Response to Reviewer(s)’ Comments:  
  
Reviewer: 1

|  |
| --- |
| Comments to the Author  This study addressed one important and interesting topic- impacts of lacking home address in evaluating exposure and quantifying relationship between exposure and health outcomes. With increasing demands and interests on highly resolved exposure data, home address is an important factor in determining exposure accurately, but this information as noted by the authors is missing in some cases. In this paper, the authors hypothesized that the missing address information is limited to the district and evaluate its impacts on health effect estimation. The study has sufficient novelty and scientific significance in supporting its publication in the journal JESEE. Some points are suggested when the authors revise the manuscript to make it clearer: |

Response: We appreciate nice summary and generous evaluation of our work. We provided our point-by-point responses to reviewer’s comments in the following.

|  |
| --- |
| Comments to the Author (continued)  The authors are suggested to briefly comment on the choice of PM10 and LBW in the introduction section, as, for example, PM2.5 is more concerned than PM10, and particles have more adverse impacts on diseases such as respiratory and cardiovascular diseases than LBW as seen in published studies. The second point is briefly mentioned in the limitation section (line 330), but unclear in the introduction and method section. |

Response: We appreciate this comment for the important implication of our study. We agree with our limitation of not including PM2.5. Although we chose PM10 to make our simulation study realistic based on our published study (Choe et al. 2018), we attempted to allow our study findings to be applicable to other pollutants including PM2.5. Specifically, we employed various and extended environmental scenarios in addition to the environmental scenario 8 directly based on the parameters obtained from our published study for PM10. In these seven environmental scenarios, we used modified parameters that represent increased or decreased impacts of local sources and/or spatial correlation, and indicate various pollutants different from regionally homogenous PM10. We added our purpose of eight environmental scenarios focusing on the applicability to different pollutants to the Introduction and Methods sections.

Line 93-99 in the change-tracked revised manuscript (italicized and underlined for the revised text):

*“**This simulation study aimed to understand the impact of incomplete address information on outdoor exposure prediction and health effect estimation.* In order to achieve the applicability and generalizability of the simulation, we designed our study based on a previous epidemiological study of long-term exposure to particulate matter with a diameter less than or equal to 10 micrometers (PM10) and low birth weight (LBW) in Seoul, Korea (28). *We designed our simulation to make our finding applicable to other pollutants than PM10 by constructing various exposure scenarios using modified simulation parameters.”*

(Line 133-147 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“Assuming that exposure to PM10 follows a Gaussian random field with spatial dependency, we generated true annual-average PM10 concentrations using mean and variance parameters at all locations (Table 1). These locations included 46,007 mothers’ homes, 37 air quality regulatory monitoring sites, 25 district governmental offices, 422 neighborhood community centers, 16,230 census tract centroids, and 610 centroids on the 1-km grid in Seoul (Figure S2). To represent possibly different spatial structures of true PM10 annual-average concentrations, we used different combinations of mean and variance parameters and constructed eight environmental scenarios (ES1-ES8). Eight combinations of parameters gave varying contributions of the mean structure, spatial variability, and non-spatial variability of PM10 to total variability (Table 1, Figure S1). While ES1-ES4 was defined based on different spatial correlation structures, ES5-ES8 additionally included different mean structures characterized by five geographic variables that were highly associated with particulate matter in Seoul(31). ES8 was constructed by the optimal parameters from our data analysis. *These various and extended environmental scenarios can also represent different pollutants other than PM10.”*

|  |
| --- |
| Comments to the Author (continued)  Ambient (outdoor) air pollution, on a regional scale, is homogeneous while indoor air pollution is more likely heterologous. But from the point of air pollution exposure, indoor air pollution contributes larger to the overall exposure. The present study did not consider indoor air pollution and potential impacts of missing home address in evaluating indoor air pollution and its health outcome estimation. Any comments on this? |

We agree with the reviewer’s point about measurement error between outdoor exposure and overall exposure including indoor exposure. However, the impact of this measurement error would not be large, as a large fraction of outdoor PM penetrates indoor environments and possibly affects human health (Chen et al. 2010). In addition, eight environmental scenarios where we assumed different conditions of spatial variability can possibly represent the impact of heterogeneous indoor air pollution. We added these points to the Discussion.

