STA723 Case Study - Group 1

Melody Jiang, Irene Ji, Keru Wu

1/20/2020

Abstract

This report was commissioned to analyze the data from the multilaboratory study on testing the effect of estrogen agonist EE and estrogen antagonist ZM on the weight of the uterus using the rat uterotrophic bioassay. The report evaluates the ability of detecting effects of EE and ZM of different labs with different protocols with OLS model. Methods of analysis include ANOVA and confidence intervals. Results of data analysed show that the bioassay is successful overall at identifying estrogenic effects of EE and anti-estrogenic effects of ZM, but dose response varies across labs and protocols. In particular, lab Huntingd failed to detect the estrogenic effects of EE, while TNO and ChungKor failed to detect the anti-estrogenic effects of ZM. It is recommended to use the protocol of immature female rats dosed by injection because its best sensitivity of detecting EE and ZM effects.

1. Introduction

2. Materials & Methods

1. Full Model
2. Effects of EE and ZM
   (a) Overall Effects of EE and ZM
   (b) Identification of EE Effects in Individual Labs
3. Inter-laboratory comparison
4. Inter-protocol comparison

3. Results

3.1 Exploratary Data Analysis

3.2 Main Results

3.3 Sensitivity Analysis

4. Discussion

This report has analyzed how DDE and PCBs relate to the risk of premature delivery. After preprocessing data (e.g. impute missing data) and dimensionality reduction (PCA), we build up different models for the data and finally adopt the Generalized Additive Model (GAM) and its Bayesian version. Our approach has advantages that it captures the nonlinearity relationship between exposures and outcomes, and it also explaines . We conclude that higher exposure to DDE and certain PCBs may be associated with the risk of premature delivery.

The first extension of our approach is to deal with different centers specially. Our model demonstrates that center 15 and center 37 may deviate from others, which genrally collect samples with higher risks. There are other ways which may perform better in dealing with centers. One can adopt a Bayesian hierarchical model which specifies different variances between centers. Another more direct extension of our GAM model is to include mixed effect. We may use Generalized Additive Mixed Model (GAMM) to consider random effect of centers and it can also be applied to other categorical variables like smoking status and race.

After we find out that DDE and PCBs are related to higher risk of preterm birth, we can examine the trend of exposures effects more accurately. Neelon and Dunson (2004) developed a Bayesian Isotonic Regression model which incorporates both non-decreasing effect and flat region problems. Using specialized autoregressive prior for piecewise linear funciton in GAM, it benefits from narrower credible intervals compared to frequentist approach and naive bayes approach.

Furthermore, interaction between chemicals also impacts human health outcomes. Collinearity among PCBs indicate the need for a general dimension reduction method or a variable selection approach. Ferrari and Dunson (2019) build up a bayesian factor model designed for interactions. High correlation between exposure levels can be explained in this flexible dimension reduction approach.