Environment International 119 (2018) 388–397

Contents lists available at ScienceDirect

Environment International

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

Association between birthweight and ambient $PM_{2.5}$ in the United States: Individually-varied susceptibility and spatial heterogeneity☆

Tao Xue[a], Tong Zhu[a,*], Yiqun Han[b]

[a] BIC-ESAT, SKL-ESPC, College of Environmental Science and Engineering, Peking University, Beijing 100871, China
[b] Analytical, Environmental and Forensic Sciences, School of Population Health & Environmental Sciences, Kings College London, United Kingdom

ARTICLE INFO

Handling Editor: Xavier Querol

Keywords:
Fine particulate matter
$PM_{2.5}$
Birthweight
Susceptibility
Infant health

ABSTRACT

The association between maternal exposure to $PM_{2.5}$ and birthweight varies geographically, which may be caused by susceptibility. Whether this population-level association is a function of mixtures of individuals with different susceptibilities is unclear. We investigated the probability distribution of individuals with different susceptibilities to $PM_{2.5}$-related birthweight change, and evaluated spatial variation of the effect across the United States (US). We estimated the individual-level susceptibility using the effect of $PM_{2.5}$ among a homogenous subpopulation, which was defined by a specific combination of modifiers. According to frequencies for all combinations, we derived the probability distribution of differential susceptibilities across the US and by states. From birth certificates across the US (1999–2004), we analyzed a total of 18,317,707 samples of singletons. Of the samples, 54–55% were assigned valid exposures, and linked to $PM_{2.5}$. The subpopulation-specific associations of $PM_{2.5}$ on birthweight change (*i.e.*, susceptibilities) ranged from negative to positive. For the first-trimester exposure, 61.4% of the associations were negative, and the mean was −1.01 g (95% confidence interval, CI: −1.63, −0.38) of birthweight change per 5 μg/m^3 increase of $PM_{2.5}$. The state-level associations varied (from −2.04 g [−2.76, −1.31] in New Hampshire to −0.30 g [−1.01, 0.41] in Texas) with demographic compositions in the US. The between-state variations of maternal race and education level were the greatest contributors to the spatial heterogeneity. Our findings may be useful to the policymaker in planning interventions for subpopulations susceptible to ambient pollution.

1. Introduction

Maternal exposure to ambient pollutants, including fine particulate matter with an aerodynamic diameter of < 2.5 μm ($PM_{2.5}$), is associated with decreased birthweight and the incidence of low birthweight (LBW) (Dadvand et al., 2013; Ebisu and Bell, 2012; Ebisu et al., 2016; Hao et al., 2016; Parker and Woodruff, 2008; Pedersen et al., 2013; Stieb et al., 2016), a risk factor for infant morbidity and mortality and development of diseases during adulthood (McCormick, 1985). However, these associations differ between studies (Dadvand et al., 2013; Sun et al., 2016) and vary geographically (Ebisu et al., 2016; Hao et al., 2016; Parker and Woodruff, 2008). Few studies have assessed the reasons underlying this heterogeneity. The effect on birthweight of a given ambient $PM_{2.5}$ level varies among subpopulations; *e.g.*, different ethnicities (Ebisu and Bell, 2012); this is termed differential susceptibility to $PM_{2.5}$ (Bell et al., 2013; Sacks et al., 2011). The fraction of susceptible individuals in the surveyed population varies among studies and geographically. Thus, a comprehensive evaluation of individual variations in susceptibility to $PM_{2.5}$ is warranted.

The terms *susceptibility*, *vulnerability*, and *sensitivity* are used interchangeably to denote inter-individual variation in the risk of adverse health outcomes per unit increment in ambient exposure to pollutants (Bell et al., 2013). Regardless of the subtle distinctions among these terms, in this study we use *susceptibility* to represent the magnitude of toxicity of air pollution to an individual. Susceptibility is dependent on internal factors (*e.g.*, genetics and underlying disease[s]), external factors (*e.g.*, socioeconomic status), and exposure patterns (*e.g.*, travel). Epidemiological studies use the term 'effect-modifier' to denote these factors or their surrogates, and subject them to interaction analyses. However, previous studies explored single effect-modifiers separately. For example, the $PM_{2.5}$-LBW association is reportedly stronger among white mothers (Ebisu and Bell, 2012). The simplicity of such studies may preclude evaluation of the different levels of susceptibility among the general population. Previous studies have reported three-way (Dubowsky et al., 2006) or higher-order interactions (Rosa et al., 2017) between the health effects of $PM_{2.5}$ and individual characteristics,

☆ Declarations of interest: none.

* Corresponding author at: Room 501, College of Environmental Science and Engineering, Peking University, Beijing 100871, China.
E-mail address: tzhu@pku.edu.cn (T. Zhu).

https://doi.org/10.1016/j.envint.2018.07.003
Received 26 February 2018; Received in revised form 27 June 2018; Accepted 3 July 2018

which indicates that multiple effect-modifiers might contribute to susceptibility both cumulatively and dependently. Additionally, few previous studies quantified the fraction of susceptible individuals among the general population, which is determined by not only the modifying effects of individual characteristics but also their joint probability distribution among the target population.

To fully characterize susceptibility to $PM_{2.5}$, we assume that the individual-specific health effect is determined by multiple effect-modifiers; thus, susceptibility may be quantified as a function of multiple variables at the level of the individual. Therefore, individuals with identical effect-modifiers have identical susceptibilities. In other words, the individual-specific effect (*i.e.*, susceptibility) can be estimated from a homogenous subpopulation. For the $PM_{2.5}$-related birthweight change, we collected 11 infant and maternal variables as effect-modifiers or their surrogates to represent the gradient variation of individual-specific susceptibilities. These variables were selected according to previous findings and data availability, and are described in the following section. In this study, we quantify the individual-specific magnitudes of susceptibility as the birthweight changes per unit increment of $PM_{2.5}$, which are estimated in homogenous subgroups categorized by the 11 variables. By combining the different susceptibilities with the probabilities of the corresponding subgroups, we derived a new statistical measure, the 'human susceptibility distribution', which reflects both the magnitude and the prevalence of susceptibility in the general population.

Although several studies have linked nationwide data of birthweight or LBW to ambient particles (*e.g.*, $PM_{2.5}$ [Hao et al., 2016; Parker and Woodruff, 2008], $PM_{2.5}$ [Ebisu and Bell, 2012], $PM_{2.5–10}$ [Ebisu et al., 2016], and PM_{10} [Parker and Woodruff, 2008]) in the United States (US), none explored susceptibility to the effect of $PM_{2.5}$. Using US birth certificates and the $PM_{2.5}$ concentrations monitored by national networks, we derived the human susceptibility distribution of the $PM_{2.5}$-related change in birthweight. Furthermore, by considering geographic variation in the proportions of susceptible individuals in the US, we assessed the spatial variation of $PM_{2.5}$-related birthweight change as a practical application of the human susceptibility distribution.

2. Methods

2.1. Study population

Birth certificate data of the contiguous US from 1999 to 2004 were obtained from the National Center for Health Statistics, Centers for Disease Control and Prevention. This database was in previous studies on the adverse effects of air pollutants on infants (*e.g.*, Hao et al., 2016). Many individual-level variables on both newborns and their mothers, such as county of residence, date of last menstrual period (LMP), and birthweight, were available during 1999–2004. Because many population characteristics could affect the susceptibility to $PM_{2.5}$-related birthweight change, we targeted the 11 individual characteristics (Table 1) used as the modeling covariates in previous studies (Ebisu and Bell, 2012; Ebisu et al., 2016).

We prepared the birth data as described previously (Ebisu and Bell, 2012; Ebisu et al., 2016). Briefly, we first excluded plural deliveries, as the target population was singleton births. Second, we assumed that the pregnancy period began 2 weeks after the LMP and was equal to the reported gestational duration. The reported month of birthdate was used to validate the estimated gestational period. We excluded birth records when the difference between the estimated delivery date and the middle day of the birth month was > 30 days. Third, based on a previous study (Alexander et al., 1996), we excluded records with impossible combinations of gestational age and birthweight. Fourth, we excluded records with missing values for any of the 11 individual characteristics. All births from California were removed because maternal status on smoking or tobacco usage was not recorded on California birth certificates. After applying the above exclusion criteria, 18,317,707 records were analyzed (Web Fig. 1).