<Reference>

Chen C, Zhao B. Review of relationship between indoor and outdoor particles: I/O ratio

(Line 343-355 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“Our study includes several limitations to be further investigated in future research. *First, we focused on ambient exposure and did not consider the impact of indoor exposure. However, this impact could be small for PM which showed relatively high infiltration compared to other pollutants 44. Besides, our application of diverse environmental scenarios including locally heterogeneous exposure may also represent indoor and/or personal exposure. Second, we created mothers’ residential addresses using census tract centroids and assumed them fixed over the simulation. Future studies that apply real addresses of participants and/or incorporate mobility should investigate the sensitivity of our findings. Third, we did not consider multi-pollutant models and correlated exposure measurement error could affect bias45. Future studies should investigate this impact in cohort-study design. Lastly, we used low birth weight and logistic regression. Future studies should confirm whether our suggestions are consistent with different health outcomes and health analysis models.”*

Reviewer: 2

|  |
| --- |
| Comments to the Author  This is a well written paper on a timely topic. Many large epidemiological studies have incomplete location information on their participants. A variety of exposure estimation approaches are available and may perform better or worse depending on what information is available. |

Response: We appreciate the positive comment of the reviewer. We provided our point-by-point responses to your comments in the following.

|  |
| --- |
| Comments to the Author (continued)  In addition, some of my minor comments my big picture question that this paper does not directly answers is if the conclusions of your epidemiology findings would change based on the models/scenarios that were employed. |

Response:

Our findings suggested that individual exposure modeled by geographic characteristics and averaged across population-representative points, as shown in universal kriging-based averaging, can provide improved accuracy in health effect estimates when complete address data are unavailable. Also, this improved accuracy was generally consistent across different environmental scenarios. We clarified this point in the Discussion.

(line 261-275 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“This study focused on the impact of limited availability of address data on health effect estimation compared to complete availability. After hypothesizing that address data availability affects health effect analysis of predicted exposure, we explored the impact of address availability on the performance of health effect estimates depending on exposure prediction methods and environmental scenarios based on the real-world example of the association between PM10 and LBW. Eight environmental scenarios represented various pollution environments related to the different contributions of geographic features and spatial dependency. Furthermore, nine prediction methods exhibited commonly applied approaches of individual exposure assessment given limited monitoring data with and without additional limitation in address data. *Our findings showed that when address data are limited, individual exposure modelled by geographic characteristics and averaged across population-representative points, as shown in UK-based averaging, can provide comparable accuracy in health effect estimation to those using complete address information. This improved accuracy was prominent compared to other exposure prediction approaches and generally consistent across different environmental scenarios.”*

|  |
| --- |
| Comments to the Author (continued)  While the title of the paper includes health analysis there are not health effects in the paper. The inclusion of the them would help the reader understand the magnitude of the effect of the potential errors. |

Response: Following the suggestion of the reviewer, we added the summary of health effect estimates to Table 2 and created the scatter plots of all health effect estimates across different prediction methods and environmental scenarios as Figure S6.

**FigureS6.** Scatter plots for logarithm of health effect estimates using true (x-axis) and predicted (y-axis) annual-average PM10 concentrations at home addresses of 46,007 mothers over the 1,000 simulation studies across major environmental scenario (ES2, ES3, ES5, ES8). Blue, red, black solid, and black dotted lines are for best-fitted, 45-degree lines, the true coefficient value, and the average of the fitted coefficient values, respectively.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Prediction Method | ES2 | ES3 | ES5 | ES8 |
| NM |  |  |  |  |
| IDWA |  |  |  |  |
| LUR |  |  |  |  |
| UK |  |  |  |  |
| AA |  |  |  |  |
| UKD |  |  |  |  |
| UKNA |  |  |  |  |
| UKCA |  |  |  |  |
| UKGA |  |  |  |  |

|  |
| --- |
| Introduction  1) Please specify that this is exposure to outdoor air pollution. Individual air pollution measurement will measure personal exposures and is a bit different than what this paper is looking at. |

Response:

We revised the text in the Introduction to clarify our focus on outdoor air pollution. In addition, we discussed the limitation of this focus and our efforts to overcome by implementing diverse environmental scenarios in the Discussion.

