

Case Study 1

Emily Gentles Phuc Nguyen Joseph Lawson

Jan 21, 2020

Case Discussion

- ▶ Data obtained from a subset of women enrolled in the CPP during pregnancy

Goal: Assess how exposure to DDE and PCBs relates to the risk of premature delivery and the relative severity thereof

Exploratory Data Analysis

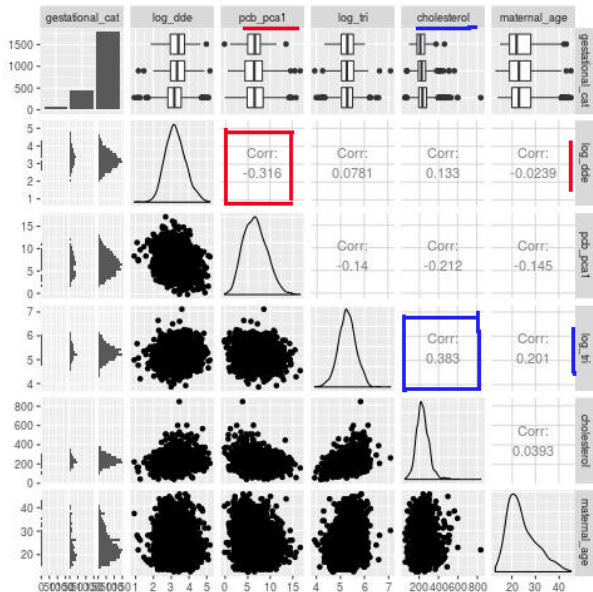
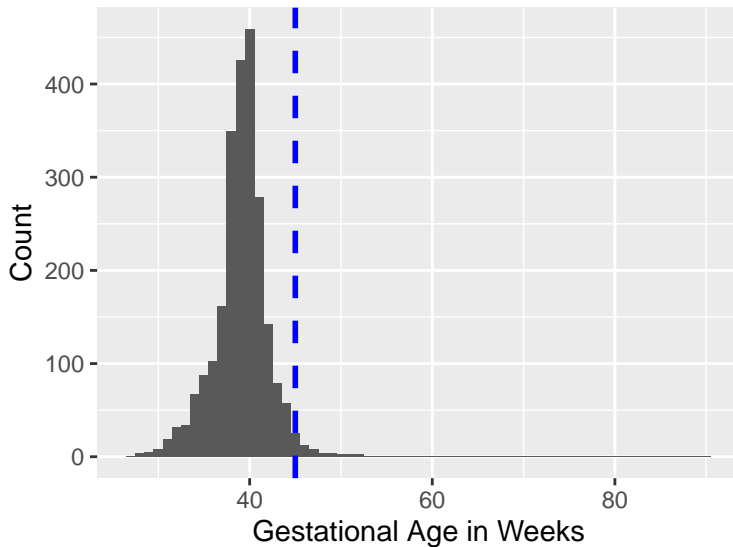
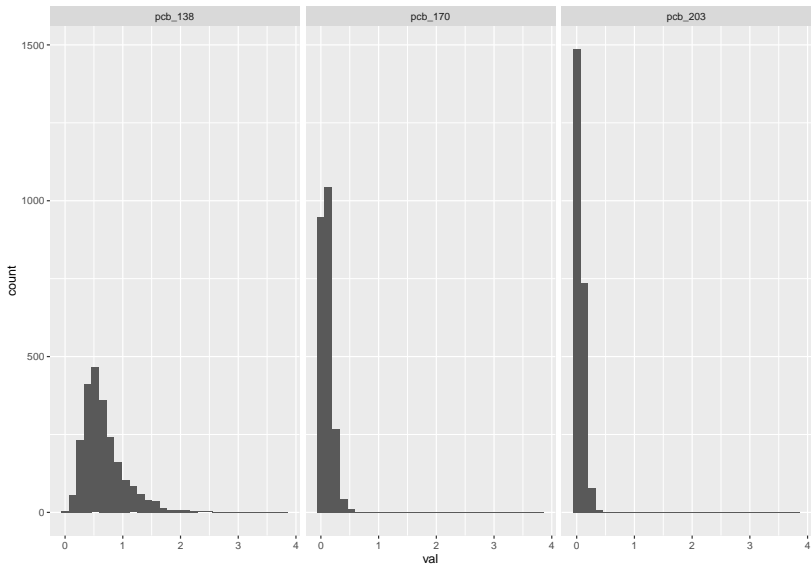
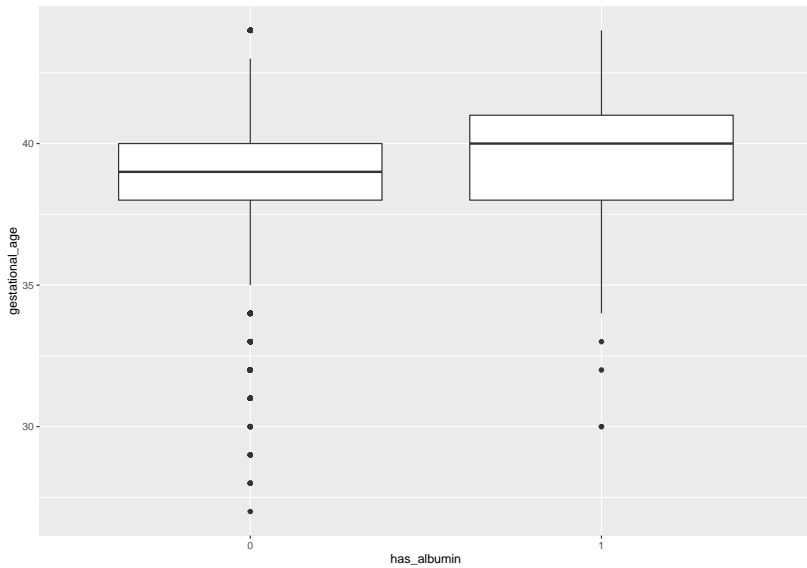


