Science of the Total Environment 630 (2018) 1–10

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Maternal exposure to ambient PM_{10} during pregnancy increases the risk of congenital heart defects: Evidence from machine learning models☆

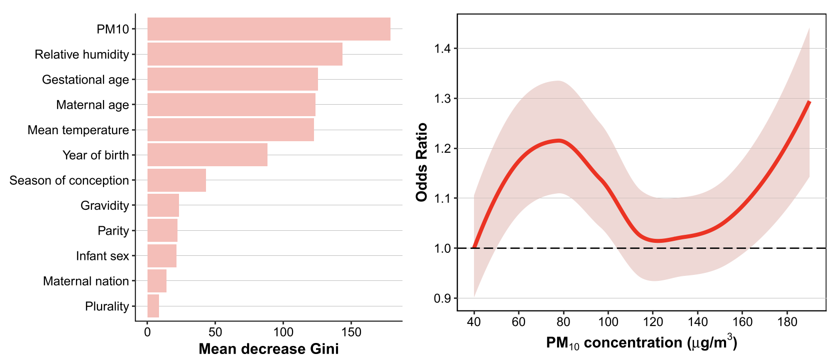
Zhoupeng Ren [a,1], Jun Zhu [b,c,1], Yanfang Gao [a,1], Qian Yin [a], Maogui Hu [a], Li Dai [b], Changfei Deng [b], Lin Yi [c], Kui Deng [c], Yanping Wang [b], Xiaohong Li [c,d,*], Jinfeng Wang [a,e,**]

[a] State Key Laboratory of Resources and Environmental Information System (LREIS), Institute of Geographic Science and Natural Resource Research, Chinese Academy of Sciences, Beijing, China
[b] National Office of Maternal and Child Health Surveillance of China, Department of Obstetrics, West China Second University Hospital, Sichuan University, Chengdu, China
[c] National Center for Birth Defect Monitoring of China, Department of Pediatrics, West China Second University Hospital, Sichuan University, Chengdu, China
[d] Key Laboratory of Birth Defects and Related Diseases of Women and Children, Sichuan University, Ministry of Education, China
[e] University of Chinese Academy of Sciences, Beijing, China

HIGHLIGHTS

- This study extends the application of machine learning model to birth outcomes and air pollution research.
- Maternal exposure to PM_{10} was identified as the primary risk factor for CHDs in two machine learning models.
- Our models consistently suggested that maternal exposure to PM_{10} can increase the risk of congenital heart defects.
- Machine learning model has better predictive performance than traditional logistic regression models.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 5 November 2017
Received in revised form 8 February 2018
Accepted 15 February 2018
Available online xxxx

Editor: Lidia Morawska

Keywords:
Congenital heart defects
Machine learning
Air pollution
Birth defects
Particulate matter

ABSTRACT

Previous research suggested an association between maternal exposure to ambient air pollutants and risk of congenital heart defects (CHDs), though the effects of particulate matter ≤10 μm in aerodynamic diameter (PM_{10}) on CHDs are inconsistent. We used two machine learning models (i.e., random forest (RF) and gradient boosting (GB)) to investigate the non-linear effects of PM_{10} exposure during the critical time window, weeks 3–8 in pregnancy, on risk of CHDs. From 2009 through 2012, we carried out a population-based birth cohort study on 39,053 live-born infants in Beijing. RF and GB models were used to calculate odds ratios for CHDs associated with increase in PM_{10} exposure, adjusting for maternal and perinatal characteristics. Maternal exposure to PM_{10} was identified as the primary risk factor for CHDs in all machine learning models. We observed a clear non-linear effect of maternal exposure to PM_{10} on CHDs risk. Compared to 40 μg m^{-3}, the following odds ratios resulted: 1) 92 μg m^{-3} [RF: 1.16 (95% CI: 1.06, 1.28); GB: 1.26 (95% CI: 1.17, 1.35)]; 2) 111 μg m^{-3} [RF: 1.04 (95% CI: 0.96, 1.14); GB: 1.04 (95% CI: 0.99, 1.08)]; 3) 124 μg m^{-3} [RF: 1.01 (95% CI: 0.94, 1.10); GB: 0.98 (95% CI: 0.93, 1.02)]; 4) 190 μg m^{-3} [RF: 1.29 (95% CI: 1.14, 1.44); GB: 1.71 (95% CI: 1.04, 2.17)]. Overall, both machine models showed an

☆ The authors declare they have no actual or potential competing financial interests.

