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: 3 Date: 10 August 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 how we have addressed the critique of the review team.

Reviewer #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: Thanks! This 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, which has becomes a standard method to conduct air pollution researches in environmental epidemiology nowadays. To control the confounding like meteorological factors, trends in the long run, season and other confounders relied on variables within time series, we approached with spline smooth function while incorporating non-linear independent variables, including time(time, time=1~365), temperature and relative humidity to contain the influences of confounders in this study. Time(time, time=1~365) as a seasonality factor was introduced in the GAM model aim at controlling the seasonality in the rate of preterm births. In this study, We mainly aim at studying the acute effects of pollutants on the fetus. The result of this study show that lag effects could not last for a long time, which just maintained on day 0 with statistical significance. With regard to cumulative effects, time span of NO₂, PM₁₀ were consistent with each other, appearing on day 0, day 3 and day 4, the strongest for NO₂ was on day 3, while on day 4 for PM₁₀ with statistical significance. when it comes to cumulative effects of SO₂, day 0 to day 4 were maintained with the maximum on day 3, all with statistical significance.

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 NO₂, SO₂, and PM₁₀.

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 12-14. The approach of selecting the model with the strongest results is referring to related literature.
(Zhang YP et al., 2008) in Taiyuan, China. 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.

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 rewrote. And the repetitions had been deleted.

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 2-3.

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: In this study, much information such as job level, education, etc did not consider for to analyze these factors should be stratified. As the sample size is not large enough number of events occur every day there will be zero in many cases, which affect the model fitting.

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 17.

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.

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 as by comparing the health index of day 0 with meteorological or ambient air pollution concentration in previous days 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, days before a 7 lag day were chose after referring to related literature.
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 to make it possible 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, pregnant women in a pregnancy, we can not separate effects of pollutants a day, 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.

Reviewer #2
The authors should give more description of their birth data. As described, the data was collected for a city, Guangzhou. Does it mean the metropolitan areas including the outskirts or surrounding counties? The basic statistics should cover the total population, live births and stillbirths, as well as brief information about the area. How complete and consistent of the birth registration and information recording/taking across the numerous (and several types of) birth hospitals within the city, and how many births occurred in hospitals and homes? Missing values of those variables in the birth certificate registration, particularly gestational age, birth weight, and how did they define “prematurity”, use <37 weeks of complete weeks gestation and/or anything else? Please state these in a little more detail and clear in the Methods section. In addition, population migration or even daily change in births may be mentioned.

Response: Thanks! The description of Guangzhou, including the metropolitan areas, the total population, live births as well as how to get the birth data, how to deal with the missing values of those variables in the birth certificate registration, and how did they define “prematurity” and so on. Please see the first paragraph of the Methods section in the revised manuscript page 4, line 23 to page 5, line 11.

In general, intrauterine growth retardation (IUGR) is considered a better measurement in assessing fetal health or adverse development status, including prematurity. Since data on birth weight is available, if complete, I would suggest adding the analysis of association between IUGR and air pollution.

IUGR is important, but is limiting, since it is a dichotomous outcome like preterm birth (i.e., <37 complete weeks of gestational age). The authors may look at gestation as a continuous variable as well in the analysis. Looking at gestation age (reduction) will provide additional information on effects of air pollution, though the authors could argue whether this is necessary.

Response: This study is mainly aim at studying the acute effects of pollutants on the fetus, intrauterine growth retardation (IUGR) are slow-acting effect, and hard to definition of the evaluation index. Birth weight and gestation age as a continuous variable, their distribution does not meet the normality, if birth weight or gestation age as a dependent variable of linear regression analysis would not be appropriate. Moreover, the article was too long to not allow us to do much in terms of analysis.

In the Methods section, the authors commented excessively on the statistical analysis (i.e. time series GAM), especially about its statistical modeling. I would suggest that one describe the method briefly, but give comment on its application to this analysis in its Discussion section. In other word, remove the last second paragraph (ie., Principal Component …….) in the Methods section.
**Response:** Thanks! We had removed the last second paragraph in the Methods section to condense the 1.2 Statistical analysis, and added a comment on its application to this analysis in its Discussion section. Please see the revised manuscript page 13, line 13-23.