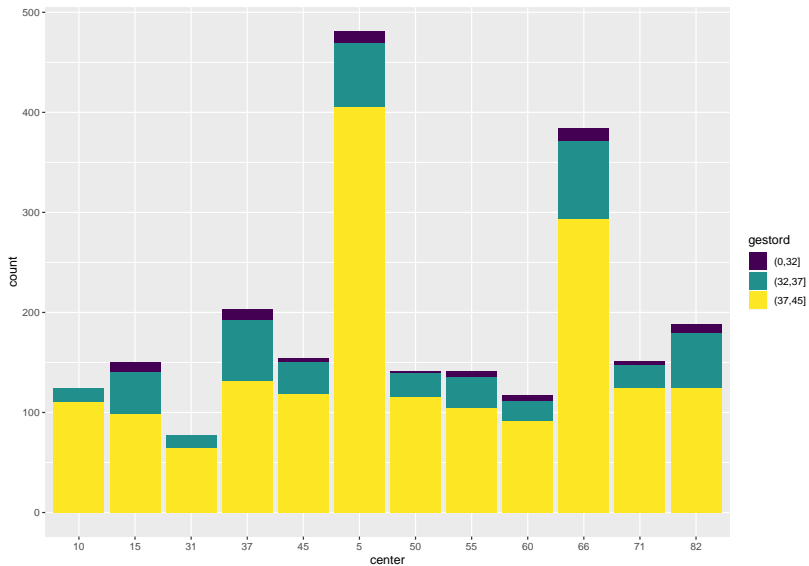
Figure 1: corr-plot

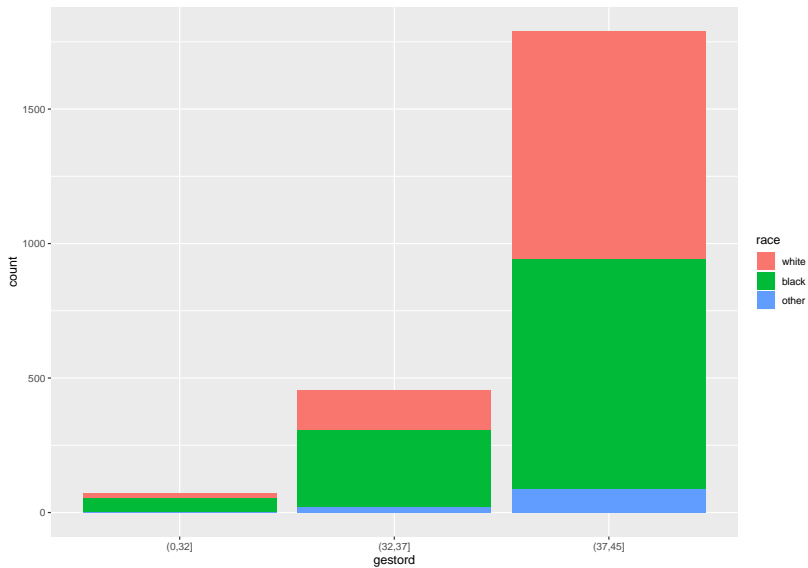
► Data issues: uncertainty, inflation, and missingness



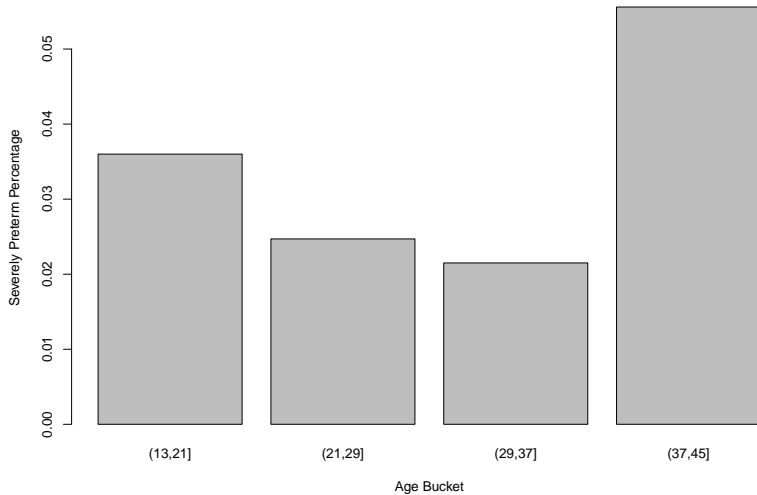








Severely Preterm Probability vs Maternal Age



Analysis

- ▶ Ordinal Logistic Regression with Term, Preterm, and Severely Preterm Gest. Categories
 - ▶ Useful interpretation in terms of risk
 - ▶ Uses naturally ordinal structure
 - ▶ Differentiates between different severities
- ▶ Impute score data with MICE to check usefulness
- ▶ Remove obs with missing PCB value
- ▶ Include blood cholesterol/triglyceride levels, as well as center and SES/Lifestyle metrics

- ▶ Model Comparison indicated ($p=0.97$) that the first principle component of the pcb_* values is sufficient.
- ▶ Indication ($p=0.74$) against including Score Variables (post imputation)
- ▶ Indication ($p=0.21$) against including Center interactions
- ▶ Indication ($p=0$) for including Center as variable
 - ▶ Indicates heterogeneity in preterm birth risk accross medical centers

- ▶ Indication ($p=0.46$) against PCB-DDE interaction effect
- ▶ Indication ($p=0.15$) (weakly) against Triglyceride interaction with PCE/DDE
 - ▶ Hypothesized that there might be because of fat-solubility of the contaminants of interest
- ▶ Indication ($p=0.02$) for inclusion of quadratic term in maternal age
 - ▶ Makes sense given heightened birth risk at (relatively) young and old ages of pregnancy (for possibly different reasons)
- ▶ Strong indication that the indicator of testing for Albumin is associated with longer gestational period on the margin

- ▶ Control Variables of Significance (marginally) in Final Model:
 - ▶ Center
 - ▶ Triglycerides (g/dL)
 - ▶ Cholesterol (g/dL)
 - ▶ Maternal Age (years)
 - ▶ Albumin (Testing Indicator)
 - ▶ Race

