Reviewer’s report

Title: Mortality Associated with Wildfire Smoke Exposure in Washington State, 2006-2017: a case-crossover study

Version: 1 Date: 29 Oct 2019

Reviewer: Colleen Reid

Reviewer's report:

Summary: This study examines the effects of wildfire smoke days on mortality for Washington state for an eight-year time period. The authors did address many of the concerns raised in the first review, although I still have a few concerns that need to be addressed (see below). The work is the first in Washington state to analyze the effects of wildfire smoke pollution on mortality which is especially important given the increasing wildfires in western North America. Another important part of this research is its focus on stratification by age groups, area-level income, cause of death, and race, which are important given the scant literature investigating differential impacts of wildfire smoke exposure on health.

Specific Points:

In the abstract, it says "in our lag analysis" but it should say in our "distributed lag analysis". However, putting in the results from both the distributed lag analysis and the individual lag analyses could be confusing to a reader who is not well versed in time-series/case-crossover epidemiological methods. I would suggest removing the distributed lag analysis results, as they only focus on individual lags rather than the combined effect across the lags, which is the standard implementation of the distributed lag analysis.

The other reviewer pointed out that Humidex has no units, which the authors said they had removed, but it is mentioned as the units of Humidex on line 154. Please remove.

Why was such a complicated method used to assign grid cells to their nearest monitoring site? (See Additional File 1 Text S1.) I am particularly wondering about why this step was used: "Each monitoring site was then assigned the ratio of its measured summer mean PM2.5 concentration to its AIRPACT-modeled summer mean PM2.5 concentration. These ratios were interpolated across the state using Empirical Bayesian Kriging (Pilz and Spöck, 2008). The interpolated ratio at each 4x4 km grid cell was then multiplied by the modeled 2014-2017 summertime mean PM2.5 to yield estimated summertime means across the state, at a 4x4 km resolution." The kriging of the ratios of measured to modeled would seem to create a spatial surface of the inverse of the performance of the model compared to the monitoring data. One question is whether the summer mean was the summer mean across the 2014-2017 for these ratios or not. Is the reason for using the model because the 75 sites do not cover the full state well enough and the model provides more spatial information? And the ratio then would have been more informative as the ratio of modeled to measured values to get a sense of how well the
model performed in locations with monitors? And then the kriged surface would be a spatial patterning of how well one would assume the model performed that could be then multiplied by the model to essentially calibrate it to the data?

Then this calibration was used to spatially assign the monitoring data to each grid cell. The authors should explain who such an elaborate modeling strategy was used and also why they used the monitoring data rather than the modeled data for assessing smoke exposed and how accurate they think this method was to link monitoring data to grid cells.

All of the information on this in the supplemental material is essential to understanding what was done. If there is room, I would put more of this into the actual manuscript with additional clarifying information about why this was done, how the authors think that this got at a good exposure estimate.

Lines 213-214: how was median household income classified for the stratified analysis?

Table 1: Does the exposure contrast mean that the referent days differ from the day of death in its classification as smoke-exposed or not? If so, and only these 31,719 were included in the analysis, shouldn't there also be statistical tests of the difference in covariates between those included and those not included? Also, if this is true that only individuals with an exposure contrast were included in the analysis, wouldn't this bias the findings? Isn't the purpose of a case-crossover to include all individuals to see if the exposure is related to the outcome (i.e., if you only include individuals in which the exposure changes, then you are more likely to find an association compared to the many that do not have a change in exposure on exposed and referent days?)

It would be good to have a map for the state of Washington showing the locations of the 75 monitoring stations.

Lines 267-269: Is this comparing wildfire days compared to non-wildfire days?

Lines 351-353: This statement is a bit misleading as none of the cited studies investigated mortality impacts of wildfire smoke, therefore they could not have found such associations whether they occurred or not. I suggest clarifying this sentence to state that none of those studies investigated mortality impacts of wildfire smoke exposure.

Line 361: Suggest clarifying what kind of threshold the authors are referring to.

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field

**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 organisation 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.