestions, and we believe that it improved our paper.

**Response to Reviewer: 3**

**Dr. David Pascall, University of Cambridge**

**Comments to the Author:**

**I have been asked to review the statistical aspects of this manuscript, and so I will limit my comments to this area.**

**This manuscript seems like the statistics have generally been applied appropriately. I have a few comments.**

1) **Please give more details on the probabilistic matching method used to link the data. There is currently not enough information here to replicate.**
   **Response:**
   We used the public-use linked mortality data which are available here (https://www.cdc.gov/nchs/data-linkage/mortality-public.htm). While the linkage methods are described here (https://www.cdc.gov/nchs/data-linkage/mortality-methods.htm). For the reader's clarity, we have also cited these sources in the manuscript.

   For clarity, linked mortality files can be loaded into statistical software. Mortality data can be merged with the NHANES survey data using the unique participant ID ("SEQN"). Mortality data has variables like eligibility, mortality status, person-months, etc. detailed information on what linked mortality files contain are provided in this dictionary

(https://www.cdc.gov/nchs/data/datalinkage/public-use-linked-mortality-files-data-dictionary.pdf)

2) **You state "Unlike the results observed in the overall cohort, individuals with diabetes and eGFR <60 mL/min/1.73m2 nwere not associated with variable obstructive lung function". I'm not sure this is true. Looking at Table 3, you see large effects in the same direction as the main cohort. These are effects are not significant, and this should be stated, but that's not the same as not associated.**
**Response:**
Thank you for pointing this out. As suggested, we have revised it.

"Although higher risk of variable obstruction was observed for individuals with diabetes and eGFR <60 mL/min/1.73m$^2$, it was statistically insignificant (OR 2.43[95% CI, 0.65-9.09], p=0.18)"

3) **I would recommend that the authors share their code to aid replicability, so that the precise coding choices that were made in the analysis can be preserved.**
**Response:**
Thank you for your recommendation regarding sharing our code to aid in replicability. After careful consideration and discussion within our team, we have decided not to share the code at this time. We appreciate your understanding and are happy to provide any additional details or clarifications about our analysis methods that might be helpful.

4) **If the above changes are made, I do not feel the need to see this manuscript again as I am happy with the methods applied. Thank you to the authors for this work!**
**Response:**
Thank you for your valuable expert opinion. We have tried our best to address your concerns and hope our revisions meet your expectations.

Response to Reviewer: 4
Dr. Floriano Amimo, Eduardo Mondlane University

Comments to the Author:

A. GENERAL COMMENTS

1. **Important data and details to support key statements and/or comply with applicable requirements/guidelines/etc., as well as the rationale of the methods, need to be provided, etc.**
**Response:**
Thank you for your careful evaluation. Our study examined the relationship between measures of kidney function (including eGFR and UACR) and impaired lung function (including PRISm and variable obstruction) using cross-sectional data from the previously validated NHANES database. Furthermore, we determined the all-cause mortality risk associated with CKD and/or various spirometry categories.

Given our study's independent and dependent variables, we employed widely accepted methods such as logistic regression and Cox regression, which are appropriate for our study design and allow for robust data analysis. In addition, we followed analytical guidelines put forward by NHANES, which involved using sample weights, clustering, and stratification to ensure that the included sample represents US population.

In summary, we found that an elevated urinary albumin-to-creatinine ratio (UACR) and albuminuria were linked to a higher risk of PRISm. Our study provided further evidence of the kidney-lung crosstalk.

As an observational cohort study, we have followed the STROBE guidelines. The relevant statement is now attached with the revised manuscript. We have also improved the abstract, Figure 1, and several passages in the methods (See lines 95-106, 131-140, 142-148, 163-171, 174-185)

2. **Implications of study limitations to internal and external validity and interpretation of the results should be discussed.**
**Response:**
Thank you for your valuable feedback. We have improved the discussion of limitations by including the following points:

Causality: The cross-sectional design prevents us from establishing causality.
Self-reported Data: The reliance on self-reported data introduces potential biases.
Residual Confounding: Residual confounding cannot be entirely ruled out.
eGFR Categories: Using only two eGFR categories (<60 and >60) might not fully capture the spectrum of kidney damage.
Small Diabetes Sample: The relatively small number of diabetic participants may influence the interpretability of our findings.
U.S. Population: Our findings apply only to U.S. adults.

We have revised the manuscript to reflect these limitations and their implications for the validity and interpretation of our results (lines 323-336).

3. **There is important lack of clarity/conciseness/consistency/rigor/etc. in the manner that several passages/methods/results/contents/etc. are presented/structured.**
**Response:**
We appreciate your valuable feedback. We have carefully improved the clarity, conciseness, consistency, and rigor in the manuscript's presentation and structure of various passages, methods, results, and content. We believe these revisions have significantly enhanced our study's overall quality and readability.

Briefly, the discussion starts by describing the main findings (first paragraph, Line 259-264), followed by the second paragraph (Lines 266-292) which effectively integrates various studies and past/current evidence to build a coherent narrative on the interplay between kidney and lung function. We then discussed potential mechanisms and

implications in the third (Lines 294-304) and fourth (Lines 306-313) paragraphs, respectively. Finally, we revised the limitations section as we identified a few limitations thanks to your careful review.

B. SPECIFIC COMMENTS

I. Abstract:

1. **What is the study design?**
   **Response:** Cross-sectional study and retrospective cohort study. Revised abstract, line 18

2. **What is the study setting?**
   **Response:** NHANES. Revised abstract, Line 19

3. **What models/etc were used to "examine the relationship" and "analyze the impact"? Why?**
   **Response:** Logistic regression (cross-sectional) and COX models (retrospective cohort). Revised abstract, Line 25-29

4. **What are the main outcome measures? Why?**
   **Response:** Revised abstract, Line 22-25

5. **"The study included 10,809 individuals,": a) Can a summary of baseline characteristics of the study participants be provided? If not, why? a1) E.g.: age mean (SD), sex/gender ratio, etc.**
   **Response:** We revised the abstract as per journal guidelines. Thus, we only described the main results in the abstract. However, characteristics (age mean (se), gender ratio) of included participants are described in the results section of the manuscript (Line 191-193)

6. **"Results showed that those with PRISm had higher WBC count, BMI, WC, HOMAIR, FPG, HbA1c, CRP, and diabetes duration compared to those with normal spirometry.": a) Can the estimates to support this be provided? If not, why?**
   **Response:** As mentioned earlier, only the main results were described in the revised abstract.

7. **Can concerns e) to h) regarding a) to d) be addressed? If not, why?**
   **a) "OR 1.10[1.01,1.21], p=0.03"**
   **b) "OR 1.72 [1.07,2.74], p=0.03"**
   **c) "OR 1.21[1.08, 1.36], p=0.002"**
   **d) "HR 3.46[1.94,6.16], p=<0.0001"**
   **e) What are "1.10", "1.72", "1.21", and "3.46"?**
   **f) What are "[1.01,1.21], "[1.07,2.74]", "[1.08, 1.36]", and "[1.94,6.16]"?**
   **g) How were the "OR" and "HR" calculated? Why?**
   **h) "p=<0.0001": Why "=<"?**

**Response:**
a) to d):
They are Odds ratios [95% confidence interval]. ORs were calculated by logistic regression. In the abstract, estimates are now presented as (OR [95%CI] =xx [xx, xx], p=xx) for clarity. Line 31-39
h):
by p=<0.0001, we meant p-value is <0.0001

8. **Can abbreviations be defined the first time they are used? If not, why?**
   a) **"WBC count, BMI, WC, HOMAIR, FPG, HbA1c, CRP,", "UACR", "CKD", etc.**

   **Response:** In the revised manuscript, the abstract does not have an above line, as we only presented the main result.

9. **How was "normal spirometry" defined in this study? Why?**
   **Response:** We defined normal spirometry as FEV1≥80%pred and FEV1/FVC ratio > 0.7. (Wan ES et al. 2021. doi:10.1001/jama.2021.20939). We have added the definition of the normal spirometry.

10. **Can these and other applicable passages be adjusted accordingly? If not, why?**
    **Response:** We have revised the abstract in the manuscript. Line 13-43

**IV. Methods:**

1. **Which reporting guideline does this study comply with? Why?**
   **b) Can the checklist of the reporting guideline indicating location of each item in the main text be provided in the supplementary materials? If not, why?**
   **Response:** STROBE. We have attached the STROBE statement

2. **What is the study design?**
   **Response:** Cross-sectional and retrospective cohort study. Line 96

3. **What is the study setting?**
   **Response:** We have rewritten the section of methods and described study settings under the “data source” subheading.

   “This cross-sectional and retrospective cohort study used data from the National Health and Nutrition Examination Survey (NHANES) cycles 2007-2008, 2009-2010, and 2011-2012, conducted by the National Center for Health Statistics (NCSH) of the Centers for Disease Control and Prevention (CDC)” Line 96-99

4. **Can citations be provided for the following and other applicable passages? If not, why?**
   **a) "a large-scale, representative survey of the non-institutionalized civilian population in the United States."**

**b) "The complex and multi-stage probability design of NHANES ensured a representative sample of the US population."**
**Response:** Thank you for pointing this out. NHANES references (#27, #28) have been cited.

5. **"Our sample comprised of 10,809 participants who were aged over 20 years, underwent a medical examination, had acceptable spirometry data (classified as A, B, or C), had available information on their estimated glomerular filtration rate (eGFR) and urine albumin-creatinine ratio (UACR), and had available mortality data":**
**a) What "mortality data"? Why?**
**Response:** The National Center for Health Statistics (NCHS) has linked data from NHANES and other surveys with death certificate records from the National Death Index (NDI), and made available the public-use linked mortality files. We used this data to ascertain the all-cause mortality status of participants in our study through the unique participant ID. We have revised the relevant passage of methods (lines 142-148) for clarity and cited the sources for more information.

6. **"The spirometry data were classified based on the quality of data collection, with "A" being the highest quality and exceeding American Thoracic Society standards, "B" meeting the standards, and "C" being potentially usable but not meeting all the standards. The sample selection process is detailed in Figure 1.":**
**a) What is the difference between "the standards" and "all the standards"? Why?**
**b) How was "quality of data collection" determined/measured? Why?**
**Response:**
**Difference between the standards:**
NHANES followed ATS criteria and standardisation of spirometry (Miller, Hankinson, Brusasco et al. 2005).
A= Exceeds ATS data collection standards: 3 acceptable curves present and 2 reproducible curves; 2 observed values within 100 ml.
B= Meets ATS data collection standards: 3 acceptable curves present and 2 reproducible curves; 2 observed values within 150 ml.
C= Potentially usable value, but does not meet all ATS standards. Estimates usually based on 2 curve results with values within 200 ml. of each other.
To aid readers, we have cited the source in the manuscript. Line 132

**Quality of the spirometry:**
"The NIOSH Division of Respiratory Disease Studies served as the NHANES Spirometry training and quality control consultant. Each MEC Health Technician received formal training and satisfactorily completed the NIOSH-approved spirometry course. Additionally, all NHANES spirometry data were reviewed by expert reviewers at the NIOSH quality control center on an ongoing basis. MEC Technicians received continuous feedback on their performance throughout the survey." More details is available at https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/SPX_G.htm.

