

Letter



Association between PM_{2.5} Chemical Constituents and Preterm Birth: The Undeniable Role of Preconception *H19* Gene Variation

Yalong Wang^{1,2}, Panpan Sun^{2,3,&}, Xinying Wang¹, Junxi Zhang^{2,3}, Xiangyu Yu¹, Jian Chai^{2,3}, Ruo Du¹,
Wenyi Liu⁴, Fangfang Yu¹, Yue Ba^{1,5}, and Guo Yu Zhou^{1,2,5,#}

Preterm birth (PTB), defined as delivery before 37 weeks of gestation, is the most common adverse pregnancy outcome^[1]. PTB is a global health concern, with an estimated 13.4 million cases in 2020¹, accounting for more than one in 10 births worldwide. Compared to full-term births, PTBs are associated with a higher risk of short- and long-term complications, including bronchopulmonary dysplasia, necrotizing enterocolitis, visual impairment, and cerebral injuries^[2]. Despite substantial research efforts to prevent PTB, the global PTB rate has shown little improvement over the past decade¹. Therefore, identifying additional risk factors remains a critical goal in preventing PTB.

Air pollutants exert harmful effects on human health^[3], and recognized as significant environmental risk factor for adverse pregnancy outcomes^[4]. Among these pollutants, fine particulate matter (PM_{2.5}) is a major contributor to adverse pregnancy outcomes, including PTB. PM_{2.5} comprises various chemical constituents—black carbon (BC), organic matter (OM), nitrate (NO₃⁻), ammonium (NH₄⁺), and sulfate (SO₄²⁻)—which differ in their toxicity levels. A limited number of studies have investigated the relationship between specific PM_{2.5} constituents and PTB risk^[4]. Although some studies have identified BC, NO₃⁻, SO₄²⁻, and NH₄⁺ as key contributors to PTB^[4], gene-environment interactions suggest that the effects of these constituents may be more complicated than previously understood.

The *H19* imprinted gene plays a critical role in fetal development. Research has shown that the rs2107425 polymorphism in *H19* is associated with *H19* mRNA expression^[5]. Reduced expression of *H19*

has been linked to decreased *in vitro* migration and invasion of extravillous trophoblasts, ultimately affecting embryonic growth^[6]. Additionally, the rs2067051 polymorphism has been associated with birth weight defects^[7]. However, only a few studies have investigated the impact of maternal preconception *H19* gene variation on the risk of PTB.

Given this, the present study focuses on two maternal single-nucleotide polymorphisms (SNPs), rs2067051 and rs2107425, with the hypothesis that preconception *H19* gene variation is associated with the risk of PTB and may modify the relationship between exposure to PM_{2.5} chemical constituents and the risk of PTB. Understanding the complex patterns of gene-environment interactions associated with PTB risk may provide important insights for genetic counseling and targeted public health interventions.

A total of 6,000 women planning pregnancies (1,000 per county) residing in six districts (Supplementary Table S1) were recruited using successive sampling. With the support from local Maternal and Child Health Hospitals, China, professional nursing staff conducted preconception examinations, collected fasting peripheral blood samples, and performed genotyping. Ultrasound imaging was performed to check pregnancy status. Among the participants, 1,874 women who conceived within one year of their preconception examination and subsequently delivered live singleton infants were included in the study. These pregnancies occurred between June 2016 and June 2018. Demographic and clinical data collected from preconception examinations and birth records

doi: [10.3967/bes2025.076](https://doi.org/10.3967/bes2025.076)