(Line 93-99 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

*“This simulation study aimed to understand the impact of incomplete address information on* outdoor *exposure prediction and health effect estimation.* In order to achieve the applicability and generalizability of the simulation, we designed our study based on a previous epidemiological study of long-term exposure to particulate matter with a diameter less than or equal to 10 micrometers (PM10) and low birth weight (LBW) in Seoul, Korea (28). *We designed our simulation to make our finding applicable to other pollutants than PM10 by constructing various exposure scenarios using modified simulation parameters.”*

(Line 343-355 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“Our study includes several limitations to be further investigated in future research. *First, we focused on ambient exposure and did not consider the impact of indoor exposure. However, this impact could be small for PM which showed relatively high infiltration compared to other pollutants 44. Besides, our application of diverse environmental scenarios including locally heterogeneous exposure may also represent indoor and/or personal exposure. Second, we created mothers’ residential addresses using census tract centroids and assumed them fixed over the simulation. Future studies that apply real addresses of participants and/or incorporate mobility should investigate the sensitivity of our findings. Third, we did not consider multi-pollutant models and correlated exposure measurement error could affect bias45. Future studies should investigate this impact in cohort-study design. Lastly, we used low birth weight and logistic regression. Future studies should confirm whether our suggestions are consistent with different health outcomes and health analysis models.”*

|  |
| --- |
| Introduction 2) I recommend having some discussion on what type of errors you would expect to see if you used incomplete address. Spatial? Temporal? How in theory may that effect the health effect estimates? This is the Discussion sections some and you have a bit in the Introduction, but some more detail may be warranted. |

Response:

We assume that the reviewer suggested to add the detailed discussion about the impact of measurement error to the Discussion instead of the Introduction. Our study focused on the spatial variability of long-term exposure, rather than temporal variability of short-term exposure, in order to make epidemiological inference in the association between long-term exposure to air pollution and human health. Given the unavailability of individual air pollution measurements, epidemiological studies of long-term air pollution applied exposure prediction models to estimate individual exposure that results in exposure measurement error over space. Previous simulation studies (Szpiro et al. 2011a; Szpiro et al. 2011b; Szpiro and Paciorek 2013; Lee et al. 2015), citied in our paper, investigated the behavior of exposure measurement error related to exposure prediction, and provided the interpretation based on Classical and Berkson measurement errors. Classical measurement error derived by incorrectly specified prediction models tended to increase bias of health effect estimates, while Berkson measurement error induced by spatial smoothing led to increased uncertainty. ~~Each of these errors, however, affected both bias and variability, different from the traditional definitions.~~ Our findings of large bias in AA, NM, and UKGA possibly suggest the impact of poor model specification affected by including non-representative locations for study subjects. Relatively large uncertainty in kriging also suggests the impact of spatial smoothing. We provided these details in the Discussion.

(Line 313-332 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“All prediction methods except for UK-based averaging generally showed underestimated health effects given limited PM10 or address data. This underestimation can be explained by exposure measurement error derived by poor characterization of individual exposure in prediction models (39,40). In our simulation, prediction methods heavily relying on a mean structure such as LUR gave greater underestimation when there is no mean structure in true exposure scenarios, while simple prediction approaches using measurements only shown as NM and IDWA gave larger underestimation when there is a mean structure. Prediction methods using a single location based on the nearest monitor (NM), or district governmental office (UKD) also gave larger negative bias than other methods. Bias was larger in AA and NM based on regulatory monitoring sites than UKD using population-representative locations. In addition, poor assessment of individual exposure can result from poor representativeness of prediction points used for averaging. Our study showed increased positive or negative bias in UKGA using grid coordinates than UKCA based on census tract centroids. Previous simulation studies reported that measurement error derived by a spatial misalignment between monitoring sites and people’s residences affected misspecification of prediction models and resulted in positive or negative bias in following health effect analysis (41–43). *Our findings of large bias in AA, NM, LUR, and UKGA possibly suggest the impact of this classical-type measurement error resulted from poor model specification. Relatively large uncertainty in kriging suggests the impact of the Berkson-type measurement error driven by spatial smoothing (42).*”

|  |
| --- |
| Methods 1) Prior to getting into the steps of the simulation it would be helpful to have a few sentences giving more of a general overview. For example, we are created a cohort with X number of births. We created X amount of exposure estimation approaches; we are going to look at the following statistics… It is all in the subsections, but it may be good to start with the big picture view to orientate the reader before delving into the details. |

Response:

We provided a general overview to the first paragraph of the Methods section. We also revised Figure S1 to include the overview summary of our data analysis, prediction models, and environmental scenarios, simulation results.

