Flexible Modeling Methods: Semiparametric-mixed Model & Functional Linear Model

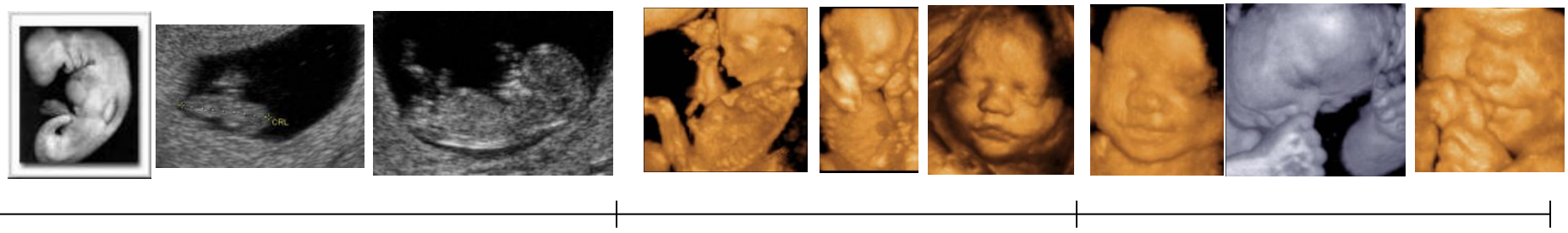
Lu Wang, PhD

Associate Professor

Department of Biostatistics

University of Michigan

- **Prenatal Exposure to PAH are clinically shown to affect fetal growth, development and survival**
- **Adverse fetal environment linked to long-term consequences**



Window of Vulnerability:

- Variable Rate of Development for different organs
- Immature Immune Functions
- Greater Exposure per Body Weight

Hypothesis:

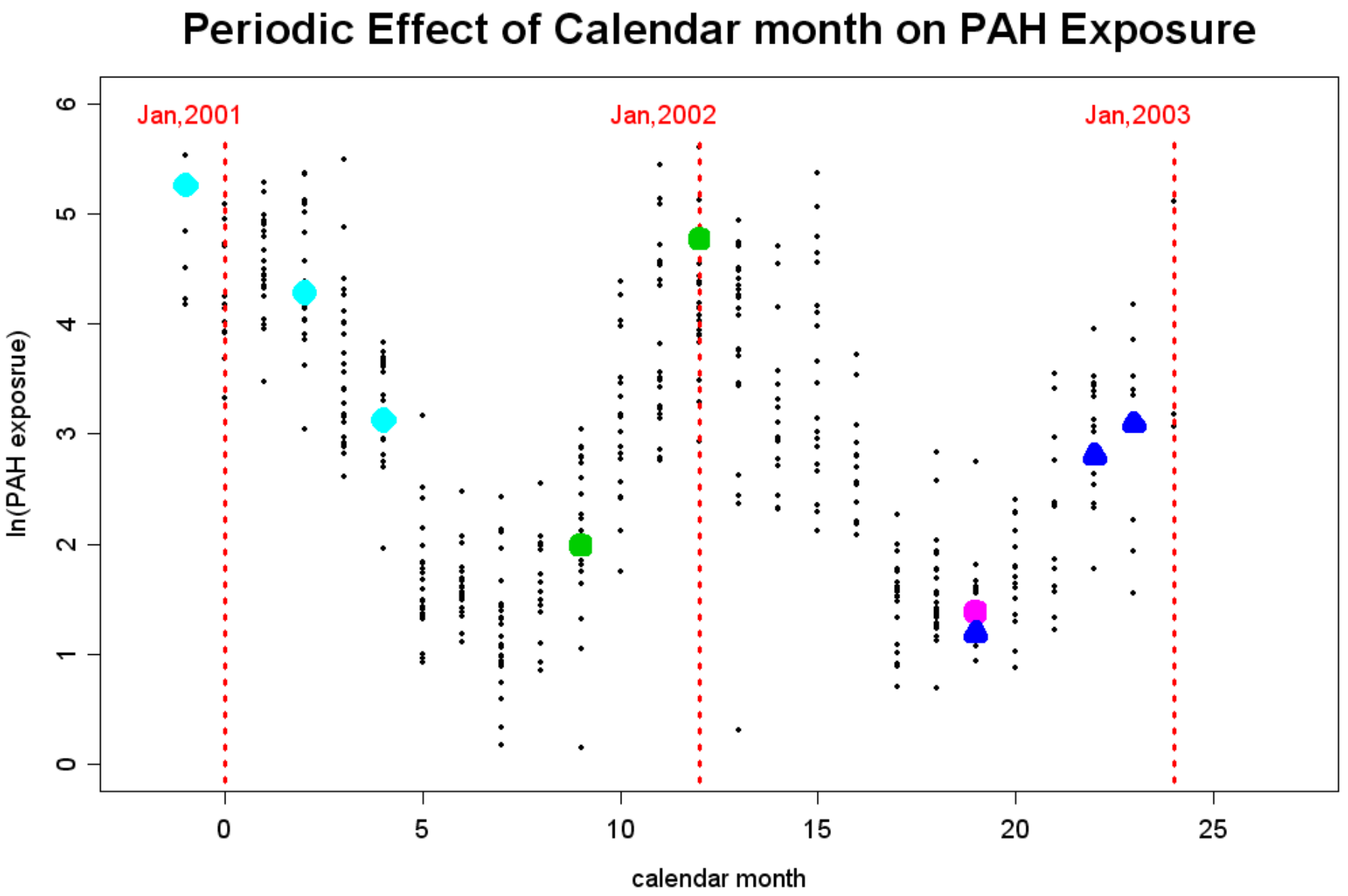
high exposure to airborne polycyclic aromatic hydrocarbons (PAHs) during the first trimester most significantly impairs fetal growth.

Birth cohort study, prenatal care clinics in Krakow, Poland (2000-2003)

- **Outcome Data**
  - **Birth weight, Birth length, Birth head circumference**
- **Exposure Data**
  - **Personal exposure to PAH**
  - **Longitudinally measured over different gestational age**

Discussion Point #1:

What are the potential challenges while you are processing the data?


Periodic Effect of Calendar month on PAH Exposure
Jan,2001
Jan,2002
Jan,2003
ln(PAH exposrue)
0
1
2
3
4
5
6
0
5
10
15
20
25
calendar month

Periodic Semiparametric Stochastic Mixed Model

- Outcome: the log of PAH measurement of the ith subject at time point t_{ij}. (i=1,…,m, j=1…n_i)
- n_i : number of observations of subject i, vary individually.
- Model:

$$\ln(PAH)_{ij} = x_{ij}^{T}\beta + f(t_{ij}) + b_i + U_i(t_{ij}) + e_{ij}$$

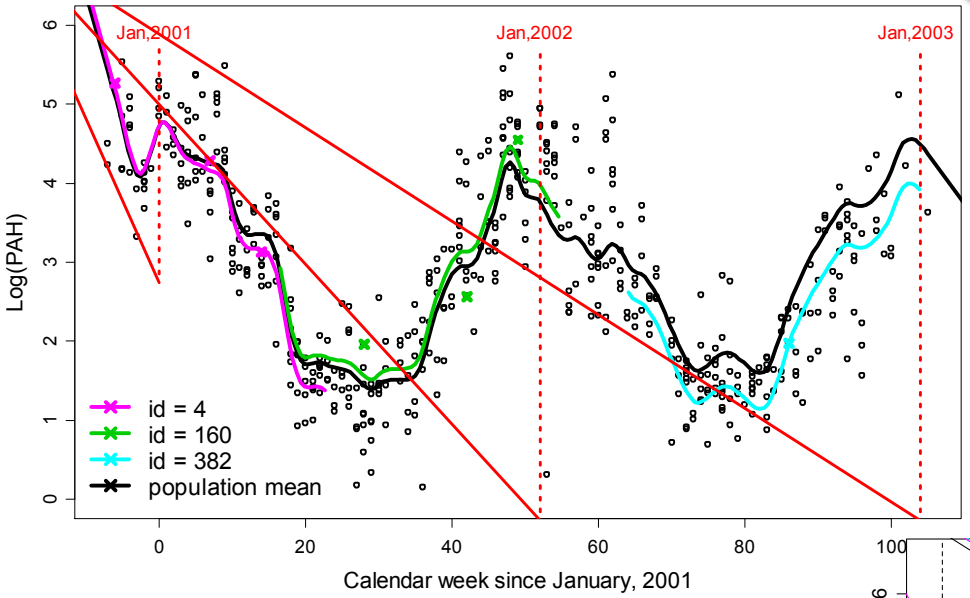
x : subject-level covariates
t : calendar month, changes within a subject
f : a smooth periodic function (period length=12)
b : subject-specific random intercept
U : mean 0 stochastic process with periodic variance
e : independent error

Discussion Point #2:

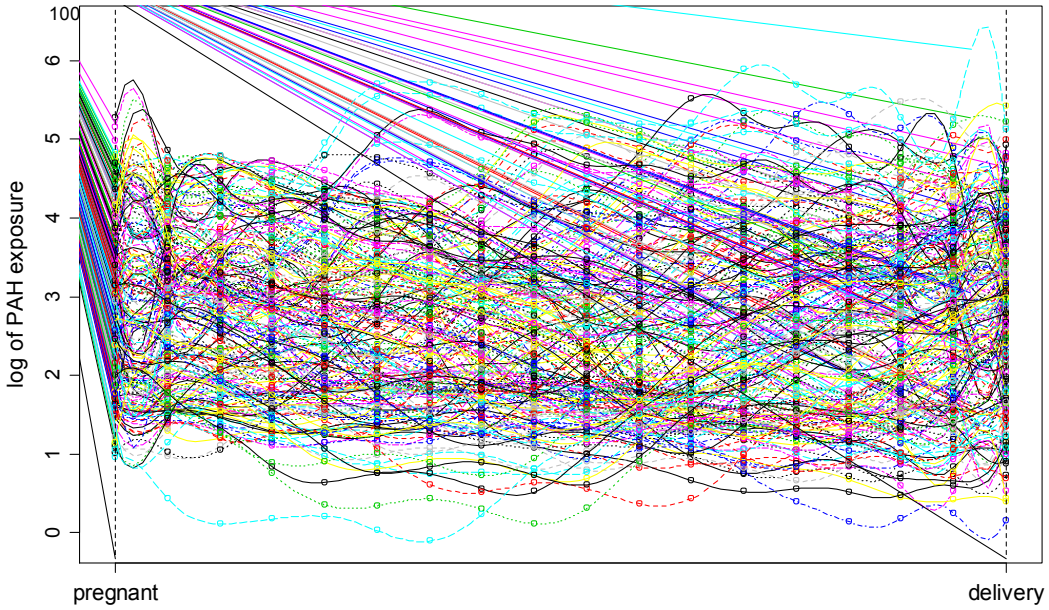
What are the unique model properties?

- Semiparametric-mixed Effect Model:

$$Y_{ij} = x_{ij}^T \beta + f(t_{ij}) + b_i + Z_{ij}^T b + e_{ij}$$



Cyclic Cubic Regression Spline



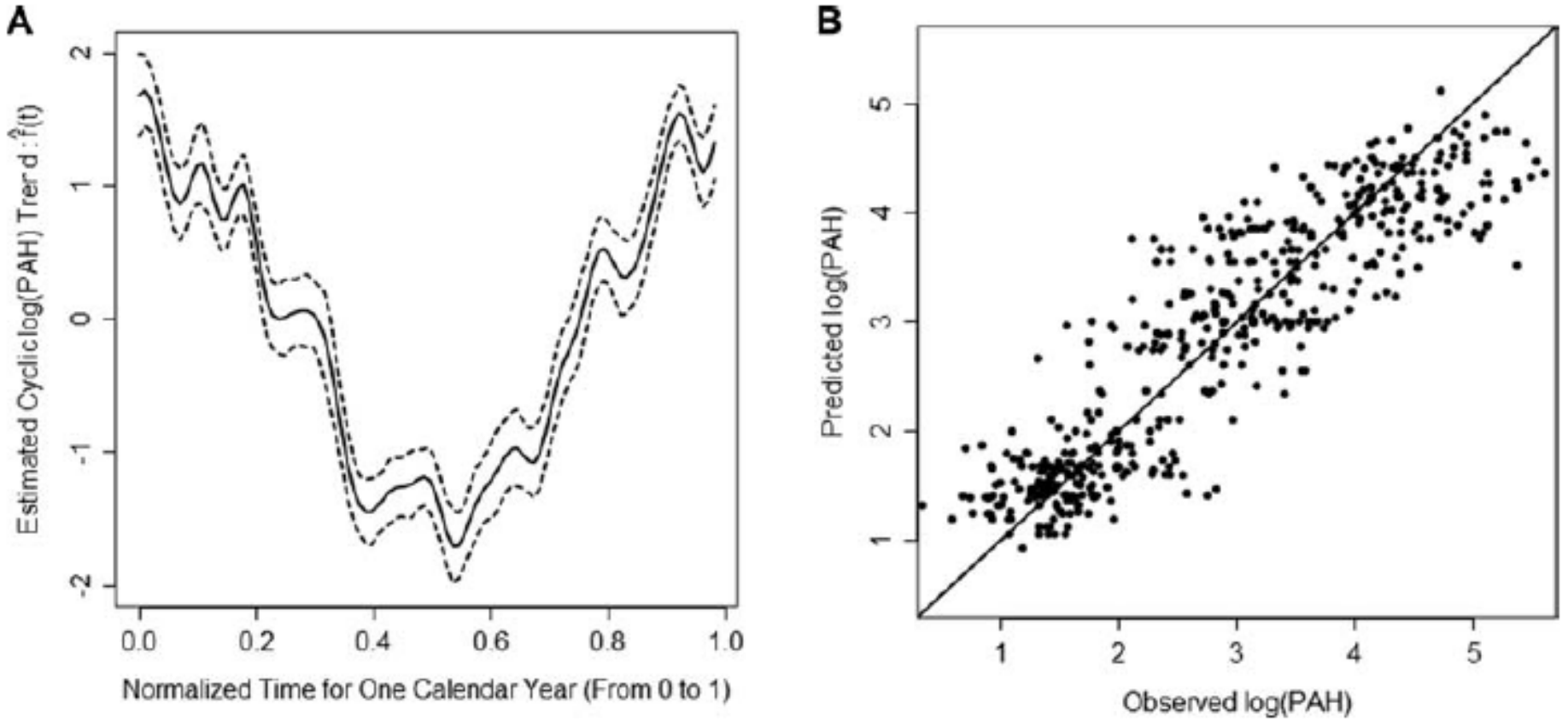


Figure 2 Estimation of the cyclic nonparametric function $f(t)$ in one cycle (— is the estimate, and ---- is the 95% confidence interval) and the predictions of individual prenatal PAH exposure during pregnancy based on semiparametric-mixed effect model. (A) Estimation of cyclic function $f(t)$ in one cycle (b) predicted versus observed PAH.

Prediction Model for Fetal Growth

- **Functional Linear Model:**

$$Z_i = \tilde{x}_i^T \tilde{\beta} + \int_0^{T_d} Y_i(s)\alpha(s)ds + \varepsilon_i$$

Z_i : birth outcome of baby i, e.g. birth weight.

$Y_i(s)$: the PAH exposure curve of the ith subject evaluated at gestational age s.

x : other covariates

epsilon : measurement error

Discussion Point #3:

What is the main model difference compared to other parametric regression models?

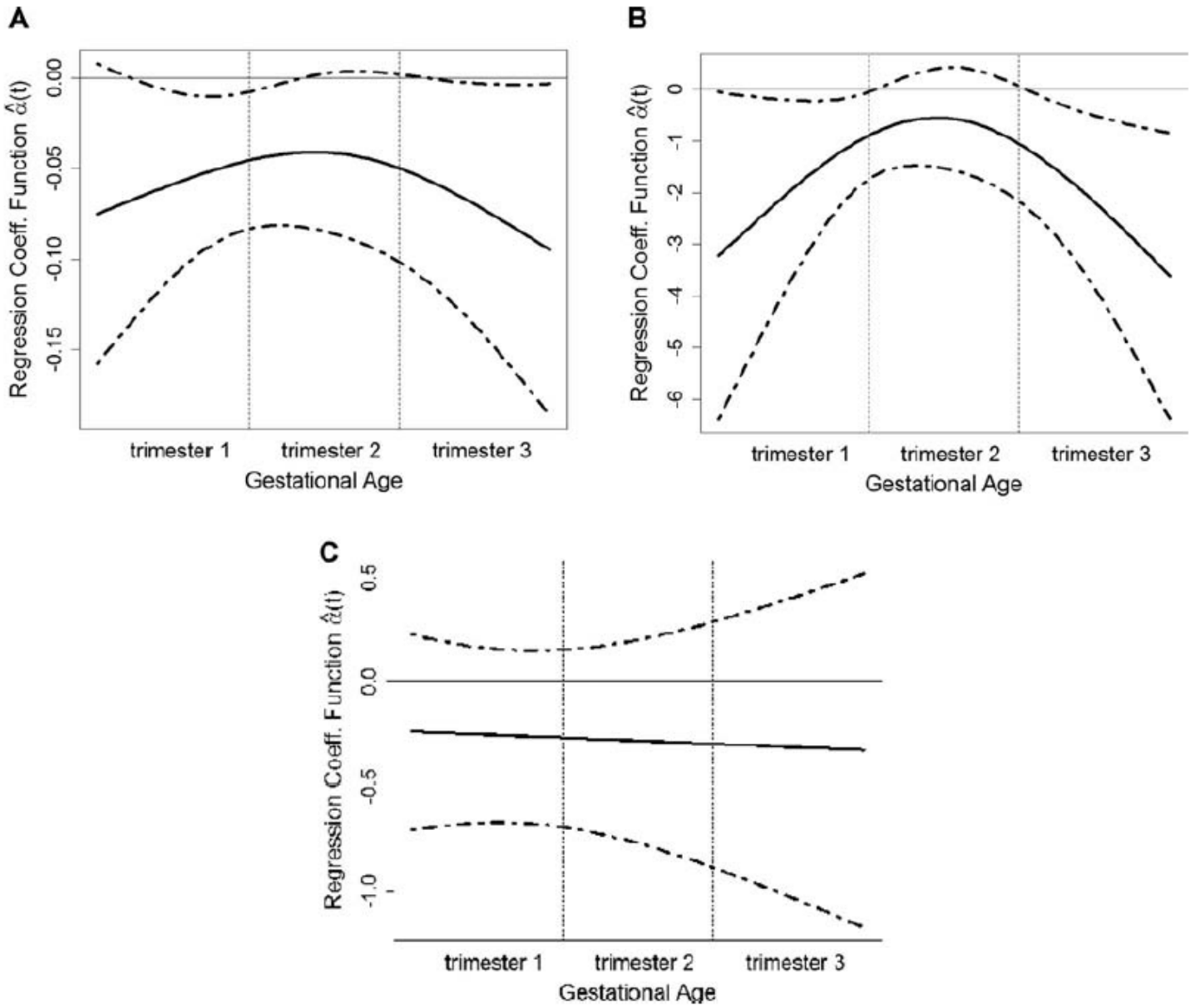


Figure 3 The estimated regression coefficient function $\alpha(t)$ in functional linear model, that is, the effect of prenatal log PAH exposure on baby's log birth weight, birth length, and birth head circumference across gestational age t. —— is the estimate and – · – · – is the 95% confidence interval. (A) log Birth Weight, (B) Birth Length, (C) Birth Head Circumference.

Table 1 Estimates of the effects of other risk factors in the functional linear models of birth outcomes.

	log(Birth Weight)		Birth Length		Birth H-C	
	coefficient	*p*-value	coefficient	*p*-value	coefficient	*p*-value
Maternal height	0.002	0.041	0.056	0.025	0.030	0.023
Prepregnancy weight	0.003	<0.001	0.046	0.003	0.027	<0.001
log(Gestational Age)	2.261	<0.001	36.37	<0.001	14.10	<0.001
Parity (yes versus no)	0.027	0.038	0.364	0.188	0.474	0.001
Newborn gender (girl versus boy)	−0.058	<0.001	−1.078	<0.001	−0.733	<0.001
Whether c-section delivery	−0.020	0.204	0.278	0.421	0.403	0.027
Whether born in summer season	0.029	0.224	0.001	0.993	0.120	0.639

Recap Statistical Inference and Findings

- **Using longitudinal semiparametric mixed effect model and functional linear models**
  - We minimize the penalized least squares objective function using a spline-based expansion of the nonparametric functions.
  - The smoothing parameters are selected using GCV criteria.
- **This study suggests that the vulnerability of fetus against high prenatal PAH exposure varies across different gestational age.**
  - Hint a couple of critical windows of vulnerability for fetal weight and height development, during which the PAH exposure yields significant impairment.
  - Thus, reducing PAH exposure during these gestational windows may help for fetal weight and length development.
  - For birth head circumference, it appears to be affected more and more detrimentally across the gestational age, but no statistical significance is found.

Other Discussion Points….

- Treat time (e.g. gestational age) as a totally continuous variable, with day as unit, not month.
- Fixed integral bounds v.s. Random? Note actually the time interval from getting pregnant to delivery varies individually.
- e.g. $$Z_i = \tilde{x}_i^T \tilde{\beta} + \int_{s_1}^{s_2} PAH_i(s)\alpha(s)ds + \varepsilon_i$$

  s_1 : time got pregnant

  s_2 : time for delivery

- Sparse of longitudinal measurements
- Potential bias from excluding non-full-term babies

Thank you!