* Correspondence to: X.H. Li, Department of Pediatrics, West China Second University Hospital, Sichuan University, Sec.3 No.17, South RenMin Road, Chengdu, China.
** Correspondence to: J.F. Wang, LREIS, Institute of Geographic Science and Natural Resource Research, Chinese Academy of Sciences, A11, Datun Road, Chaoyang District, Beijing 100101, China.
E-mail addresses: iiaoong@163.com (X. Li), wangjf@lreis.ac.cn (J. Wang).
[1] Contributed equally to this study.

https://doi.org/10.1016/j.scitotenv.2018.02.181
0048-9697/

association between maternal exposure to ambient PM_{10} and CHDs in Beijing, highlighting the need for non-linear methods to investigate dose-response relationships.

1. Introduction

Over the past two decades, from 1990 to 2013, the under-5 mortality rate in China has decreased by 78% (Wang et al., 2016). Between 1990 and 2015, China achieved the Millennium Development Goal 4 (MDG 4) target of reducing the under-5 mortality rate by two-thirds (He et al., 2017). The substantial reduction in the under-5 mortality was mainly attributed to successful prevention of malnutrition and infectious diseases, such as pneumonia, diarrhea, and malaria (Rudan et al., 2010; Wang et al., 2014). Due to these declines, birth defects became the main cause of under-5 mortality, which accounts for 20% of deaths considering all causes of death (Cui et al., 2016). For example, a new study of cause-specific under-5 mortality in 2015 (He et al., 2017) reported that congenital abnormalities (35,700 deaths) and preterm birth complications (30,900 deaths) were the leading cause of child deaths in China. Of the congenital abnormalities, congenital heart defects (CHDs) were the leading cause of infant deaths and accounted for approximately 27% of all birth defects in China (Zhang et al., 2016a). Moreover, a recent report of birth defects prevention in China showed that CHDs cases increased rapidly throughout the past decade (Zhang et al., 2016a), but few studies have provided sufficient evidence to interpret this upward trend.

Although the etiology of the majority of CHDs is not fully understood, both genetic and environmental factors are likely contributors to the occurrence of CHDs (Dadvand et al., 2011; Zhang et al., 2016a). Epidemiological studies of maternal exposure during pregnancy to ambient air pollution including particulate matter (e.g., PM_{10}) have shown associations with birth defects, such as premature birth (Padula et al., 2014), low birth weight (Bell et al., 2007) and CHDs (Agay-Shay et al., 2013). However, the available evidence linking PM_{10} exposure to CHDs is still limited and controversial. For example, Agay-Shay et al. (2013) found that maternal exposure to high PM_{10} concentrations significantly increased the risk of critical CHDs in offspring. However, positive but non-significant associations (Zhang et al., 2016a), or even negative associations (Schembari et al., 2014), between increased PM_{10} exposure and CHDs (e.g., ventricular septal defect) were also reported. These inconsistencies distort understanding of the mechanism of ambient PM_{10} exposure and CHDs. Therefore, it is essential to examine the hypothesis that maternal exposure to PM_{10} in early pregnancy will elevate the risk of CHDs.

