

# Peer Review File

Article Information: <https://dx.doi.org/10.21037/tlcr-21-601>

## Reviewer A

**Comment 1:** Page 3, line 1: the authors have listed high body mass index as one of the risk factors of lung cancer, although current evidence does not support this association. Please see the IARC (DOI: 10.1056/NEJMs1606602) and WCRF/AICR ([https://www.wcrf.org/wp-content/uploads/2021/01/Body-fatness-and-weight-gain\\_0.pdf](https://www.wcrf.org/wp-content/uploads/2021/01/Body-fatness-and-weight-gain_0.pdf)) reports.

**Reply:** Thank you for highlighting that high BMI is not a risk factor of lung cancer and the IARC references that mention that evidence for an association between BMI and lung cancer is inadequate due to insufficient data or inconsistent results. In the PLCO dataset (as well as other epidemiologic studies), low BMI has been associated with lung cancer incidence and obesity has a paradoxical role in lung cancer. We have changed the language in the paragraph to say that we included BMI and the other risk factors as covariates in our model because they have been shown to be associated with lung cancer risk in the PLCO dataset specifically.

**Changes in text:** Page 5 Lines 19-20

**Comment 2:** Page 3, lines 4-7: please revise the statement on results of the meta-analysis, because in that study, there was no statistically significant difference between the risk among men (RR: 1.11; 95%CI: 0.92-1.35) and women (RR: 1.18; 95%CI: 1.10-1.28), or between individuals with a BMI of  $\geq 25$  kg/m<sup>2</sup> (1.12; 95%CI: 1.01-1.25) and with a BMI  $< 25$  kg/m<sup>2</sup> (1.08; 95%CI: 0.99-1.18), or by proportion of current smokers in individual cohorts (RR: 1.16; 95%CI: 1.04-1.28 vs. 1.23; 95%CI: 1.04-1.44 in cohorts with  $< 20\%$  and  $\geq 20\%$  current smoking prevalence, respectively). Perhaps a more important result from that meta-analysis was that there was no association in high-quality studies (RR: 1.03; 95%CI: 0.99-1.23), but the association was significant in studies labeled as low-quality studies (1.18; 95%CI: 1.06-1.31).

**Reply:** We have revised the language to more accurately present the findings of the meta-analysis and have also highlighted the interesting finding that an association between diabetes and lung cancer was found in low-quality studies, but not high-quality studies.

**Changes in Text:** Page 3 lines 16-23, Page 4 Lines 1-4.

**Comment 3:** Page 4, line 39: please replace “5 pack years” with “5 pack-years”.

**Reply:** We have made the correction in the text.

**Changes in the text:** Page 7 Line 8

**Comment 4:** Page 4, line 40 to page 5, line 8: no need to highlight unadjusted results. About 80% of lung cancer are attributable to known risk factors, so without adjusting for those risk factors, any association can be spurious.

**Reply:** We now only present the unadjusted results for the main study question and deleted the stratified analyses.

**Changes in the text: Page 7 Lines 10-17.**

**Comment 5:** Page 5, lines 9-20: this paragraph can be shortened.

**Reply:** We have shortened this paragraph by modifying the description of the data from the stratified analyses.

**Changes in the text: Page 7 lines 10-17.**

**Comment 6:** Page 5, paragraph starting from line 21 (and the corresponding section in the Methods): it is unclear to me why the authors have chosen lung adenocarcinoma as the reference group; why not simply consider PLCO participants without a diagnosis of lung cancer as the reference group, and compare histological subtypes of lung cancer (including adenocarcinoma) with this group? However, it may not matter much here, as in this study, diabetes is not associated with lung cancer, and there is no statistically significant difference in the association among histological subtypes.

**Reply:** Thank you for bringing this up. We wanted to compare different subtypes to each other as squamous cell carcinoma and small cell carcinoma have been shown in prior studies to have a stronger association with diabetes in prior studies. To compare distributions of lung cancer types in patients with and without diabetes, we focused on lung cancer cases. We edited the methods section to clarify this further.

**Changes in the text: Page 6, Lines 14-15**

**Comment 7:** Discussion: the authors could shorten the discussion and remove the redundant statements; for example, the following sentence could be shortened: “Additionally, the study had rigorous methods for ascertaining lung cancer diagnoses, which should have led to highly valid information about lung cancer incidence among study participants.”