To simplify, we only included the most important details about the spirometry data in the manuscript. However, we have cited the sources (ATS criteria) for those needing more details.
"Quality of the spirometry have been described elsewhere (https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/SPX_G.htm)." Line 135

7. **"Following Ford et al.(27), participants were categorized as current smokers if they had smoked 100 or more cigarettes in their lifetime and reported smoking currently, former smokers if they had smoked 100 or more cigarettes but had quit, and never smokers if they had not smoked 100 or more cigarettes in their lifetime":**
**a) "had quit" when? Why?**
**b) How was the uncertainty in the count of "cigarettes" smoked accounted for analytically? Why?**
**Response:** Using a response from the following two interview questions smoking status defined:
-Smoked at least 100 cigarettes in life? (Yes/No)
-Do you now smoke cigarettes? (Everyday/Someday/Not at all etc.)

Current smoker: smoked ≥ 100 cigarettes during his or her lifetime and reported smoking currently.
Former smoker: smoked ≥ 100 cigarettes during his or her lifetime but reported having stopped smoking.
Never smoker: not smoked ≥ 100 cigarettes during his or her lifetime

Due to the self-reported nature of this variable, we have recognised this limitation. However, our results should remain sound as these definitions of smoking variable have been used previously. (Navaneethan et al. 2016. doi: 10.1053/j.ajkd.2016.03.415, Ford ES et al. Chest 2013. doi: 10.1378/chest.12-1135)

"Third, some covariates used in the study depended on the self-report by the participant, which may produce biases such as recall and social desirability bias" lines 327-329

8. **"Diabetes was diagnosed using criteria such as the use of insulin or oral hypoglycemic agents, fasting plasma glucose levels of 126 mg/dL or higher, or glycated hemoglobin levels of 6.5% or higher."**
**a) How was the risk of misdiagnosis of "Diabetes" accounted for analytically? Why?**
**b) What guideline does the approach used to diagnose "Diabetes" comply with? Why?**
**Response:** NHANES study team consists of a physician, medical and health technicians, as well as dietary and health interviewers. In addition to the self-reported data, we used laboratory data to determine diabetes status (ADA guidelines, A1C ⩾6.5% or FPG ⩾126 mg/dL). Therefore, we may have reduced some bias. Although we did not account for misdiagnosis analytically, our result should remain sound.

9. **"Participants were considered hypertensive if they had systolic blood pressure of 140 mm Hg or higher and diastolic blood pressure of 90 mm Hg or higher, or ..."**
**a) Under which conditions? On a single day?**
**a1) E.g.: see: ?**
**b) What guideline does the approach used to diagnose hypertension comply with? Why?**
**Response:** "After resting quietly in a seated position for 5 minutes and once the participants maximum inflation level (MIL) has been determined, three consecutive blood pressure readings are obtained. If a blood pressure measurement is interrupted or incomplete, a fourth attempt may be made. All BP determinations (systolic and diastolic) are taken in the mobile examination center (MEC). Participants with any of the following on both arms were excluded from the exam: rashes, gauze dressings, casts, edema, paralysis, tubes, open sores or wounds, withered arms, a-v shunts, radical mastectomy, or if BP cuff does not fit on the arm". (https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/BPX_G.htm). We used the mean of available readings. The 140/90 mmHg criteria refer to the International Society of Hypertension guideline, often used in current clinical practice.

10. **"Logistic regression was employed to assess the odds of PRISm (impaired spirometry) and variable obstructive lung function given declining estimated glomerular filtration rate (eGFR) and log-transformed urine albumin-to-creatinine ratio (UACR).":**
**a) How was the potential correlations between these outcomes accounted for analytically? If not, Why?; a1) E.g.: see: ; etc.; b) How was the "odds" of "variable obstructive lung function given declining estimated glomerular filtration rate (eGFR) and log-transformed urine albumin-to-creatinine ratio (UACR)." estimated?; b1) "log-transformed" why?**
**Response:** Thank you for your questions. We appreciate the opportunity to provide further clarification.

a) a1) Potential correlations between outcomes:

In logistic regression models, predictors were either continuous or categorical, and dependent variables were dichotomous. Our analysis did not specifically account for potential correlations between the outcomes. We acknowledge that correlations between outcomes could influence the results and will consider this in future analyses.

b) Estimating the "odds":

We apologise for the carelessness; we originally wanted to write "given increasing eGFR" (per unit increase). Relevant phrase has been revised.
The odds of spirometry findings (normal, PRISm, VO) per unit increase in eGFR and log-transformed UACR were estimated using logistic regression models. These models allow us to calculate the odds ratios, which describe the relationship between the predictors

(eGFR and UACR) and the likelihood of having PRISm and VO compared with normal lung function.

b1) Log-transformed UACR:

The urine albumin-to-creatinine ratio (UACR) was log-transformed to normalise its distribution to address skewness. Log transformation helps in meeting the assumptions of the logistic regression model, particularly the assumption of linearity between the log odds of the outcome and the predictor variables.

11. **Is "Logistic regression" adequate for the data?; a) Why?; a2) How was this tested?**
**Response:** We believe logistic regression was appropriate for our analysis because it is well-suited for modelling the relationship between a predictor and outcomes, allowing us to estimate the probability of each outcome based on the predictor variable.

We did not conduct formal statistical tests to specifically assess the adequacy of logistic regression for our data. However, the theoretical foundation and common use of logistic regression in similar contexts support its suitability for our analysis. We ensured that the model assumptions were met and the results were interpretable and consistent with our research objectives.

12. **How were the models validated?**
**Response:** We did not conduct formal validation. However, we ensured that logistic regression was a theoretically appropriate choice for our predictor/s and outcomes. We acknowledge the importance of the validation steps. In future work, we will incorporate formal validation techniques to assess the robustness and generalizability of our models.

13. **Can the results of model calibration and variable selection be provided in the supplementary materials? If not, why?**
**Response:** We did not perform formal variable selection procedures in this study as we pre-specified models. As a result, we do not have detailed results from these processes to include in the supplementary materials. We acknowledge that variable selection enhance the robustness and interpretability of the models. We will consider this in our future work.

14. **Can the results of model goodness of fit be provided in the supplementary materials? If not, why?**
**Response:** We did not perform this analysis. However, as mentioned earlier, we ensured that the data were adequate for the logistic regression model and that assumptions were met. We appreciate an insightful comment. We will incorporate this, as well as the above-mentioned statistical issues, in our future work.

15. **"Categorical analyses were carried out in three categories of UACR (<30, 30-300, and ≥300 mg/g) and two categories of eGFR (>60 and <60 mL/min per 1.73 m2), where the UACR <30 mg/g and eGFR >60 mL/min per 1.73 m2 categories were used as the reference groups for comparison. "**

**a) What is the rationale of these and other categories created for this study? b) Can relevant sources be provided to support these and other analytical choices?**
**Response:**

a) rationale for using UACR categories:

Because the standard cutoff for normal UACR is 30mg/g, >30-300mg/g for microalbuminuria, and > 300mg/g for macroalbuminuria. we have cited the reference to support categorization. To aid readability, we have changed microalbuminuria to "moderately increased albuminuria" and macroalbuminuria to "severely increased albuminuria".

b) Rationale for using eGFR categories:
due to very low number participants in some categories, we could not use more standard eGFR categories (G1-G5). Thus, we used the next best option, which was >60 mL/min per 1.73 $m^{2,}$ (normal kidney function with mild loss of function) and <60 mL/min per 1.73 $m^2$ (suggestive of kidney function loss).

16. **How was data quality verified and ensured? Why? a) Who verified data quality? b) Using what tools? b1) Can these be described?**
**Response:** We used data from the National Health and Nutrition Examination Survey (NHANES), which is a highly reputable and widely used dataset. The NHANES program is conducted by the National Center for Health Statistics (NCHS) and is known for its rigorous data collection and quality assurance procedures(https://www.cdc.gov/nchs/data/nhanes/nhanes_release_policy.pdf). More details about the data collection and process can be found in this pdf (https://www.cdc.gov/nchs/data/series/sr_01/sr01_056.pdf)under the "Data Collection and Processing" section. We have cited the source of information in the manuscript (ref #27, #28).

17. **When were the data retrieved from the sources?**
**Response:** In November 2022.

18. **How were the data retrieved from the sources? a) By whom? b) Using what tools? b1) Can these be described?**
**Response:** Data were downloaded directly from NHANES website and loaded into the R by the first author.

19. **"The Cox proportional hazards model was utilized to calculate hazard ratios for PRISm and variable obstructive lung function groups to examine the relationships between chronic kidney disease, lung function, and mortality.":**
**a) Were model assumptions (e.g., (i) proportional hazards, (ii) non-informative censoring, (iii) independent survival times for each observation, (iv) linear effect of predictor variables on the log hazard, (v) values of predictor variables for individuals constant over time, etc) verified? How? If not, why?**

**a1) Can the results be provided in the supplementary materials? If not, why?**
**b) For each model, what variable was used as:**
**b1) Follow up time (for right censored) or starting time (for interval censored data)?**
**b2) Ending time? Why?**
**b3) Event variable? Why?**
**b4) Predictor variable? Why?**
**c) What type of censoring was used? Why?**
**d) Is "Cox proportional hazards model" adequate for the data? d1) Why? d2) How was this tested? d3) Were parametric models tested? How? e) E.g.: see: ; ; ; ; ; etc.**
**Response:**

a) a1) d): Thank you for highlighting this. We tested the COX assumption using Schoenfeld residuals. In all models, p-values were greater than 0.05 for all covariates as well as in the global test. We apologize for leaving out the important details. We added the following line:
"We examined the proportional hazards assumption using Schoenfeld residuals; assumptions were satisfied in all models (global p-value >0.05)" Line 18-183. However, we as team feel that it is not necessary to add them to the supplementary material:

b) b1) b2) b3) b4): We used follow-up time that was calculated from the date of the interview to the death date (acquired by NDI linkage) or the end of the mortality period, December 31, 2019. The event variable was all-cause mortality, and predictors were spirometry findings and/or CKD status. We have improved this point (Lines 174-181)

c): Right censored. because participants were assumed alive if there was no death record in the NDI/ death was not confirmed. Right-censoring allows us to account for participants whose complete survival time is not observed within the study period. This ensures that the analysis accurately reflects the survival experience of the cohort up to the point of censoring.

