nature sustainability

ARTICLES

https://doi.org/10.1038/s41893-019-0387-y

Air pollution-induced missed abortion risk for pregnancies

Liqiang Zhang[1,6*], Weiwei Liu[2,6*], Kun Hou[1,6], Jintai Lin[3*], Chenghu Zhou[4*], Xiaohua Tong[5], Ziye Wang[1], Yuebin Wang[1], Yanxiao Jiang[1], Ziwei Wang[3], Yibo Zheng[1], Yonglian Lan[2], Suhong Liu[1], Ruijing Ni[3], Mengyao Liu[3] and Panpan Zhu[1]

Fetus death risk reduction is included in the United Nations Sustainable Development Goals. However, little is known about how missed abortion in the first trimester (MAFT) is related to maternal air pollution exposure. We quantify the link between air pollution exposure and MAFT in Beijing, China, a region with severe MAFT and air quality problems. We analyse the records of 255,668 pregnant women from 2009 to 2017 and contrast them with maternal exposure to air pollutants (particulate matter $PM_{2.5}$, SO_2, O_3 and CO). We adjust for confounding factors such as sociodemographic characteristics, spatial autocorrelation and ambient temperature. We find that, for all four pollutants, an increased risk of MAFT is associated with rises in pollutant concentrations and the adjusted odds ratios (ORs) of these associations increase with higher concentrations. For example, the adjusted OR of MAFT risk for a 10.0 μg m^{-3} increase in SO_2 exposure is between 1.29 and 1.41 at concentrations of 7.1–19.5 μg m^{-3}; it drops to 1.17 below this range and rises to 1.52 above it at higher SO_2 concentrations. This means that the risk increase is not linear but becomes more severe the higher the pollutant concentration. The findings provide evidence linking fetus disease burden and maternal air pollution exposure.

Poor air quality is a leading cause of global disease burden[1,2]. Considerable evidence has consistently indicated that maternal exposure to air pollution contributes to increased risks of adverse birth outcomes such as low birthweight[3], preterm birth[3,4], gestational hypertension and preeclampsia[5,6], and may also affect maternal health during pregnancy and over the course of a woman's life. Approximately 28% of pregnant women risk a loss of pregnancy in developed countries[7]. Missed abortion in the first trimester (within 14 weeks of gestation; MAFT), which is characterized by the arrest of embryonic or fetal development[8], is a common complication of pregnancy. MAFT may occur in up to 15% of all clinically recognized pregnancies, especially in developing countries[9–11]. Determining whether or not the risk of MAFT responds to air quality conditions is important, as a pregnancy loss is devastating for the expectant parents[12].

Few quantitative studies exist that explore how maternal air pollution exposure affects the MAFT risk. Several studies[13,14] have been carried out in high-income countries with relatively good air quality. A study[15] conducted in Tianjin, China using the official air pollution monitoring data from 2001 to 2007 found that there were possible adverse impacts of air pollution on pregnancy outcomes; however, data during this period are unreliable[16]. In contrast, data collected since 2013, when the government enforced strict regulations to ensure air quality, are more reliable. Quantifying the relationship between maternal exposure to air pollutants and MAFT requires detailed, difficult-to-obtain information on personal exposures and confounders for a wide range of pollution exposures.

We investigated how the MAFT risk varies with the level of maternal ambient air pollution exposure using air pollution measurements and clinical data from pregnant women living in Beijing, China. Air pollutants considered for the study included particulate matter (PM) with diameter below 2.5 μm ($PM_{2.5}$), sulfur dioxide (SO_2), ozone (O_3) and carbon monoxide (CO). Given its size, Beijing has diverse terrains (Supplementary Fig. 1) and a considerable range of air quality conditions across space and time[17]. The spatial distribution of daily mean concentrations of $PM_{2.5}$ in Beijing, averaged from 2008 to 2017, indicates that the temporally average daily (Supplementary Fig. 2a) and maximum daily (Supplementary Fig. 2b) concentrations exceeded 100.0 μg m^{-3} at several locations, although the minimum concentrations (Supplementary Fig. 2c) were below 6.0 μg m^{-3} in many places. Daily mean concentrations of ambient SO_2, O_3 and CO in Beijing also showed large spatiotemporal variabilities (Supplementary Figs. 3–5). In addition, while Beijing is a well-developed region, it still has large rural areas with relatively low household income, such as those in Fangshan District (Supplementary Fig. 1). Considering the above factors, the exposure–response relationship between air pollution and MAFT risk derived from Beijing may be representative of the general situation in China.

We collected the clinical records of 255,668 pregnant women in Beijing from 2009 to 2017. The dataset contained information on maternal education level, occupation, residence and working places, and last menstrual date. Following earlier work[18], we computed the air pollutant exposure level of each pregnant woman on the basis of measurements at the nearest air monitoring stations from her

[1]State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing, China. [2]Beijing Obstetrics and Gynecology Hospital, Capital Medical University, Beijing, China. [3]Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China. [4]State Key Laboratory of Resources and Environment Information System, Institute of Geographical Science and Natural Resources, Chinese Academy of Sciences, Beijing, China. [5]School of Surveying and Geo-informatics, Tongji University, Shanghai, China. [6]These authors contributed equally: Liqiang Zhang, Weiwei Liu, Kun Hou. *e-mail: zhanglq@bnu.edu.cn; lffhornet01@163.com; linjt@pku.edu.cn; zhouch@lreis.ac.cn