**Reply:** We have edited the Discussion section and have either removed or shortened some sentences to make the section more concise.

**Changes in the text: Page 8 Line 16, Page 9 Lines 1-8, 10, 19-23.**

**Comment 8:** Table 2 and Table 3: I think these 2 tables could be combined; as the number of individuals included in adjusted analysis is not very different from those included in the unadjusted one, numbers included in adjusted analysis could be removed from the table (they might be shown in a supplementary table, if necessary). Then, unadjusted and adjusted RRs (95% CIs) could be shown side by side in the combined table. Please replace “Number of patients” and “Number” in the table headers to “Number of PLCO participants”. Also, “Lung Cancer Incidence Rate Ratio for Diabetes” is a bit odd. In the footnotes, replace “pack year” with “pack-years of smoking (continuous or categories? Please also add this information here or in the Methods)”.

**Reply:** We have combined Tables 2 and 3, which we believe now improves the presentation of our data. We also removed the numbers included in the adjusted analysis and added information about the number in the footnote along with covariates. We also edited our table headers to “Number of PLCO Participants” and “Incidence Rate Ratio”. We have corrected our footnote to include pack-years of smoking and for numeric variables (age, BMI, pack-years of smoking)- have indicated that they are continuous in the table footnotes, as well as the Methods section.

**Changes in the text: Page 5, Line 19. Edited Table 3, deleted Table 2 (now all Table 2 together)**

## **Reviewer B**

**Comment 1:** 1. General comments. 1) This study investigated whether diabetes is associated with lung cancer incidence using a large cohort data. The authors reported the null association between diabetes and lung cancer incidence. And this null association is consistent after stratifying according to lung cancer type. 2) This study is well-designed and well-written. Logical flow from the hypothesis to conclusion is reasonable. However, I think several modifications are required to publish the research.

**Reply:** The reviewer summarized the study.

**Comment 2:** 2. Detailed comments. 1) In Materials and Methods section, annual chest x-ray is used to screen lung cancer. However, in high risk individuals who smoked more than 15 pack-years, USPSTF recommend low-dose chest CT. I think screening methods significantly impact the detection rate of lung cancer.

**Reply:** Current lung cancer screening guidelines recommend low-dose computed tomography (CT) in high risk smokers, which has improved the detection rate of lung cancer at earlier stages of disease. The PLCO study was conducted prior to the USPSTF recommendations for low-dose CT screening for lung cancer, so CT-screened patients were not included in the study and would not have influenced our results. We agree it is important to highlight the current lung cancer screening guidelines and that the PLCO does not include CT-screened patients and have now mentioned this in our manuscript in the Methods section. Studying the influence of diabetes on a selected CT-screened cohort of patients would be an interesting future study.

**Changes in text:** Page 4 Lines 22-23, Page 5 Line 1

**Comment 3:** 2) Anti-diabetic medication such as metformin affects cancer incidence. Markers such as glycated hemoglobin or glucose levels that indicate how diabetes is well-controlled and anti-diabetic medication should be adjusted to investigate the association between diabetes and lung cancer incidence.

**Reply:** We agree that more data regarding diabetes severity (with laboratory data like glucose levels and glycated hemoglobin) and diabetes medications are important for further investigating the relationship between diabetes and lung cancer incidence. Unfortunately, the PLCO dataset does not report detailed information regarding diabetes severity and medication history. In the limitations section of our Discussion section, we noted the limitations of lack of medication data and duration of diabetes and now have also included lack of diabetes severity data as a limitation. Despite these limitations, strengths of the PLCO dataset include large patient sample size, robust measurement of smoking history and other risk factors for lung cancer, and rigorous cancer diagnosis ascertainment. Investigating whether or not a diabetes diagnosis increases risk (even with limited diabetes data) in a dataset with these strengths is valuable for knowing if diabetes more generally should be considered as a risk factor for lung cancer. We would be very interested in future studies that investigate how diabetes severity, medications, insulin resistance, and abdominal obesity influence lung cancer risk.

**Changes in text:** Page 10 Lines 5-6.

**Comment 4:** 3) In Table 4, why did you simultaneously adjust for pack years and packs per day? Pack-years contain the meaning of packs per day.

**Reply:** We wrote packs per day in error and have corrected the description of covariates.

**Changes in text: Table 4 footnote**