

# Assessing Effects of Exposures to DDE and PCBs on Premature Delivery via Ordinal Logistic Regression

Raphael Morsomme   Rihui Ou   Alessandro Zito

Case Study 1 - Stat 723

January 20, 2020

# Overview

- 1 Introduction
- 2 Data
- 3 Model (I) - Ordinal Logistic Regression
- 4 Model (II) - Bayesian Ordinal Logistic Regression
- 5 Results
- 6 Conclusions

# Introduction

- **Framework:**

*Dichlorodiphenyldichloroethylene (DDE) and Polychlorinated Biphenyls (PCBs) are chemicals that persist in the environment and get stored in fatty deposits in the human tissues.*

⇒ Potential adverse effect on health

- **Question:**

*Is exposure to DDE and PCBs associated with a higher chance of premature delivery in pregnant women?*

## Pregnancy timeline

- **Dangerous preterm:** delivery at 34 weeks or before (when main organs are underdeveloped)
- **Preterm:** delivery between 35 and 37 week
- **At term:** delivery after 37 weeks

Data collected by 12 centers contained gestational age (in weeks) of the mother, the DDE and PCBs concentration, socio-economic info and scores (race, occupation, education, income), amount of triglycerides and cholesterol in blood and smoking status.

## Preprocessing:

- Drop obs. with gestational age  $> 45$  (the world record)
- Standardize and average levels of PCBs<sup>1</sup>

$$PCB_i = \frac{1}{11} \sum_{j=1}^{11} \frac{PCB_{ij} - mean_i(PCB_{ij})}{sd_i(PCB_{ij})}$$

- Mean impute of occupation, education and income scores
- Aggregate race into  $race = 1$  if white and  $race = 0$  if non-white

⇒ Total obs. = **2336**

---

<sup>1</sup>This avoids the correlation between the PCBs. See the appendix.

- Our dependent variable is:

$$gestgroup_i = \begin{cases} 0 & \text{if Dangerous preterm} \\ 1 & \text{if Preterm} \\ 2 & \text{if At term} \end{cases}$$

- To account for triglycerides and cholesterol, we introduce an **adjusted measure for PCB and DDE** by:
  - 1 Computing total lipids using Phillips et al.(1989) and Bernert et al.(2007) formula

$$lipid_i = 2.27 * cholesterol_i + triglycerides_i + 0.623$$

- 2 Setting<sup>2</sup>

$$adjDDE_i = \frac{DDE_i}{\log(lipid_i)} \quad adjPCB_i = \frac{PCB_i}{\log(lipid_i)}$$

---

<sup>2</sup>The choice of the log comes from a Box-Cox analysis of the log-likelihood, as in Li, Longnecker and Dunson (2013)

# EDA (I) - Exposures and gestational groups by race

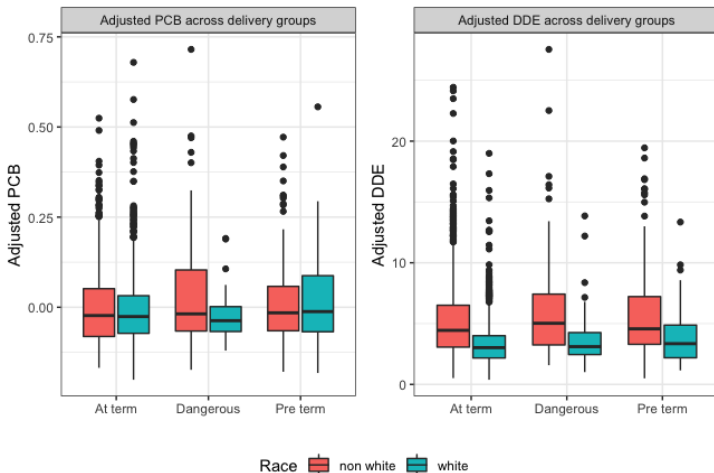


Figure: Relationship between delivery group and adjusted exposures, by race

# EDA (II) - Exposure across centers

# Model (I) - Ordinal Logistic Regression

We run the following ordinal logistic regression model:

$$\text{logit}(P(\text{gestgroup} \leq j)) = \beta_{0j} - \mathbf{X}\boldsymbol{\beta} + \epsilon$$

where  $j = 0, 1, 2$  corresponds to the outcome level, and  $\mathbf{X}$  contains:

- *adjDDE*, *adjPCB*, *race*, *center*, *smoke*, the 3 scores [main effects]
- $(DDE + PCB) * (race + center)$  [interactions].

AIC-based backward variable selection:

- Maintain *DDE*, *PCB*, ...,  $(PCB + DDE) * race$
- Drop  $(DDE + PCB) * center$

Model assumptions are checked in the appendix.



# Model (II) - Bayesian Ordinal Logistic Regression

# Results

# Conclusions

# Appendix (I) - More EDA

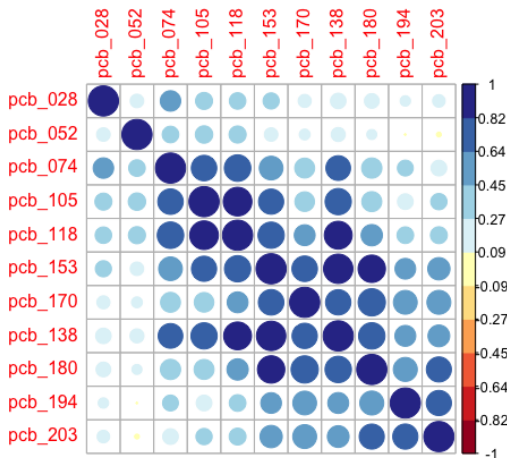


Figure: Correlation plot across PCBs

# Frequentist Model Checking

We can check the assumption of the (frequentist) ordinal logistic model by looking at the Surrogate residuals. **ADD CITATION HERE**

If the model assumptions are correct, then the surrogate residuals  $R_S$  will have three properties:

- 1  $E(R_S|X) = 0$
- 2  $Var(R_S|X) = c$ , the conditional variance of  $R_S$  is constant
- 3 The empirical distribution of  $R_S$  resembles an explicit distribution that is related to the link function  $G^{-1}(\cdot)$ . Specifically,  
 $R_S \sim G(c + \int u dG(u))$ .

# Frequentist Model Checking

Assumptions (i) and (ii) are checked with the Surrogate residuals plot. Both are satisfied in this case.

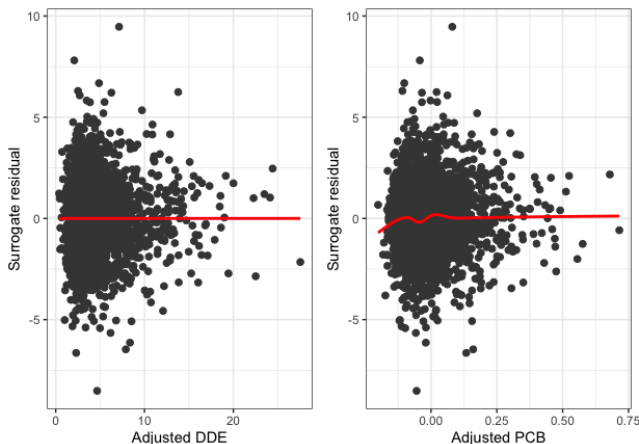


Figure: Surrogate residuals of DDE and PCB

# Frequentist Model Checking

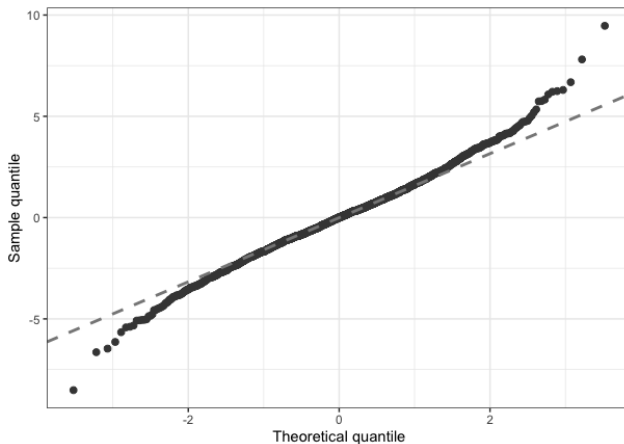


Figure: QQ plot of the Surrogate residuals