(Line 101-110 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“Our simulation procedure consists of four steps (Figure S1): 1) exploratory data analyses to obtain parameters for the underlying distributions of PM10 and LBW; 2) generation of true PM10 exposure and LBW status; 3) application of incomplete addresses and estimation of mothers’ exposures; and 4) health effect estimation of LBW for PM10 and comparison of the performance of health effect estimates by complete and incomplete addresses. *We constructed eight environmental scenarios representing different distributions of air pollution of 46,007 mothers, and nine exposure prediction methods under either complete or incomplete residential addresses.* The following sub-sections provide detailed information on each step. Further details including formulas are provided in the Supplementary Information.”

.

Figure S1 Summary flow charts of four steps in our simulation study

|  |
| --- |
| Methods  2) If possible, it would help to have a table of the different exposure prediction models similar to table 1. It could either be in the main results or supplemental. |

Response: We created a new supplemental table (Table S1) that summarizes the key characteristics of eight prediction models.

**Table S1.** Exposure assessment approaches of nine exposure prediction methods in our simulation study

|  |  |  |
| --- | --- | --- |
| .Prediction method | Assessment of individual exposure | Address Availability |
| Nearest monitor (NM) | PM concentration measured at the nearest regulatory monitoring site to mother’s home address | Complete |
| Inverse distance weighted average (IDWA) | Average concentration across regulatory monitoring sites weighted by inverse squared Euclidean distance from mother’s home | Complete |
| Land use regression (LUR) | Concentration at mother’s home predicted based on regression model including geographic predictorsa | Complete |
| Universal kriging (UK) | Concentration at mother’s home predicted by a geostatistical method including geographic predictors and spatial correlation | Complete |
| Area averaging (AA) | Concentration averaged across all monitoring sties in a district | Incomplete |
| Universal kriging - District (UKD) | Concentration predicted at a district governmental office by UK | Incomplete |
| Universal kriging - Neighborhood Average (UKNA) | District-specific concentration averaged over UK predictions at neighborhood community centers | Incomplete |
| Universal kriging - Census Track Average (UKCA) | District-specific concentration averaged over UK predictions at census tract centroids | Incomplete |
| Universal kriging - Grid Average (UKGA) | District-specific concentration averaged over UK predictions at 1-km grid coordinates | Incomplete |

a Length of major road within 100-m buffer; Proportion of water surface land use within 500m; Number of construction companies within 1000m; Distance to the nearest bus stop; Number of employees in construction industries within 100m)

|  |
| --- |
| Results 1) It would be helpful for the authors to out into contexts some of the results. The big question is whether the conclusions of the epidemiological study would change depending on which exposure estimation approach you would use. There are not health effect estimates in this paper so it hard to tell the differences that the errors are really making. Are the health effects estimates really different from one another? In other words, would their confidence intervals overlap one another in addition to the true value. A plot of the health effects may make this clearer. |

Response: Following the suggestion of the reviewer, we created Figure S4 that includes scatter plots of effect estimates between different ESs and prediction methods, added the average relative risks (RRs: exponentiated beta coefficients) over 1,000 simulations to Table 2, and summarized our finding in the Results section. As our simulation provides 1,000 effect estimates, we summarized their confidence intervals using coverage probability (CP) that indicates how much percent of the 95% confidence intervals include the true effect and looked at whether CP approaches the nominal 0.95. We revised our text to clarify the concept of coverage probability in the Methods. Furthermore, in order to clarify the meaning of coverage probability, we added a description for the horizontal line of 0.95 to the CP bar charts in Figure S7.

(Line 190-201 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

Using true and predicted PM10 as well as true LBW status of mothers, we estimated the health effects of LBW for PM10 using logistic regression. Then, we repeated the whole procedure from exposure generation to health effect estimation 1,000 times, and computed properties of health effect estimates over 1,000 simulations as bias, root mean square error (RMSE), average standard error (ASE), coverage probability (CP), and true positive rate (TPR). *CP was computed as the proportion of the simulations where the 95% confidence intervals of health effect estimates contain the true effect.* TPR was the ratio of the number of simulations that provide significantly positive health effect estimates for each predicted PM10 (p-value < 0.05) to those for true PM10. While bias, RMSE, ASE, and CP aim to evaluate the accuracy or uncertainty of the estimates, TPR can provide the insight into statistical power. Finally, we compared the health effect estimate properties between complete and incomplete addresses depending on the exposure prediction method and pollution environment.