2.2. Exposure assessment

Daily values of $PM_{2.5}$ from January 1998 to December 2004 were obtained from the US Environmental Protection Agency Air Quality System network (Web Fig. 1). We first assigned and averaged the monitored $PM_{2.5}$ levels by county, and prepared the exposure values during the entire pregnancy and each of the three trimesters (first trimester, 1–13 weeks; second trimester, 14–26 weeks; and third trimester, 27 weeks to delivery) as described previously (Ebisu and Bell, 2012). Briefly, we first calculated weekly averages and then derived the exposure value during the entire pregnancy or each trimester based on the estimated period of pregnancy for each birth, if > 75% of the weekly values were available. We obtained temperature data from January 1998 to December 2004 from the National Climatic Data Center, and transformed the weekly averages of monitored temperature into county-level averages during the entire pregnancy or each trimester, analogously.

2.3. Statistical analysis

The statistical analysis procedure is shown in Fig. 1. Briefly, among all valid birth records, those assigned with environmental exposures ($PM_{2.5}$ and temperature) were used as regression samples to link birthweight to maternal exposure to $PM_{2.5}$. We used a batch of separate regressions to derive subpopulation-specific susceptibilities. We also used all valid records to estimate the frequencies of the subpopulations among the total study population. The two results were combined to estimate the human susceptibility distribution.

First, we generated subpopulation indexes (i) based on 10,368 ($2^7 \times 3^4$) combinations of the population characteristics ($\mathbf{\Omega}_i$), which included 7 binary and 4 trinary variables. Among them, 8632 were present in the population according to valid records of singleton births. Only a portion of the valid birth certificates (those used as regression samples) could be simultaneously assigned levels of exposure to both $PM_{2.5}$ and temperature, due to the incomplete spatiotemporal coverage of the monitoring networks. The sample sizes are summarized in Table 1.

We then attempted to link birthweight to maternal exposure to $PM_{2.5}$ during the entire pregnancy and each trimester, independently within each subpopulation, with adjustments for three spline terms to control for the nonlinear confounding effects of temperature, the centroid coordinates of county of residence, and the temporal index (Fig. 1: Eq. (1)). To model seasonality, 4 degrees of freedom per year were utilized in the spline term of temporal index. For some subpopulations, the regression sample size (n_i) might be too small to generate a robust estimate of the effect of $PM_{2.5}$. The statistical power of the regression model (Fig. 1: Eq. (1)) increased to > 0.8 at a significance level of 0.05 when the sample size was > 1800 (Web Fig. 2). Therefore, we combined the subpopulations with sample sizes < 1800, and estimated the subpopulation-specific effects using a mixed-effects model with a random slope and a random intercept (Fig. 1: Eq. (2)). In this model, the subpopulation-specific effects ($\beta_0 + \beta'_i$) were assumed to be normally distributed with a mean value of β_0 and a standard deviation of ${\sigma_\beta}^2$. By combining the results of Eqs. (1) and (2), we derived the subpopulation-specific effects ($\beta_i \mid \mathbf{\Omega}_i$), which were used to quantify the variation in susceptibility according to the population characteristics ($\mathbf{\Omega}_i$).

Next, we approximated the probability that an individual belonged to a specific subpopulation ($p_i \mid \mathbf{\Omega}_i$) using the frequency of the subpopulation among all valid records of births. By considering the probability distribution of subpopulations ($p_i \mid \mathbf{\Omega}_i$) and the susceptibilities of these subpopulations to the $PM_{2.5}$-related birthweight change ($\beta_i \mid \mathbf{\Omega}_i$), we assessed the human susceptibility distribution for maternal exposure during the entire pregnancy and each trimester according to Eq. (3) (Fig. 1).