**BIOST 2050: Longitudinal and Clustered Data Analysis**

**Homework Assignment #5**

**Due: Friday, November 3, 2023**

Answer the questions and justify your answers. Note that you will receive a major deduction if you answer a question by giving only the software output without justifying your answer.

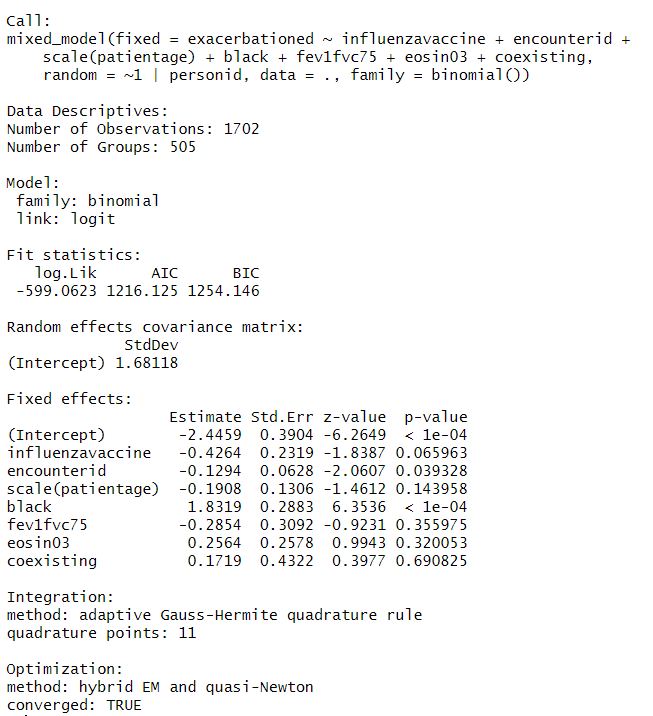
Until now, most pediatric studies examining the relationship between asthma and COVID-19 have been ecological and provided limited insights. To enrich this understanding, an observational study was conducted to evaluate the association at an individual level. The study population consists of children and adolescents aged 2 to 21 years who were diagnosed and hospitalized with asthma. The primary research questions were: (1) whether receiving an influenza vaccination is associated with improved asthma outcomes, specifically in the context of emergency department (ED) visit due to asthma exacerbation, and (2) whether racial disparities impact this association.

The asthma dataset contains variables specifically tailored to address the research questions. Descriptions of these variables can be found in the table provided below.

|  |  |
| --- | --- |
| **Variable name** | **Description** |
| personid | patient’s unique identification number |
| encounterid | hospital/clinic visit (1, 2, 3, 4, 5, 6, 7) |
| patientage | patient’s age (in years) |
| eosinophils | lab result for eosinophil counts in blood. eosinophile is a type of white blood cell |
| eosin03 | eosinophils >=0.3/L (0=no, 1=yes) |
| fev1fvc | ratio of FEV1/FVC, which is a pulmonary function test |
| fev1fvc75 | fev1fvc >=75% (0=no, 1=yes) |
| black | patient’s race/ethnicity (0=white, 1=black) |
| influenzavaccine | whether the patient had a flu shot for that season (0=no, 1=yes) |
| coexisting | whether the patient has other respiratory disorders coexisting (0=no, 1=yes) |
| exacerbationed | Emergency Department (ED) visit due to asthma exacerbation (0=no, 1=yes) |