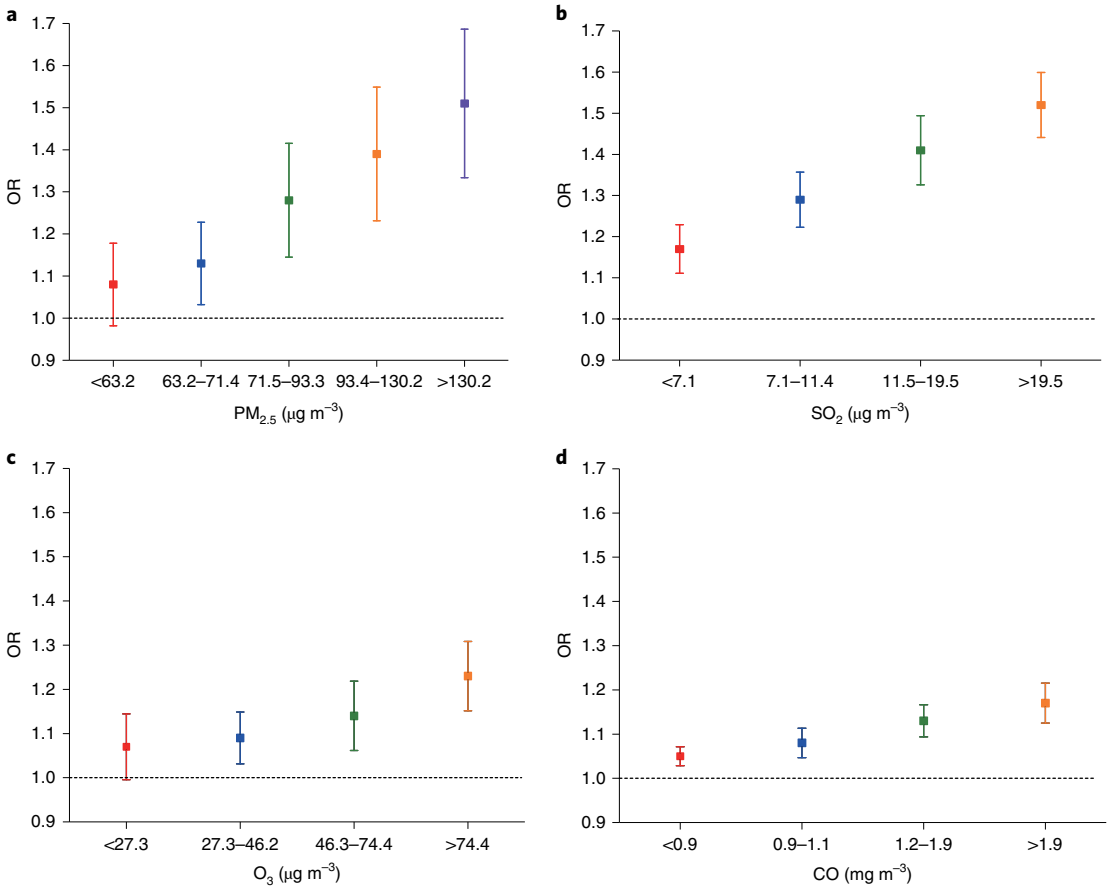

Fig. 1 | The ORs and 95% CIs of MAFT associated with maternal exposure to each pollutant in Phase 4. a–d, The OR of MAFT with respect to $PM_{2.5}$ (**a**), SO_2 (**b**), O_3 (**c**) and CO (**d**) exposure. Confounders were controlled here.

residential and working places (Methods). This choice, as opposed to previous studies that only considered the distance to residence place[19–21], was preferred because most pregnant women in the first trimester still went to work.

We used a logistic regression model and a restricted cubic spline model to explore the quantitative relationship between the MAFT risk and maternal exposure to air pollutants, including $PM_{2.5}$, SO_2, O_3 and CO, among the 255,668 pregnant women. Potential confounders, including maternal age at pregnancy, occupation, spatial dependence and ambient air temperature were controlled in the models. We performed several robustness analyses: (1) we tested whether the exposure–response association was characterized by different lag periods; (2) we assessed the exposure–response relationship in individual concentration ranges for each air pollutant; and (3) we conducted a restricted cubic spline analysis to characterize the morphology of the nonlinear association between air pollutant exposure and the MAFT risk. Model results are reported as odds ratios (ORs)[22] and their 95% confidence intervals (CIs).

Results

We grouped pregnant women by age at conception (five groups), occupation (two groups) and air temperature (four groups). Among the participating pregnant women, 17,497 (6.8%) experienced MAFT. We took the Bayes factor as a measure of evidence for the association between MAFT in different subgroups and air pollution. As a summary measure, the Bayes factor gave an alternative to the *P*-value for the ranking of associations or for the flagging of associations as significant[23]. The Bayes factor $BF_{10} > 30$ represents strong evidence for the associations between the MAFT occurring in the subgroups and air pollution[24,25] (Supplementary Table 1). Women older than 39 years at conception or female farmers and blue-collar workers had higher percentages of MAFT than their counterparts did. In all groups, maternal exposure to each air pollutant was associated with the risk of MAFT.

Associations between MAFT risk and air pollution exposure. We used a logistic regression model to calculate the ORs and 95% CIs for the association between MAFT risk and exposure to each pollutant. We adjusted for potential confounders including maternal age, occupation, spatial dependence and ambient temperature. As we estimated the association between $PM_{2.5}$ and MAFT risk, we did not control for other air pollutants (SO_2, O_3 and CO) in the model. The main reason is that the period of the available $PM_{2.5}$ data (from 2008 to 2017) was different from those of other pollutants (from June 2014 to December 2017). Moreover, the correlation between $PM_{2.5}$ and other two pollutants (SO_2 and O_3) is not high (Supplementary Table 2), thus controlling for these two pollutants had a little effect (Supplementary Table 3). As we estimated the association between O_3 and the MAFT risk, we controlled for CO and SO_2 but not $PM_{2.5}$, given the strong correlation between $PM_{2.5}$ and CO (Supplementary Tables 2 and 4). Similarly, as we estimated the association between SO_2 (CO) and the MAFT risk, we controlled for CO (SO_2) and O_3 but not $PM_{2.5}$.