20. **"To analyze the data, two models were created: Model 1 included demographic variables such as age, sex, ethnicity, education, and annual household income, while Model 2 (full adjustment) further included body mass index, hypertension, smoking, cardiovascular disease, and diabetes mellitus." a) How was multicollinearity among predictors accounted for analytically? Why?**
**Response:** Thank you for your insightful comment. In our analysis, we checked multicollinearity using car::vif() in R. in all models VIF was below 2, indicating very low multicollinearity.

21. **"The determination of mortality status was established through a probabilistic matching process between the NHANES dataset and death certificate records obtained from the National Death Index." a) How was this validated? Why?**

**Response:** We have revised this line for clarity and cited the source of detailed information.

"The National Center for Health Statistics (NCHS) has linked data from NHANES and other surveys with death certificate records from the National Death Index (NDI), and made available the public-use linked mortality files (https://www.cdc.gov/nchs/data-linkage/mortality-public.htm, accessed June 5, 2024)."

Previous studies using NHANES data to investigate mortality have used linked mortality files. (i.e., Chen C et al. BMJ. doi: 10.1136/bmj.l5584. PMID: 31619383).

22. **Can these and other applicable passages be rewritten to maximize clarity/rigor/conciseness/relevance? If not, why?**
**a) "In the period between 2007 and 2012, the NHANES conducted pulmonary function tests on all adult participants. However, individuals experiencing chest pain, difficulties with forceful expiration, use of supplemental oxygen, recent surgeries on the eye, chest, or abdomen, recent heart attack, stroke, tuberculosis exposure, coughing up of blood, or a history of detached retina, collapsed lung, or aneurysm were excluded from the study(29)."**
**a1) E.g.: Then, how is the "data"" used in the study "representative survey of the non-institutionalized civilian population in the United States."**
**a2) E.g.: How was this accounted for analytically? Why?**
**b) "The estimated glomerular filtration rate (eGFR) was calculated using the Chronic Kidney Disease Epidemiology Collaboration equation(28)."**
**b1) "equation(28)" or equation (28)? Why?**
**b2) How? Why?**
**c) "The eGFR was also divided into two categories: less than 60 mL/min/1.73m2 and greater than 60 mL/min/1.73m2."**
**c1) E.g.: "m2" or m^2? Why?**
**d) Etc.**
**Response:**

a1) a2):

NHANES instructs researchers to account for complex survey designs and ensure a representative sample of the US population. It is done by using sample weights and adjusting for clusters and strata. Therefore, even though some participants are excluded from the examination, the final sample will still represent the US population.

"In line with the instructions for using NHANES data, we used the sample weights, clustering, and stratification whenever feasible to account for the complex survey design using the "survey" R package (version 4.1-2)"

b) b1) b2):

We used CKD-EPI to calculate the eGFR because it is widely used in similar research. We have cited the reference.

GFR = 141 × min (Scr/κ, 1)α × max(Scr/κ, 1)-1.209 × 0.993Age × 1.018 [if female] × 1.159 [if black]

Scr = serum creatinine,
κ =0.7 for females and 0.9 for males,
α = -0.329 for females and -0.411 for males,
min = the minimum of Scr/κ or 1,
max = the maximum of Scr/κ or 1.

c) c1) d):

Thank you. It should be $m^2$. We have corrected the typos.

23. **Can the models used be represented mathematically, at least in the supplementary materials? If not, why?**
**Response:** Thank you for your suggestion. While the models used, such as logistic regression and Cox proportional hazards models, can certainly be represented mathematically, we have chosen not to include the mathematical equations in the supplementary materials. These models are widely used and familiar to the target audience, and the mathematical formulations are standard and well-documented in the literature. Including the equations in the supplementary materials would not provide additional insights or clarity for the readers. However, we are happy to provide any specific mathematical details upon request or if deemed necessary by the editorial board.

24. **"The interviews collected self-reported information on demographics, socioeconomic status, and health conditions, while the medical examinations included a range of physiological measurements and laboratory tests performed by highly trained medical staff. "**
**a) How was the risk of social desirability bias accounted for analytically? Why?**
**b) How was the risk of recall bias accounted for analytically? Why?**
**c) Etc.**
**Response:** Thank you for your comment. We acknowledge the limitations associated with self-reported information, particularly the risks of social desirability and recall biases. In our study, we have added a line in the limitations section stating:

"Third, some covariates used in the study depended on self-report by the participant, which may introduce biases such as recall and social desirability bias."

25. **"The complex and multi-stage probability design of NHANES ensured a representative sample of the US population."**
a) How was this accounted for in the analysis/models? Why?
**Response:** As mentioned earlier, to ensure the complex survey design of NHANES, we used sample weights and adjusted for clusters and strata, as per NHANES instructions.

**26. Is the study design consistent with survival analysis?**

**a) E.g.: see: ; ; ; etc.**

**Response:** Thank you for your question. Our study design is indeed consistent with survival analysis methods. We have employed survival analysis techniques such as the proportional hazards model, commonly applied in studies involving time-to-event outcomes.

**27. How was multiple testing correction conducted? Why?**

**a) E.g.: see: ; ; etc.**

**Response:**

We did not conduct multiple testing correction in our analysis. Given the specific research question and the exploratory nature of some analyses, we opted not to apply multiple testing correction methods. While multiple testing correction can be valuable in certain contexts, such as confirmatory analyses or when conducting a large number of hypothesis tests, we deemed it appropriate to present uncorrected p-values alongside careful interpretation of results in the context of our study objectives.

**28. Was sensitivity analysis conducted? If not, why?**

**a) Can the results be discussed in the main text and detailed in the supplementary materials? If not, why?**

**Response:**

We did not conduct the sensitivity analysis. for sensitivity analysis, we planned to further exclude participants with lung cancer to test the robustness of our result. However, we found that participants with lung cancer were extremely low, just 15-20 in one cycle (due to non-response to the question). Thus, we did not further analyse our data. However, we did analyse the all-cause mortality risk, while changing the reference category. Compared with PRISm without CKD, those with CKD had significantly higher risk of all-cause death, further indicating that having both PRISm and CKD compared to just PRISm further increases the mortality risk. The table below has been added as Supplementary Table 3:

| | | Entire cohort | | Diabetes[a] | |
|---|---|---|---|---|---|
| | | HR [95% CI] | *p* value | HR [95% CI] | *p* value |
| PRISm without CKD | | | reference | | |
| PRISm with CKD | Model 1 | **2.60(1.60,4.23)** | **0.002** | **3.40(1.63,7.07)** | **0.001** |
| | Model 2 | **2.23(1.33,3.74)** | **0.0001** | **2.91(1.33,6.39)** | **0.008** |
| VO without CKD | | | reference | | |
| VO with CKD | Model 1 | 0.66(0.22,2.01) | 0.47 | 2.79(0.23,34.63) | 0.42 |
| | Model 2 | 0.57(0.19,1.75) | 0.33 | 2.47(0.17,34.87) | 0.50 |

PRISm, preserved ratio impaired spirometry; VO, variable obstruction; Model 1: adjusted by age, sex, eth, education, annual household income; Model 2: adjusted by age, sex, eth, education, annual household income, BMI, hypertension, smoking, cardiovascular disease, diabetes mellitus

**29. What is the definition of "U.S. adults" in the study? Why?**

**Response:**

"U.S. adults" refers to individuals who are citizens of the United States of America. This definition is used in our study because NHANES specifically recruits participants from the United States, including both adults and children. As our study focuses on adult participants, we referred to them as "U.S. adults".

**30.** Can these and other applicable passages be adjusted accordingly? If not, why?

**Response:**

Thank you for your in-depth review of our manuscript. We tried our best to address your comments and revised the manuscript accordingly, and we believe that our manuscript improved significantly as a result.

Reviewer: 1

Competing interests of Reviewer: No

Reviewer: 2

Competing interests of Reviewer: None

Reviewer: 3

Competing interests of Reviewer: NA

Reviewer: 4

Competing interests of Reviewer: None

---

## VERSION 2 - REVIEW

---

| | |
|---|---|
| **Reviewer** | **3** |
| **Name** | **Pascall, David** |
| **Affiliation** | **University of Cambridge, MRC Biostatistics Unit** |
| **Date** | **20-Jun-2024** |
| **COI** | **NA** |

---

Thank you for the hard work addressing the reviewers comments. I am satisfied this is ready for publication.

| | |
|---|---|
| **Reviewer** | **4** |
| **Name** | **Amimo, Floriano** |
| **Affiliation** | **Eduardo Mondlane University, Faculty of Medicine** |
| **Date** | **11-Aug-2024** |
| **COI** | **None** |

A. GENERAL COMMENTS

1. There is progress; however, there are still several important issues that need to be addressed.

2. Important data/details to support key statements and/or comply with applicable requirements/guidelines/etc. and/or address concerns raised are still not provided, at least not clearly.

3. There is important lack of clarity/rigor/consistency/conciseness/etc. in the manner that several passages/methods/contents/limitations/etc. are presented.

B. SPECIFIC COMMENTS

I. Abstract:

I.I. Pending concerns:

3. What models/etc were used to "examine the relationship" and "analyze the impact"? Why? {Peer review round 2: see a) to e)}:

a) Authors' response: "Logistic regression (cross-sectional) and COX models (retrospective cohort). Revised abstract, Line 25-29"

b) Relevant changes have been partially implemented by the authors.

c) This reviewer recommends the authors to consider making applicable changes to indicate clearly for what design each model was used.

d) Can relevant changes be implemented in the revised manuscript? If not, why?

e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

4. What are the main outcome measures? Why? {Peer review round 2: see a) to f)}:

a) Authors' response: "Revised abstract, Line 22-25"

b) Relevant changes have been partially implemented by the authors.

c) This reviewer recommends the authors to consider separating exposures from outcomes for clarity, e.g., by creating a separate "Exposures" subsection.

d) Additionally, no clear distinction is made between "Primary" and "secondary" outcomes; then why write "Primary and secondary outcome measures"?

e) Can relevant changes be implemented in the revised manuscript? If not, why?

f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

10. Can these and other applicable passages be adjusted accordingly? If not, why?

I.II. New concerns:

1. Can these and other applicable passages be rewritten to maximize clarity/rigor/conciseness? If not, why?

a) "(normal spirometry= 9503, PRISm=951, variable obstruction=355)"

2. Can these and other applicable passages be adjusted accordingly? If not, why?

IV. Methods:

IV.I. Pending concerns:

1. Which reporting guideline does this study comply with? Why? {Peer review round 2: see c) to g)}:

b) Can the checklist of the reporting guideline indicating location of each item in the main text be provided in the supplementary materials? If not, why?

c) Authors' response: "STROBE. We have attached the STROBE statement"

d) Relevant changes have been partially implemented by the authors.

e) This reviewer recommends the authors to consider adding a passage that indicates clearly that "This study complies with ..." or similar and cite the Checklist.

f) Can relevant changes be implemented in the revised manuscript? If not, why?

g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

2. What is the study design? {Peer review round 2: see a) to e)}:

a) Authors' response: "Cross-sectional and retrospective cohort study."