1. Department of Environmental Health & Environment and Health Innovation Team, School of Public Health, Zhengzhou University, Zhengzhou, Henan, 450001, China; 2. National Health Commission Key Laboratory of Birth Defects Prevention, Zhengzhou, Henan, 450002, China; 3. Institute of Reproductive Health, Henan Academy of Innovations in Medical Science, Zhengzhou, Henan, 450002, China; 4. Medical Department, The Third Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, 450052, China; 5. Yellow River Institute for Ecological Protection & Regional Coordinated Development, Zhengzhou University, Zhengzhou, Henan, 450001, China

included mother's age, preconception body mass index (BMI, kg/m²), passive exposure to smoke (defined as exposure for at least 15 minutes per day over a period of at least one month), education level (middle school or below, and high school or above), season of conception (warm season: April–September; cold season: October–March), residential address, date of the last menstrual period, parity (primiparous and multiparous), date of delivery, gestational age, and infant sex (boy or girl). After excluding participants who were not genotyped due to insufficient DNA sample size or genotyping failure ($n = 25$), urban residents ($n = 1$), and women with gestational ages outside the viable range (< 20 or > 44 weeks, $n = 2$), a total of 1,846 mother-infant pairs were included in the study (Supplementary Table S1).

Daily average estimates of PM_{2.5}, BC, OM, NO₃⁻, NH₄⁺, and SO₄²⁻ were obtained at a 10 km² resolution using data from the Tracking Air Pollution, China (TAP) (<http://tapdata.org.cn>)^[8,9]. The concentrations of PM_{2.5} constituents from TAP were highly consistent with the observed measurements, with correlation coefficients exceeding 0.67 for all species^[8,9]. To assess maternal exposure, the geographical coordinates (latitude and longitude) of the mothers' residential addresses were extracted using Baidu Maps API (<https://lbsyun.baidu.com/>). Meteorological data, including ambient temperature (T, °C), relative humidity (RH, %), and concentrations of other air pollutants (PM₁₀, inhalable particles; CO, carbon monoxide; SO₂, sulfur dioxide; NO₂, nitrogen dioxide; O₃, ozone), were obtained as described in a previous study.¹⁰ Exposure windows were calculated based on the last menstrual period and the date of delivery. Four exposure periods were defined: first trimester (Trimester 1, weeks 1–12), second trimester (Trimester 2, weeks 13–27), third trimester (Trimester 3, ≥ 28 weeks), and the entire pregnancy.

Two SNPs in the *H19* gene (rs2067051 and rs2107425, Supplementary Table S2) were selected using Haploview software, with inclusion criteria of minor allele frequency > 0.05 and low linkage imbalance (LD, $r^2 < 0.8$). Genomic DNA was extracted from maternal peripheral blood samples using the Blood Genomic DNA Extraction Kit (Shanghai Genaray Biotechnology Corporation, Shanghai, China). Genotypes of the selected SNPs were determined using the Improved Multiple Ligase Detection Reaction (iMLDRTM) technique (Shanghai Genesky Biotechnologies Inc., Shanghai, China), which utilizes double ligation and multiplex fluorescence Polymerase Chain Reaction (PCR).

Ligation reactions were conducted using an ABI3730XL thermal cycler (Supplementary Tables S3–S4). All samples met quality control standards, with SNP call rates $> 95\%$, Hardy-Weinberg equilibrium test $P > 0.05$, and no amplification was observed in negative controls.

Descriptive statistics (percentages, means, and standard deviations) were calculated for the mother-infant pairs. Chi-square (χ^2) test and Student's *t*-test were used to compare differences between PTB and term birth groups. Pearson's correlation analysis was used to assess the correlations among T, RH, PM_{2.5}, and its constituents. Logistic regression models were used to evaluate the association between PM_{2.5}, its constituents, *H19* gene variants (rs2067051 and rs2107425), and PTB risk. Additionally, a weighted quantile sum (WQS) regression model was employed to estimate the overall combined impact of PM_{2.5} constituents on PTB risk and identify the key contributing pollutants. Subgroup analyses were conducted based on the SNP genotypes to examine the potential effects of modifications by *H19* variants. Interaction terms (Pollutants * SNP) were included in the logistic regression models assessing PM_{2.5} constituents and PTB associations. Furthermore, generalized multifactor dimensionality reduction (GMDR) was used to investigate gene-environment interactions between *H19* variants and PM_{2.5} constituents concerning PTB risk, using 10-fold cross-validation (<http://ibi.zju.edu.cn/software/>). Binary variables for PM_{2.5} and its constituents were created using their medians as cutoff points. Model performance was assessed by testing the balanced accuracy (TBA), cross-validation consistency (CVC), and *P*-values. TBA measured the accuracy of interaction models in predicting PTB status; CVC-reflected model stability and *P*-values assessed statistical significance. Known risk factors and potential confounders were adjusted for as covariates in the models, including maternal age, preconception BMI, season of conception, education level, exposure to smoke, parity, infant sex, T, and RH. Variable assignments are provided in Supplementary Table S5.