We investigated the correlation between maternal exposures in different time periods and the incidence of MAFT, to test the effect of time lag between pollution exposure and MAFT. Seven

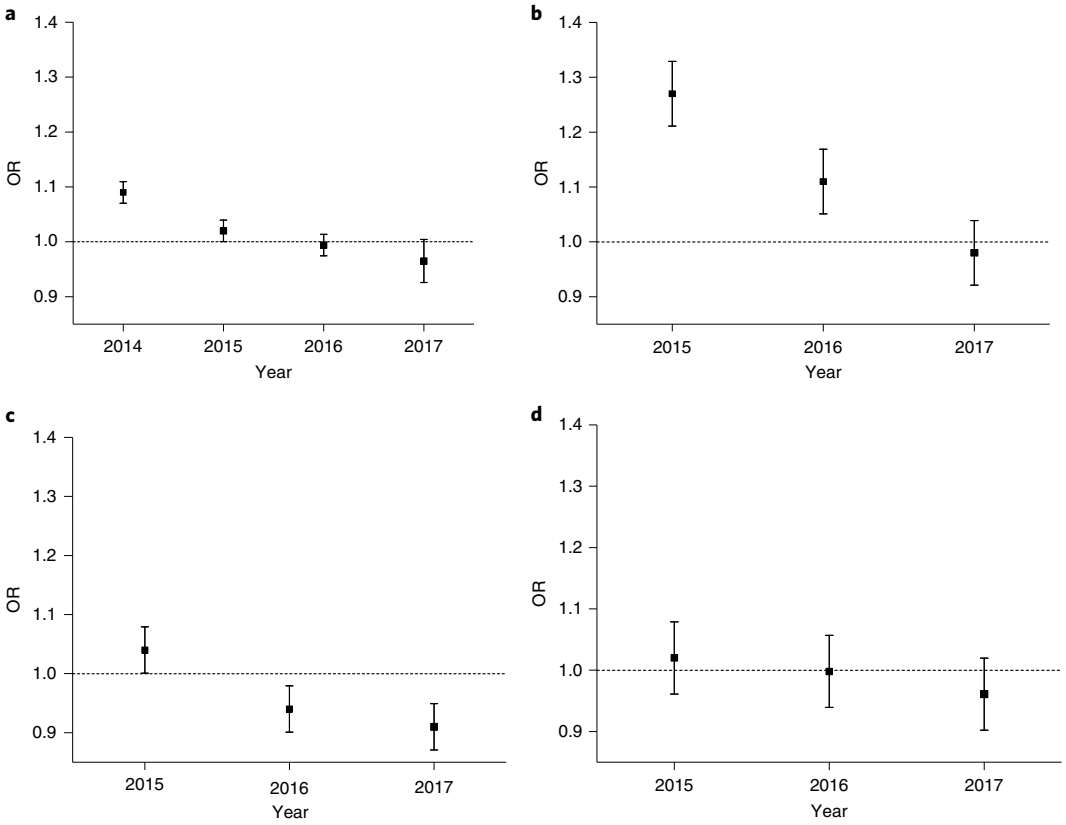

Fig. 2 | The ORs and 95% CIs of the MAFT risks associated with maternal exposure to different average annual air pollutant concentrations. a–d, Association of the MAFT with a 10.0 μg m^{-3} increase in $PM_{2.5}$ exposure from 2014 to 2017 (**a**), a 10.0 μg m^{-3} increase in SO_2 exposure from 2015 to 2017 (**b**), a 10.0 μg m^{-3} increase in O_3 exposure from 2015 to 2017 (**c**) and a 1.0 mg m^{-3} increase in CO exposure from 2015 to 2017 (**d**). Note that since the datasets of SO_2, O_3 and CO before June 2014 contained many missing values, for these pollutants we only used the pollutant data from June 2014 to December 2017 and thus the MAFT data from 2015 to 2017.

time windows of exposures (Phases 1–7; Methods) were considered. The risk of MAFT was associated with $PM_{2.5}$ exposure in all phases (Supplementary Fig. 6a). A 10.0 μg m^{-3} increase in the temporally averaged $PM_{2.5}$ concentration in Phase 4 (the period from 90 d before the first day of the last menstrual period (LMP) to the last exposure time (LET) of the pregnant women with MAFT; see Methods for definition of LET) was associated with the greatest risk of MAFT, when compared to the ORs from other phases. The MAFT risk also changed in different time windows for SO_2, O_3 and CO exposures and each had a peak risk with respect to Phase 4 (Supplementary Fig. 6). In the following, exposure to each air pollutant averaged in Phase 4 was used for further analysis.

Due to lag in air pollutant exposure, $PM_{2.5}$ data for the years 2008–2017 (Methods) were combined with the clinical data from 2009–2017 for analysis. Maternal $PM_{2.5}$ exposure in Phase 4 was categorized into five concentration ranges separated by the 25th, 50th, 75th and 95th percentiles of the $PM_{2.5}$ concentrations: <63.2 μg m^{-3}, 63.2–71.4 μg m^{-3}, 71.5–93.3 μg m^{-3}, 93.4–130.2 μg m^{-3} and >130.2 μg m^{-3}. An increase in ambient $PM_{2.5}$ concentrations was significantly associated with an increased MAFT risk (Fig. 1). For a 10.0 μg m^{-3} increase in $PM_{2.5}$ concentrations, the OR of MAFT (after adjusting for the confounders) was 1.08 (95% CI, 0.98–1.18) for $PM_{2.5}$ concentrations <63.2 μg m^{-3}, 1.13 (95% CI, 1.03–1.23) for $PM_{2.5}$ of 63.2–71.4 μg m^{-3}, 1.28 (95% CI, 1.14–1.42) for $PM_{2.5}$ of 71.5–93.3 μg m^{-3}, 1.39 (95% CI, 1.23–1.55) for $PM_{2.5}$ of 93.4–130.2 μg m^{-3} and 1.51 (95% CI, 1.33–1.69) for $PM_{2.5}$ >130.2 μg m^{-3}. The increase of the OR with increasing $PM_{2.5}$ concentrations was evident. Results based on $PM_{2.5}$ data over the period June 2014–December 2017 were similar (Supplementary Table 5).

