Reviewer’s report

Title: Cancer Prevalence among Flight Attendants Compared to the General Population

Version: 0 Date: 02 Jan 2018

Reviewer: Candice Johnson

Reviewer's report:

Cancer diagnoses among flight crew compared to the general population

The authors compare the prevalence of cancer diagnoses between a survey of flight attendants and NHANES, finding that all cancer types are more prevalent in the flight attendant cohort.

Major comments:

1. Add the NHANES population to Supplemental Table 1 so we can see how the two populations compare in terms of sociodemographic variables and cancer risk factors. It should also be included in the main manuscript (not as a supplement) because this table will help us understand the potential for residual confounding. This table also needs to be expanded to include risk factors for the cancers investigated, such as parity, age at first birth, age at menarche, family history of cancer, or any other important risk factor available.

2. The comparison population was restricted based on educational attainment, household income, and current employment status, and the same needs to be done to the flight attendant cohort. Currently, flight attendants who have left work because of a cancer diagnosis are still eligible for the study but NHANES participants in the same situation are excluded. This would artificially increase the SPR.
3. The study design implicitly includes cancer survivors and excludes people who died from cancer. Because of this, an increased SPR could just mean that flight attendants have a better chance of surviving cancer than the general population, not that they have a higher incidence of cancer. This could be due to health insurance access, paid leave policies, or other benefits that this workforce might enjoy compared to the majority of U.S. workers, or to sociodemographic differences that lead to better survivorship, such as living in or near an urban center (a large proportion of the study population likely lives near a hub). This needs to be added to the discussion.

4. The general U.S. population is likely not a good comparison group for a cohort of flight attendants, and confounding could be a problem. Although no details on the NHANES population are provided in the manuscript, there is a mention that the racial/ethnic compositions are strikingly different (75% vs. 43% non-Hispanic white), which suggests there could be large differences in cancer incidence and survival between the two groups if other variables differ as well. There is evidence that BMI and parity (for example) differ substantially between flight attendants and the general population. The sensitivity analysis restricting to non-Hispanic whites and the exclusion of NHANES participants without high school degrees and below a certain income/poverty ratio certainly help (particularly if the same exclusions are applied to the FAHS cohort), but there is still potential for residual confounding that should be acknowledged. Providing data on the NHANES population (see comment above) would also help to determine how different the two populations are. This problem needs to be acknowledged in the discussion.

5. A previous study (reference 24) found that differences between parity and age at first birth between flight attendants and the general population could explain the increased risk for breast cancer seen for flight attendants. At very least, adjustment for parity (and age at first birth if available) seems warranted in the models, along with reporting the distributions of these variables in the FAHS and NHANES populations.

6. Discussion (line 204): the limitation that the same set of covariates was used for every model can be easily addressed by re-running the models with the appropriate covariates instead of waiting for the next paper in the series to do the adjustment. Is there a reason this cannot be done now?
7. Table 2: there are fewer cancer cases in Table 2 than in Table 1. The exclusions that have been made need to be described in the manuscript (number of participants and reasons for exclusion). If these exclusions were for missing data, include the counts of individuals with missing data for each variable in Supplemental Table 1.

8. Supplemental Table 2: based on the participant counts, the NHANES population was not age-restricted for this analysis. For a flight attendant to be working in 1988, they would have to be, at the latest, born in 1970 (18 years old). Thus the FAHS population in this analysis would have been 45+ years old in 2015. The NHANES population must therefore also be restricted to participants 45+ years old for this analysis, to avoid artificially decreasing the prevalence of cancer in the comparison population by including younger (18-44) people who have a lower prevalence of cancer.

Minor comments:

1. Was there information on the age at which the cancer was diagnosed in FAHS and NHANES, and could cancers diagnosed before age 20 (for example), which are unlikely to be occupationally-related, be excluded?

2. Methods (line 111): define "high levels of historical occupational secondhand smoke". Is this based solely on working in 1988 or earlier, or was this based on a question about exposure?

3. Add additional waves of NHANES data. Doing so would be unlikely to cause substantial time effects but would increase the sample size in the comparison group, producing more precise estimates. The NHANES samples sizes are very small (<10 cases for 4 types of cancer), making interpretation of results difficult.

4. The analysis of "job tenure to age 45" should be restricted to participants in FAHS and NHANES who are 45+ years old to avoid mixing in participants who are less than 45 years old and have not yet had the opportunity to accrue as much exposure as those who are above 45 years old. Ideally the whole job tenure analysis would have some sort of similar restriction to
avoid all the low risk (young) participants clustering in the low exposure (low job tenure) part of the distribution - in the model, it is impossible for the low risk flight attendants to have high exposure. Adjusting for age likely does not fix this, as it seems to be a problem of non-positivity, not a problem of confounding. This should either be addressed analytically, or if this is not possible, in the discussion.

5. Discussion (line 218): why is it stated that these findings are applicable to pregnant women? These results do not seem to be related to pregnancy.

6. Discussion (line 153) and Conclusion (line 224): this is not the largest study of cancer rates among flight crew relative to the general population, nor is it the largest study of cancer rates among flight crew. At least two of the studies cited (references 23 and 24) have included over 6,000 flight crew.

7. Table 1: the column header "NHANES count, weighted" should be "unweighted". If the number of cancer cases had been weighted to the U.S. population, this number would have been in the thousands.

8. Table 2: where are the results for uterine, cervical, and gastrointestinal cancers?

**Level of interest**

Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

**Quality of written English**

Please indicate the quality of language in the manuscript:

Acceptable

**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