Final Model

$$GestCategory \sim Center + \log(dde) + \log PCB1 \quad (1)$$

$$+ \log(trigl.) + Poly(MaternalAge, 2) \quad (2)$$

$$+ Smoking + \log(Cholest.) + AlbuminTested \quad (3)$$

$$+ race \quad (4)$$

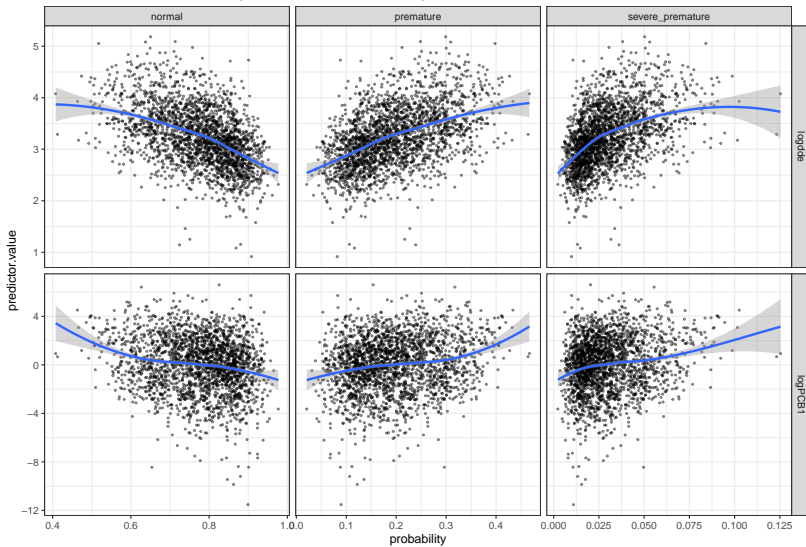
Results

- ▶ (log) DDE and PCB are (collectively) significantly associated with preterm delivery likelihood even when adjusting for other factors
- ▶ Interpretation that 1% increase in DDE is approximately associated with a -0.18% change, and a unit increase in logPCB1 with a -0.06 change, in the log odds of of a gestational age occurrence below a given ordinal threshold.
 - ▶ Note for interpretability that the PCB coefficient is relative to the (standardized/centered) first principle component of the pcb_* measurements
- ▶ Testing inclusion of both DDE and PCB vs control indicate ($p=0.005$) that they provide increased explanatory power vs control, but marginally in the full model each is just at the edge of significance, as seen in the following table:

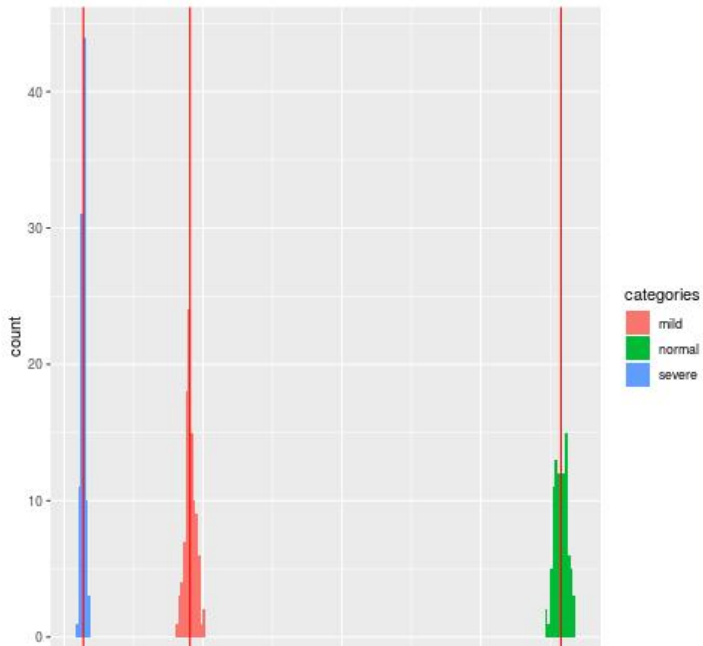
	Coef Est	2.5 %	97.5 %
log(dde)	-0.183	-0.388	0.022
logPCB1	-0.056	-0.113	0.000

- Note the overlap of the confidence intervals with 0.

Fitted Probabilities vs DDE (ug/dL)/PCB (princ comp.) (log scale)



Sensitivity Analysis



Sensitivity Analysis