SO_2, O_3 and CO data were available from 2014 to 2017 (Methods), thus we assessed the correlation between MAFT and each of SO_2, O_3 and CO over the same period. SO_2 concentration exposures ranged from 2.6 μg m^{-3} to 44.0 μg m^{-3}. Maternal SO_2 exposure in Phase 4 was categorized into four concentration ranges separated by the 25th, 50th and 75th percentiles of the SO_2 concentrations. For a 10.0 μg m^{-3} increase in SO_2 exposure, the ORs for SO_2 exposure in Phase 4 were 1.17 (95% CI, 1.10–1.22) for SO_2 concentrations <7.1 μg m^{-3}, increasing to 1.29 (1.22–1.36) for SO_2 of 7.1–11.4 μg m^{-3}, 1.41 (1.33–1.49) for SO_2 of 11.5–19.5 μg m^{-3} and 1.52 (1.44–1.60) for SO_2 concentrations >19.5 μg m^{-3} (Fig. 1b).

Maternal O_3 exposure in Phase 4 was categorized into four concentration ranges separated by the 25th, 50th and 75th percentiles of the O_3 concentrations. For a 10.0 μg m^{-3} increase in O_3 exposure, the ORs for O_3 exposure in Phase 4 were 1.07 (95% CI, 1.00–1.14) for O_3 concentrations <27.3 μg m^{-3}, increasing to 1.09 (1.03–1.15) for O_3 of 27.3–46.2 μg m^{-3}, 1.14 (1.06–1.22) for O_3 of 46.3–74.4 μg m^{-3} and 1.23 (1.15–1.31) for O_3 >74.4 μg m^{-3} (Fig. 1c).

Maternal CO exposure in Phase 4 was categorized into four concentration ranges separated by the 25th, 50th and 75th percentiles of the CO concentrations. For a 1.0 mg m^{-3} increase in CO exposure, the ORs for CO exposure in Phase 4 were 1.05 (95% CI, 1.03–1.07) for CO concentrations <0.9 mg m^{-3}, increasing to 1.08 (1.05–1.11) for CO of 0.9–1.1 mg m^{-3}, 1.13 (1.09–1.17) for CO of 1.2–1.9 mg m^{-3} and 1.17 (1.12–1.22) for CO >1.9 mg m^{-3} (Fig. 1d).

We further combined a restricted cubic spline model and a logistic regression model to construct the OR curve for MAFT and maternal pollution exposure for each pollutant, after other factors were controlled (Supplementary Fig. 7). The relationship between pollution exposure and the risk of MAFT was enhanced with increased pollutant concentrations, consistent with the above (logistic model based) findings.

Since 2013, the Chinese government has issued new rules to reduce atmospheric pollution. Air pollutant concentrations have substantially decreased since 2014 (Supplementary Fig. 8)[16]. The MAFT risk has also decreased since 2013 (Fig. 2), which further suggests the strong quantitative link between maternal air pollution exposure and MAFT risk.

Maternal characteristics, pollution exposure and risk of MAFT. Pregnant women with different sociodemographic status might be exposed to different air pollution levels and, therefore, could be subject to different MAFT risks. Older pregnant women (>39 years old), female farmers and blue-collar workers and those conceiving in low temperature (<5 °C) had higher ORs for the risk of MAFT associated with air pollution exposure when compared to their counterparts (for example, pregnant women aged 25–39 years) (Supplementary Fig. 9). In China, farmers and blue-collar workers usually had a low socioeconomic status and engaged in outdoor work[21]. Female farmers or female blue-collar workers were more exposed to ambient air pollution and thus subjected to a higher OR and a higher MAFT risk than office workers (as discussed above, the OR increases with increasing pollutant concentrations). This result was consistent with the previous finding suggesting that wealthier households were better able to avoid the adverse health impacts of hazardous environmental exposures[26–28].

Discussion

On the basis of a large record of maternal clinical data and a broad range of air pollution concentrations, this study has demonstrated a quantitative association between ambient air pollution exposure and risk of MAFT. Previous studies have also indicated that maternal long-term exposure to air pollution may mean a higher likelihood of abortion/miscarriage, stillbirth and birth defects[7,29,30].

We investigated several possible causal mechanisms to explain this linkage. Maternal long-term exposure to $PM_{2.5}$ allows the pollutant to cross the maternal–fetal blood barrier and ultimately perturb fetal growth and development[7,31]. Pollutants entering the bloodstream of a fetus might interact with its tissue components to produce pathological effects[17], leading to irreversible damage to the dividing cells of the fetus and triggering hypoxic harm or immunomediated injury during critical periods of development[32,33]. Air pollution-induced placental epigenetic alterations were observed during all trimesters of pregnancy[34]. This suggests that maternal exposure to air pollution might damage placental functions. Previous studies have shown that perturbations in the maternal environment could be transmitted to the fetus by changes in placental functions[31] and that ambient environmental insults on placenta had negative effects on the developing fetus[35].

In addition, poor air quality was significantly associated with the amount of polycyclic aromatic hydrocarbons (which can be absorbed by or adhere to $PM_{2.5}$) bound to DNA in both maternal and fetal cord white blood cells[36]. Mothers exposed to air pollution were more likely to have chromosomal abnormalities[37]. Therefore, maternal long-term exposure to air pollution increased the chances of abortion/miscarriage, stillbirth and birth defects.

Furthermore, toxicants could pass through the placenta and attack the developing fetus by potentially inducing alterations in immune competence[38]. CO might interfere with metabolic and transport function of the placenta and, after crossing the placental barrier, collect at higher concentrations in the fetus than in the mother[39]. Moreover, ambient CO was associated with carboxyhaemoglobin (COHb) and nucleated red blood cells[40]. Redundant COHb in mothers might cause fetal hypoxia, which could lead to fetal death[41].