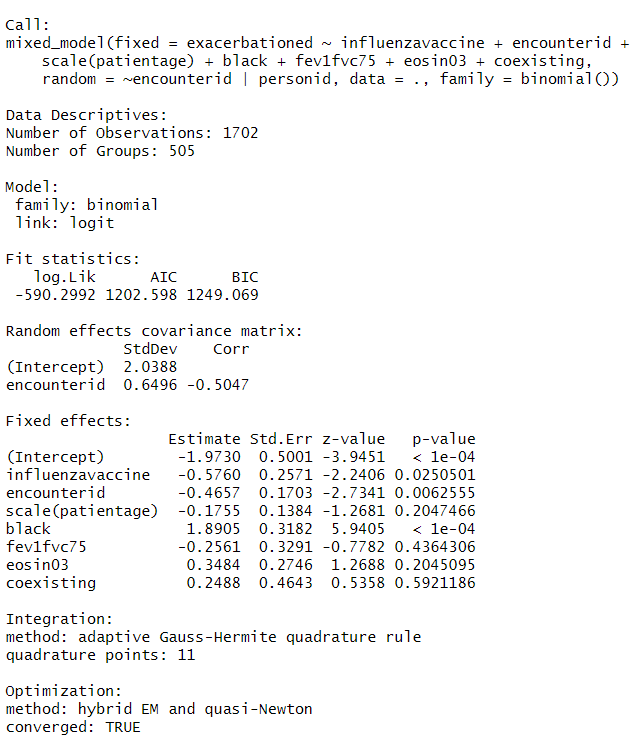
1. Compare the following three models to investigate whether receiving an influenza vaccination is associated with Emergency Department (ED) visit due to asthma exacerbation: the random-intercepts model, the random-coefficients model (random coefficients for “encounterid” only) with independent random components, and the random-coefficients model (random coefficients for “encounterid” only) with potentially correlated random components. For all models, adjusting covariates include patient’s age, race, a ratio of FEV1/FVC less than vs. greater than or equal to 75%, eosinophils less than vs. greater than or equal to 0.3/L, and whether has other co-existing respiratory disorders. Please provide formal statistical tests in your model comparisons. (40 pts)

## Model 1: Random-intercepts model



The model demonstrates, conditional on the same random intercept, statistically significant associations between `exacerbationed` and having race == black. The model also indicates that the intercept (i.e., 0 age & the absence of any of the categorical factors) is significantly different from 0. Additionally, advancements in encounter ID are statistically significant for a given random intercept.

## Model 2: Random coefficients model (random coefficients for 'encounterid' only)



Given the same random intercept and random coefficient for encounterid, the intercept (i.e., 0 age, '0th' encounterid, and the absence of any of the categorical factors), influenza vaccine, encounterid, and black race are statistically significantly associated with exacerbationed.

