Response to referee #1

This paper presented a methodology/protocol of an epidemiological study during two large air pollution monitoring campaigns (APHH) in both urban and peri-urban areas in Beijing during two seasons (winter and summer). The author elaborated on the study design for both exposure and health measurement. The design of this study is quite complex in terms of the usage of portable monitors for personal exposures, in coordination with the intensive monitoring campaign period, and comprehensive examinations of health outcomes. The protocol shows the strength of combining the panel study with monitoring campaigns which provide the potential to investigate the health effect of detailed chemical compositions and biological mechanisms. It would be useful for other researchers to carry out further studies. Although it's a protocol paper, it's still necessary to present some preliminary results to show the quality of the data collection and general information of the study. The current result part has summarized the information of participation, and the calculation of sample size, but it would be helpful to add more results to both exposure and health measurements.

Overall: Thank you for your comments and useful suggestions to improve our manuscript. We acknowledged that despite this is an overview paper of AIRLESS project with the focus on methodology, more results are warranted for the readers to understand the studied participant, including their demographic characteristics, and the levels of exposure and health outcomes. Therefore, we have restructured the paper and added more preliminary results accordingly. The word has also been revised for a clear expression throughout the manuscript. Please find below the point-to-point response.

Comment 1: For the health part, it's important to know the basic demographic statistics of the participants in both urban and peri-urban sites. E.g. the attended clinical visits, distributions of age, gender, socioeconomic status, baseline exposure status, etc. Is there any significant difference between the two groups? Are you going to compare to the two groups of participants or treat them as two different cohorts? In addition, it would be useful to have some descriptive results related to the measurements of health outcomes.

Thank you for this suggestion. We have added a table (Table 4 in revised version) to summarize the demographic characteristics of urban and peri-urban participants. Basic health outcomes (BMI, WHR, hypertension status) were also included in Table 4. For the other detailed biomarkers, we plan to included in the following papers. Regarding to the exposure in the two sites, we added two figures (Figure 6 and 9 in revised version) to show the ambient and personal exposures during the campaigns in both sites.

We also revised the manuscript regarding to the results of these analysis, which is described in subsection 3.1 "Demographics characteristics of urban and peri-urban participants", 3.2 "Ambient concentration of $PM_{2.5}$ during study periods", and 3.5 "Seasonal and spatial pattern of the difference between personal and ambient exposure".

Generally, although we managed to keep the recruited urban and peri-urban participants in this study balanced in gender and hypertension ratio, we did observe significant differences in many ways (e.g. social-economic status, ambient air pollution levels, health conditions, etc). Given the similarities and difference between urban and peri-urban sites, we would examine the health effect while using sites as a modifier to see the difference, and will also investigate the reasons behindthe difference.

Comment 2: For personal exposure, it's crucial to have some results to validate the performance of PAM with reference instruments. How did you calibrate the instruments, and how well they agree with the reference instruments? What's the measurement range and error? What's the performance of PAM for different microenvironments (i.e. indoor and outdoors)? It's also important to know the completeness of personal exposure monitoring (e.g. how many validate days for the personal dataset, etc), as carrying personal monitors for 7 days is not common in an epidemiological study which would cause a lot burden.

Thank you for this comment. We understand the importance of the validation of personal monitor used in this study. Our previous paper (Chatzidiakou et al., 2019) has elaborated the corresponding information in detail, including the measurement performance (range and error), and the validation also considered under different conditions and microenvironment. We added a paragraph in the subsection of method (2.5 Personal exposure in revised manuscript) to summarize the key results from that paper, detailed as below:

"The characterisation of the performance of the air quality sensors integrated in the PAM is presented in a previous publication (Chatzidiakou et al., 2019). Briefly, all PAMs were calibrated in two outdoor co-location deployments at the urban PKU site next to reference instrumentation for one month after the winter and summer deployments to participants. The performance of the NO_2 and $PM_{2.5}$ sensors was additionally characterised in an indoor microenvironment next to commercial instruments. Overall, the air pollution sensors showed high reproducibility (mean R^2= 0.93, min–max: 0.80–1.00) and excellent agreement with standard instrumentation (R^2>0.84 for all sensors in winter, while R^2>0.71 in summer). Further work (Chatzidiakou et al., 2020) showed that the error of the PAM was negligible compared with the error introduced when deriving exposure metrics from fixed ambient monitoring stations close to the participants' residential addresses. Hence, novel sensing technologies such as the ones used here are suitable for collecting highly resolved personal exposure measurements in large-scale health studies."

Regarding to the completion of the personal measurement, all participants have completed 3548 personal-days measurements (~3.5 million observations in 1-min time resolution). The participants showed high compliance with the protocol, with a mean capture rate of personal data of > 86%.

We added a new sub-section (3.3 Completion of personal exposure during study periods) and a figure (Figure 7 in revised manuscript) to illustrate the results.