Since there was strong collinearity between $PM_{2.5}$ and CO (Supplementary Tables 2 and 4), we could not separate the individual effects of these two pollutants. Although we were able to adjust for many known risk factors for MAFT that would confound the association, residual confounding cannot be ruled out, as it is possible that other factors we were unable to control for, such as traffic-related noise, may be associated with pregnancy outcome. The impacts of indoor air pollution on MAFT were not studied due to the lack of indoor pollution data, although indoor and ambient pollution (type and severity) are highly correlated.

Associating air pollution with the spatial–temporal variability in MAFT enhances scientific and policy understanding of pregnant women's health in developing countries[42]. Our findings uncovered potential opportunities to prevent or reduce harmful pregnancy outcomes by proactive measures before pregnancy. Meanwhile, our study helped us understand the relationship between air pollution exposure and a spectrum of reproductive outcomes.

Pregnant women or those who want to become pregnant, must protect themselves from air pollution exposure not only for their own health but also for the health of their fetuses. China is an aging society and our study provides an additional motivation for the country to reduce ambient air pollution for the sake of enhancing the birth rate. Although ambient air pollution has reduced in China in recent years[16], pollution levels are still high and must reduce further for many reasons, including reducing MAFT. Future work should explore the human health benefits from air pollution mitigation through modelling a wide range of environmental conditions using more data sources including land-use and land-cover change data[43].

Methods

Maternal clinical dataset. We collected, processed and selected maternal clinical data as explained below.

Collection of clinical data. We collected clinical data of 260,231 pregnant women in Beijing, China. Pregnancy outcomes were classified according to the International Classification of Diseases, 10th revision[44].

Validation of the dataset. Here we presented an independent validation of the dataset. Specifically, a midwife familiar with clinical coding techniques randomly selected 926 maternity case notes and then compared these case notes with those recorded in the dataset. The comparison results were re-checked by an obstetrician who had undergone training in clinical coding.

The 926 maternity case notes were randomly sampled at the maternity hospital wards (H-1, H-2, H-3 and H-4). Samples of 235, 247, 203 and 241 singleton deliveries were selected from these wards, respectively.

Two pregnancy outcome metrics were adopted to measure the quality of the data. The first metric was percentage agreement between the contents of ten selected fields recorded in the dataset and their counterpart data extracted from the maternity notes. The second metric uses Cohen's Kappa to assess the consistency between the two datasets.

For the 926 cases examined, the contents of the ten fields were consistent between the data and case notes, with the percentage agreement exceeding 95% for all fields and hospitals (Supplementary Table 4). The values of Kappa were also much larger than 0.6, indicating high agreement[45]. The validation result provided confidence in the reliability of the dataset.

Data screening for MAFT based on gestational age. We only considered fetal loss occurring at <98 d (14 weeks) of gestation in this study.

Gestational age was computed as the number of days between the date of the LMP and the date of fetal mortality. It was difficult to determine gestational age precisely since fetal mortality might happen weeks before it was found. In this study, we estimated the gestational age as follows.

We assumed the date of fetal mortality to be 2 weeks before the abortion date, when the number of weeks from the LMP to the abortion date was >7 weeks. Otherwise, we assumed the date of fetal mortality to be the abortion date, due to fetal heart not being monitored in the first 5 weeks of the first trimester. The abortion procedure was done on the date a fetus was determined dead for almost

all cases, except in rare occasions in which clinical complications and other issues led to a 1–2 d delay in the abortion procedure. Ultrasound was used to determine whether a fetus was alive or dead.

Pregnant women in Beijing usually had one prenatal care visit every month in the first trimester. They were checked by ultrasound when maternal or fetus anomalies were suspected based on regular, non-ultrasound examinations. In addition, vaginal bleeding often occurred after fetuses died. When pregnant women found vaginal bleeding, they usually went to hospital and were checked by ultrasound to determine whether their fetuses had died. The above two aspects helped to detect fetal death as early as possible and decreased the bias in the gestational age estimation. Moreover, since we assessed the relationship between air pollution and MAFT using clinical data over 9 years (from 2009 to 2017), the effect of errors in estimated gestational ages could be reduced.

Very early losses like biochemical pregnancy and stillbirth were beyond the scope of our research, thus pregnant women whose fetal gestational ages were less than 40 d or were more than 98 d from the LMP were excluded.

Other data screening. Pregnant women were asked to provide the address they had lived at for the longest in the half-year before and during pregnancy. Women were excluded from the current study if there were no records of their addresses. Those with irregular menstruation cycles or with a history of miscarriage were also excluded; the dataset recorded the miscarriage dates. We did not account for maternal smoking status, since most Chinese women do not smoke, especially before and during pregnancy.

Basic statistics of finally selected data. After all the aforementioned exclusions, data for a total of 255,668 women in Beijing from 2009 to 2017 were valid for analysis (Supplementary Table 5).

Air pollution data. We used air pollution data from 34 air quality monitoring sites (see their spatial distribution in Supplementary Fig. 10). Hourly measurements were established in 2013 and maintained by the Ministry of Ecology and Environment (MEE, formally the Ministry of Environmental Protection). Air pollution measurements at the MEE sites followed the official measurement standards[46]. For SO_2, O_3 and CO measurements, there were few valid data before June 2014, thus for these pollutants we only used data from June 2014 to December 2017 in this study. Daily mean pollutant concentrations were derived from hourly data of each day and the average of daily mean values during the exposure phase was used as the exposed concentration. Measured pollutants included PM_{10}, $PM_{2.5}$, SO_2, CO, NO_2 and O_3. The NO_2 measurements were not used here due to concerns regarding contamination by other nitrogen pollutants[47]. We excluded the PM_{10} data, which contained many missing values.