Although the time series modelling is not new, its hypothesis and usefulness of this statistical analysis of association between prematurity or fetal development and air pollution would be of help if they could be stated clearly in the text.  
**Response:** Thanks! We have a detailed explanation in the 1.2 Statistical analysis. Please see the first three paragraphs of the Methods section in the revised manuscript page 6, line 1 to page 8, line 4.

Suggest adding a graph or table showing the variations/fluctuations in preterm births and IUGR over the time.  
**Response:** Thanks! We had added figure 1 showing the variations/fluctuations in preterm births over the time. Please see the revised manuscript page 20, line 3.

A brief description of air pollution exposure data may be necessary. For example, the authors can specify for the daily max whether they averaged all the hourly values and then took the max of these hourly values or they averaged all the daily values and took the max.  
**Response:** Thanks! We have made a brief description of air pollution exposure data. Please see the second paragraph of the Methods section in the revised manuscript page 5, line 10-19.

Obviously, English is the authors’ second language. The authors should first check their expressions, wording and spelling more carefully before they submit their revision.  
**Response:** Thanks! We had the paper edited in English by professional. Some errors had been corrected. Some sentences had been rewrote. And the repetitions had been deleted.

**Here were the details:** (please notice that all the “page” and “line” appeared in the following text indicated the “page” and “line” of the original manuscript not the revised manuscript!)

Page2, line 13  
*The words* “revising the” were corrected to “the revision of the”

Page3, line 3  
*We deleted the word* “of”

Page3, line 3  
*The word* “But” were corrected to “However,”

Page3, line 5  
*The words* “are still” were corrected to “remain”

Page3, line 10
The words “as well as” were corrected to “We also”  
Page3, line 10
The words “co-linearity” were corrected to “the collinearity”  
Page3, line 10
We added the word “taken”  
Page3, line 10
The word “record” were corrected to “records”  
Page3, line 13
The word “concentration” were corrected to “concentrations”  
Page3, line 16
We added the word “and”  
Page3, line 18
The word “were” were corrected to “was”  
Page3, line 18
We added the word “cases”  
Page3, line 19
The word “concentration” were corrected to “concentrations”  
Page3, line 19
We added the word “and”  
Page3, line 23
The words “RR as” were corrected to “relative risks (RRs) of”  
Page3, line 25
The words “adjusted collinearity” were corrected to “adjusting the collinearity”  
Page3, line 27
The words “RR as” were corrected to “RRs of”  
Page4, line 1
The sentence “and RR of the strongest cumulative effects of the three air pollutants was interpreted as” were corrected to “and RRs of the three air pollutants, at their strongest cumulative effects, were”

Page 4, line 6

The sentence “There is evidence showing that daily concentration of NO\textsubscript{2}, PM\textsubscript{10}, SO\textsubscript{2} was positively associated with the preterm rate in Guangzhou, China.” were corrected to “There is evidence showing that daily concentrations of NO\textsubscript{2}, PM\textsubscript{10} and SO\textsubscript{2} are positively associated with the preterm rate in Guangzhou, China.”

Page 4, line 9- Page 5, line 20.

We rewrote the background: “Air pollution affects the health of children and the elderly, and remains a major public health concern that has gathered increasing attention in the past few years. Evidence has shown that air pollution is associated with an increased risk rate of adverse pregnancy outcomes (Glinianaia et al., 2004; Maisonet M et al., 2001). International survey data showed a 7-10% premature rate (Chien PF et al., 1997) in developed countries, and 9-12% in United States in recent years, displaying an upward trend (Ponce NA et al., 2005). A survey in China indicated a 5-15% preterm rate (Gong JH, 1990). 15% of preterm babies die in neonatal period. In addition to fatal malformations, 70% of neonatal deaths and 75% of neonatal complications are associated with premature birth. Complications include respiratory diseases, intracerebral hemorrhage, infections and dysplasia. Compared with full-term babies, premature babies suffer greater exposure to cerebral palsy, amblyopia, deafness and mental retardation (Le J, 2001). It has been found that prematurity was not only a significant reason in neonatal deaths, but also a substantial contributor to diabetes mellitus, coronary heart disease and hypertension in adulthood (Mikkola K et al., 2009). Hence, seeking the causes of preterm birth is of vital importance to public health.

Previous researches on possible risk factors of premature birth were mainly focused on socioeconomics, level of education, smoking and drinking behavior during pregnancy, intrauterine infections, multiple births, history of abortion and preterm birth, genital abnormality, pregnancy-induced hypertension, risky sexual behavior, etc. Recently, researchers in China and abroad began to realize the connection between air pollution and the occurrence of prematurity. America, Canada, Australia, Lithuania and other countries have carried out researches in this respect. The results suggested that exposure to air pollutants such as NO\textsubscript{2}, PM\textsubscript{10} and SO\textsubscript{2} during pregnancy was possibly related to premature birth (Sagiv SK et al., 2005; Liu S et al., 2003; Hansen C et al., 2006; Maroziene L and Grazuleviciene R, 2002). Meanwhile, cities in China such as Beijing, Taiyuan and Taibei have also conducted such studies, revealing that the increased concentration of air pollutants such as NO\textsubscript{2}, PM\textsubscript{10} and SO\textsubscript{2} presents as a risk attributing to premature birth (Xu X et al., 1995; Zhang YP et al., 2008; and Tsai SS et al, 2004). This study adopted the Generalized Additive Model (GAM) extended Poisson regression model to quantitatively evaluate the effects of ambient air pollutants –NO\textsubscript{2}, PM\textsubscript{10} and SO\textsubscript{2} – on the preterm rate of newborns by analyzing the time-series data of air pollution, meteorological factors, preterm in Guangdong Province in 2007.”

Page 5, line 22 to page 6, line 7.

We rewrote the 1.1 Data sources into the 1.1 Data gather of the Methods section: “All live births were obtained from the live birth database including all obstetric institutions of
Guangzhou City in 2007, where the live birth database system was established. Guangzhou, composed of 10 districts and 2 satellite cities, has an urban area of over 7434.4 square kilometers and a metropolitan area population of 9.8 million at the beginning of 2007. There is a total of 142,312 births in Guangzhou City over the entire study period, including 9,083 preterm births. Gestational age was computed as the number of weeks between the date of the last menstrual period (LMP) and the date of birth. For birth records missing the date of the LMP, the clinical estimate of gestation was used. Eligible births with gestational ages <37 weeks were considered preterm. Twin pregnancy and multiple pregnancy will be excluded from this study. After exclusions, 7,836 of 9,083 preterm births (86.3%) were eligible for analysis. The number of preterm births was tallied for each day in 2007.

The daily mean concentrations of air pollutants nitrogen dioxide \((\text{NO}_2)\), particulate matter less than or equal to 10 microns \((\text{PM}_{10})\) and sulfur dioxide \((\text{SO}_2)\) in 2007 were collected from the Environmental Monitoring Center of Guangzhou city. The daily concentrations of each pollutant were averaged from the available monitoring results of nine fixed-site stations under China National Quality Control located in the urban areas of Guangzhou. We collected the 24-hour average concentrations for \(\text{PM}_{10}\), \(\text{SO}_2\) and \(\text{NO}_2\). In the calculation of each 24-hour average concentration, it is required to have at least 75% of the one-hour values on that particular day. If a station had more than 25% of the values missing for the whole period of analysis, the entire station would be excluded from the analysis.

To allow adjustment for the possible effect of weather on preterm birth, daily average temperature (℃) and relative humidity (%) data were collected from Guangzhou Meteorological Bureau. The weather data were measured at a fix-site station located in Yuexiu District of Guangzhou.

We rewrote the 1.2 Statistical analysis: “Given the total population, daily premature birth represents a small probability event. As a kind of time-series data, its distribution approximately follows the Poisson distribution (Sagiv S K and Mendola P, 2005). To determine the influence of air pollution on premature birth, analysis should be carried out in time-series Generalized Additive Model (GAM) extended Poisson regression (Hastie T and Tibshirani R, 1990), which expands the traditional Generalized Log-Linear Model. Besides fitting common linear subjects. Complicated non-linear variables of induced variables were incorporated in different functions of additive operation. Since first introduced by Schwartz J (Schwartz J and Dockery DW, 1996), time-series Generalized Additive Model (GAM) extended Poisson regression became a standard method to conduct air pollution researches in environmental epidemiology. The formula is explained in detail as follows:

\[
\log[E(Y_t)] = \alpha + \beta Z_t + S(\text{time, df}) + S(\text{temperature, df}) + S(\text{relative humidity, df}) + \text{DOW (day of week)}
\]

In the formula, \(Y_t\) represents daily number of preterm babies, \(E(Y_t)\) — expected value of daily number of preterm babies, \(\alpha\) — residual, \(\beta\) — regression coefficient, \(Z_t\) — concentration of air pollution or accumulated average concentration over several days, \(S(\text{time, df})\) — time spline smooth function, \(S(\text{temperature, df})\) — temperature spline smooth function, \(S(\text{relative humidity, df})\) — relative humidity spline smooth function, \(\text{DOW (day of week)}\) — dummy variable.

To study the impact of air pollution on preterm birth, it is essential to control confounding factors such as meteorology, long-term trends, season and other factors which affect the
variables within the time-series. In this study, we first built basic models on the daily numbers of preterm births which did not include the air pollution variables. We approached with the spline smooth function while incorporating time-independent variables, including time (time, time=1−365), temperature and relative humidity, in order to reduce the influence of confounders, and accommodate non-linear and non-monotonic patterns between preterm birth and time/weather conditions, creating a flexible modeling tool (Hastie T and Tibshirani R, 1995). Meanwhile, dummy variable was also used to control the effects of “day of the week” (DOW). Residuals of each model were examined to check whether there were discernible patterns and autocorrelation by means of residual plots and partial autocorrelation function plots, respectively (Jiang LL and Zhang YH, 2007).

After the establishment of basic models, we introduced the pollutant variables into the models and analyzed their effects on preterm births. The number of gestations at risk for preterm birth was used as an offset. Generalized cross-validation (GCV) scores were used to compare the relative quality of the incidence of preterm predictions across these non-nested models and verify how well the models fit the data (Hastie T and Tibshirani R, 1995). Taking into consideration of the delayed effects of air pollutants, the model also included the lag effects. The purpose of the study was to investigate the effects of air pollutants on health over a short-time period, days before a 7 lag day were chosen after referring to relevant literature (Anderson HR et al., 1996; Jong TL et al., 2000). Lag day is specifically used in time-series analysis (Joel S et al., 1993), to compare the health index of day 0 with meteorological or ambient air pollution concentration in previous days to conduct regression analysis in order to observe the impacts on health in the future. 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 to make it possible to quantify the influence of air pollutants on premature birth. Moreover, sensitivity analysis was also conducted within the established models. While carrying out the sensitivity test, multiple air pollutants model was fitted to observe the stability of single air pollutant model, and to compare lag and cumulative effects to analyze the stability of air pollutants’ effects.

“Principal Component” was introduced into the study while establishing the dose-response model of multiple ambient air pollutants’ health effects in order to exclude the impacts of collinearity (Zhong XN and Zhou YR, 2007). The composite latent variable (Principal Component) able to stand for variable information of original air pollutants by principal component analysis was substituted into the time-series Generalized Additive Model (GAM) extended Poisson regression model. The variables were also fitted into the linear model of principal components, meanwhile transforming the regression coefficient $\beta$ of the principal components into the regression coefficient $b$ of the original air pollutants to calculate the relative risk and CI so as to quantify the influence of each air pollutant on preterm birth in the multiple air pollutants model.

All the above statistical analyses were conducted using SAS 8.2 and S-Plus 2000.”
make an independent and specific composite latent variable (Principal Component), with extracted variation information of original index to establish equation of linear regression of the logarithmic latent variable and induced variable, and then converts the latent variable into the original independent variable. (Zhong XN and Zhou YR, 2007).

Page 9, line 14.

We Split the first paragraph in the Results section into two paragraphs and added a figure 1 in page 20, line 3:

“Descriptive statistical results of premature births
Figure 1 shows the variations in preterm births over the entire study period. The rate of preterm newborns was 21.47 cases/day in 2007, the quartile range was 8 cases/day. P₀, P₂₅, P₅₀, P₇₅, P₁₀₀ represented as 7.00, 17.00, 21.00, 25.00, 39.00 cases/day.
Descriptive statistical results of air pollutants and meteorological factors
Table 1 shows the daily average concentrations of NO₂, PM₁₀, and SO₂ in Guangzhou were 61.04 μg/m³, 82.51 μg/m³, 51.67 μg/m³ respectively, while daily average temperature in 2007 was (23.87±5.76)°C (n=365), P₀, P₂₅, P₅₀, P₇₅, P₁₀₀ corresponds to 8.00, 19.60, 24.50, 28.77, 32.65°C, the quartile range was 9.17°C, relative humidity was (67.55±10.26)% (n=365), P₀, P₂₅, P₅₀, P₇₅, P₁₀₀ were 40.00%, 62.50%, 67.50%, 75.00%, 88.00% respectively. The quartile range was 12.50%.”

Page 10, line 3

We added the word “and”

Page 10, line 4

The word “was” were corrected to “were”

Page 10, line 4

We added the word “and”

Page 10, line 6

The word “suggestd” were corrected to “suggest”

Page 10, line 6

The word “the” were corrected to “a”

Page 10, line 7

The word “of” were corrected to “in”

Page 10, line 8-13.

The sentences “Correlation of three air pollutants indicated a strong statistical significance (P<0.01) with the strongest correlation between NO₂ and PM₁₀, r value being 0.8533, then between NO₂ and SO₂, r value being 0.8440. As to temperature and relative humidity, r value was 0.1924 with statistical significance (P<0.01). Correlation analysis of the above indexes showed possible collinearity for the independent variables.” was corrected to “Correlation of three air pollutants indicated a strong statistical significance (P<0.01) with the strongest correlation being between NO₂ and PM₁₀, r value being 0.8533, then between NO₂ and SO₂, r value being 0.8440.
As for temperature and relative humidity, the r value was 0.1924 and statistically significant (P<0.01). Correlation analysis of the above indexes showed possible collinearity for the independent variables.”

Page 10, line 14
The word “of” were corrected to “using”

Page 10, line 15
The words “fitting to” were corrected to “and aiming to”

Page 10, line 21 - Page 11, line 2.
The sentence “While comparing the results of lag effects and cumulative effects, we found that lag effects could not last for a long time, which just maintained on day 0 with statistical significance(P<0.05). With regard to cumulative effects, time span of NO2, PM10 were consistent with each other, appearing on day 0, day 3 and day 4; the strongest effect for NO2 was on day 3, while strongest for PM10 was on day 4 with statistical significance(P<0.05). When it comes to cumulative effects of SO2, day 0 to day 4 were maintained with the maximum on day 3, all with statistical significance(P<0.05).” was corrected to “While comparing the results of lag effects and cumulative effects, we found that lag effects did not last for a long time, with only day 0 resulting in statistical significance (P<0.05). In regard to cumulative effects, the time spans of NO2 and PM10 were consistent with each other, appearing on day 0, day 3 and day 4; the strongest effect of NO2 was on day 3, while the strongest effect of PM10 was on day 4, also with statistical significance (P<0.05). When it came to the cumulative effects of SO2, the effects were maintained from day 0 to day 4, with maximum effect on day 3, all with statistical significance (P<0.05).”

Page 11, line 6
The word “with” were corrected to “to”

Page 11, line 15
The word “for” were corrected to “of”

Page 11, line 18
The words “in ways of” were corrected to “according to the”

Page 11, line 19
We added the word “we also”

Page 11, line 21
The word “of” were corrected to “using”

Page 12, line 6
The words “by adding it into the GAM model linear model was fitted” were corrected to “By adding it into the GAM model, the linear model was fitted”

Page 12, line 6-8
Table 7 displays that results of effects on day 0 by GAM model of different air pollutants combinations suggest after adjusted collinearity within air pollutants index by principal component analysis. were corrected to “Table 7 displays the Gam model results for the effects on day 0 of different air pollutants combinations. After adjusting the collinearity within air pollutants index by principal component analysis.”

In triple model, after adjustment by principal component analysis, the relative risk of NO\textsubscript{2} influencing preterm birth was 1.0185 (95%CI: 1.0056~1.0313), PM\textsubscript{10} was 1.0215 (95%CI: 1.0066~1.0365), SO\textsubscript{2} was 1.0326 (95%CI: 1.0101~1.0552). were corrected to “In the triple model, after adjustment by principal component analysis, the relative risk of NO\textsubscript{2} influencing preterm birth was 1.0185 (95%CI: 1.0056~1.0313), PM\textsubscript{10} was 1.0215 (95%CI: 1.0066~1.0365), SO\textsubscript{2} was 1.0326 (95%CI: 1.0101~1.0552).”

We could see from table 9 that the strongest cumulative effects of different air pollutants combination in the GAM model, only after adjustment of collinearity of the air pollutants index by principal component analysis indicate a transformation from non-statistical to statistical significance for both double model and triple model.” were corrected to “We could see from table 9 that after adjusted the collinearity of air pollutants index by means of principal component analysis, the strongest effects for GAM model of different air pollutants combination indicated a transformation from non-statistical to statistical significance for double model and triple model in terms of the best cumulative effects of NO\textsubscript{2}, PM\textsubscript{10} and SO\textsubscript{2}.”

An increase concentration of 100 µg/m\textsuperscript{3} for the air pollutants, RR of NO\textsubscript{2}, PM\textsubscript{10}, SO\textsubscript{2} on day 0 represented as 1.0542 (95%CI: 1.0080~1.1003), 1.0688 (95%CI: 1.0074~1.1301), 1.1298 (95%CI: 1.0480~1.2116). were corrected to “An increased concentration of 100µg/m\textsuperscript{3} of air pollutants NO\textsubscript{2}, PM\textsubscript{10}, SO\textsubscript{2} on day 0, corresponded to RR of 1.0542 (95%CI: 1.0080~1.1003), 1.0688 (95%CI: 1.0074~1.1301), 1.1298 (95%CI: 1.0480~1.2116) respectively.”

The risk extent for the study was relatively lower than that of the research results carried out by Sagiv (Sagiv SK et al., 2005, United States), (or) Canada’s Liu (Liu S et al.,
2003), Australia Hansen (Hansen C et al., 2006), and Lithuania Maroziene (Maroziene L and Grazuleviciene R, 2002) and Taiwan Tsai (Tsai SS et al., 2004), Beijing Xu (Xu X et al., 1995) and Taiyuan Zhang Yan-ping (Zhang YP et al., 2008) in China. ” was corrected to “The extent of risk resulted in this study was relatively lower than that of the research results by United States’ Sagiv (Sagiv SK et al., 2005), Canada’s Liu (Liu S et al., 2003), Australia Hansen (Hansen C et al., 2006), and Lithuania Maroziene (Maroziene L and Grazuleviciene R, 2002) and Taiwan Tsai (Tsai SS et al., 2004), Beijing Xu (Xu X et al., 1995) and Taiyuan Zhang (Zhang YP et al., 2008) in China. ”

Page 13, line 23
The word “It” were corrected to “This”

Page 14, line 4
The word “just” were corrected to “only”

Page 14, line 4-8
The word “index” were corrected to “indexes”

The sentences “In order to solve the problem, principal component analysis was adopted to adjust the collinearities of different air pollutants index in the multiple air pollutants model. By doing this, not only did the evaluation of the composite effects of different air pollutants become easier, but also the collinearities among multiple air pollutants were solved.” was rewrote to “In order to solve the problem, principal component analysis was adopted to adjust the collinearity of different air pollutants indexes in the multiple air pollutants model. Principal Component analysis is a multivariate statistical analysis method. which combines three air pollutant indexes by means of an appropriate linear model to make an independent and specific composite latent variable (Principal Component), with extracted variation information of original index to establish equation of linear regression of the logarithmic latent variable and dependent variable, and then converts the latent variable into the original independent variable. By doing this, not only did the evaluation of the composite effects of different air pollutants become easier, but also the issue of collinearity among multiple air pollutants was resolved.”

Page 14, line 14
The words “Mechanisms of effect air pollution” were corrected to “Mechanisms of the effect of air pollution”

Page 14, line 20
The words “in” were corrected to “during”

Page 15, line 3-6.
The sentences “it is difficult to be evaluated by simple linear relationship model. This study was conducted with non-parametric smooth function to control the confounding factors such as time trends, season, weather, so it’s more powerful on evaluating the relationship than the traditional methods.” was rewrote to “it is difficult to allow for evaluation when applying simple linear relationship model. This study was conducted with non-parametric smooth function
to control the confounding factors such as time trends, season and weather, allowing a more powerful way of evaluating the relationship than traditional methods.”

Page 15, line 7

The words “Sensitivity test of the model suggested” were corrected to “Sensitivity testing of the model suggest”

Page 15, line 15

We added the word “and”

Page 15, line 24

We added the word “and”

Page 16, line 2

The words “haven’t studied” were corrected to “not”

Page 16, line 6

The words “In spite of” were corrected to “Despite the”

Page 16, line 7-14

The sentences “Although the absolute increase amount in risk is relatively small, thousands of pregnant women could been exposed to high levels of air pollution in long-term period. Therefore, their public health significance can not be ignored. The studies regarding the impact of air pollution on preterm deliveries are still rare in China. This study explored the potential exposure-reaction between preterm birth and air pollutants such as NO₂, PM₁₀ and SO₂, aiming to provide scientific facts to help relevant departments to make air pollution control decisions, and to decrease adverse pregnancy outcomes.” was rewrote to “Although the absolute increase of risk is relatively small, thousands of pregnant women could have been exposed to high levels of air pollution in a long-term period. Therefore, the public health significance can not be ignored. Studies regarding the impact of air pollution on preterm deliveries are still rare in China. This study explored the potential exposure-reaction between preterm birth and air pollutants such as NO₂, PM₁₀ and SO₂, and aims to provide scientific facts to help relevant departments in their decision-making regarding air pollution control, and decreasing adverse pregnancy outcomes.”

Page 16, line 16- 23

We rewrote the conclusions: “In summary, this paper has examined that the concentrations of the NO₂, PM₁₀ and SO₂ of air pollutants contributed to occurrence of preterm birth in Guangzhou city, and has proved that the three air pollutants have dose-response reactions in terms of neonatal prematurity, through analyzing a single air pollutant model and a multi-air pollutants GAM model. Although there were limitations in this study, it provides the fact that air pollution plays a critical role in prematurity. Thus, this study can highlight such importances to policy-makers as they make decisions which aim to control air pollution and decrease preterm birth.”

Page 17, line 2

We deleted the word “any”
The words “study protocol design development, data collection” were corrected to “development of the study’s protocol design, data collection”

The words “was” were corrected to “were”

The word “revising” were corrected to “the revision of”

The word “of” were corrected to “from”

The words “mean、std、min、P25、P50、P75、max、P75-P25” were corrected to “Mean、Std、P0、P25、P50、P75、P100、P75-P25” in the table 1.

The words “Lag00、Lag01、Lag02、Lag03、Lag04、Lag05、Lag06、Lag07” were corrected to “Avg0、Avg1、Avg2、Avg3、Avg4、Avg5、Avg6、Avg7” in the table 3-5.

The words “NO2、PM10、SO2” were corrected to “NO2(0)、PM10(0)、SO2(0)” in the table 6.

The words “NO2(03)、PM10(04)、SO2(03)” were corrected to “NO2(3)、PM10(4)、SO2(3)” in the table 8.

The words “NO2(03)、PM10(04)、SO2(03)” were corrected to “NO2(Avg3)、PM10(Avg4)、SO2(Avg3)” in the table 9.

The words “NO2、PM10、SO2” were corrected to “NO2、PM10、SO2” in the text respectively.

We all participated sufficiently in the study conception or design, data analysis or interpretation, and drafting or revision of manuscript, so that every one of us takes 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 upon which the results and conclusions were based. Also, 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: Guangdong Women & Children Hospital, 43 Zhanxi Road, Guangzhou City, Guangdong Province, 510010, P.R. China
E-mail: zqgfrost@126.com

Best regards