To assess the robustness of findings, a dual-pollutant model ("PM_{2.5} constituent + co-pollutant") was applied to examine the stability of the associations between PM_{2.5} constituents and PTB risk. Sensitivity analyses were conducted by restricting the population to women aged 18–35 (excluding adolescents and older mothers) and excluding parity as a covariate due to missing data ($\approx 15\%$).

GMDR analyses were conducted using GMDR software version 1.0, and all other statistical analyses were performed using R software version 4.3.3. Interaction effects were considered statistically significant at $P < 0.10$, while $P < 0.05$ was considered statistically significant for all other analyses.

Among the 1,846 mother-infant pairs enrolled in this study, 94 infants were considered with PTB. The mean maternal age and preconception BMI were 26.4 ± 3.66 years and 22.3 ± 3.50 kg/m², respectively. The PTB group had a higher proportion of male infants compared to the term birth group ($P < 0.05$) (Supplementary Table S1). Additionally, the frequency of the T allele of rs2107425 was significantly higher among mothers in the PTB group ($P < 0.05$) (Supplementary Table S1).

The average concentrations of PM_{2.5} and its constituents during the entire pregnancy were as follows: BC, OM, NO₃⁻, NH₄⁺, and SO₄²⁻ were 67.34, 2.78, 14.69, 16.20, and 11.88 µg/m³, respectively (Supplementary Figure S1). Furthermore, the average RH and T were 62.94% and 15.40°C, respectively. At all pregnancy stages, PM_{2.5} and its constituents showed high positive correlations with each other and were negatively correlated with RH and T ($P < 0.05$) (Supplementary Figure S1).

Overall, exposure to PM_{2.5} and its constituents was associated with an increased risk of PTB (Table 1). Adjusted model estimates indicated that each 10.0 µg/m³ increase in PM_{2.5} concentration was associated with a 0.47-fold increase in PTB risk (95% confidence interval [CI]: 1.19–1.83). For the entire pregnancy, exposure to BC (2.57, 95% CI: 1.52–4.43), OM (1.20, 95% CI: 1.08–1.33), NO₃⁻ (1.15, 95% CI: 1.06–1.26), NH₄⁺ (1.26, 95% CI: 1.10–1.46), and SO₄²⁻ (1.28, 95% CI: 1.13–1.46) was significantly associated with increased PTB risk. Further, these associations remained stable in the dual-pollutant model (Supplementary Table S6).

The results of the WQS regression analysis further supported the contribution of five constituents (BC, OM, NO₃⁻, NH₄⁺, and SO₄²⁻) to the risk of PTB (entire pregnancy: 1.68, 95% CI: 1.19–2.36; trimester 1: 1.51, 95% CI: 1.07–2.14; trimester 3: 1.98, 95% CI: 1.25–3.12) (Supplementary Table S7). Among these, SO₄²⁻ and BC were identified as the major contributors, with SO₄²⁻ playing a particularly significant role (Figure 1). SO₄²⁻ is a secondary pollutant primarily originating from industrial activities, power generation, and vehicular emissions^[4], while BC results mainly from gasoline

combustion, such as traffic exhaust emissions^[4]. These findings provide pertinent evidence supporting PTB prevention through traffic pollution control and reduction of coal combustion in factories.

Logistic regression analysis showed a significantly increased risk of PTB among mothers carrying the rs2107425 CT genotype (2.60, 95% CI: 1.50–4.81) or CT+TT genotypes (2.42, 95% CI: 1.42–4.42), compared to those with the CC genotype (Supplementary Table S8). However, rs2067051 was not significantly associated with the risk of PTB (Supplementary Table S8).

Furthermore, rs2107425 significantly interacted with PM_{2.5} and its constituents (BC, OM, NO₃⁻, NH₄⁺, and SO₄²⁻) (Figure 2 and Supplementary Figure S2). Mothers carrying the rs2107425 CT genotype were more susceptible to the effects of PM_{2.5} constituent exposure throughout pregnancy than those carrying the CC genotype ($P_{\text{interaction}} < 0.10$). Stratified analysis based on maternal genotypes revealed that mothers carrying the rs2067051 C allele or rs2107425 T allele had an increased risk of PTB associated with PM_{2.5} and its constituents, particularly among those with the rs2067051 CC or rs2107425 CT genotypes ($P < 0.05$). Sensitivity analyses that excluded adolescents and older mothers (Supplementary Tables S9 and Supplementary Figure S3) or did not adjust for parity (Supplementary Tables S9 and Supplementary Figure S4) yielded consistent results.

A GMDR analysis was performed to further evaluate gene-environment interactions. Interaction models, such as PM_{2.5}*rs2107425*rs2067051 and SO₄²⁻*rs2107425, showed significant differentiation (all $P = 0.0547$), indicating potential gene-environment interactions between PM_{2.5} constituents and these two SNPs. Based on TBA and CVC, the PM_{2.5}*rs2107425*rs2067051 model was identified as optimal (TBA, 58.15%; CVC, 10/10) (Supplementary Table S10). The risk distributions of the Pollutants*rs2107425*rs2067051 models are shown in Supplementary Figure S5. Particularly, mothers carrying the rs2107425 CT genotype, exposed to PM_{2.5} concentrations above the median (>50th percentile), and carrying the rs2067051 CC genotype had the highest PTB risk and the highest summed score in the optimal model. Similarly, the SO₄²⁻*rs2107425 model showed the highest PTB risk when high SO₄²⁻ exposure was combined with the rs2107425 CT genotype (Supplementary Figure S5). Based on the interaction analysis, we speculate that mothers carrying the rs2107425 CT and rs2067051

CC genotypes may be more vulnerable to the effects of PM_{2.5}, ultimately leading to impaired fetal development and an increased risk of PTB. However, further experimental studies are required to explore the mechanisms underlying this interaction.

This study had several limitations. Although the five constituents analyzed in this study do not

represent the entirety of PM_{2.5}, their significant contribution to the total PM_{2.5}-mass supports the relevance of our findings. The exposure assessment was solely based on the maternal residential address during pregnancy, without accounting for maternal mobility, which may have introduced potential exposure misclassification. Given the regional and

Table 1. Association between risk of PTB and exposure to PM_{2.5} constituents.

	Crude		Adjusted	
	OR (95% CI)	P	OR (95% CI)	P
PM_{2.5}				
Entire pregnancy	1.09 (0.94, 1.27)	0.270	1.47 (1.19, 1.83)	< 0.001
Trimester 1	1.05 (0.96, 1.13)	0.281	1.23 (1.07, 1.41)	0.003
Trimester 2	1.01 (0.91, 1.11)	0.878	1.16 (0.98, 1.37)	0.079
Trimester 3	1.07 (0.97, 1.18)	0.164	1.29 (1.10, 1.52)	0.002
BC				
Entire pregnancy	1.17 (0.81, 1.71)	0.400	2.57 (1.52, 4.43)	< 0.001
Trimester 1	1.12 (0.92, 1.33)	0.243	1.62 (1.19, 2.20)	0.002
Trimester 2	0.97 (0.76, 1.21)	0.762	1.19 (0.80, 1.75)	0.393
Trimester 3	1.13 (0.92, 1.38)	0.242	1.66 (1.16, 2.36)	0.006
OM				
Entire pregnancy	1.03 (0.96, 1.11)	0.460	1.20 (1.08, 1.33)	< 0.001
Trimester 1	1.01 (0.98, 1.06)	0.321	1.10 (1.03, 1.17)	0.003
Trimester 2	0.99 (0.95, 1.04)	0.871	1.05 (0.97, 1.13)	0.266
Trimester 3	1.02 (0.98, 1.07)	0.343	1.10 (1.02, 1.18)	0.015
NO₃⁻				
Entire pregnancy	1.02 (0.96, 1.08)	0.531	1.15 (1.06, 1.26)	0.001
Trimester 1	1.01 (0.98, 1.04)	0.451	1.07 (1.02, 1.13)	0.008
Trimester 2	0.99 (0.97, 1.03)	0.960	1.05 (0.99, 1.12)	0.118
Trimester 3	1.02 (0.99, 1.06)	0.253	1.09 (1.02, 1.16)	0.007
NH₄⁺				
Entire pregnancy	1.03 (0.94, 1.14)	0.513	1.26 (1.10, 1.46)	0.001
Trimester 1	1.02 (0.97, 1.07)	0.416	1.12 (1.03, 1.21)	0.009
Trimester 2	0.99 (0.94, 1.06)	0.894	1.08 (0.97, 1.20)	0.164
Trimester 3	1.03 (0.97, 1.09)	0.283	1.14 (1.03, 1.26)	0.010
SO₄²⁻				
Entire pregnancy	1.09 (0.99, 1.21)	0.089	1.28 (1.13, 1.46)	< 0.001
Trimester 1	1.05 (0.98, 1.11)	0.149	1.15 (1.05, 1.25)	0.002
Trimester 2	1.02 (0.95, 1.10)	0.592	1.11 (1.01, 1.23)	0.042
Trimester 3	1.07 (0.99, 1.15)	0.058	1.19 (1.08, 1.32)	< 0.001

Note. Odds ratio (OR) and 95% confidence interval (95% CI) was estimated per 10.0 ug/m³ in PM_{2.5} and 1.0 ug/m³ increment in BC, OM, NO₃⁻, NH₄⁺, and SO₄²⁻. PTB, preterm birth; PM_{2.5}, fine particulate matter; BC, black carbon; OM, organic matter; NO₃⁻, nitrate; NH₄⁺, ammonium; SO₄²⁻, sulfate.

population-specific limitations of this study, future studies should aim to validate these findings in a more diverse population. Additionally, we could not adjust for socioeconomic factors and the health status of a broader cohort of pregnant women due to data access restrictions. It is also important to consider that individuals with higher socioeconomic status may have greater health awareness and, consequently, might be more likely to participate in preconception health examinations, potentially introducing a selection bias. However, the biological mechanisms underlying these associations remain unclear. Further animal studies are required to elucidate these mechanistic pathways and validate our findings. Additionally, there are inherent biases in calculating gestational age based on the last menstrual period; a more accurate method, such as ultrasound imaging, will be employed in future studies. Finally, this study did not explore the critical windows of vulnerability during pregnancy, which will be addressed in future studies.

In summary, exposure to PM_{2.5} and its constituents—particularly SO₄²⁻ and BC—is associated with an increased risk of PTB. In addition, mothers carrying the rs2107425 CT genotype were more susceptible to PM_{2.5} and its constituents compared to those with the CC genotype. These findings underscore the importance of PM_{2.5} control measures and highlight the gene-environment interaction pathway involving H19 gene polymorphisms.

Funding This work was supported by the National Natural Science Foundation of China (82003401; 81972981), the Scientific and Technological Project of Henan Province (222102310150, China), and the Open Research Fund of the National Health Commission Key Laboratory of Birth Defects Prevention (ZD202205, China).

Competing interests The authors declare that they have no conflict of interest.

Ethics All participants were thoroughly briefed on

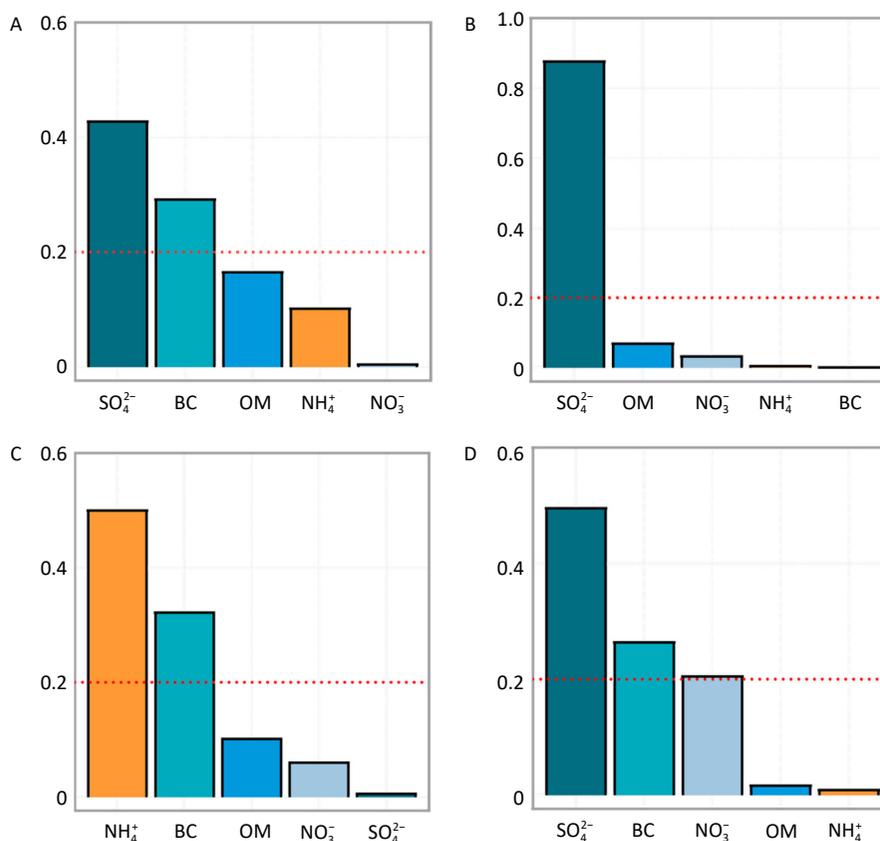


Figure 1. Weighted Quantile Sum (WQS) regression—Estimated weights of PM_{2.5} constituents associated with preterm birth (PTB). PM_{2.5}, fine particulate matter; BC, black carbon; OM, organic matter; NO₃⁻, nitrate; NH₄⁺, ammonium; SO₄²⁻, sulfate. A-D represented the entire pregnancy, trimester 1, trimester 2, and trimester 3, respectively.

the research protocol and provided written informed consent. Health examinations and the collection of biological samples were approved by the Ethics Committee of the Henan Academy of Population and Family Planning Science and Technology. The use of data and detection of biomarkers were approved by the Ethical Committees of the Henan Institute of Reproductive Health Science and Technology (HIRHST-IRB2020-02) and Zhengzhou University

(ZZUIRB2021-67).

Authors' Contributions Writing – original draft and data analysis: WANG Ya Long, SUN Pan Pan; Funding acquisition: SUN Pan Pan, BA Yue, ZHOU Guo Yu; Data collection: SUN Pan Pan, ZHANG Jun Xi, CHAI Jian; Writing - Review & Editing: WANG Xin Ying, YU Xiang Yu, ZHANG Jun Xi, DU Ruo, LIU Wen Yi, YU Fang Fang, BA Yue, CHAI Jian, ZHOU Guo Yu; Conceptualization, Methodology: ZHOU Guo Yu,