To extend the $PM_{2.5}$ data time period previous to June 2014, we made use of the long-term measurements taken by the US Embassy (http://www.stateair.net/web/historical/1/1.html; accessed 12 July 2018). $PM_{2.5}$ measurements at the US Embassy site used the beta-attenuation instrument. The US Embassy data were shown to be consistent with the MEE measurements[16]. For data at each MEE site from June 2014 to December 2017, we established a linear relationship with the US Embassy data on an hourly basis. Measurements at each MEE site were consistent with those at the US Embassy site: R^2 ranged from 0.47 to 0.93 with an average of 0.70 (Supplementary Table 6). This consistency allowed us to apply the linear relationship to prior periods when there were no MEE measurements, as done here. Our further test using only data from June 2014 to 2017 suggested a similar association between $PM_{2.5}$ exposure and MAFT (Supplementary Table 7), which supported our use of US Embassy data for earlier times.

Meteorological measurement data. Three-hourly data for air temperature at 2 m above ground were taken from the meteorological measurement station near the southwestern Fourth Ring Road of Beijing (Supplementary Fig. 10). Data at this station were reported to the World Meteorological Organization and maintained at the US National Oceanic and Atmospheric Administration National Centers for Environment Information (https://www.ncdc.noaa.gov/isd/data-access; accessed 24 October 2018). Daily mean air temperature was derived from 3-hourly data.

Interpolation of missing air pollutants and air temperature data. There was large diurnal variation in each air pollutant and meteorological variable. To fill in the missing meteorological or air pollution data to accurately quantify the exposure level of each pregnant woman, we interpolated the missing values using the same interpolation methods as in our previous study[48].

Maternal exposure to air pollutants. Maternal residential and working district addresses before or at the period of conception and air quality monitoring stations were geocoded to obtain their latitudes and longitudes. Most of the participating pregnant women did not change their residences and working places before the first trimester of pregnancy. We estimated maternal exposure to air pollution by attributing representative concentrations provided by the air quality monitoring stations closest to the maternal residence and working place after geolocalization.

Approximately 86% of the pregnant women provided the working addresses. Since most pregnant women in the first trimester still went to work in China, the women who provided work addresses were assumed to work; therefore we estimated pollutant concentrations based on both their residence and work addresses. For each pregnant woman, the exposure concentration of each air pollutant C_d was computed as $C_d = (C_{dw}/3) + (2C_{dr}/3)$, where C_{dw} and C_{dr} denote air pollutant concentrations at the air monitoring stations closest to the maternal working and residential addresses, respectively. The weights (1/3 and 2/3) approximately accounted for the times a pregnant woman spent at work and at home. For the other 14% of pregnant women who did not provide work addresses, we assumed that they did not go to work and we only used their residential addresses to estimate the pollution exposure.

To define the period of pollution exposure, we first determined the LET of pregnant women before MAFT. We took the date of the LMP plus gestational age as the LET of the pregnant woman. We examined whether and how the exposure–response association was affected by different time periods of pollution exposure. We examined seven time windows of maternal exposure to air pollution (Phases 1–7), each from 0, 30, 60, 90, 120, 150 or 180 d before the first day of the LMP to the LET. We calculated the mean daily concentrations (the average of 24-h average across multiple days) of the pollutants in different periods (Phase 1, Phase 2 and so on) during which a pregnant woman was exposed. We showed that exposure in Phase 4 (from 90 d before the LMP to the LET) had strongest association with MAFT and these data were used for the analyses.

Spatial generalized additive model. Spatial autocorrelation was considered in this study. Supplementary Fig. 11 shows that most pregnant women with MAFT were clustered in densely populated areas where air pollution was high.

The spatial generalized additive model was used to account for variation.

$$g(u) = \beta_0 + s(\mathrm{PM}_{2.5}, \mathrm{df}_1) + s(\mathrm{SO}_2, \mathrm{df}_2) + s(\mathrm{O}_3, \mathrm{df}_3) + s(\mathrm{CO}, \mathrm{df}_4) + s(\mathrm{lat}, \mathrm{long}) \quad (1)$$

where $u = \mathrm{E}(Y)$ is the mathematical expectation of Y and $Y(t = 1, 2, 3, \ldots, n)$ denotes the set of the participating pregnant woman; g is a monotonic link function of u; β_0 is the intercept; df represents the degree of freedom and is used to control the impact of various pollutants; (lat, long) denotes the location of the residence or working place of a pregnant women; s is a smooth function; and s(lat, long) denotes the impact of the spatial autocorrelation on the MAFT risk.

Supplementary Fig. 12 illustrates the partial residuals of s(lat, long) through controlling the spatial distribution of the pregnant women in the generalized additive model. It is noted that the spatial distribution of the pregnant women in the lower right corner of the domain had large residual differences and a large degree of aggregation, which indicated that the data may have the spatial autocorrelation.

Logistic regression and restricted cubic spline regression. We used a logistic regression model to evaluate the risk factors that influence MAFT. The results of the analysis showed that the mother's age, mother's occupation, ambient air temperature and maternal exposure to each of the pollutants $PM_{2.5}$, SO_2, CO and O_3 were correlated with the MAFT risk.

Potential confounding factors were controlled in the final logistic regression model. In addition, when associating each pollutant with MAFT, other pollutants were controlled in the logistic regression model[30]. Taking into consideration the possibility that the OR might be influenced by spatial dependence among participating pregnant women, the logistic regression model was formulated as

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \gamma(\mathrm{lat}) + \delta(\mathrm{long}) \quad (2)$$

where P denotes the probability of the MAFT risk and β_0 is a constant term; β_1, β_2, …, β_7 are the regression coefficients of the independent variables X_1 to X_7 (X_1 denotes the mother's age at conception, X_2 the mother's occupation, X_3 the ambient air temperature, X_4 is the ambient $PM_{2.5}$ concentration, X_5 is the ambient SO_2 concentration, X_6 is the ambient O_3 concentration and X_7 is the ambient CO concentration). The OR value of each independent variable is $\mathrm{OR}_i = \exp(\beta_i)$; γ and δ are the coefficients; (lat, long) is the geographical location of the maternal residence or work place.

