Reviewer’s report

Title: Smoking and prostate cancer: a life course analysis

Version: 1 Date: 02 Nov 2017

Reviewer: Shama Virani

Reviewer’s report:

Smoking and prostate cancer: a life course analysis

This manuscript assesses the gold standards of measuring smoking exposures in relation to prostate cancer and provides an alternative that takes changes in smoking intensity over the lifetime into account. This is a useful paper that can contribute to the literature with the following comments addressed:

1- page 4: this is the first instance where differentiation is used to describe "risk". The word risk is being misused in this context. Cases already have cancer and this is not a survival study. Therefore, you should compare the term differentiated with "aggressive", not risk. For example, well differentiated tumors are less aggressive (not low risk!). Poorly differentiated tumors are highly aggressive. This needs to be addressed throughout the manuscript

2- page 6 line 100: the calculation for the cumulative smoking lifespan index included dividing "by 20". Why 20? Is there a reference for this calculation?

3- page 6 line 119-121. The last sentence is very unclear.

4- page 7, line 142: smoking intensity needs to be defined here. What are the units? What time frame is considered (if any)?

5- Page 8 line 171-172: Thus far you have provided three groups when describing the Gleason score: [well differentiated or low-risk (Gleason ≤6), moderately differentiated (Gleason 7), or poorly differentiated (Gleason ≥8)]. Yet when providing results of how many cases were classified for each, you don't provide information for each group. Seeing 73.6% of your cases classified as poorly differentiated or highly aggressive (not high risk) concerns the reader that there may not be enough variability in your dataset to assess all variables mentioned

5- page 10 line 190: by providing information in the text that was found in your study, you are, in fact, providing data. "Data not shown" is not needed

6- page 10, lines 193-195: The patterns you identify are crucial to this manuscript and need to be better defined. What is "low" intensity? When "followed by an increase" how much of an increase is relevant? What does "higher smoking intensity" mean? These phrases need to be quantified if they are to be useful in future studies.
7- page 10 Figure 1: These figures are unclear. Because of the black lines, it is difficult to discern the patterns mentioned. I recommend you change the color of the lines based on the pattern they represent. Also, there seems to be other patterns present (a subset that increases a lot and then divides into different directions. What kind of pattern do these people fit into? There needs to be a more quantitative measure of these patterns to justify their use.

8- Table 2: The purpose of table is unclear. What is the relevance of understanding these variables in controls only? It would be more informative to see these results directly compared to cases.

9- Table 4: footnote a) should be split into two footnotes, it is difficult to tell which models are being referred to.

10- page 14 line 249: it is more appropriate to use "probability" than "possibility".

11- page 14 lines 251-258: see comment about about using "risk". Please fix the grammar in this paragraph.

12- Table 5: what is "all" referring to? All Gleason scores? So essentially you are testing to see if smoking status/cessation time is associated with PC overall? This is unclear.

Are the methods appropriate and well described?  
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?  
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?  
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?  
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English  
Please indicate the quality of language in the manuscript:
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal