

Assessment of critical exposure and outcome windows in time-to-event analysis with application to air pollution and preterm birth study

HOWARD H. CHANG*

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, USA
howard.chang@emory.edu

JOSHUA L. WARREN

Department of Biostatistics, Yale University, New Haven, CT 06510, USA

LNYDSEY A. DARROW

Department of Epidemiology, Emory University, Atlanta, GA 30322, USA

BRIAN J. REICH

Department of Statistics, North Carolina State University, Raleigh, NC 27695, USA

LANCE A. WALLER

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, USA

SUMMARY

In reproductive epidemiology, there is a growing interest to examine associations between air pollution exposure during pregnancy and the risk of preterm birth (PTB). One important research objective is to identify critical periods of exposure and estimate the associated effects at different stages of pregnancy. However, population studies have reported inconsistent findings. This may be due to limitations from the standard analytic approach of treating PTB as a binary outcome without considering time-varying exposures together over the course of pregnancy. To address this research gap, we present a Bayesian hierarchical model for conducting a comprehensive examination of gestational air pollution exposure by estimating the joint effects of weekly exposures during different vulnerable periods. Our model also treats PTB as a time-to-event outcome to address the challenge of different exposure lengths among ongoing pregnancies. The proposed model is applied to a dataset of geocoded birth records in the Atlanta metropolitan area between 1999–2005 to examine the risk of PTB associated with gestational exposure to ambient fine particulate matter $< 2.5 \mu\text{m}$ in aerodynamic diameter ($\text{PM}_{2.5}$). We find positive associations between $\text{PM}_{2.5}$ exposure during early and mid-pregnancy, and evidence that associations are stronger for PTBs occurring around week 30.

Keywords: Air pollution; Fine particulate matter; Joint effects; Preterm birth; Time-dependent exposure; Time-to-event model; Time-varying effect

*To whom correspondence should be addressed..

1. INTRODUCTION

Epidemiological evidence on the health effects of outdoor air pollution plays an important role in setting regulatory standards and protecting public health. This has led to development of statistical modeling approaches motivated by specific analytic challenges arising from different health outcomes and study designs. There exists increasing interest in examining the potential link between a mother's exposure to air pollution during pregnancy and the risk of adverse birth outcomes. Preterm birth (PTB), defined as gestational age at delivery of <37 weeks, is the leading cause of neonatal mortality and morbidity (Goldenberg *and others*, 2008; Saigal and Doyle, 2008). In addition to health care expenses, PTB is also associated with long-term physical, cognitive, and developmental problems (Moster *and others*, 2008).

This paper describes a flexible modeling approach for estimating associations between PTB and a time-varying environmental exposure, such as air pollution, during pregnancy. The proposed model is motivated by the need to address two scientific questions of significant interest in environmental reproductive epidemiology. First, is PTB associated with critical gestational *exposure periods* that may correspond to specific stages of fetal development? For example, is high exposure that occurs during the first trimester more adverse than that occurring during the third trimester? Second, do exposures have different impacts on different *outcome periods*? For example, are ongoing pregnancies more vulnerable to the acute effect of air pollution as they progress closer to 37th week? There are differences in risk factors for earlier (e.g. < 32 weeks) versus later PTB (e.g. 34–36 weeks) and air pollution may only increase the risk of PTB occurring at certain stages of pregnancy. These findings on exposure and outcome windows can potentially provide crucial information on the biological mechanisms by which air pollution affects adversely on birth outcomes, as well as help develop public health messages to minimize harmful exposures during pregnancies.

The hypothesis on critical exposure windows can be examined by considering time-varying coefficients across gestational length; while the hypothesis on outcome windows considers PTB as a spectrum of outcomes, defined by gestational week at birth, rather than as a dichotomous outcome (Suh *and others*, 2009). We propose an approach to address these two scientific questions simultaneously by treating gestational age as time-to-event data (Chang *and others*, 2013). Here, we consider a discrete-time domain in weeks because gestational age is typically reported as the number of completed weeks. We then model the hazard rate at week t as a function of weekly air pollution exposure following conception until week t and allow the exposure effects to vary both across exposure weeks, as well as across outcomes weeks. The main advantage of a time-to-event approach is that it can effectively accommodate the differences in exposure length among pregnancies of different gestational ages. By considering PTB at different gestational ages, we are also able to distinguish between risks associated with early versus late PTB, which is important because each additional week of gestation during this critical time period leads to better health outcomes for the child.

The idea of estimating time-varying effects of time-varying exposures in a health model is similar to the distributed lag approach commonly used in time series studies of air pollution and health to examine associations between daily adverse health outcome counts and exposure to air pollution on the previous days (Schwartz, 2000). To stabilize estimation and borrow information across exposure days, several constraints have been proposed, including penalized splines (Zanobetti *and others*, 2000), Gaussian processes (Heaton and Peng, 2012), autoregressive priors (Fahrmeir and Knorr-Held, 1997), and hierarchical structures (Welty *and others*, 2009). Recently, the distributed lag model has been applied to PTB for identifying critical exposure window by Warren *and others* (2012). In this paper, we extend previous work by considering distributed exposures that vary in two dimensions (i.e. exposure weeks and outcome weeks). Specifically, we allow the associations between exposure at each week of pregnancy and the risk of PTB to vary over the course of the pregnancy. Our model is also related to survival models that consider time-dependent exposures or time-varying (dynamic) effects (McKeague and Tighiouart, 2000;

Haneuse *and others*, 2008; Buchholz and Sauerbrei, 2011; He *and others*, 2012); however, to our knowledge no method has been developed to examine whether the time-varying effects differ across outcome periods.

The proposed model is applied to a birth cohort in the Atlanta metropolitan area. We examine association between PTB and ambient fine particulate matter < 2.5 μm in aerodynamic diameter ($\text{PM}_{2.5}$). $\text{PM}_{2.5}$ is a complex mixture of solid and liquid particles suspended in the atmosphere, and has been linked to increased risks of various adverse health outcomes (US EPA, 2009). The remainder of this paper is organized as follows. Section 2 describes the motivating Atlanta air pollution and birth record data. Section 3 describes the distributed exposure discrete-time survival model for PTB. In Section 4, we describe a simulation study that evaluates the proposed model's performance in identifying exposure and outcome windows. Results from the health analysis are given in Section 5. Finally, discussion and future work appear in Section 6.

2. DATA

Birth records were obtained from the Office of Health Research and Policy, Georgia Division of Public Health. We included only singleton births without major structural birth defects conceived from January 1, 1999 to December 31, 2005. Gestational age in weeks was determined by the reported date of last menstrual period. We defined the study population using conception dates instead of birth date to avoid fixed-cohort bias when examining time-varying exposures (Strand *and others*, 2011b). Maternal residential address at delivery was geocoded to the census blockgroup level and this study included geocoded births within the 5-county Atlanta metropolitan area (Clayton, Cobb, DeKalb, Fulton, and Gwinnett). We restricted the analysis to Hispanic, non-Hispanic black, non-Hispanic white, and Asian mothers between the age of 15–44. We only considered PTB that reached at least 27 weeks of gestation.

Daily measurements of ambient $\text{PM}_{2.5}$ concentrations were obtained from 14 air quality monitors within the study region. Daily $\text{PM}_{2.5}$ concentrations showed strong spatial homogeneity, with a median between-monitor correlation of 0.85. We also obtained daily average temperature from the National Climatic Data Center, and census tract-level percent household below poverty from the US 2000 Census. Each geocoded pregnancy was linked to the closest $\text{PM}_{2.5}$ monitor within a 10 km radius to minimize spatial exposure measurement error (median distance of 5.15 km). We defined the exposure for week t as the average pollution levels on the date that week t was completed and the 6 days leading up to it. Weekly average temperature was defined similarly. Daily air pollution measurements contained missing values (2.5%). For weeks without any $\text{PM}_{2.5}$ measurements, the averages of the weeks before and after were used. Births with 2 or more consecutive missing weeks were excluded.

3. METHODS

3.1 Distributed exposure time-to-event model for PTB

We view gestational age as discrete-time survival data where each pregnancy enters the risk set at the 27th week. Full-term births of at least 37 weeks are censored at 36 weeks because they are no longer at risk of being preterm. Let $\pi_{ig} \in [0, 1]$ denote the hazard rate of birth during outcome (at-risk) window week g for pregnancy i . Denote x_{ik} the average air pollution exposure for pregnancy i during gestational week k where $k \leq g_i$. We model the discrete event hazard rate using probit regression,

$$\pi_{ig} = \Phi \left[h(g) + \sum_{k=1}^g \theta_{g,k} x_{ik} + \delta' \mathbf{z}_i(g) \right], \quad (3.1)$$

samples. Time-varying lagged exposures can also be examined. For example, $\sum_{k=g}^{g-3} \theta_{g,k}$ represents the risk due to air pollution exposures experienced 1-month prior at gestational age week g .

Subsequently, β_g can also be combined across different outcome windows. For example, under a fixed-effect meta-analysis framework, we can define the overall first trimester exposure effects across gestational week 27 until 36 as $[\mathbf{1}'V_{\beta}^{-1}\mathbf{1}]^{-1}\mathbf{1}'V_{\beta}^{-1}\boldsymbol{\beta}$, where $\boldsymbol{\beta} = (\beta_{27}, \beta_{27}, \dots, \beta_{36})$ and V_{β} is the corresponding variance–covariance matrix. We note that we do not sum over g because it combines risks due to the same exposure that are conditioned on different baseline hazard and at different time points. Instead, we believe it is more intuitive to summarize the effect over outcome weeks as an average, weighted by its uncertainty. The ability to estimate air pollution risks across different gestational ages is particularly useful for examining whether the impacts of air pollution exposure differ between early and late PTBs.

4. SIMULATION STUDY

We carried out a simulation study to evaluate the benefits of borrowing information across exposure and outcome windows in a distributed exposure analysis. We generated 20 replicates of simulated air pollution exposures and gestational age for 50 000 births as follows. First, for each birth i , we generated weekly exposures x_{ik} from a multivariate normal distribution with marginal means of 17.2 and marginal variances of 43. We assumed the correlation between weeks k and k' to be $e^{-0.69|k-k'|}$. These values were chosen to reflect the temporal variability and correlation observed in the Atlanta data. Then gestational age t_i for pregnancy i was generated with probability $P(t_i = g) = \pi_{ig} \times \prod_{t' < g} [1 - \pi_i(t')]$ where $\pi_{ig} = \Phi[\hat{h}(g) + \sum_{k=1}^g \theta_{g,k} x_{ik}]$ for $g \leq 36$. We estimated the baseline hazard function $\hat{h}(g)$ using the Atlanta birth record data. We allowed the gestational age of each birth to be subject to misclassification. Based on previous studies on the disagreement between gestational age defined using clinical estimate or last menstrual period (Mustafa and David, 2001), we assumed the error between observed and true gestational age to be at most 1, 2, and 3 weeks with probability 0.8, 0.1, and 0.1, respectively.

We considered 5 scenarios of different exposure and outcome windows.

- Scenario 1: For all g , $\theta_{g,k} \neq 0$ if $14 \leq k \leq 26$.
- Scenario 2: For all g , $\theta_{g,k} \neq 0$ if $14 \leq k \leq 19$.
- Scenario 3: For $27 \leq g \leq 31$, $\theta_{g,k} \neq 0$ if $14 \leq k \leq 26$.
- Scenario 4: For all g , $\theta_{g,k} \neq 0$ if $g - 3 \leq k \leq g$.
- Scenario 5: For $32 \leq g \leq 36$, $\theta_{g,k} \neq 0$ if $g - 3 \leq k \leq g$.

In Scenario 1, we assumed a fixed-length 13-week exposure window between gestational weeks 14–26 (second trimester). This effect is identical for all ongoing pregnancies during the entire at-risk window (weeks 27–36). In Scenario 2, we assumed a shorter 6-week exposure during weeks 14–21, common across all at-risk weeks. In Scenario 3, we assumed the 13-week exposure from Scenario 1 only has an effect for pregnancies early during the at-risk window. In Scenario 4, we assumed a 4-week lagged exposure that is common across all at-risk weeks. Finally, Scenario 5 assumed the lagged effect is only present for pregnancies that have reached gestational age 32. We assigned each non-zero $\theta_{g,k} = 0.025$, corresponding to an approximate 1.05 increase in hazard ratio per interquartile range (IQR) increase in weekly exposure.

For each simulated dataset, we fitted the time-to-event model with the following four covariance structure specifications for $\theta_{g,k}$: (1) the proposed dynamic model, (2) an *exchangeable* model by setting $\alpha = 0$ that does not borrow information across outcome weeks, (3) an *independent* random effect model where $\theta_{g,k} \stackrel{i.i.d.}{\sim} N(0, \tau^2)$ that does not borrow information across both exposure and outcome weeks,

and (4) a *fixed* prior structure ($\alpha = 1$ and $\psi_g = \mathbf{0}$ for all $g < 36$) that assumes identical weekly exposure effects across outcome weeks. The *fixed* prior also represents an extension of the distributed lag model of Warren and others (2012) to a time-to-event setting.

Table 1 gives the relative changes in root mean-square error (RMSE), 95% posterior interval (PI) length, and changes in deviance information criterion (DIC) and effective degrees of freedom (pD) (Spiegelhalter and others, 2002). We found that the dynamic model consistently out-performed the independent and the exchangeable model in terms of RMSE and PI length. The gain in performance is greater for longer exposure window (Scenario 1). The difference between the dynamic and the exchangeable model is less pronounced for shorter fixed-length exposure windows (Scenario 2). Importantly, performance of the dynamic model does not degrade in situations where the exposure effects are only present for some outcome weeks (Scenarios 3 and 5). Finally, the dynamic model consistently results in smaller DIC and smaller effective degrees of freedom, which indicate improved model fit and the presence of additional shrinkage across outcome/exposure weeks.

The fixed prior resulted in a large reduction in 95% PI length because it utilizes data from all outcome weeks. Compare with the dynamic prior, it under-performs when the exposure effects vary across outcome weeks (Scenarios 3 and 4). For Scenario 5, where the 4-week lag effects are only present during outcome weeks 31–36, the fixed and dynamic priors perform similarly, likely because there is considerable overlapping exposure effects across outcome weeks. Figure S1 of supplementary material available at

Table 1. Simulation study results for estimating weekly $PM_{2.5}$ effects across outcome weeks: relative change in RMSE, length of 95% PIs, and change in DIC and effective degrees of freedom (pD), averaged across 20 simulated replicate datasets

	Exposure weeks	Outcome weeks	Prior structure	Relative Δ RMSE	Relative Δ PI _{95%} length	Δ DIC	Δ pD
Scenario 1	14–26	27–36	Dynamic		Reference		
			Independent	2.13 (0.17)	3.39 (0.16)	121 (19)	130 (12)
			Exchangeable	1.41 (0.08)	1.53 (0.09)	46 (11)	36 (12)
			Fixed	0.88 (0.04)	0.66 (0.04)	–2 (10)	–22 (3)
Scenario 2	14–19	27–36	Dynamic		Reference		
			Independent	1.64 (0.13)	2.74 (0.26)	57 (15)	69 (5)
			Exchangeable	1.14 (0.09)	1.95 (0.23)	38 (14)	44 (6)
			Fixed	0.87 (0.07)	0.63 (0.06)	–3 (9)	–18 (5)
Scenario 3	14–26	27–31	Dynamic		Reference		
			Independent	1.74 (0.13)	2.78 (0.28)	70 (16)	76 (7)
			Exchangeable	1.31 (0.16)	1.52 (0.32)	31 (13)	27 (14)
			Fixed	1.62 (0.12)	0.43 (0.06)	106 (21)	–28 (6)
Scenario 4	4-wk lag	27–36	Dynamic		Reference		
			Independent	1.34 (0.09)	2.75 (0.30)	23 (9)	46 (5)
			Exchangeable	1.32 (0.09)	2.63 (0.38)	21 (8)	45 (5)
			Fixed	1.13 (0.11)	0.41 (0.08)	–1 (9)	–18 (4)
Scenario 5	4-wk lag	31–36	Dynamic		Reference		
			Independent	1.25 (0.09)	2.85 (0.45)	17 (10)	44 (6)
			Exchangeable	1.24 (0.13)	2.68 (0.61)	15 (7)	40 (9)
			Fixed	0.93 (0.08)	0.45 (0.08)	–6 (6)	–16 (5)

The standard deviations across replicate datasets are given in parentheses.

Biostatistics online plot the estimated weekly effects, averaged across replicates, for outcome weeks 28, 32, and 36. Visually, Figure S1 of supplementary material available at *Biostatistics* online demonstrates the dynamic prior's ability to capture the underlying true effects, as well as the smoothing effect compared to the other priors.

We also compared the cumulative effects obtained from combining weekly effects to that estimated by first averaging weekly exposure over a pre-defined window. Specifically, we fitted a time-to-event model where the exposures were averaged across the known true at-risk and outcome windows. This corresponds to the standard exposure assessment approach taken by [Chang, Reich, and others \(2012\)](#). This allows us to examine the model's performance associated with using weekly exposures when the true association spans across several weeks. Simulation results are given in Table S1 of supplementary material available at *Biostatistics* online. Compared with an analysis where the true exposure windows are known, the cumulative effects estimated by combining weekly effects were associated with a slight negative bias. This is likely due to shrinkage and the use of a Gaussian process to approximate the true step risk function. However, the dynamic model consistently out-performs the other 2 prior structures. Table S1 of supplementary material available at *Biostatistics* online also indicates that the use of averaged exposure can lead to larger estimation uncertainty.

To examine how the different models perform in detecting risk, we calculated the sensitivity and specificity (Table 2). The dynamic model has better sensitivity compared with the exchangeable and the independent prior structures. The dynamic model does suffer from a slightly reduced specificity compared

Table 2. Simulation study results for the sensitivity and specificity of detecting a non-null weekly $PM_{2.5}$ effects, averaged across 20 replicates

	Exposure weeks	Outcome weeks	Prior structure	Sensitivity	Specificity	95% PI coverage
Scenario 1	14–26	27–36	Dynamic	1.00 (0.00)	0.90 (0.03)	0.85 (0.03)
			Exchangeable	0.85 (0.04)	0.96 (0.02)	0.88 (0.03)
			Independent	0.16 (0.03)	0.99 (0.01)	0.98 (0.01)
			Fixed	1.00 (0.00)	0.85 (0.05)	0.78 (0.05)
Scenario 2	14–19	27–36	Dynamic	0.96 (0.04)	0.93 (0.03)	0.86 (0.03)
			Exchangeable	0.28 (0.08)	1.00 (0.00)	0.94 (0.02)
			Independent	0.06 (0.03)	1.00 (0.00)	0.98 (0.01)
			Fixed	1.00 (0.00)	0.89 (0.03)	0.77 (0.04)
Scenario 3	14–26	27–31	Dynamic	0.98 (0.02)	0.91 (0.03)	0.89 (0.04)
			Exchangeable	0.60 (0.21)	0.98 (0.02)	0.92 (0.04)
			Independent	0.03 (0.02)	1.00 (0.00)	0.99 (0.01)
			Fixed	0.95 (0.06)	0.69 (0.07)	0.54 (0.07)
Scenario 4	4-wk lag	27–36	Dynamic	0.70 (0.18)	0.94 (0.02)	0.84 (0.03)
			Exchangeable	0.03 (0.04)	1.00 (0.00)	0.94 (0.02)
			Independent	0.01 (0.02)	1.00 (0.00)	0.94 (0.02)
			Fixed	0.87 (0.11)	0.90 (0.04)	0.79 (0.04)
Scenario 5	4-wk lag	31–36	Dynamic	0.80 (0.19)	0.95 (0.03)	0.90 (0.03)
			Exchangeable	0.07 (0.12)	1.00 (0.00)	0.97 (0.01)
			Independent	0.02 (0.03)	1.00 (0.00)	0.97 (0.01)
			Fixed	0.99 (0.02)	0.91 (0.03)	0.86 (0.03)

The overall empirical coverage probability of the 95% PI is also given. The standard deviations across replicate datasets are given in parentheses.

with the exchangeable and the independent priors; however these false positives occur near the boundary of the true risk window and across outcome weeks. Finally, the dynamic prior out-performs the fixed prior in specificity, especially when the exposure weeks differ across outcome weeks (Scenarios 3). Table 2 also gives the empirical 95% PIs coverage probability. For the dynamic and fixed priors, coverage is slightly under due to the presence of exposure measurement error.

5. ANALYSIS OF THE ATLANTA BIRTH COHORT, 1999–2005

The study included 175 891 pregnancies (10.6% preterm). Weekly $\text{PM}_{2.5}$ exposures, averaged across all pregnancies, were similar for weeks 1–37, ranging 17.1–17.4 $\mu\text{g}/\text{m}^3$. The variability in weekly exposures in terms of IQR was also similar across weeks, ranging from 7.34 to 7.63 $\mu\text{g}/\text{m}^3$. This is due to the relatively small seasonal trend in the number of ongoing pregnancies in our study population. Weekly $\text{PM}_{2.5}$ exposures showed moderate temporal correlation with a median lag-1 auto-correlation of 0.50 across pregnancies. For outcome weeks 27–36, the corresponding PTB numbers are 303, 320, 383, 439, 563, 830, 1291, 2362, 4134, and 8023. We fitted the distributed exposure time-to-event model with a dynamic prior for weekly exposures. The list of confounders included in the model and their associations are given in Section 2 of supplementary material available at *Biostatistics* online.

5.1 Results

For small probabilities, the Gaussian distribution function can be approximated by an exponential function: $\Phi(z) \approx \exp(0.136 + 1.927z)$. The unadjusted hazard rates of PTB during outcome weeks 27–26 were <0.05 . Therefore, to aid in the interpretation of probit regression coefficients, we present the posteriors of $\exp(1.927 \times \theta_{g,k})$ as the approximate relative risk (RR) of PTB due to a unit increase in the exposure. Here an RR of 1 indicates the covariate is not associated with PTB risk. Figure 1 shows the posterior means and 95% PI of the RRs associated with different exposure weeks by outcome weeks. The PI width decreases for later outcome weeks due to the increasing number of PTBs. Overall we found relatively

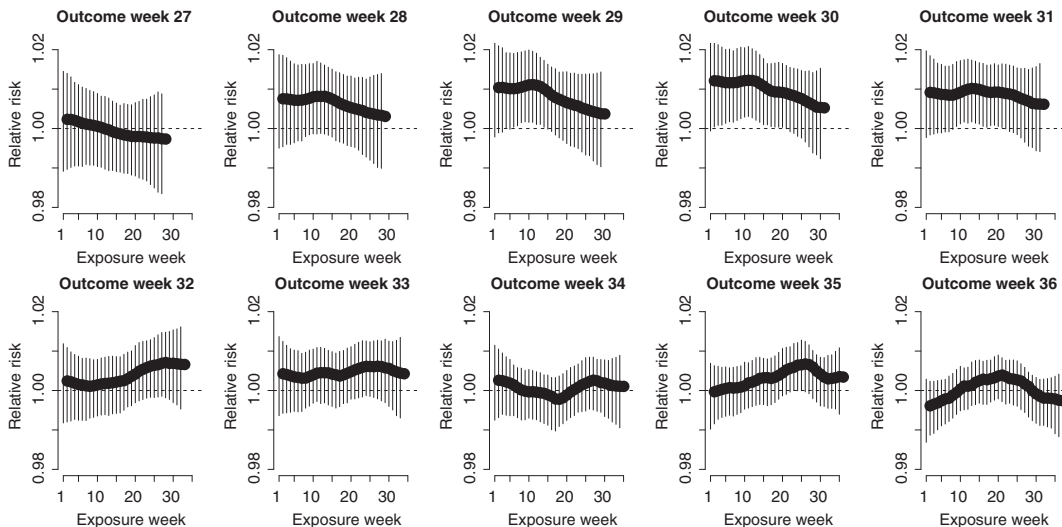


Fig. 1. Posterior means and 95% PIs of the RRs for PTB associated with an IQR ($7.4 \mu\text{g}/\text{m}^3$) increase in weekly $\text{PM}_{2.5}$ exposure.

smooth time-varying effects within each outcome week. Figure 2 shows the posterior probability of $\theta_{g,k} > 0$. There is evidence of positive associations with exposures occurring during between early and mid-pregnancy, and these associations are only present when an ongoing pregnancy was between weeks 28–31 (among very PTB). We also found weak evidence that late trimester 2 exposures were associated with PTB occurring after week 32 (late PTB). We also fitted the health data using the *fixed* prior. This resulted in a slight increase in DIC (1.7). Figure S3 of supplementary material available at *Biostatistics* online shows the posterior means and 95% PI's of the weekly exposure effects. While increased risks were found around exposure week 20, the effects were attenuated compared with that from the dynamic model, and none of the 95% PI excluded 1.

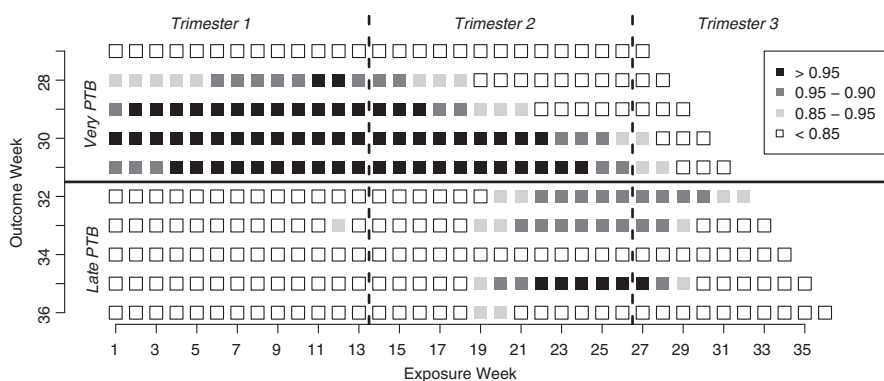


Fig. 2. Posterior probabilities of weekly $\text{PM}_{2.5}$ effects being greater than zero: $P(\theta_{g,k} > 0)$.

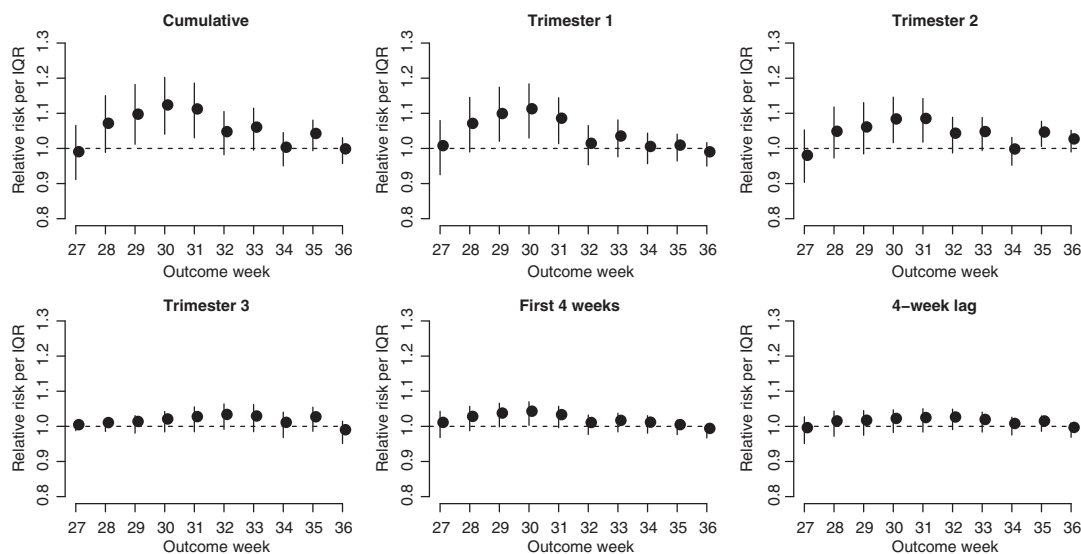


Fig. 3. Posterior means and 95% PIs of the RRs for PTB per IQR increase in different metrics of exposure to $\text{PM}_{2.5}$. The IQR for the exposures are given in Table 3.

We then examined cumulative effects as described in Section 3.2. Figure 3 shows the RR for PTB associated with different windows of exposure commonly utilized in the literature. We examined 3 time-invariant exposures: trimester 1 (weeks 1–13); trimester 2 (weeks 14–27); and the 4 weeks after conception. Given that an ongoing pregnancy completed gestational week t , we also examined 3 time-varying exposures: cumulative (week 1 to t); trimester 3 (week 27 to t); and a 4-week lag (week $t-3$ to t). The estimates are scaled by an IQR of the exposure metric across the study population given in Table 3. Based on the aggregated exposures, we found evidence of positive associations between PTB risk and $PM_{2.5}$ levels during trimester 1 and trimester 2. These associations were also more pronounced for PTB occurring between 29 and 31 weeks.

Finally, Table 3 gives the RRs associated with each aggregated exposure window that are further pooled across outcome weeks. We considered pooling across the entire outcome window (weeks 27–36), as well as across very early PTB (weeks 27–31) and late PTB (weeks 32–36). We also re-analyzed the health data following the approach in Chang, Reich, and others (2012) where the weekly exposures were first averaged across the at-risk windows. These results are referred to as the *standard* approach. The estimated risks from these two approaches are very similar, indicating that the proposed approach can indeed recover long-term effects. There is also evidence that standard approach resulted in smaller risk estimates, particularly for longer average windows.

From Table 3, the strongest associations were found for cumulative and trimester 1 exposures among very early PTB. Specifically, an IQR increase in exposure was associated with an RR of 1.07 (95% PI: 1.04–1.12) and 1.07 (95% PI: 1.03–1.11), for cumulative and trimester 1 exposures, respectively. We also

Table 3. Posterior means and 95% PIs of the pooled RRs for PTB per IQR increase in different metrics of exposure to $PM_{2.5}$

Exposure window	IQR ($\mu\text{g}/\text{m}^3$)	Outcome window	Dynamic distributed lag RR (95% PI)	Standard approach RR (95% CI)
Cumulative	3.11	VPTB	1.07 (1.04, 1.12)	1.04 (1.02, 1.07)
		LPTB	1.01 (0.99, 1.04)	1.01 (0.99, 1.03)
		PTB	1.03 (1.00, 1.06)	1.02 (1.00, 1.03)
Trimester 1	5.22	VPTB	1.07 (1.03, 1.11)	1.06 (1.03, 1.08)
		LPTB	1.00 (0.97, 1.02)	1.00 (0.98, 1.01)
		PTB	1.01 (0.99, 1.04)	1.01 (0.99, 1.03)
Trimester 2	5.14	VPTB	1.05 (1.02, 1.09)	1.03 (1.01, 1.06)
		LPTB	1.02 (1.00, 1.05)	1.01 (0.99, 1.03)
		PTB	1.03 (1.01, 1.05)	1.02 (1.00, 1.03)
Trimester 3	5.43	VPTB	1.00 (0.99, 1.00)	0.99 (0.97, 1.01)
		LPTB	1.01 (0.99, 1.03)	1.00 (0.99, 1.02)
		PTB	1.00 (0.99, 1.01)	1.00 (0.98, 1.01)
First 4 weeks	5.92	VPTB	1.02 (1.00, 1.04)	1.03 (1.01, 1.05)
		LPTB	1.00 (0.98, 1.02)	1.00 (0.98, 1.01)
		PTB	1.00 (0.99, 1.02)	1.00 (0.99, 1.01)
4-week lag	6.04	VPTB	1.01 (0.99, 1.03)	0.98 (0.96, 1.01)
		LPTB	1.00 (0.99, 1.02)	1.00 (0.99, 1.01)
		PTB	1.00 (0.99, 1.02)	1.00 (0.98, 1.01)

Estimates were pooled across all PTB outcome weeks (week 27–36), as well as across very PTB (week 27–31, VPTB) and late PTB (week 32–36, LPTB). Under the standard approach, estimates and 95% confidence interval (CI) were estimated by a time-to-event model where the weekly exposures were first averaged across the exposure window.

found associations for trimester 2 among all PTBs. We note that the above results cannot be obtained from the standard logistic regression model with average $PM_{2.5}$ levels over trimesters as a predictor. Specifically, the standard approach can be viewed as a special case of the distributed exposure model where weekly risks are assumed to be constant throughout the trimester. By averaging over exposure weeks, the standard approach would also have less power to detect an effect if the true critical window of exposure is shorter than a single trimester.

6. DISCUSSION

There exists broad literature indicating that fetuses are more vulnerable to environmental exposures than adults due to their developing physiological systems. Despite many population studies on gestational exposure to air pollution and reduced gestational length, there is no consensus on the exact exposure windows, if any, that increase the risk of prematurity. Previous studies often define a limited set of *a priori* exposure windows based on biological hypotheses (e.g. the first 4 weeks following conception for implantation and placentation) or convenience (e.g. trimester-wide averages). The standard one-exposure-at-a-time approach has raised concerns regarding selective reporting (Bosetti and others, 2010). This paper presents a distributed exposure approach to estimate joint effects of weekly exposures on the risk of PTB, as well as their interactions with gestational age. This framework is motivated by the need of a more comprehensive and flexible analytic approach to examine specific hypothesis on exposure and outcome windows.

In our analysis of a 7-year birth cohort in Atlanta, Georgia, we found evidence that indicates exposure to $PM_{2.5}$ during early and mid-pregnancy was associated with increased risk of PTB. There is stronger evidence that long-term sustained exposure to high levels of ambient $PM_{2.5}$ was positively associated with PTB for early preterm pregnancies. Our estimated RRs for cumulative and first trimester exposures are consistent in magnitude with previous studies using different statistical models for PTB (Bosetti and others, 2010). More importantly, our results contribute to the robustness of first and second trimester exposure association because the estimates were automatically adjusted by $PM_{2.5}$ exposure during other periods, a sensitivity analysis rarely conducted in previous studies. While all weekly exposures are included, residual confounding between weekly effects may be still present due to shrinkage across them. For PTB analysis, the proposed model can also be applied to other time-varying environmental exposures such as ambient temperature, as well as longitudinal maternal characteristics that are typically recorded in a prospective birth cohort study (e.g. blood pressure, weight gain, amount of exercise).

We found the strongest associations between non-acute $PM_{2.5}$ exposures and PTB for ongoing pregnancies before week 32. Several factors may contribute to this observation. First, births near the 37th week cutoff may be subjected to higher misclassification error due to uncertainty in the gestational age. Specifically, the reported last menstrual period used in our analysis is likely to suffer from recall bias. Mustafa and David (2001) found that the concordance between gestational age obtained using reported last menstrual period and clinical estimates based on ultrasound are 78% for a 1-week difference and 87% for a 2-week difference. A second potential explanation is the *displacement* hypotheses or healthy survivor bias where exposure to $PM_{2.5}$ early in the pregnancy will shorten the gestational age of pregnancies at-risk of being preterm, depleting the pool of late ongoing pregnancies vulnerable to air pollution. This is similar to the mortality displacement (harvesting) hypothesis for the acute health effects of air pollution (Zanobetti and others, 2000).

The proposed model and application have several limitations that warrant future development and sensitivity analyses. First, our time-to-event model assumes common effects across individuals. There is increasing interest in examining effect modifications of the associations between air pollution and birth outcomes, particularly for identifying susceptible populations (Généreux and others, 2008). Second, outdoor air pollution concentration was used as a surrogate for personal exposure. Variations in the ambient versus personal exposures across individuals, for example, due to residential mobility and time

activity pattern, may result in differential risks (Bell and Belanger, 2012). The health effect estimated using ambient concentration may be different from that obtained from using personal exposure estimates (Chang, Fuentes, and others, 2012; Sarnat and others, 2014). Finally, in simulations, we found that error in gestation age leads to an underestimation of the health risks. Uncertainty in gestational age is an unavoidable challenge in birth outcome analysis, and methods to account for exposure and outcome measurement errors are needed to ensure study reproducibility and improve our ability to quantify the health impacts of ambient air pollution.

SUPPLEMENTARY MATERIAL

Supplementary material is available online at <http://biostatistics.oxfordjournals.org>

ACKNOWLEDGMENTS

Conflict of Interest: None declared.

FUNDING

This study was partially supported USEPA grant R834799 (Southern Center for Air Pollution & Epidemiology) and NIH grant R21ES022795-01A1. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA and NIH. Further, USEPA and NIH do not endorse the purchase of any commercial products or services mentioned in the publication.

REFERENCES

- BELL, M. L. AND BELANGER, K. (2012). Review of research on residential mobility during pregnancy: consequences for assessment of prenatal environmental exposures. *Journal of Exposure Science and Environmental Epidemiology* **22**, 429–438.
- BOSETTI, C., NIEUWENHUIJSEN, M. J., GALLUS, S., CIPRIANI, S., LA VECCHIA, C. AND PARAZZINI, F. (2010). Ambient particulate matter and preterm birth or birth weight: a review of the literature. *Archives of Toxicology* **348**, 447–460.
- BUCHHOLZ, A. AND SAUERBREI, W. (2011). Comparison of procedures to assess non-linear and time-varying effects in multivariable models for survival data. *Biometrical Journal* **53**, 308–331.
- CHANG, H. H., FUENTES M. AND FREY, H. C. (2012). Time series analysis of personal exposure to ambient PM_{2.5} and mortality using an exposure simulator. *Journal of Exposure Science and Environmental Epidemiology* **22**, 483–488.
- CHANG, H. H., REICH, J. B. AND MIRANDA, M. L. (2012). Time-to-event analysis of fine particle air pollution and preterm birth: results from North Carolina, 2001–2005. *American Journal of Epidemiology* **175**, 91–98.
- CHANG, H. H., REICH, J. B. AND MIRANDA, M. L. (2013). A spatial time-to-event approach for estimating associations between air pollution and preterm birth. *Journal of the Royal Statistical Society Series C* **62**, 167–179.
- FAHRMEIR, L. AND KNORR-HELD, L. (1997). Dynamic discrete time duration models: estimation via Markov chain Monte Carlo. *Sociological Methodology* **27**, 417–452.
- GÉNÉREUX, M., AUGER, A., GONEAU, M. AND DAINIEL, M. (2008). Neighbourhood socioeconomic status, maternal education and adverse birth outcomes among mothers living near highways. *Journal of Epidemiology & Community Health* **62**, 695–700.
- GOLDENBERG, R. L., CULHANE, J. F., IAMS, J. D. AND ROMERO, R. (2008). Epidemiology and causes of preterm birth. *Lancet* **9606**, 75–94.
- HANEUSE, S. J., RUDSER, K. D. AND GILLEN, D. L. (2008). The separation of timescales in Bayesian survival modeling of the time-varying effect of a time-dependent exposure. *Biostatistics* **9**, 400–410.
- HE, J., MCGEE, D. L. AND NIU, X. (2012). Application of the Bayesian dynamic survival model in medicine. *Statistics in Medicine* **29**, 347–360.

- HEATON, M. J. AND PENG, R. D. (2012). Flexible distributed lag models using random functions with application to estimating mortality displacement from heat-related deaths. *Journal of Agricultural, Biological, and Environmental Statistics* **17**, 313–331.
- MCKEAGUE, I. W. AND TIGHIOUART, M. (2000). Bayesian estimators for conditional hazard functions. *Biometrics* **56**, 1007–1015.
- MOSTER, D., LIE, R. T. AND MARKESTAD, T. (2008). Long-term medical and social consequences of preterm birth. *New England Journal of Medicine* **359**, 262–273.
- MUSTAFA, G. AND DAVID, R. (2001). Comparative accuracy of clinical estimate versus menstrual gestational age in computerized birth certificates. *Public Health Report* **116**, 15–21.
- R CORE TEAM. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- SAIGAL, S. AND DOYLE, L. W. (2008). An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet* **371**, 261–269.
- SARNAT, S. E., SARNAT, J. A., MULHOLLAND, J., ISAKOV, V., OZKAYNAK, H., CHANG, H. H., KLEIN, M. AND TOLBERT, P. E. (2014). Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. *Journal of Exposure Science and Environmental Epidemiology* **23**, 598–605.
- SCHWARTZ, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology* **11**, 320–326.
- SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. AND VAN DER LIND, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B* **64**, 583–639.
- STRAND, L., BARNETT, A. AND TONG, S. (2011b). Methodological challenges when estimating the effects of season and seasonal exposures on birth outcomes. *BMC Medical Research Methodology* **11**, doi:10.1186/1471-2288-11-49.
- SUH, Y. J., KIM, H., SEO, J. H., PARK, H., KIM, Y. J., HONG, Y. C. AND HA, E. H. Different effects of PM₁₀ exposure on preterm birth by gestational period estimated from time-dependent survival analyses. *International Archives of Occupational and Environmental Health* **82**, 613–621.
- U.S. EPA. Integrated Science Assessment for Particulate Matter (Final Report). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-08/139F.
- WARREN, J., FUENTES, M., HERRING, A. AND LANGLOIS, P. (2012). Spatial-temporal modeling of the association between air pollution exposure and preterm birth: identifying critical windows of exposure. *Biometrics* **68**, 1157–1167.
- WELTY, L. J., PENG, R. D., ZEGER, S. L. AND DOMINICI, F. (2009). Bayesian distributed lag models: estimating effects of particulate matter air pollution on daily mortality. *Biometrics* **65**, 282–291.
- ZANOBETTI, A., WAND, M. P., SCHWARTZ, J. AND RYAN, L. M. (2000). Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics* **1**, 179–292.

[Received March 27, 2014; revised September 8, 2014; accepted for publication December 15, 2014]