We also used the restricted cubic spline regression model[49] to help associate air pollution exposure and the MAFT risk. In the restricted cubic regression spline, the function was linear in two intervals, $[t_0, t_1]$ and $[t_{k-1}, t_k]$, of a predicting variable, so the restricted cubic regression spline, RCS(X), can be described as

$$\mathrm{RCS}(x, k) = \sum_{i=1}^{k-1} \beta_i S_i(X) \quad (3)$$

with

$$S_1(x) = x$$

$$S_i(x) = (x - t_{i-1})_+^3 - \frac{(x - t_{k-1})_+^3 (t_k - t_{i-1})}{t_k - t_{k-1}} + \frac{(x - t_k)_+^3 (t_{k-1} - t_{i-1})}{t_k - t_{k-1}} \quad \text{if } i \geq 2$$

$$(x - t_{i-1})_+^3 = \begin{cases} (x - t_{i-1})^3 & \text{if } x \geq t_{i-1} \\ 0 & \text{else} \end{cases}$$

where k denotes the number of the nodes and S_i denotes the spline function; x is the value of a continuous exposure X, i is an integer and t denotes the endpoint of each interval. We selected four nodes representing the 25th, 50th, 75th and 95th percentiles of $PM_{2.5}$ concentrations and three nodes representing the 25th, 50th and 75th percentiles of SO_2, O_3, CO concentrations, respectively.

The logistic regression model could be combined with the restricted cubic splines to deal with the nonlinear relationship between the response variables and independent variables. We combined the two models to assess the exposure–response relationship. We associated MAFT and each air pollutant ($PM_{2.5}$, SO_2, CO and O_3) separately. Using the spline function RCS(x) to replace the independent variable x in equation (2), we estimated the nonlinear relationship between the exposure concentration of each air pollutant and the MAFT risk through equation (4).

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \cdots + \sum_{i=1}^{k-1} \beta_i S_i(X) + \gamma(\text{lat}) + \delta(\text{long}) \qquad (4)$$

Correlations between air pollutants. We computed the correlations between individual air pollutants, on the basis of daily mean air pollution concentration data from June 2014 to December 2017 that were obtained from 34 air pollution monitoring stations in Beijing (Supplementary Table 2). We found that $PM_{2.5}$ and CO had the strongest correlation. We constructed a multivariate linear model to further analyse the multicollinearity between pollutants. The variance inflation factor of $PM_{2.5}$ was larger than 10.0, suggesting that $PM_{2.5}$ had collinearity with other pollutants (primarily CO) (Supplementary Table 8). Similarly, the variance inflation factor of CO was close to 10.0, reflecting its high collinearity with $PM_{2.5}$.

Data availability

The collected data are available from the corresponding authors on reasonable request.

Code availability

The source code is available from the corresponding authors on reasonable request. It is copyrighted by Beijing Normal University and Beijing Obstetrics and Gynecology Hospital and is to be used only for educational and research purposes. Any commercial use is prohibited.

Received: 9 January 2019; Accepted: 27 August 2019;
Published online: 14 October 2019