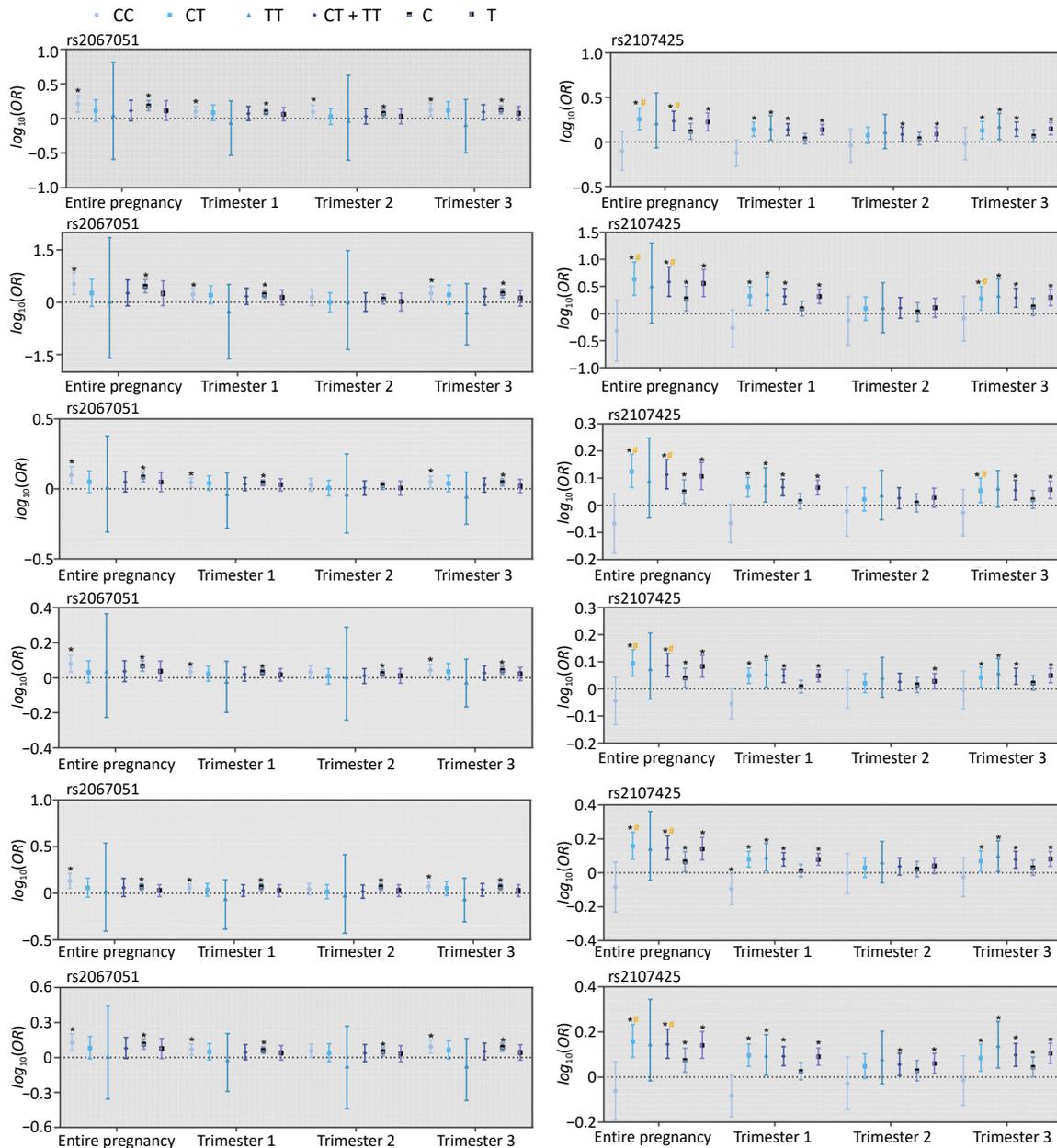


Figure 2. Association between preterm birth (PTB) risk and constituents of $PM_{2.5}$ stratified by maternal genotypes (rs2067051, rs2107425). * denoted a $P < 0.05$. # is $P_{interaction} < 0.10$, compared with the CC genotype or C allele in rs2067051 or rs2107025. $PM_{2.5}$, fine particulate matter; BC, black carbon; OM, organic matter; NO_3^- , nitrate; NH_4^+ , ammonium; SO_4^{2-} , sulfate.