## Model 3: Random coefficients model (random coefficients for 'encounterid' only) with potentially correlated random components

```{r}

library(geepack)

library(nlme)

library(gee)

library(lme4)

gm3 <- asthma %>%

glmer(exacerbationed ~ influenzavaccine + encounterid + patientage + black + fev1fvc75 + eosin03 + coexisting + (1 + encounterid | personid),

data = .,

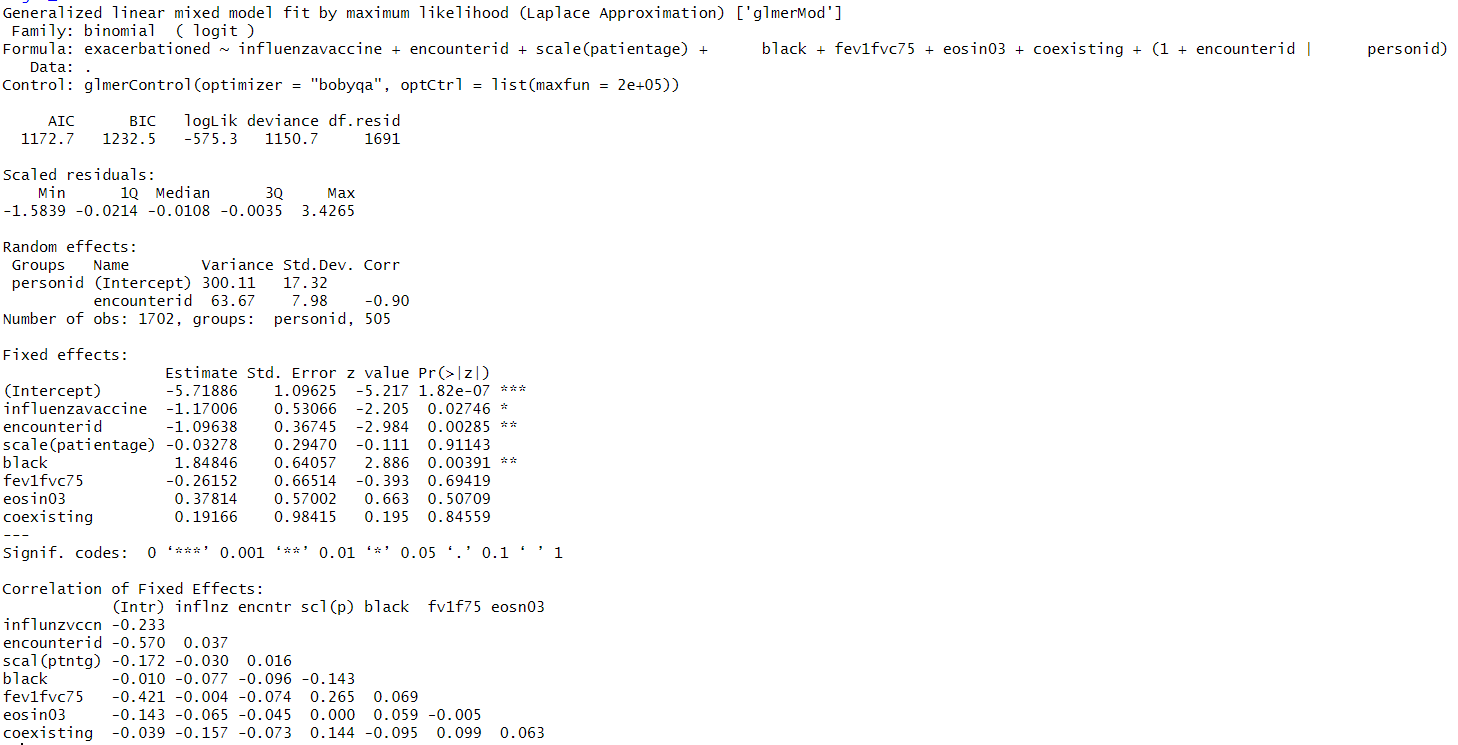
family = 'binomial',

control = glmerControl(optimizer="bobyqa", optCtrl=list(maxfun=2e5)))

gm3\_sum <- summary(gm3)

gm3\_sum

```



Given the same random effects, the intercept, influenza vaccine, encounterid, and black race are all statistically significantly related to exacerbationed.

## Compare 3 models

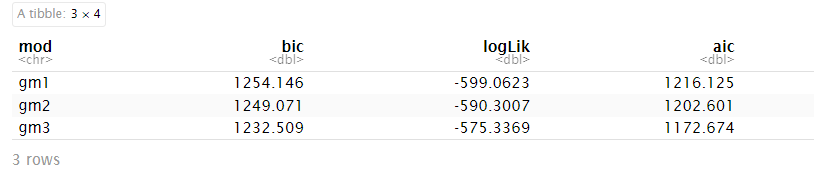
```{r}

tibble(mod = c("gm1", 'gm2', 'gm3'), bic = c(gm1\_sum$BIC, gm2\_sum$BIC, gm3\_sum$AICtab['BIC']),

logLik = c(gm1\_sum$logLik, gm2\_sum$logLik, gm3\_sum$AICtab['logLik']),

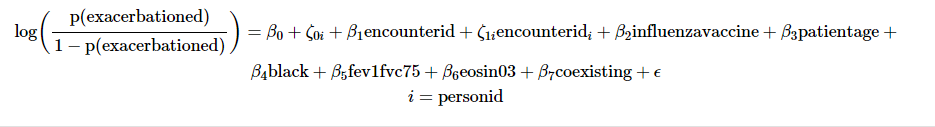
aic = c(gm1\_sum$AIC, gm2\_sum$AIC, gm3\_sum$AICtab['AIC']))