The impact of ambient air pollution on health has become a major concern for policy makers and scientists in China in recent years. Although China's ambient air pollution problem has captured global attention, the scientific evidence on whether maternal exposure to unhealthy air will increase the risks of birth defects is still sparse (Jiang et al., 2007; Qian et al., 2016; Rich et al., 2015; Zhao et al., 2015), particularly for CHDs (Zhang et al., 2016a, 2016b; Jin et al., 2015). Ambient air pollution in China is different from that in developed countries due to the higher magnitude and longer duration of exposure (Guan et al., 2016). Air quality monitoring observations that in Beijing from the 2008–2012 showed PM_{10} concentrations varied from <10 μg/m^3 to >500 μg/m^3 (Hu et al., 2013). Consequently, previous studies in developed countries that were based on relatively lower PM_{10} concentrations could not provide conclusive evidence for developing countries to use for predictive purposes.

In this study, we used the random forest (RF) and gradient boosting (GB) models to investigate the effects of PM_{10} exposure on risk of CHDs during the critical time window for cardiac development in pregnancy, specifically weeks 3–8. Using live births and CHDs records from mothers residing in Beijing during their pregnancies, we hypothesized that increased PM_{10} concentrations would be associated with increased risk of CHDs in a developing country.

2. Methods

2.1. Study population

The CHDs data used in this study were obtained from a population-based birth defects surveillance system in China. This system was built in 2007, which included 64 surveillance counties located in 30 provinces (except for Tibet). The surveillance system uses three-tiered data collection networks to capture/measure neighborhoods: a) in urban areas, Committee (1)-Street (2)-District (3), are used; b) in rural areas, Village (1)-Town (2)-County (3) are used. Trained doctors in the village/neighborhood Committee Health Center are in charge of data collection. The pregnant women lived in the surveillance sites more than one year. Perinatal births (including stillbirth, live birth, and therapeutic induced labor) were included when their mother lived the monitoring area longer than one year. Annually 300,000 live births (in urban and rural areas) aged ≥28 gestational weeks that was determined by mother's last menstruation were under surveillance during 2009 to 2012. Babies weighed at least 1000 g at birth were included (even if gestational record information was lacking). Live-born neonates were visited, respectively, at 0–6, 7–27, and 28–42 days after birth, by the trained doctors at the township or street health care center, consistent with the *National Technical* Specification for Newborn Visit and National Work Specification for Maternal Health Care enacted by the Health Ministry of China. Any babies with congenital anomalies diagnosed within 42 days after birth should be reported. Basic information such as maternal age, maternal nation, parity and infant weight, and birth date was recorded.

Birth records from 2009 to 2012 in Beijing from the two population surveillance districts were selected in terms of the availability of PM_{10} data in China and diagnosis level for CHDs. Fig. 1 shows the location of the two districts. Xicheng and Huairou represent the urban and suburban populations, respectively. Cases were classified according to the *International Classification of Diseases, Tenth Revision* (ICD-10). ICD-10 categorizes CHD if any of the following subtypes were confirmed: Q20-Q26 (ICD-10 codes). Data provided include year of birth, infant sex, maternal age, district of residence, parity, *hukou* information, gravidity, season of conception, plurality, gestational age, as well as maternal nation. Gestational age was included because some prior studies have found preterm birth can increase the risk of CHDs (Laas et al., 2012; Wogu et al., 2014). *Hukou* information describes whether the pregnant woman was a Beijing native or an immigrant who moved from other places. This variable categorized the study population into three groups: 1) native; 2) immigrant who lived here (Beijing) shorter than one year; 3) immigrant who lived here (Beijing) longer than one year. Births with missing maternal age, *hukou*, and exposure information were excluded from analyses.

2.2. Ambient PM_{10} concentrations and exposure assignment

Ambient PM_{10} observations were obtained from our previous study in Beijing (Hu et al., 2013). The daily average PM_{10} concentrations were estimated by Air Pollution Index and corresponding main pollutant type obtained from Beijing Environmental Protection Bureau. PM_{10} data were monitored and recorded hourly from four monitoring stations (*Dongsi*, *Tiantan*, *Guanyuan*, *Wanshouxigong*) in the Xicheng district (urban) and three monitoring stations (*Huairouzhen*, *Miyunshuiku*, *Miyunzhen*) in the Huairou district (suburban), respectively (Fig. 1).