b) Relevant changes have been partially implemented by the authors.

c) This reviewer recommends the authors to consider adding relevant details as follows:

c1) Change "Data source" to "Study design and data source" or create a separate "Study design" subsection.

c2) Indicate clearly to what end each design was used.

d) Can relevant changes be implemented in the revised manuscript? If not, why?

e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

6. "The spirometry data were classified based on the quality of data collection, with "A" being the highest quality and exceeding American Thoracic Society standards, "B" meeting the standards, and "C" being potentially usable but not meeting all the standards. The sample selection process is detailed in Figure 1.": {Peer review round 2: see c) to g)}:

a) What is the difference between "the standards" and "all the standards"? Why?

b) How was "quality of data collection" determined/measured? Why?

c) Authors' response: "Difference between the standards: NHANES followed ATS criteria and standardisation of spirometry (Miller, Hankinson, Brusasco et al. 2005). A= Exceeds ATS data collection standards: 3 acceptable curves present and 2 reproducible curves; 2 observed values within 100 ml. B= Meets ATS data collection standards: 3 acceptable curves present and 2 reproducible curves; 2 observed values within 150 ml. C= Potentially usable value, but does not meet all ATS standards. Estimates usually based on 2 curve results with values within 200 ml. of each other. To aid readers, we have cited the source in the manuscript. Line 132"; "Quality of the spirometry: “The NIOSH Division of Respiratory Disease Studies served as the NHANES Spirometry training and quality control consultant. Each MEC Health Technician received formal training and satisfactorily completed the NIOSH-approved spirometry course. Additionally, all NHANES spirometry data were reviewed by expert reviewers at the NIOSH quality control center on an ongoing basis. MEC Technicians received continuous feedback on their performance throughout the survey.” More details is available at https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/SPX_G.htm. To simplify, we only included the most important details about the spirometry data in the manuscript. However, we have cited the sources (ATS criteria) for those needing more details. “Quality of the spirometry have been described elsewhere (https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/SPX_G.htm).” Line 135"

d) Relevant changes have been partially implemented by the authors.

e) This reviewer recommends the authors to consider adding relevant details as follows:

e1) Add a brief description, just a few words, of what they mean with "standards" and "all the standards" for clarity, e.g., by replacing X and Y as follows: "[... “B” meeting the standards (that is, X) ... “C” ... not meeting all the standards (that is, Y), ...]".

e2) Cite <https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/SPX_G.htm>, <https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/SPX_G.htm>, and other URLs provided in the main text in line with standard citation and referencing rules (to mitigate the effect of link rot in the main text, etc.).

f) Can relevant changes be implemented in the revised manuscript? If not, why?

g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

9. "Participants were considered hypertensive if they had systolic blood pressure of 140 mm Hg or higher and diastolic blood pressure of 90 mm Hg or higher, or ...": {Peer review round 2: see c) to g)}:

a) Under which conditions? On a single day?

a1) E.g.: see: <https://www.who.int/news-room/fact-sheets/detail/hypertension>?

b) What guideline does the approach used to diagnose hypertension comply with? Why?

c) Authors' response: "“After resting quietly in a seated position for 5 minutes and once the participants maximum inflation level (MIL) has been determined, three consecutive blood pressure readings are obtained. If a blood pressure measurement is interrupted or incomplete, a fourth attempt may be made. All BP determinations (systolic and diastolic) are taken in the mobile examination center (MEC). Participants with any of the following on both arms were excluded from the exam: rashes, gauze dressings, casts, edema, paralysis, tubes, open sores or wounds, withered arms, a-v shunts, radical mastectomy, or if BP cuff does not fit on the arm”. (https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/BPX_G.htm). We used the mean of available readings. The 140/90 mmHg criteria refer to the International Society of Hypertension guideline, often used in current clinical practice"

d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

e) This reviewer recommends the authors to consider adding d1) in the revised manuscript and/or supplementary materials as applicable and citing relevant sources:

e1) A summary of their response to this reviewer.

f) Can relevant changes be implemented in the revised manuscript? If not, why?

g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

10. "Logistic regression was employed to assess the odds of PRISm (impaired spirometry) and variable obstructive lung function given declining estimated glomerular filtration rate (eGFR) and log-transformed urine albumin-to-creatinine ratio (UACR).": {Peer review round 2: see c) to g)}:

a) How was the potential correlations between these outcomes accounted for analytically? If not, Why?

a1) E.g.: see: <https://doi.org/10.1002/(sici)1097-0258(19990815)18:15%3C2011::aid-sim169%3E3.0.co;2-8>; etc.

b1) "log-transformed" why?

c) Authors' response: "a) a1) Potential correlations between outcomes: In logistic regression models, predictors were either continuous or categorical, and dependent variables were dichotomous. Our analysis did not specifically account for potential correlations between the outcomes. We acknowledge that correlations between outcomes could influence the results and will consider this in future analyses."'; "b1) Log-transformed UACR: The urine albumin-to-creatinine ratio (UACR) was log-transformed to normalise its distribution to address skewness. Log transformation helps in meeting the assumptions of the logistic regression model, particularly the assumption of linearity between the log odds of the outcome and the predictor variables."

d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

e) This reviewer recommends the authors to consider adding e1) in the Discussion section and e2) in the Methods section:

e1) A rigorous discussion of the authors' response to a) as a study limitation, including the potential implications to the validity and interpretation of the study results.

e2) A summary of the authors' response to b1) "[... to meet the linearity assumption of binary logistic regression]" or related.

f) Can relevant changes be implemented in the revised manuscript? If not, why?

g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

11. Is "Logistic regression" adequate for the data? {Peer review round 2: see b) to f)}:

a) Why?

a2) How was this tested?

b) Authors' response: "We believe logistic regression was appropriate for our analysis because it is well- suited for modelling the relationship between a predictor and outcomes, allowing us to estimate the probability of each outcome based on the predictor variable."; "We did not conduct formal statistical tests to specifically assess the adequacy of logistic regression for our data. However, the theoretical foundation and common use of logistic regression in similar contexts support its suitability for our analysis. We ensured that the model assumptions were met and the results were interpretable and consistent with our research objectives."

c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

d) This reviewer recommends the authors to consider adding d1) in in the Discussion section:

d1) A rigorous discussion of the authors' response to a) and a2) as study limitations, including the potential implications to the validity and interpretation of the study results.

e) Can relevant changes be implemented in the revised manuscript? If not, why?

f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

12. How were the models validated? {Peer review round 2: see a) to e)}:

a) Authors' response: "We did not conduct formal validation. However, we ensured that logistic regression was a theoretically appropriate choice for our predictor/s and outcomes. We acknowledge the importance of the validation steps. In future work, we will incorporate formal validation techniques to assess the robustness and generalizability of our models."

b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:

c1) A rigorous discussion of the authors' response as study limitations, including the potential implications to the validity and interpretation of the study results.

d) Can relevant changes be implemented in the revised manuscript? If not, why?

e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

13. Can the results of model calibration and variable selection be provided in the supplementary materials? If not, why? {Peer review round 2: see a) to e)}:

a) Authors' response: "We did not perform formal variable selection procedures in this study as we pre- specified models. As a result, we do not have detailed results from these processes to include in the supplementary materials. We acknowledge that variable selection enhance the robustness and interpretability of the models. We will consider this in our future work."

b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:

c1) A rigorous discussion of the authors' response as study limitations, including the potential implications to the validity and interpretation of the study results.

d) Can relevant changes be implemented in the revised manuscript? If not, why?

e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

14. Can the results of model goodness of fit be provided in the supplementary materials? If not, why? {Peer review round 2: see a) to e)}:

a) Authors' response: "We did not perform this analysis. However, as mentioned earlier, we ensured that the data were adequate for the logistic regression model and that assumptions were met. We appreciate an insightful comment. We will incorporate this, as well as the above-mentioned statistical issues, in our future work."

b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:

c1) A rigorous discussion of the authors' response as study limitations, including the potential implications to the validity and interpretation of the study results.

d) Can relevant changes be implemented in the revised manuscript? If not, why?

e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

15. "Categorical analyses were carried out in three categories of UACR (<30, 30-300, and ≥300 mg/g) and two categories of eGFR (>60 and <60 mL/min per 1.73 m2), where the UACR <30 mg/g and eGFR >60 mL/min per 1.73 m2 categories were used as the reference groups for comparison." {Peer review round 2: see c) to g)}:

a) What is the rationale of these and other categories created for this study?

b) Can relevant sources be provided to support these and other analytical choices?

c) Authors' response: "a) rationale for using UACR categories: Because the standard cutoff for normal UACR is 30mg/g, >30-300mg/g for microalbuminuria, and > 300mg/g for macroalbuminuria. we have cited the reference to support categorization. To aid readability, we have changed microalbuminuria to “moderately increased albuminuria” and macroalbuminuria to “severely increased albuminuria”."; "b) Rationale for using eGFR categories: due to very low number participants in some categories, we could not use more standard eGFR categories (G1-G5). Thus, we used the next best option, which was >60 mL/min per 1.73 m2, (normal kidney function with mild loss of function) and <60 mL/min per 1.73 m2 (suggestive of kidney function loss)."

d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

e) This reviewer recommends the authors to consider adding e1) and e2) in the Methods section:

e1) A summary of the authors' response to a) and b).

e2) Citation of relevant sources.

f) Can relevant changes be implemented in the revised manuscript? If not, why?

g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

16. How was data quality verified and ensured? Why? {Peer review round 2: see c) to g)}:

a) Who verified data quality?

b) Using what tools?

b1) Can these be described?

c) Authors' response: "We used data from the National Health and Nutrition Examination Survey (NHANES), which is a highly reputable and widely used dataset. The NHANES program is conducted by the National Center for Health Statistics (NCHS) and is known for its rigorous data collection and quality assurance procedures(https://www.cdc.gov/nchs/data/nhanes/nhanes_release_policy.pdf). More details about the data collection and process can be found in this pdf (https://www.cdc.gov/nchs/data/series/sr_01/sr01_056.pdf)under the “Data Collection and Processing” section. We have cited the source of information in the manuscript (ref #27, #28)."

d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

e) This reviewer recommends the authors to consider adding e1) and e2) in the Methods section:

e1) A summary of the authors' response to a), b), and b1).

e2) Citation of relevant sources.

f) Can relevant changes be implemented in the revised manuscript? If not, why?

g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

17. When were the data retrieved from the sources? {Peer review round 2: see a) to e)}:

a) Authors' response: "In November 2022."

b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:

c1) A summary of the authors' response.

d) Can relevant changes be implemented in the revised manuscript? If not, why?

e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

18. How were the data retrieved from the sources? {Peer review round 2: see c) to g)}:

a) By whom?

b) Using what tools?

b1) Can these be described?

c) Authors' response: "Data were downloaded directly from NHANES website and loaded into the R by the first author."