```

x

gm3 (random coefficient with possible covariance) is the best performing model of the 3 according to BIC, AIC, and log-likelihood

1. Write down the model specification of the best model from question #1. (15 pts)



1. For the best model, assess whether racial disparities exist in the association between receiving an influenza vaccination and ED visit due to asthma exacerbation. (15 pts)

```{r}

gm4 <- asthma %>%

glmer(exacerbationed ~ influenzavaccine\*black + encounterid + scale(patientage) + fev1fvc75 + eosin03 + coexisting + (1 + encounterid | personid),

data = .,

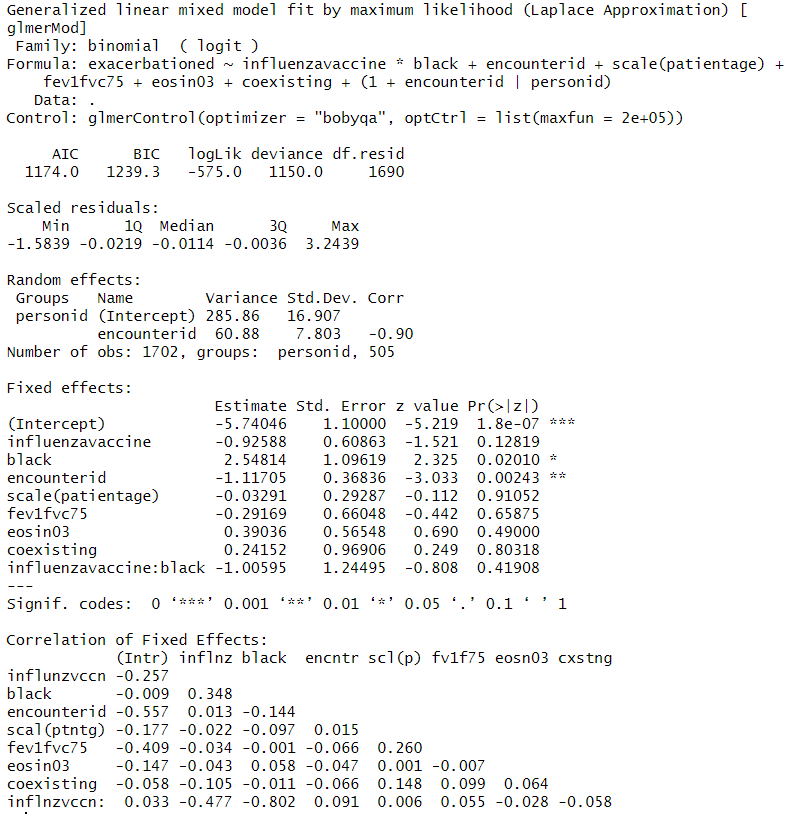
family = 'binomial',

control = glmerControl(optimizer="bobyqa", optCtrl=list(maxfun=2e5)))

gm4\_sum <- summary(gm4)

gm4\_sum

```



There is no significant interaction effect between black vs. other race and influenza vaccine for a given set of random effects. Additionally, this renders the main effect of influenza vaccine, given a random effect, insignificant. Additionally, the model including the interaction term performs worse than the model without:

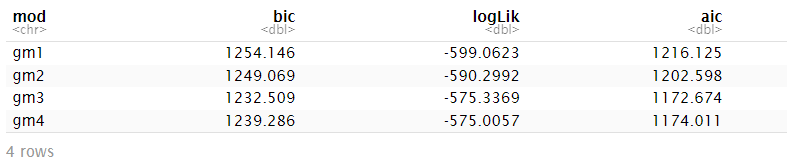
```{r}

tibble(mod = c("gm1", 'gm2', 'gm3', 'gm4'), bic = c(gm1\_sum$BIC, gm2\_sum$BIC, gm3\_sum$AICtab['BIC'], gm4\_sum$AICtab['BIC']),

logLik = c(gm1\_sum$logLik, gm2\_sum$logLik, gm3\_sum$AICtab['logLik'], gm4\_sum$AICtab['logLik']),

aic = c(gm1\_sum$AIC, gm2\_sum$AIC, gm3\_sum$AICtab['AIC'], gm4\_sum$AICtab['AIC']))