(Line 221-233 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

Performance of effect estimates of LBW for true and predicted annual-average PM10 concentrations became worse when data availability for PM10 or address was limited. *Table 2 shows the average relative risk, as well as bias, RMSE, ASW, and CP of health effect estimates in four ESs including ES2, ES3, ES5, and ES8 where different exposure environments are more distinct*: Tables S3 and S4, and Figures S4 and S5 show all eight scenarios. Bias and RMSE tended to increase using predicted exposure compared to true exposure, while there was a slightly larger increase with incomplete addresses than with complete addresses. *Larger bias is also seen in scatter plots of health effect estimates of predicted exposures against those of true exposures in Figure S6.* Performance varied more across different prediction methods and environmental scenarios under complete addresses than incomplete addresses. Regardless of address availability, *CPs were close to 0.95 (Figure S7). TPR was generally lower with incomplete addresses than complete addresses (Figure S8)*.

|  |
| --- |
| Discussion 1) If the CP coverage is high for all approaches and scenarios is it worth the time and effort to used some of the more complex exposure models? The bias may cause under/overestimation of the results and a wider error will potentially make the health effect estimate insignificant. |

Response:

We agree that CPs close to 95% nominal level across all prediction models were derived by large uncertainty which possibly affect insignificant health effect estimates. To assess the significance of health effect estimates, we computed the true positive rate (TPR) as the rate of the simulations that provide significant effect estimates. Although TPRs are relatively small across all prediction methods as implied by large uncertainty and high CP, we found some differences across different prediction approaches. We clarified the implication of CPs close to 95% and our finding of TPR in the Discussion.

(Line 292-312 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“Our findings generally showed that kriging-based approaches gave good performance in health effect estimates consistently across different air pollution environments, when individual air pollution measurements are not available. While UK showed better performance compared to other prediction approaches when complete address data are available, UK averaging approaches outperformed with individual address data limited to the district. A possible explanation is that UK modelled by using both mean and variance structures well characterizes air pollution conditions at people’s residences even when there is no mean structure (38). In addition, employment of population-representative locations and the following averaging process under the unavailability of precise residential addresses possibly minimized the impact of exposure misclassification. Bias was the smallest and also non-systematic as opposed to other prediction methods that consistently gave negative bias. Out of three UK averaging approaches, UKCA based on UK predictions at census tract centroids gave the lowest RMSE and ASE which were comparable to those of other prediction approaches under the complete address condition. UKCA also showed comparable TPR to those with complete addresses, while it was less likely to detect statistically significant health effect estimates overall with incomplete addresses. However, the benefit of UK-based averaging could be reduced, when we use predictions at the locations including those poorly represented for population as shown in UKGA. *CPs close to 95% nominal level across all prediction models might have induced statistically non-significant health effect estimates. However, our finding of the true positive rate distinctively higher in kriging-based approaches compared to others indicates the advantage of kriging.*”

|  |
| --- |
| Discussion  2) There is a paper by Dionisio et al. “A simulation study to quantify the impacts of exposure measurement error on air pollution health risk estimates in copollutant time-series models in Environment Health that conducted a similar study. The authors may want to review it. |

Response:

We appreciate the suggestion of a relevant study. We added our discussion for the possible impact of correlated pollutants on our findings to the Discussion with the citation of Dionisio et al. 2016. As our study focused on the health effect of long-term exposure at the individual level as opposed to Dionisio et al. 2016 dealing with on the health effect of short-term exposure at the population level, our discussion was restricted to measurement error derived by co-pollutants rather than other sources.

(line 343-355 in the change-tracked revised manuscript) (italicized and underlined for the revised text):

“Our study includes several limitations to be further investigated in future research. *First, we focused on ambient exposure and did not consider the impact of indoor exposure. However, this impact could be small for PM which showed relatively high infiltration compared to other pollutants 44. Besides, our application of diverse environmental scenarios including locally heterogeneous exposure may also represent indoor and/or personal exposure. Second, we created mothers’ residential addresses using census tract centroids and assumed them fixed over the simulation. Future studies that apply real addresses of participants and/or incorporate mobility should investigate the sensitivity of our findings. Third, we did not consider multi-pollutant models and correlated exposure measurement error could affect bias45. Future studies should investigate this impact in cohort-study design. Lastly, we used low birth weight and logistic regression. Future studies should confirm whether our suggestions are consistent with different health outcomes and health analysis models.**”*