WANG Ya Long.

Acknowledgements We sincerely thank all those who volunteered to participate in this study as well as to the staff of the Maternal and Child Health Hospital who participated in this study. We also thank WU Jing Jing, HU Yue, LI Qing Yuan, WEI Tian Tian, ZENG Yu Ting, CHENG Yi, QIN Qin Yang, and MAIREWULANA bola for their constructive reviews of this article.

Data sharing The authors do not have permission to share these data directly; however, it can be accessed by submitting an application to the Henan Institute of Reproductive Health Science and Technology, Zhengzhou, Henan, China. The Supplementary Materials will be available in www.besjournal.com.

[&]These authors contributed equally to this work.

[#]Correspondence should be addressed to ZHOU Guo Yu, associate professor, Ph.D, Tel: 15515772923; E-mail: zhouguoyu@zzu.edu.cn

Biographical note of the first author: WANG Ya Long, male, born in 2000, master of epidemiology and health statistics, focusing on the relationship between air pollution and adverse pregnancy outcome

Biographical notes of the first authors: WANG Ya Long, male, born in 2000, master of epidemiology and health statistics, focusing on the relationship between air pollution and adverse pregnancy outcome; SUN Pan Pan, female, born in 1990, attending physician, focusing on reproductive health and prevention of birth defects

Received: February 9, 2025;

Accepted: May 20, 2025

REFERENCES

1. Ohuma EO, Moller AB, Bradley E, et al. National, regional, and global estimates of preterm birth in 2020, with trends from 2010: a systematic analysis. *Lancet*, 2023; 402, 1261–71.
2. Platt MJ. Outcomes in preterm infants. *Public Health*, 2014; 128, 399–403.
3. Dehghani S, Moshfeghinia R, Ramezani M, et al. Exposure to air pollution and risk of ovarian cancer: a review. *Rev Environ Health*, 2022; 38, 439–50.
4. Cai J, Zhao Y, Kan JL, et al. Prenatal exposure to specific PM_{2.5} chemical constituents and preterm birth in China: a nationwide cohort study. *Environ Sci Technol*, 2020; 54, 14494–501.
5. Harati-Sadegh M, Kohan L, Teimoori B, et al. The effects of placental long noncoding RNA H19 polymorphisms and promoter methylation on H19 expression in association with preeclampsia susceptibility. *IUBMB Life*, 2020; 72, 413–25.
6. Zuckerwise L, Li J, Lu LG, et al. H19 long noncoding RNA alters trophoblast cell migration and invasion by regulating TβR3 in placenta with fetal growth restriction. *Oncotarget*, 2016; 7, 38398–407.
7. Adkins RM, Somes G, Morrison JC, et al. Association of birth weight with polymorphisms in the *IGF2*, *H19* and *IGF2R* Genes. *Pediatr Res*, 2010; 68, 429–34.
8. Geng GN, Zhang Q, Tong D, et al. Chemical composition of ambient PM_{2.5} over China and relationship to precursor emissions during 2005–2012. *Atmos Chem Phys*, 2017; 17, 9187–203.
9. Liu SG, Geng GN, Xiao QY, et al. Tracking daily concentrations of PM_{2.5} chemical composition in China since 2000. *Environ Sci Technol*, 2022; 56, 16517–27.
10. Wang QM, Wang YL, Sun PP, et al. Preconception mitochondrial DNA copy number plays a crucial role in linking prenatal air pollution with the risk of preterm birth. *Int J Hyg Environ Health*, 2024; 262, 114441.