Comment 3: A summary of key air pollutants in both urban and rural sites during the health campaign periods in two seasons would lead the readers with a better understanding of the background AP settings, which can be useful to compare with other health studies around the world.

Thank you for this comment, we agreed with you that it's important to have a general idea of the pollution level for this study. We added a figure (Figure 7 in revised manuscript) using PM2.5 as an example to show the ambient level in both urban and rural site during the monitoring campaign.

A paragraph (subsection 3.2 Ambient concentration of PM2.5 during study periods) was also added in the manuscript for description.

"Figure 6 shows the ambient $PM_{2.5}$ concentration during AIRLESS campaigns in winter and summer with a comparison between sites. A clear seasonal trend with a large variation of ambient $PM_{2.5}$ concentration was observed. Specifically, during winter, the mean (SD) daily concentrations were 132.3 (104.8) µg m^{-3} and 87.4 (79.0) µg m^{-3} in the peri-urban and urban site respectively, which were significantly higher than the corresponding concentrations in summer as 35.2 (15.0) and 45.1 (20.8)

µg m^{-3}. The degraded ambient air quality and several high PM$_{2.5}$ pollution events in winter were due to the greater stagnation and weak southerly circulation suggested by synoptic-scale meteorological analysis (Shi et al., 2019). The number of days with concentrations exceeding the Chinese standard of 75 µg m^{-3} was 29 and 19 during winter in peri-urban and urban sites respectively. The PM$_{2.5}$ concentration in the urban area was constantly lower than the peri-urban site during winter, but the trend was opposite in summer."

Comment 4: Comparison of personal exposure with examples from certain participants, and capture rate of personal exposure data.

We appreciated this comment. An example plot of personal exposure to multiple pollutants of a selected case participant (U123) was added to our manuscript (Figure 8 in revised manuscript). The plot also includes the ambient concentration of the corresponding pollutant, along with the time-activity (i.e. indoor vs. outdoor) to illustrate the difference between personal and ambient exposure.

Accordingly, a paragraph (subsection 3.4 An illustrative example of exposure misclassification) was added in results section, as followed:

"A representative participant was selected to illustrate the concept of exposure misclassification in Figure 8. Personal exposure measurements of example participant U123 participating during the winter campaign are compared with data from the closest monitoring station to the participant's home location (< 5 km). The time-activity model (Section 2.6) determined when the participant was located at home (top row). The personal CO, NO, NO$_2$ and PM$_{2.5}$ concentrations regularly exceeded the outdoor levels, indicating that strong indoor emission sources (such as a gas stove) operated at regular times. The sources caused personal exposures up to 10 times higher than the ambient pollution levels. When no emission sources were active, the indoor CO and NO concentrations approached the outdoor concentrations, whereas the NO$_2$, O$_3$ and PM$_{2.5}$ were much lower than the outdoor concentrations, indicating the presence of indoor chemical sinks. In the case of ozone particularly, the personal indoor exposures were up to 25 times lower than the ambient concentrations, due to the high indoor reactivity of the pollutant. "

Comment 5: The application of personal measurements is increasing in the epidemiological studies, it's good to add some reviews on the progress in this area to highlight the advantage of this design.
Thank you for your suggestions. We have revised in the discussion section regarding the advantage of using PAM along with time activity model to exhibit the difference between ambient and personal exposure.

"Firstly, the study deployed a state-of-the-art and validated personal air pollution monitor to improve the personal exposure assessment to multiple pollutants. The high compliance rate of the participants with the study protocol highlighted the feasibility of collecting personal exposure data at high spatiotemporal resolution matched with detailed health assessments. The preliminary results highlight a clear difference between personal and ambient exposure driven by individual activity patterns, meteorological factors and the built environment. In line with previous literature, we show the large biases arising from the use of ambient measurements to represent personal exposure in most epidemiological studies, and the potential of novel sensor technologies to revolutionise future human-based studies.

Secondly, time-activity-location patterns of individuals are important determinants of personal exposure but due to the relative difficulty of collecting such information, they have rarely been taken into account by air pollution epidemiology. For the relatively sedentary participants of this panel study,

the home environment was the major contributor to overall exposure, and an important modifier of personal concentrations for all investigated air pollutant species. Exposure differences between the two panels were attributed partly to the variation in domestic energy use. For instance, in winter the urban building stock in China relies on centralised gas heating system, while traditional biomass and coal stoves remain the key emission source for heating and cooking in peri-urban areas. However, the exposure variability between participants was larger than the variability between the two groups, stressing the need to go beyond current methodologies to estimate population exposures."

References:

1 Chatzidiakou, L., Krause, A., Popoola, O. A., Di Antonio, A., Kellaway, M., Han, Y., ... & Fan, Y. (2019). Characterising low-cost sensors in highly portable platforms to quantify personal exposure in diverse environments. *Atmospheric measurement techniques*, *12*(8), 4643.