P8106 Midterm - Report

Group 2: Kate Colvin (KAC2301), Jeong Yun (Lizy) Choi (JC6452), and

Introduction

In this project, our team explored the dataset collected from a study on evaluating antibody responses to a newly authorized vaccine. The primary outcome of interest is the log-transformed antibody level measured via dried blood spots. The dataset includes a range of demographic and clinical predictors such as age, gender, race/ethnicity, smoking status, BMI, chronic conditions, and time since vaccination.

Our goal is to develop a predictive model that characterizes how these factors influence antibody responses and asses how well this model generalizes to a new independent dataset collected at a later time point. By doing so, we hope to identify key predictors of antibody levels and evaluate the robustness/generalizability of our model across different dataset.

Exploratory Analysis

notes

for hist - more recent vaccines = higher antibody level ?

for scatter plots - testing slopes more flat?

Model Training

Results

Appendix