Author's response to reviews

Title: Effects of air pollution on neonatal prematurity in Guangzhou of China, a time-series study

Authors:

Qingguo Zhao (zqqfrost@126.com)
Zhijiang Liang (liangzhijiang07@163.com)
Shijuan Tao (victoriataojuan@yahoo.cn)
Juan Zhu (litte-nurse@163.com)
Yukai Du (duyukai100@yahoo.com.cn)

Version: 4 Date: 23 October 2010

Author's response to reviews: see over
Dear Editor:

On behalf of all authors, I would like to thank you and the reviewers for constructive comments and suggestions on our manuscript. The manuscript has now been revised according to the editor’s and reviewers’ suggestions/comments. Below is the detail of our response to the concerns.

Comment 1
The authors did not account for the number of gestations “at risk” for preterm birth. See the pair of recent articles by Darrow et al. in Epidemiology.
Essentially, the problem is that the daily count of preterm births will be a function of the number of gestations “at risk” for preterm birth. If there is seasonality in births then there will also be seasonality in the rate of preterm births. If this happens when pollutant concentrations are high then it will appear that pollution causes preterm birth. Note that the smooth function of time included in the GAM does not account for this issue.

Response: Thank you for your good advices! Firstly, there is no doubt that the article of Darrow et al is a classic study between ambient air pollution and preterm birth. But the problem is that we have only one year birth data at present, it is not appropriate to follow the analysis of Darrow et al article. Furthermore, Guangzhou is subtropical regions, where seasonal change is not obvious. If we could gather the data of seasonal change of many years in the future, we would do our study continually.

Secondly, Our study adopted Generalized Additive Model (GAM) extended Poisson regression model to quantitatively evaluate the acute effects of ambient air pollutants as NO2, PM10, SO2 on preterm rate of newborns. Before carrying out this study, we have referred to following two literatures:


We approached with the spline smooth function while incorporating time-independent variables, including calendar time, temperature and relative humidity, to control for the nonlinear confounding effects of trend. Seasonality and weather, which can accommodate non-linear and non-monotonic patterns between preterm birth and time/weather conditions, creating a flexible modelling tool [1]. Meanwhile, dummy variable was also used to control the effects of “day of the week” (DOW). The non-parametric flexibility of GAMs has resulted in their widespread use in time-series studies to adjust for the nonlinear confounding effects of seasonality and trend [2-9].

References:
Comment 2
The approach for “final model selection” was based on AIC, which is a practice typically avoided when the goal of the analysis is to control for confounding. Also, the readers need to know what covariates are in each of the final models. The approach of selecting the model with the strongest results is also a generally discouraged practice in epidemiology. It’s difficult to evaluate the models without knowing which terms were in them. I am interested to see how different the models were for NO$_2$, SO$_2$, and PM$_{10}$.

Response: Thanks! Base on your advices, Generalized cross-validation (GCV) scores were used to compare the relative quality of the incidence of preterm predictions across these non-nested models and how well the models fit the data. Please see the revised manuscript page 7, line 9-11. We selected the model with the strongest results referring to related literature (Zhang YP et al., 2008) in Taiyuan, China. The statistical analysis method in the methods section showed the detail of formula. Table 3, table 4 and table 5 showed the differences of three pollutants in single-pollutant models, and table 7 and table 9 showed the differences of three pollutants in multi-pollutant models.
Comment 3
There are several grammatical edits needed throughout, but overall I could follow the writing fine. There were also places where it says things like “Text for this sub-section…”

Response: Thanks! We had the paper edited in English by professional. Some errors had been corrected. Some sentences had been rewritten. And the repetitions had been deleted.

Comment 4
I am interested to know the size of the underlying cohort – how many births were included?

Response: There are a total of 142,312 births in Guangzhou City over the entire study period, including 9,083 preterm births. Please see the revised manuscript page 5, line 3-4.

Comment 5
Much information was gathered about job level, education, etc. Was any of this information included in the regression model? They play into the issue of seasonality and the tendency for low SES groups and high SES groups to have different seasonal patterns of births.

Response: we didn’t get information about job level, education, etc. These factors need be stratified to analyze. As the sample size is not large enough, number of events occurring every day could be zero in many cases, which affect the model fitting.

Comment 6
In the statistical model the authors write “alpha – residual error(?)” I don’t understand this…typically alpha is an intercept.

Response: Thanks! The words “residual error” were corrected to “residual” In a GAM model, residual is equivalent to intercept. Please see the revised manuscript page 6, line 15.

Comment 7
The issue with flat noise is applicable to linear models. The author's model is a Poisson model, and so there is no homoskedasticity assumption – the residuals will get bigger as the counts get bigger. I don’t believe than an attempt to achieve “flat noise” makes sense…that said I could be off-base here.

Response: Thanks! We had deleted the sentence “Residual analysis was conducted within the final models, drawing time-series tables and residual error self-related diagrams of each model to determine whether the residual error was randomly stable to be in conformity with flat noise, if not, model was adjusted accordingly until met the characteristics of flat noise.” page 8, line 11-14 in the original manuscript.

Comment 8
Although I know it’s not the author’s fault (the data are what they are), I don’t understand a lag 0 effect for air pollution on preterm delivery…and lag 0 was by far the strongest effect for all three pollutants investigated. Preterm delivery isn’t an instantaneous event. It generally takes a while between the start of labor and the actual delivery of the child. It would seem that at lag 0 (meaning air pollution levels on the day of the birth) that the woman would already be in the hospital. The authors should comment on this in the Discussion.

**Response:** Lag day is specifically used in time-series analysis (Joel S et al., 1993), defined by comparing the health index of day 0 with meteorological or ambient air pollution concentration in previous days, and to conduct regression analysis in order to observe the impacts on health in the future. The purpose of the study was to investigate the effects of air pollutants on health over a short-time period; a 7 lag day was chose referring to related literature (Anderson HR et al., 1996; Jong TL et al., 2000). We introduced the concentration of air pollutants on day 0, one day ago, seven days ago (Lag0-Lag7) or lag moving average (Avg0-Avg7) into the model one by one to calculate the relative risk and CI by the regression coefficient $\beta$ of air pollutants, and to quantify the influences of air pollutants on premature birth. Lag effect is only quantitative study of acute effects of pollutants in a fundamental way, in fact, for pregnant women in a pregnancy, we can not separate effects of pollutants a day, but to evaluate the cumulative effect of the health effects of pollutants on the fetus is more practical. The model sensitivity tests in this study also proved that.

We all participated sufficiently in the study of conception or design, data analysis or interpretation, and drafting or revising of manuscript, so that every one of us took responsibility for the validity and objectivity of the entire study, and has approved the final version of the manuscript. If requested, we agree to provide a copy of the original data which the results and conclusions of the study were based on. Anyway, we claim that none of the material in the paper has been published or is under consideration for publication elsewhere.

Zhao Qingguo is the corresponding author and his address and other information are as follow:
Address: College of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China;

Guangdong Women & Children Hospital, 43 Zhanxi Road, Guangzhou City, Guangdong Province, 510010, P.R.China
E-mail: zqgfrost@126.com

Best regards