References

1. Cohen, A. J. et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study. *Lancet* **389**, 1907–1918 (2015).
2. Jing, H. et al. Health impact of China's air pollution prevention and control action plan: an analysis of national air quality monitoring and mortality data. *Lancet Planet. Health* **2**, e313–e323 (2018).
3. Shah, P. S. et al. Air pollution and birth outcomes: a systematic review. *Environ. Int.* **37**, 498–516 (2011).
4. Malley, C. S. et al. Preterm birth associated with maternal fine particulate matter exposure: a global, regional and national assessment. *Environ. Int.* **101**, 173–182 (2017).
5. Brook, R. D. et al. Insights into the mechanisms and mediators of the effects of air pollution exposure on blood pressure and vascular function in healthy humans. *Hypertension* **54**, 659–667 (2009).
6. Dadvand, P. et al. Ambient air pollution and preeclampsia: a spatiotemporal analysis. *Environ. Health Perspect.* **121**, 1365–1371 (2013).
7. Ha, S. et al. Ambient air pollution and the risk of pregnancy loss: a prospective cohort study. *Fertil. Steril.* **109**, 148–153 (2017).
8. Gemzell-Danielsson, K. et al. Misoprostol to treat missed abortion in the first trimester. *Int. J. Gynaecol. Obstet.* **99**, S182–S185 (2007).
9. Steer, C. et al. Spontaneous abortion rates after natural and assisted conception. *BMJ* **299**, 1317–1318 (1989).
10. Daya, S. in *Textbook of Gynecology* (ed. Copeland, L. J.) 227–271 (WB Saunder, 2000).
11. Wood, S. Medical management of missed abortion: a randomized clinical trial. *Obstet. Gynecol.* **99**, 563–566 (2002).
12. Farah, S. & Lucy, K. Intrauterine fetal death. *Obstet. Gynaecol. Reprod. Med.* **19**, 1–6 (2009).
13. Green, R. S. et al. Residential exposure to traffic and spontaneous abortion. *Environ. Health Perspect.* **117**, 1939–1944 (2009).
14. Defranco, E. et al. Air pollution and stillbirth risk: exposure to airborne particulate matter during pregnancy is associated with fetal death. *PLoS ONE* **10**, e0120594 (2015).
15. Hou, H. Y. et al. Does ambient air pollutants increase the risk of fetal loss? A case-control study. *Arch. Gynecol. Obstet.* **289**, 285–291 (2014).
16. Liang, X. et al. $PM_{2.5}$ data reliability, consistency, and air quality assessment in five Chinese cities. *J. Geophys. Res. Atmos.* **121**, 10220–10236 (2016).
17. Ji, J. Air pollution and China's ageing society. *Lancet Public Health* **3**, e457–e458 (2018).
18. Bowe, B. et al. The 2016 global and national burden of diabetes mellitus attributable to $PM_{2.5}$, air pollution. *Lancet Planet. Health* **2**, e301–e312 (2018).
19. Chen, H. Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: a population-based cohort study. *Lancet* **389**, 718–726 (2017).
20. Smith, R. et al. Impact of London's road traffic air and noise pollution on birth weight: retrospective population based cohort study. *BMJ* **359**, j5299 (2017).
21. Wang, Y. et al. Association of long-term exposure to airborne particulate matter of 1 μm or less with preterm birth in China. *JAMA Pediatr.* **8**, e174872 (2018).
22. Mahalingaiah, S. et al. Perimenarchal air pollution exposure and menstrual disorders. *Hum. Reprod.* **33**, 512–519 (2018).
23. Wakefield, J. Bayes factors for genome-wide association studies: comparison with *P*-values. *Genet. Epidemiol.* **33**, 79–86 (2009).
24. Lee, M. D. & Wagenmakers, E.-J. *Bayesian Cognitive Modeling: A Practical Course* (Cambridge Univ. Press, 2013).
25. Schönbrodt, F. D. & Wagenmakers, E. J. Bayes factor design analysis: planning for compelling evidence. *Psychon. Bull. Rev.* **25**, 128–142 (2018).
26. Patt, A. G. et al. Estimating least-developed countries' vulnerability to climate-related extreme events over the next 50 years. *Proc. Natl Acad. Sci. USA* **107**, 1333–1337 (2010).
27. Smith, K. R. et al. in *Climate Change 2014: Impacts, Adaptation, and Vulnerability* (eds Field, C. B. et al.) 709–794 (Cambridge Univ. Press, 2014).
28. Weinberg, C. R., Moledor, E. & Baird, D. D. Is there a seasonal pattern in risk of early-pregnancy loss. *Epidemiology* **5**, 484–489 (1994).
29. Ciaula, A. & Bilancia, M. Relationships between mild PM_{10} and ozone urban air levels and spontaneous abortion: clues for primary prevention. *Int. J. Environ. Health Res.* **25**, 640–655 (2015).
30. Enkhmaa, D. et al. Seasonal ambient air pollution correlates strongly with spontaneous abortion in Mongolia. *BMC Pregnancy Childbirth* **14**, 146 (2014).
31. Slama, R. et al. Meeting report: atmospheric pollution and human reproduction. *Environ. Health Perspect.* **116**, 791–798 (2008).
32. Faiz, A. S. et al. Does ambient air pollution trigger stillbirth? *Epidemiology* **24**, 538–544 (2013).
33. Siddika, N. et al. Prenatal ambient air pollution exposure and the risk of stillbirth: systematic review and meta-analysis of the empirical evidence. *Occup. Environ. Med.* **73**, 573–581 (2016).
34. Tsamou., M. et al. Air pollution-induced placental epigenetic alterations in early life: a candidate miRNA approach. *Epigenetics* **13**, 135–146 (2016).
35. Stillerman, K. P. et al. Environmental exposures and adverse pregnancy outcomes: a review of the science. *Reprod. Sci.* **15**, 631–650 (2008).
36. Perera, F. P. et al. Molecular epidemiologic research on the effects of environmental pollutants on the fetus. *Environ. Health Perspect.* **107**(suppl 3), 451–460 (1999).
37. Byrne, J. et al. Morphology of early fetal deaths and their chromosomal characteristics. *Teratology* **32**, 297–315 (1985).
38. Šrám, R. J. et al. Ambient air pollution and pregnancy outcomes: a review of the literature. *Environ. Health Perspect.* **113**, 375–382 (2005).
39. Hardy, K. R. & Thom, S. R. Pathophysiology and treatment of carbon monoxide poisoning. *J. Toxicol. Clin. Toxicol.* **32**, 613–629 (1994).
40. Ziaei, S. et al. Effects of carbon monoxide air pollution in pregnancy on neonatal nucleated red blood cells. *Paediatr. Perinat. Epidemiol.* **19**, 27–30 (2005).
41. Singh, J. & Scott, L. H. Threshold for carbon monoxide induced fetotoxicity. *Teratology* **30**, 253–257 (1984).
42. Grace, K. Considering climate in studies of fertility and reproductive health in poor countries. *Nat. Clim. Change* **7**, 479–485 (2017).
43. Li, X. et al. A new global land-use and land-cover change product at a 1-km resolution for 2010–2100 based on human–environment interactions. *Ann. Am. Assoc. Geogr.* **107**, 1040–1059 (2017).
44. *International Statistical Classification of Diseases and Related Health Problems* 10th revision (WHO, 2007).
45. Altman D. G. *Practical Statistics for Medical Research* (Chapman and Hall, 1991).
46. *Technical Regulation on Ambient Air Quality Index (on Trial).* (Ministry of Environmental Protection of the People's Republic of China, 2012).
47. Liu, M. Y. et al. Spatiotemporal variability of NO_2 and $PM_{2.5}$ over Eastern China: observational and model analyses with a novel statistical method. *Atmos. Chem. Phys. Discuss.* **18**, 12933–12952 (2018).
48. Zhang, L. et al. Air pollution exposure associates with increased risk of neonatal jaundice. *Nat. Commun.* **10**, 6741 (2019).
49. Desquilbet, L. & Mariotti, F. Dose–response analyses using restricted cubic spline functions in public health research. *Stat. Med.* **29**, 1037–1057 (2010).

Acknowledgements

This work was carried out with the support of the National Key Research and Development Program of China (grant no. 2018YFC0213600), the National Natural Science Foundation of China (grant nos. 41775115 and 41371324) and the Beijing Natural Science Foundation (grant no. 7173258).

Author contributions

L.Z. and W.L. jointly designed the study, collected data, performed the analysis and wrote the manuscript. K.H. contributed to the model of the paper. J.L. contributed to the research framework, provided air pollution and meteorological data, contributed to results analysis and edited the paper. C.Z. and X.T. improved the research framework and edited the paper. Z.Y.W., Y.B.W. and Y.L. organized the neonatal and maternal datasets. Z.W.W., R.N. and M.L. generated the air pollution and meteorological datasets. Y.J., Y.Z., S.L. and P.Z. developed the maps and edited figures.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41893-019-0387-y.

Correspondence and requests for materials should be addressed to L.Z., W.L., J.L. or C.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.