d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

e) This reviewer recommends the authors to consider adding e1) and e2) in the Methods section:

e1) A summary of the authors' response to a), b), and b1).

e2) Citation of relevant sources.

f) Can relevant changes be implemented in the revised manuscript? If not, why?

g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

19. "The Cox proportional hazards model was utilized to calculate hazard ratios for PRISm and variable obstructive lung function groups to examine the relationships between chronic kidney disease, lung function, and mortality." {Peer review round 2: see f) to k)}:

a) Were model assumptions (e.g., (i) proportional hazards, (ii) non-informative censoring, (iii) independent survival times for each observation, (iv) linear effect of predictor variables on the log hazard, (v) values of predictor variables for individuals constant over time, etc) verified? How? If not, why?

a1) Can the results be provided in the supplementary materials? If not, why?

c) What type of censoring was used? Why?

f) Authors' response: "a) a1) d): Thank you for highlighting this. We tested the COX assumption using Schoenfeld residuals. In all models, p-values were greater than 0.05 for all covariates as well as in the global test. We apologize for leaving out the important details. We added the following line: “We examined the proportional hazards assumption using Schoenfeld residuals; assumptions were satisfied in all models (global p-value >0.05)” Line 18-183. However, we as team feel that it is not necessary to add them to the supplementary material:"; "c): Right censored. because participants were assumed alive if there was no death record in the NDI/ death was not confirmed. Right-censoring allows us to account for participants whose complete survival time is not observed within the study period. This ensures that the analysis accurately reflects the survival experience of the cohort up to the point of censoring."

g) A response and relevant changes have been partially provided by the authors.

h) However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

i) This reviewer recommends the authors to consider adding i1), i2), and i3) in the Methods section and/or Supplementary materials, or i4) in the Discussion section:

i1) How the assumptions (ii) to (v) in a) were tested (note that the authors have provided details only on (i) PH assumption).

i2) A summary of the authors' response to c).

i3) Citation of relevant sources.

i4) A rigorous discussion as study limitations if assumptions (ii) to (v) were not tested, including the potential implications to the validity and interpretation of the study results.

j) Can relevant changes be implemented in the revised manuscript? If not, why?

k) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

20. "To analyze the data, two models were created: Model 1 included demographic variables such as age, sex, ethnicity, education, and annual household income, while Model 2 (full adjustment) further included body mass index, hypertension, smoking, cardiovascular disease, and diabetes mellitus." {Peer review round 2: see b) to f)}:

a) How was multicollinearity among predictors accounted for analytically? Why?

b) Authors' response: "Thank you for your insightful comment. In our analysis, we checked multicollinearity using car::vif() in R. in all models VIF was below 2, indicating very low multicollinearity."

c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

d) This reviewer recommends the authors to consider adding d1) in the Methods section and/or Supplementary materials:

d1) A summary of the authors' response.

e) Can relevant changes be implemented in the revised manuscript? If not, why?

f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

22. Can these and other applicable passages be rewritten to maximize clarity/rigor/conciseness/relevance? If not, why? {Peer review round 2: see d) to h)}:

a) "In the period between 2007 and 2012, the NHANES conducted pulmonary function tests on all adult participants. However, individuals experiencing chest pain, difficulties with forceful expiration, use of supplemental oxygen, recent surgeries on the eye, chest, or abdomen, recent heart attack, stroke, tuberculosis exposure, coughing up of blood, or a history of detached retina, collapsed lung, or aneurysm were excluded from the study(29)."

a1) E.g.: Then, how is the "data"" used in the study "representative survey of the non-institutionalized civilian population in the United States."

a2) E.g.: How was this accounted for analytically? Why?

d) Authors' response: "a1) a2): NHANES instructs researchers to account for complex survey designs and ensure a representative sample of the US population. It is done by using sample weights and adjusting for clusters and strata. Therefore, even though some participants are excluded from the examination, the final sample will still represent the US population. “In line with the instructions for using NHANES data, we used the sample weights, clustering, and stratification whenever feasible to account for the complex survey design using the “survey” R package (version 4.1-2)”"

e) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

f) This reviewer recommends the authors to consider adding f1) and f2) in the Methods section and/or Supplementary materials:

f1) A summary of the authors' response to a), a1), and a2).

f2) Citation of relevant sources.

g) Can relevant changes be implemented in the revised manuscript? If not, why?

h) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

27. How was multiple testing correction conducted? Why? {Peer review round 2: see b) to f)}:

a) E.g.: see: <https://doi.org/10.1038/nbt1209-1135>; <https://doi.org/10.1038/s41467-019-09941-0>; etc.

b) Authors' response: "We did not conduct multiple testing correction in our analysis. Given the specific research question and the exploratory nature of some analyses, we opted not to apply multiple testing correction methods. While multiple testing correction can be valuable in certain contexts, such as confirmatory analyses or when conducting a large number of hypothesis tests, we deemed it appropriate to present uncorrected p-values alongside careful interpretation of results in the context of our study objectives."

c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

d) This reviewer recommends the authors to consider adding d1) in the Discussion section:

d1) A rigorous discussion of the the authors' response as a study limitation, including the potential implications to the validity and interpretation of the study results.

e) Can relevant changes be implemented in the revised manuscript? If not, why?

f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

28. Was sensitivity analysis conducted? If not, why? {Peer review round 2: see b) to f)}:

a) Can the results be discussed in the main text and detailed in the supplementary materials? If not, why?

b) Authors' response: "We did not conduct the sensitivity analysis. for sensitivity analysis, we planned to further exclude participants with lung cancer to test the robustness of our result. However, we found that participants with lung cancer were extremely low, just 15-20 in one cycle (due to non- response to the question). Thus, we did not further analyse our data. However, we did analyse the all-cause mortality risk, while changing the reference category. Compared with PRISm without CKD, those with CKD had significantly higher risk of all-cause death, further indicating that having both PRISm and CKD compared to just PRISm further increases the mortality risk. The table below has been added as Supplementary Table 3: ... household income, BMI, hypertension, smoking, cardiovascular disease, diabetes mellitus"

c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

d) This reviewer recommends the authors to consider adding d1) in the Discussion section:

d1) A rigorous discussion of the the authors' response as a study limitation, including the potential implications to the validity and interpretation of the study results.

e) Can relevant changes be implemented in the revised manuscript? If not, why?

f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

29. What is the definition of "U.S. adults" in the study? Why? {Peer review round 2: see a) to f)}:

a) Authors' response: ""U.S. adults" refers to individuals who are citizens of the United States of America. This definition is used in our study because NHANES specifically recruits participants from the United States, including both adults and children. As our study focuses on adult participants, we referred to them as "U.S. adults”."

b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?

d) This reviewer recommends the authors to consider adding d1) and d2) in the Methods section and/or Supplementary materials:

d1) A summary of the authors' response taking into account d3).

d2) Citation of relevant sources.

d3) Comment on <https://doi.org/10.1161/CIRCOUTCOMES.119.006215> (3rd paragraph); etc.

e) Can relevant changes be implemented in the revised manuscript? If not, why?

f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?

30. Can these and other applicable passages be adjusted accordingly? If not, why?

IV.II. New concerns:

1. This reviewer recommends the authors to consider implementing relevant changes in the subheadings as follows:

a) Add "Assessment of ..." or related before subheadings c).

b) Change subheading d) to "Ascertainment of mortality status" or related.

c) "Pulmonary function", "Kidney function", "Confounders"

d) "Mortality data"

2. Is passage a) not part of the "Data source" subsection? If not, why?

a) "The National Center for Health Statistics (NCHS) has linked data from NHANES and other surveys with death certificate records from the National Death Index (NDI), and made

available the public-use linked mortality files (https://www.cdc.gov/nchs/data-linkage/mortality-public.htm, accessed June 5, 2024)"

3. Can these and other applicable passages be rewritten to maximize clarity/rigor/conciseness/relevance? If not, why?

a) "For the cross-sectional study, we used logistic regression models to assess the odds of PRISm (impaired spirometry) and variable obstructive lung function associated with per unit increase estimated glomerular filtration rate (eGFR) and log-transformed urine albumin-to-creatinine ratio (UACR)"

4. This reviewer recommends the authors to ensure that contents are provided in the appropriate subsections.

5. "We examined the proportional hazards assumption using Schoenfeld residuals; assumptions were satisfied in all models (global p-value >0.05)"

a) Can the actual p-value be provided here and in other applicable passages? If not, why?

6. Can these and other applicable passages be adjusted accordingly? If not, why?

---

## VERSION 2 - AUTHOR RESPONSE

Please find attached response document.

**Reviewer: 3**

**Dr. David Pascall, University of Cambridge**

**Comments to the Author:**

**1. Thank you for the hard work addressing the reviewers comments. I am satisfied this is ready for publication.**

**Response:**

We sincerely appreciate the time and effort you dedicated to reviewing our manuscript. Your insightful feedback has greatly contributed to strengthening our work. Thank you.

**Reviewer: 4**

**Dr. Floriano Amimo, Eduardo Mondlane University**

**Comments to the Author:**

**A. GENERAL COMMENTS**

**1. There is progress; however, there are still several important issues that need to be addressed.**

**2. Important data/details to support key statements and/or comply with applicable requirements/guidelines/etc. and/or address concerns raised are still not provided, at least not clearly.**

**3. There is important lack of clarity/rigor/consistency/conciseness/etc. in the manner that several passages/methods/contents/limitations/etc. are presented.**

**Response:**

1) 2) 3): We sincerely apologize if the revised lines and changes were not clearly indicated in our previous submission. In this revision, we have made a good effort to clearly mark all changes and provide their precise locations within the manuscript. In addition to page and line numbers, we have highlighted all the changes in yellow. We are encouraged by your acknowledgement of the progress made thus far.

We have carefully addressed the remaining issues you identified, ensuring that all necessary data and details are now clearly presented to support key statements and meet applicable guidelines.

Thanks to your recommendations, we were able to make revisions to improve the presentation of related passages.

We appreciate your thoughtful recommendations and the time you invested in reviewing our manuscript. We have revised the manuscript accordingly and hope these changes will satisfactorily address your remaining concerns and further enhance the quality of the work.

**B. SPECIFIC COMMENTS**

**I. Abstract:**

**I.I. Pending concerns:**

**3. What models/etc were used to "examine the relationship" and "analyze the impact"? Why? {Peer review round 2: see a) to e)}:**

**a) Authors' response: "Logistic regression (cross-sectional) and COX models (retrospective cohort). Revised abstract, Line 25-29"**

**b) Relevant changes have been partially implemented by the authors.**

**c) This reviewer recommends the authors to consider making applicable changes to indicate clearly for what design each model was used.**

**d) Can relevant changes be implemented in the revised manuscript? If not, why?**

**e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

c) d) In the latest revised abstract, we clearly mentioned in what design which models were used [see response to e)]

e) **Page 2, lines 27 to 31.**

"In the cross-sectional analysis, multivariate logistic regression models were used to assess the relationship between kidney function measures and spirometry findings. In the retrospective cohort analysis, Cox proportional hazards models were employed to evaluate the impact of having PRISm or VO, combined with CKD, on all-cause mortality."

**4. What are the main outcome measures? Why? {Peer review round 2: see a) to f)}:**

**a) Authors' response: "Revised abstract, Line 22-25"**

**b) Relevant changes have been partially implemented by the authors.**

**c) This reviewer recommends the authors to consider separating exposures from outcomes for clarity, e.g., by creating a separate "Exposures" subsection.**

**d) Additionally, no clear distinction is made between "Primary" and "secondary" outcomes; then why write "Primary and secondary outcome measures"?**

**e) Can relevant changes be implemented in the revised manuscript? If not, why?**

**f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

c) d) e) We apologize for the confusion. We removed "Primary and secondary outcomes" and created a section "Exposure and Outcome Measures". Under that, exposures and outcomes are mentioned first, followed by definitions of spirometry findings and statistical methods used in each study design.

f) **Page 2, lines 23 to 31.**

Here is the revised portion: "Kidney function measures, including estimated glomerular filtration rate (eGFR) and urinary albumin to creatinine ratio (UACR), were considered exposure variables. PRISm and VO were outcome variables, with PRISm defined as an FEV1 <80% predicted and an FEV1/FVC ratio ⩾0.7, and VO defined as an FEV1/FVC ratio <0.7 pre-bronchodilator and ⩾0.7 post-bronchodilator. In the cross-sectional analysis, multivariate logistic regression models were used to assess the relationship between kidney function measures and spirometry findings. In the retrospective cohort analysis, Cox proportional hazards models were employed to evaluate the impact of having PRISm or VO, combined with CKD, on all-cause mortality."

**10. Can these and other applicable passages be adjusted accordingly? If not, why?**

**Response:**

Relevant passages have been revised according to the recommendations above. Thank you. We have included the location of the changes as well as relevant revised text under our responses.

**I.II. New concerns:**

**1. Can these and other applicable passages be rewritten to maximize clarity/rigor/conciseness? If not, why?**

**a) "(normal spirometry= 9503, PRISm=951, variable obstruction=355)"**

**2. Can these and other applicable passages be adjusted accordingly? If not, why?**

**Response:**

1a) 2) We have revised the "Participants" section of the abstract as requested, which clearly and correctly lists the participants and their respective categories. Changes were implemented on **page 2, lines 20 to 22** as following:

"A total of 10,809 participants aged over 20 years were included in this study: 9,503 with normal spirometry, 951 with preserved ratio impaired spirometry (PRISm), and 355 with variable obstruction (VO)"

**IV. Methods:**

**IV.I. Pending concerns:**

**1. Which reporting guideline does this study comply with? Why? {Peer review round 2: see c) to g)}:**

**b) Can the checklist of the reporting guideline indicating location of each item in the main text be provided in the supplementary materials? If not, why?**

**c) Authors' response: "STROBE. We have attached the STROBE statement"**

**d) Relevant changes have been partially implemented by the authors.**

**e) This reviewer recommends the authors to consider adding a passage that indicates clearly that "This study complies with ..." or similar and cite the Checklist.**

**f) Can relevant changes be implemented in the revised manuscript? If not, why?**

**g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

e) f) Thank you for pointing this out. We added the checklist in the supplemental table file. the following passage was added and the STROBE reference and checklist were cited:

"In addition, this study complies with the guidelines for reporting cross-sectional studies as

specified in the Strengthening the Reporting of Observational Studies in Epidemiology guidelines (Supplemental Table 1).(31)"

g) **Page 6, lines 115 to 117**

**2. What is the study design? {Peer review round 2: see a) to e)}:**

**a) Authors' response: "Cross-sectional and retrospective cohort study."**

**b) Relevant changes have been partially implemented by the authors.**

**c) This reviewer recommends the authors to consider adding relevant details as follows:**

**c1) Change "Data source" to "Study design and data source" or create a separate "Study design" subsection.**

**c2) Indicate clearly to what end each design was used.**

**d) Can relevant changes be implemented in the revised manuscript? If not, why?**

**e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

c1) c2) d) We realised that "Study design and participants" would be better. In the latest manuscript, this subsection introduces NHANES design (with references) followed by participant selection.

We think that the latest manuscript "Study design and participants" is clear and concise. We clearly described the selection of study sample used for cross-sectional and cohort analysis and cited Figure 1 which illustrates the same but in detail (**see page 5 lines 110 to 111 and page 6, lines 112 -114**). Moreover, a detailed description of the statistical analysis done in each design is described under the "statistical analysis" subsection. **Page 8 lines 176 to 180** ("In the cross-sectional analysis…"), and **page 8 lines 184 to 188** ("In the retrospective cohort analysis…")

e) Page and line numbers are provided above.

**6. "The spirometry data were classified based on the quality of data collection, with "A" being the highest quality and exceeding American Thoracic Society standards, "B" meeting the standards, and "C" being potentially usable but not meeting all the standards. The sample selection process is detailed in Figure 1.": {Peer review round 2: see c) to g)}:**

**a) What is the difference between "the standards" and "all the standards"? Why?**

**b) How was "quality of data collection" determined/measured? Why?**

**c) Authors' response: "Difference between the standards: NHANES followed ATS criteria and standardisation of spirometry (Miller, Hankinson, Brusasco et al. 2005) ….."**

**d) Relevant changes have been partially implemented by the authors.**

**e) This reviewer recommends the authors to consider adding relevant details as follows:**

**e1) Add a brief description, just a few words, of what they mean with "standards" and "all the standards" for clarity, e.g., by replacing X and Y as follows: "[... "B" meeting the standards (that is, X) ... "C" ... not meeting all the standards (that is, Y), ...]".**

**e2) Cite , , and other URLs provided in the main text in line with standard citation and referencing rules (to mitigate the effect of link rot in the main text, etc.).**

**f) Can relevant changes be implemented in the revised manuscript? If not, why?**

**g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

e1) revised according to the recommendation. Thank you. See response to g)

e2) Duly revised

g) **Page 6, line 137 to l38**. All URLs are now cited with citation and referencing rules.

here is the relevant revised portion: "…."B" meeting the standards (that is, adequate technical quality and reproducibility), "C" being potentially usable but not meeting all the standards (that is, some technical issues or lack of reproducibility)…."

**9. "Participants were considered hypertensive if they had systolic blood pressure of 140 mm Hg or higher and diastolic blood pressure of 90 mm Hg or higher, or ...": {Peer review round 2: see c) to g)}:**

**a) Under which conditions? On a single day?**

**a1) E.g.: see: ?**

**b) What guideline does the approach used to diagnose hypertension comply with? Why?**

**c) Authors' response: ""After resting quietly in a seated position for 5 minutes …"**

**d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**e) This reviewer recommends the authors to consider adding d1) in the revised manuscript and/or supplementary materials as applicable and citing relevant sources:**

**e1) A summary of their response to this reviewer.**

**f) Can relevant changes be implemented in the revised manuscript? If not, why?**

**g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

e) e1) f) As suggested, we have incorporated our previous response and the source was cited.

g) **Page 7, lines 162 to 168**

"In a mobile examination center (MEC), three consecutive blood pressure readings were taken. The average of the readings was used in this study. Participants were considered hypertensive if they had a systolic blood pressure of 140 mm Hg or higher and diastolic blood pressure of 90 mm Hg or higher or if they were using anti-hypertensive medication. The 140/90 mmHg threshold is based on the International Society of Hypertension guidelines(41), which are commonly applied in clinical practice."

**10. "Logistic regression was employed to assess the odds of PRISm (impaired spirometry) and variable obstructive lung function given declining estimated glomerular filtration rate (eGFR) and log-transformed urine albumin-to-creatinine ratio (UACR).": {Peer review round 2: see c) to g)}:**

**a) How was the potential correlations between these outcomes accounted for analytically? If not, Why?**

**a1) E.g.: see: ; etc.**

**b1) "log-transformed" why?**

**c) Authors' response: "a) a1) Potential correlations between outcomes: In logistic regression models, predictors were either continuous or categorical, and dependent variables were dichotomous. Our analysis did not specifically account for potential correlations between the outcomes. We acknowledge that correlations between outcomes could influence the results and will consider this in future analyses."'; "b1) Log-transformed UACR: The urine albumin-to-creatinine ratio (UACR) was log-transformed to normalise its distribution to address skewness. Log transformation helps in meeting the assumptions of the logistic regression model, particularly the assumption of linearity between the log odds of the outcome and the predictor variables."**

**d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**e) This reviewer recommends the authors to consider adding e1) in the Discussion section and e2) in the Methods section:**

**e1) A rigorous discussion of the authors' response to a) as a study limitation, including the potential implications to the validity and interpretation of the study results.**

**e2) A summary of the authors' response to b1) "[... to meet the linearity assumption of binary logistic regression]" or related.**

**f) Can relevant changes be implemented in the revised manuscript? If not, why?**

**g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

e1) We apologise for the confusion in our previous response. What we intended to convey was that we utilized logistic regression with a binary outcome/dependent variable (e.g., PRISm vs. No PRISm). As such, testing for correlation between these binary outcomes is not applicable in this context. When we mentioned acknowledging the suggestion, we meant that it is something we should consider for future work. We sincerely apologize for any misunderstanding and for not making this point clearer earlier.

e2) This was revised according to the recommendation. Please see the response to g) below.

g) **Page 8, line 179-180 for e2)**

"UACR was log-transformed to address skewness and to meet the assumptions of the logistic regression model"

**11. Is "Logistic regression" adequate for the data? {Peer review round 2: see b) to f)}:**

**a) Why?**

**a2) How was this tested?**

**b) Authors' response: "We believe logistic regression was appropriate for our analysis because …"**

**c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**d) This reviewer recommends the authors to consider adding d1) in in the Discussion section:**

**d1) A rigorous discussion of the authors' response to a) and a2) as study limitations, including the potential implications to the validity and interpretation of the study results.**

**e) Can relevant changes be implemented in the revised manuscript? If not, why?**

**f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

d1) e) This, along with other statistical limitations was discussed in the Discussion section. While we did not explicitly mention “adequacy”, the limitation of goodness of fit covers this issue.

f) **Page 14, lines 356 to 359.**

here is the relevant passage: “From a statistical point of view, some limitations should also be acknowledged. In logistic regression models, we did not conduct formal assessments of validity, variable selection, or goodness of fit. This may influence our findings' robustness and generalizability, leading to potential biases. However, this method has been used in similar NHANES studies.”

**12. How were the models validated? {Peer review round 2: see a) to e)}:**

**a) Authors' response: "We did not conduct formal validation. However, we ensured that logistic regression was a theoretically appropriate …."**

**b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:**

**c1) A rigorous discussion of the authors' response as study limitations, including the potential implications to the validity and interpretation of the study results.**

**d) Can relevant changes be implemented in the revised manuscript? If not, why?**

**e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

c1) d) This, along with other statistical limitations was discussed in the Discussion section

e) **Page 14, lines 356 to 359.**

here is the relevant passage: “From a statistical point of view, some limitations should also be acknowledged. In logistic regression models, we did not conduct formal assessments of validity, variable selection, or goodness of fit. This may influence our findings' robustness and generalizability, leading to potential biases. However, this method has been used in similar NHANES studies.”

**13. Can the results of model calibration and variable selection be provided in the supplementary materials? If not, why? {Peer review round 2: see a) to e)}:**

**a) Authors' response: "We did not perform formal variable selection procedures in this study as we pre- specified models …"**

**b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:**

**c1) A rigorous discussion of the authors' response as study limitations, including the potential implications to the validity and interpretation of the study results.**

**d) Can relevant changes be implemented in the revised manuscript? If not, why?**

**e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

c1) d) This, along with other statistical limitations was discussed in the Discussion section

e) **Page 14, lines 356 to 359.**

here is the relevant passage: "From a statistical point of view, some limitations should also be acknowledged. In logistic regression models, we did not conduct formal assessments of validity, variable selection, or goodness of fit. This may influence our findings' robustness and generalizability, leading to potential biases. However, this method has been used in similar NHANES studies."

**14. Can the results of model goodness of fit be provided in the supplementary materials? If not, why? {Peer review round 2: see a) to e)}:**

**a) Authors' response: "We did not perform this analysis. However, as mentioned earlier…"**

**b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:**

**c1) A rigorous discussion of the authors' response as study limitations, including the potential implications to the validity and interpretation of the study results.**

**d) Can relevant changes be implemented in the revised manuscript? If not, why?**

**e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

c1) d) This, along with other statistical limitations was discussed in the Discussion section

e) **Page 14, lines 356 to 359.**

here is the relevant passage: "From a statistical point of view, some limitations should also be acknowledged. In logistic regression models, we did not conduct formal assessments of validity, variable selection, or goodness of fit. This may influence our findings' robustness and generalizability, leading to potential biases. However, this method has been used in similar NHANES studies."

**15. "Categorical analyses were carried out in three categories of UACR (<30, 30-300, and ≥300 mg/g) and two categories of eGFR (>60 and <60 mL/min per 1.73 m2), where the UACR <30 mg/g and eGFR >60 mL/min per 1.73 m2 categories were used as the reference groups for comparison." {Peer review round 2: see c) to g)}:**

**a) What is the rationale of these and other categories created for this study?**

**b) Can relevant sources be provided to support these and other analytical choices?**

**c) Authors' response: "a) rationale for using UACR categories: Because the standard cutoff for normal UACR is 30mg/g, …"**

**d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**e) This reviewer recommends the authors to consider adding e1) and e2) in the Methods section:**

**e1) A summary of the authors' response to a) and b).**

**e2) Citation of relevant sources.**

**f) Can relevant changes be implemented in the revised manuscript? If not, why?**

**g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

e1) e2) f) We included the rationale for using categories that we used and cited the source. Thank you.

g) **Page 6 lines 123-124 and line 127. reference #34**

"The UACR was further categorized into three groups based on the standard cutoffs…"
"We could not use standard eGFR categories (G1-G5) due to insufficient data"

**16. How was data quality verified and ensured? Why? {Peer review round 2: see c) to g)}:**

**a) Who verified data quality?**

**b) Using what tools?**

**b1) Can these be described?**

**c) Authors' response: "We used data from the National Health and Nutrition Examination Survey (NHANES) …"**

**d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**e) This reviewer recommends the authors to consider adding e1) and e2) in the Methods section:**

**e1) A summary of the authors' response to a), b), and b1).**

**e2) Citation of relevant sources.**

**f) Can relevant changes be implemented in the revised manuscript? If not, why?**

**g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

e1) We have already described the NHANES design with sources for detailed information (see page 5 lines 99-100). However, we have added a line to support data quality and validity {see response to g)}

e2) In addition, a source related to NHANES plan and operations has been cited (reference #28)

f) duly revised

g) **Page 5, line 107**

"Moreover, NHANES data has been used in many studies, proving its validity and quality. (25, 29-31)"

**17. When were the data retrieved from the sources? {Peer review round 2: see a) to e)}:**

**a) Authors' response: "In November 2022."**

**b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**c) This reviewer recommends the authors to consider adding c1) in in the Discussion section:**

**c1) A summary of the authors' response.**

**d) Can relevant changes be implemented in the revised manuscript? If not, why?**

**e) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

c1) d) We believe that mentioning the time of data retrieval in the Discussion section is unnecessary. To enhance clarity, we have added the e) in the Methods section, along with the appropriate source citation.

e) **Page 7, lines 173-174**

"Participants' data were downloaded from the NHANES website in November 2022 by the first author."

**18. How were the data retrieved from the sources? {Peer review round 2: see c) to g)}:**

**a) By whom?**

**b) Using what tools?**

**b1) Can these be described?**

**c) Authors' response: "Data were downloaded directly from NHANES website and loaded into the R by the first author."**

**d) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**e) This reviewer recommends the authors to consider adding e1) and e2) in the Methods section:**

**e1) A summary of the authors' response to a), b), and b1).**

**e2) Citation of relevant sources.**

**f) Can relevant changes be implemented in the revised manuscript? If not, why?**

**g) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

e) f) As mentioned above, we have added the g) in the Methods section, along with the appropriate source citation.

g) **Page 7, lines 173-174**

"Participants' data were downloaded from the NHANES website in November 2022 by the first author."

**19. "The Cox proportional hazards model was utilized to calculate hazard ratios for PRISm and variable obstructive lung function groups to examine the relationships between chronic kidney disease, lung function, and mortality." {Peer review round 2: see f) to k)}:**

**a) Were model assumptions (e.g., (i) proportional hazards, (ii) non-informative censoring, (iii) independent survival times for each observation, (iv) linear effect of predictor variables on the log hazard, (v) values of predictor variables for individuals constant over time, etc) verified? How? If not, why?**

**a1) Can the results be provided in the supplementary materials? If not, why?**

**c) What type of censoring was used? Why?**

**f) Authors' response: "a) a1) d): Thank you for highlighting this. We tested the COX assumption using Schoenfeld residuals. …"**

**g) A response and relevant changes have been partially provided by the authors.**

**h) However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**i) This reviewer recommends the authors to consider adding i1), i2), and i3) in the Methods section and/or Supplementary materials, or i4) in the Discussion section:**

**i1) How the assumptions (ii) to (v) in a) were tested (note that the authors have provided details only on (i) PH assumption).**

**i2) A summary of the authors' response to c).**

**i3) Citation of relevant sources.**

**i4) A rigorous discussion as study limitations if assumptions (ii) to (v) were not tested, including the potential implications to the validity and interpretation of the study results.**

**j) Can relevant changes be implemented in the revised manuscript? If not, why?**

**k) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

a1) results of Schoenfeld residuals of all models has been added to the supplementary materials {see supplemental table 5, 6 or response to the comment IV. II. 5}

i) i2) i4) Since we did not perform statistical tests for these, we have recognized this limitation and described the potential impact on the study findings. Although we tested for the fundamental assumption, which is proportionality.

j) k) **Page 14, lines 359 to 361**

“In COX regression, we only tested the fundamental assumption of proportional hazard. Violations of other assumptions could lead to biased estimates and affect the validity of findings.”

**20. "To analyze the data, two models were created: Model 1 included demographic variables such as age, sex, ethnicity, education, and annual household income, while Model 2 (full adjustment) further included body mass index, hypertension, smoking, cardiovascular disease, and diabetes mellitus." {Peer review round 2: see b) to f)}:**

**a) How was multicollinearity among predictors accounted for analytically? Why?**

**b) Authors' response: "Thank you for your insightful comment. In our analysis, we checked multicollinearity using car::vif() in R. in all models VIF was below 2, indicating very low multicollinearity."**

**c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**d) This reviewer recommends the authors to consider adding d1) in the Methods section and/or Supplementary materials:**

**d1) A summary of the authors' response.**

**e) Can relevant changes be implemented in the revised manuscript? If not, why?**

**f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

d) d1) e) Thank you. We included this part in the methods section

f) **Page 8, lines 195 to 197**

“In regression models, we checked multicollinearity using vif() function of “car” package (version 3.1-2) in R, where the variance inflation factor (VIF) was below 2, indicating very low multicollinearity.”

**22. Can these and other applicable passages be rewritten to maximize clarity/rigor/conciseness/relevance? If not, why? {Peer review round 2: see d) to h)}:**

**a) "In the period between 2007 and 2012, the NHANES conducted pulmonary function tests on all adult participants. However, individuals experiencing chest pain, difficulties with forceful expiration, use of supplemental oxygen, recent surgeries on the eye, chest, or abdomen, recent heart attack, stroke, tuberculosis exposure, coughing up of blood, or a history of detached retina, collapsed lung, or aneurysm were excluded from the study(29)."**

**a1) E.g.: Then, how is the "data"" used in the study "representative survey of the non-institutionalized civilian population in the United States."**

**a2) E.g.: How was this accounted for analytically? Why?**

**d) Authors' response: "a1) a2): NHANES instructs researchers …"**

**e) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**f) This reviewer recommends the authors to consider adding f1) and f2) in the Methods section and/or Supplementary materials:**

**f1) A summary of the authors' response to a), a1), and a2).**

**f2) Citation of relevant sources.**

**g) Can relevant changes be implemented in the revised manuscript? If not, why?**

**h) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

f1) f2) we apologize for not highlighting the relevant passage. In the methods section, we had mentioned that we used sample weights etc. per NHANES analytical guidelines with source citation (page 8, line 169). We have added a few words to reflect our previous response {see response to g) h)}.

g) h) **Page 7-8, lines 172 to 173**

"This approach allows the **final sample** to be **representative of the U.S. population**, **despite exclusions**."

**27. How was multiple testing correction conducted? Why? {Peer review round 2: see b) to f)}:**

**a) E.g.: see: ; ; etc.**

**b) Authors' response: "We did not conduct multiple testing correction in our analysis. Given the specific research question ...."**

**c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**d) This reviewer recommends the authors to consider adding d1) in the Discussion section:**

**d1) A rigorous discussion of the the authors' response as a study limitation, including the potential implications to the validity and interpretation of the study results.**

**e) Can relevant changes be implemented in the revised manuscript? If not, why?**

**f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

d) d1) Thank you. we have added this in limitations {see e) f)}

e) f) **Page 14, lines 361 to 363**

"Finally, we did not apply multiple testing correction, which may impact the interpretability of findings. As our analysis was exploratory, we presented uncorrected p-values, aligning with our study objectives"

**28. Was sensitivity analysis conducted? If not, why? {Peer review round 2: see b) to f)}:**

**a) Can the results be discussed in the main text and detailed in the supplementary materials? If not, why?**

**b) Authors' response: "We did not conduct the sensitivity analysis. for sensitivity analysis, we planned to …"**

**c) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**d) This reviewer recommends the authors to consider adding d1) in the Discussion section:**

**d1) A rigorous discussion of the the authors' response as a study limitation, including the potential implications to the validity and interpretation of the study results.**

**e) Can relevant changes be implemented in the revised manuscript? If not, why?**

**f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

d) d1) We have included this issue in limitations. Thank you. please see response to e) f) below

e) f) **Page 14, lines 352 to 355**

"Moreover, we could not conduct the planned sensitivity analysis to exclude participants with lung cancer due to high non-response rates. However, we did assess the robustness of our findings by analysing all-cause mortality risk with alternative reference categories (supplemental table 4)."

**29. What is the definition of "U.S. adults" in the study? Why? {Peer review round 2: see a) to f)}:**

**a) Authors' response: "U.S. adults" refers ..."**

**b) A response has been partially provided by the authors. However, this reviewer was unable to locate any change in the revised manuscript aimed at addressing this issue fully. Why?**

**d) This reviewer recommends the authors to consider adding d1) and d2) in the Methods section and/or Supplementary materials:**

**d1) A summary of the authors' response taking into account d3).**

**d2) Citation of relevant sources.**

**d3) Comment on (3rd paragraph); etc.**

**e) Can relevant changes be implemented in the revised manuscript? If not, why?**

**f) Can the page and line numbers where the change implemented by the authors in response to this specific concern can be efficiently located by the reviewers in the revised version be provided clearly? If not, why?**

**Response:**

d) d1) d2) d3) there was only one place where the term "U.S. adults" was used. We changed it to "U.S. population" for clarity (page 13 line 341). Nonetheless, we have cleared this issue by improving the relevant passage {see response to f)}. This will help a reader understand that our study population come from United States.

f) **Page 5, lines 99 to 101**

"This was a cross-sectional and retrospective cohort study using the National Health and Nutrition Examination Survey (NHANES), a national survey of children and adults in the United States conducted by the National Center for Health Statistics of the Centers for Disease Control and Prevention."

**30. Can these and other applicable passages be adjusted accordingly? If not, why?**

**Response:**

We are grateful for the reviewers' crucial recommendations. We appreciate the hard work the reviewer has put into reviewing our manuscript. We considered all the recommendations and revised the applicable passages accordingly. We hope this round addressed all of the reviewers' concerns. Thank you.

**IV.II. New concerns:**

**1. This reviewer recommends the authors to consider implementing relevant changes in the subheadings as follows:**

**a) Add "Assessment of ..." or related before subheadings c).**

**b) Change subheading d) to "Ascertainment of mortality status" or related.**

**c) "Pulmonary function", "Kidney function", "Confounders"**

**d) "Mortality data"**

**Response:**

a) b) c) Duly revised accordingly

**2. Is passage a) not part of the "Data source" subsection? If not, why?**

**a) "The National Center for Health Statistics (NCHS) has linked data from NHANES and other surveys with death certificate records from the National Death Index (NDI), and made available the public-use linked mortality files (https://www.cdc.gov/nchs/data-linkage/mortality-public.htm, accessed June 5, 2024)"**

**Response:**

We agree with you. However, in light of other concerns, in the latest manuscript "data source" was changed to "Study design and participants". Since there is a separate subsection for mortality ascertainment, details regarding mortality status should be in this subsection. Therefore, a) fits better under the "Ascertainment of mortality status".

**3. Can these and other applicable passages be rewritten to maximize clarity/rigor/conciseness/relevance? If not, why?**

**a) "For the cross-sectional study, we used logistic regression models to assess the odds of PRISm (impaired spirometry) and variable obstructive lung function associated with per unit increase estimated glomerular filtration rate (eGFR) and log-transformed urine albumin-to-creatinine ratio (UACR)"**

**Response:**

We agree that a) could be written better. Thus, we have revised the relevant passage. Please see **page 8, lines 176 to 179**. Thank you.

"In the cross-sectional analysis, we employed multivariate logistic regression models to evaluate the association between kidney function indicators—including eGFR (per unit increase) and log-transformed UACR—and spirometry outcomes, specifically PRISm and VO."

**4. This reviewer recommends the authors to ensure that contents are provided in the appropriate subsections.**

**Response:**

We apologise for the oversight. As mentioned above, we changed “Data source” to “Study design and participants”, which briefly introduces NHANES design with adequate references for further details, specific NHANES data cycles used and why, and the population selection process. **Please see page 5, lines 98 to 117**. Then the method section goes on to describe kidney function assessment, pulmonary function assessment, confounders details, and mortality status ascertainment. Finally, we improved the statistical analyses section by creating paragraphs according to content to improve readability and clarity.

**5. "We examined the proportional hazards assumption using Schoenfeld residuals; assumptions were satisfied in all models (global p-value >0.05)"**

**a) Can the actual p-value be provided here and in other applicable passages? If not, why?**

**Response:**

a) We have provided the output of PH test below. however, we believe it will be too lengthy to report p-value in the manuscript. Therefore, we have included the p-values of the models below and in the supplementary material.

Output of Schoenfeld residual test:

Supplemental Table 5. Proportion hazard assumption test (Schoenfeld residuals) of Table 1

| Entire Cohort Model 1 | | | | Entire Cohort Model 2 | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Chi-Square** | **df** | **p-value** | **Variable** | **Chi-Square** | **df** | **p-value** |
| spiro | 0.000405 | 2 | 1 | spiro | 4.17E-04 | 2 | 1 |
| age | 0.001312 | 1 | 0.97 | age | 1.51E-03 | 1 | 0.97 |
| sex | 0.000316 | 1 | 0.99 | sex | 3.10E-04 | 1 | 0.99 |
| eth | 0.000134 | 4 | 1 | eth | 1.09E-04 | 4 | 1 |
| edu | 0.00053 | 2 | 1 | edu | 5.92E-04 | 2 | 1 |
| AHI | 0.00047 | 1 | 0.98 | AHI | 4.94E-04 | 1 | 0.98 |
| GLOBAL | 0.002716 | 11 | 1 | bmi | 3.16E-05 | 1 | 1 |
| | | | | Hypertension | 5.12E-04 | 1 | 0.98 |
| | | | | smoke1 | 1.67E-04 | 2 | 1 |
| | | | | CVD | 3.73E-04 | 1 | 0.98 |
| | | | | DM | 1.16E-04 | 1 | 0.99 |
| | | | | GLOBAL | 3.15E-03 | 17 | 1 |

| Diabetes Cohort Model 1 | | | | Diabetes Cohort Model 2 | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Chi-Square** | **df** | **p-value** | **Variable** | **Chi-Square** | **df** | **p-value** |
| spiro | 2.50E-03 | 2 | 1 | spiro | 2.77E-03 | 2 | 1 |
| age | 1.13E-09 | 1 | 1 | age | 2.16E-07 | 1 | 1 |
| sex | 2.07E-06 | 1 | 1 | sex | 1.12E-06 | 1 | 1 |
| eth | 1.34E-03 | 4 | 1 | eth | 1.12E-03 | 4 | 1 |
| edu | 1.22E-03 | 2 | 1 | edu | 1.37E-03 | 2 | 1 |
| AHI | 4.34E-05 | 1 | 0.99 | AHI | 1.03E-04 | 1 | 0.99 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| GLOBAL | 4.95E-03 | 11 | 1 | bmi | 5.26E-04 | 1 | 0.98 |
| | | | | Hypertension | 1.71E-03 | 1 | 0.97 |
| | | | | smoke1 | 5.61E-04 | 2 | 1 |
| | | | | CVD | 1.37E-04 | 1 | 0.99 |
| | | | | GLOBAL | 8.68E-03 | 16 | 1 |

Supplemental Table 6. Proportion hazard assumption test (Schoenfeld residuals) of supplemental table 3

| Entire Cohort Model 1 | | | | Entire Cohort Model 2 | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Chi-Square** | **df** | **p-value** | **Variable** | **Chi-Square** | **df** | **p-value** |
| CKDspiro | 0.000616 | 5 | 1.00 | CKDspiro | 6.53E-04 | 5 | 1.00 |
| age | 0.001299 | 1 | 0.97 | age | 1.48E-03 | 1 | 0.97 |
| sex | 0.000296 | 1 | 0.99 | sex | 3.05E-04 | 1 | 0.99 |
| eth | 0.000122 | 4 | 1.00 | eth | 1.01E-04 | 4 | 1.00 |
| edu | 0.000555 | 2 | 1.00 | edu | 5.98E-04 | 2 | 1.00 |
| AHI | 0.000509 | 1 | 0.98 | AHI | 5.18E-04 | 1 | 0.98 |
| GLOBAL | 0.00297 | 14 | 1.00 | bmi | 4.18E-05 | 1 | 0.99 |
| | | | | Hypertension | 5.50E-04 | 1 | 0.98 |
| | | | | smoke | 1.68E-04 | 2 | 1.00 |
| | | | | CVD | 4.06E-04 | 1 | 0.98 |
| | | | | DM | 1.32E-04 | 1 | 0.99 |
| | | | | GLOBAL | 3.38E-03 | 20 | 1.00 |

| Diabetes Cohort Model 1 | | | | Diabetes Cohort Model 2 | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Chi-Square** | **df** | **p-value** | **Variable** | **Chi-Square** | **df** | **p-value** |
| CKDspiro | 2.45E-03 | 5 | 1.00 | CKDspiro | 2.82E-03 | 5 | 1.00 |
| age | 1.52E-08 | 1 | 1.00 | age | 2.84E-07 | 1 | 1.00 |
| sex | 1.81E-06 | 1 | 1.00 | sex | 3.25E-06 | 1 | 1.00 |
| eth | 1.27E-03 | 4 | 1.00 | eth | 1.03E-03 | 4 | 1.00 |
| edu | 1.38E-03 | 2 | 1.00 | edu | 1.29E-03 | 2 | 1.00 |
| AHI | 1.14E-04 | 1 | 0.99 | AHI | 2.04E-04 | 1 | 0.99 |
| GLOBAL | 5.06E-03 | 14 | 1.00 | bmi | 5.89E-04 | 1 | 0.98 |
| | | | | Hypertension | 1.78E-03 | 1 | 0.97 |
| | | | | smoke | 5.45E-04 | 2 | 1.00 |
| | | | | CVD | 1.63E-04 | 1 | 0.99 |
| | | | | GLOBAL | 8.85E-03 | 19 | 1.00 |

**6. Can these and other applicable passages be adjusted accordingly? If not, why?**

**Response:**

Thank you for your valuable suggestions. We have carefully addressed the new concerns and made the necessary revisions to the paper.

Reviewer: 3

Competing interests of Reviewer: NA

Reviewer: 4

Competing interests of Reviewer: None