```



1. Fit a GLM model with GEE to assess the effect of receiving an influenza vaccination on ED due to asthma exacerbation, adjusting for the same covariates as in question #3 and using an exchangeable working correlation structure. Does this effect look similar to the one obtained from model in question #3? If not, please provide a possible explanation for the differences. (30 pts)

```{r}

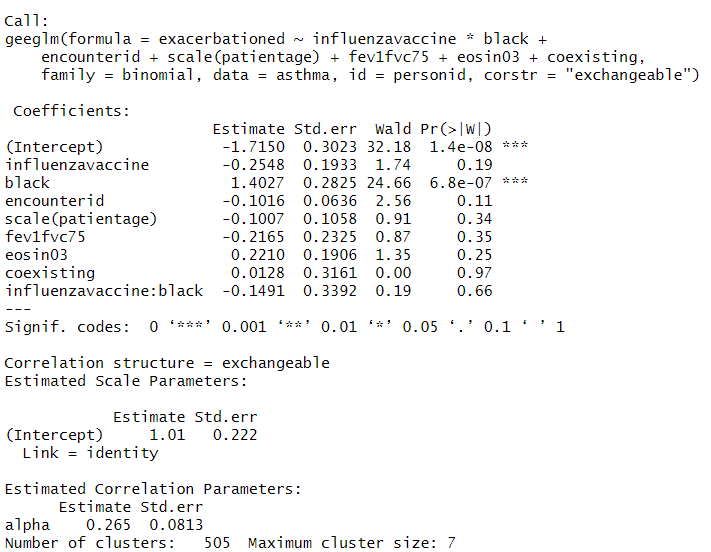
gee\_exc <- geeglm(exacerbationed ~ influenzavaccine\*black + encounterid + scale(patientage) + fev1fvc75 + eosin03 + coexisting,

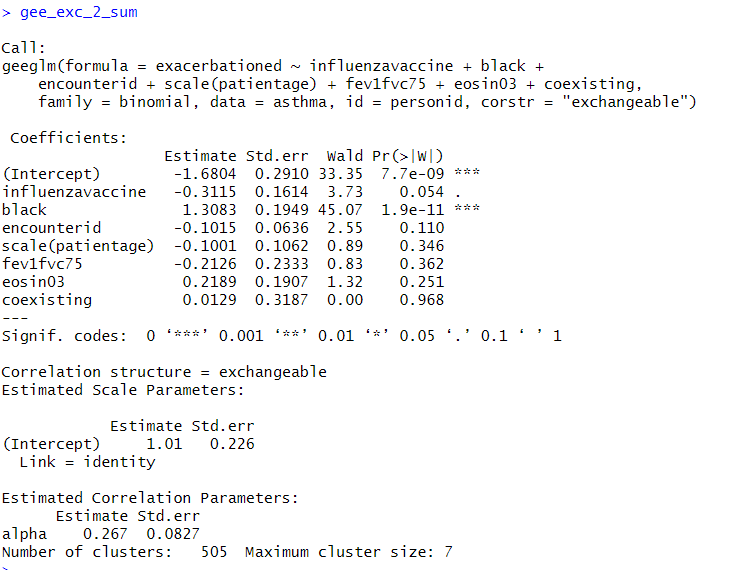
data = asthma, id = personid, family = binomial, corstr = "exchangeable")

gee\_exc\_sum <- gee\_exc %>% summary()

gee\_exc\_sum

```





The effect of influenza vaccine in the GEE is still negative but lower in magnitude compared to the GLMM effect. Neither of the effects in either model are statistically significantly different from zero. When not including the interaction effect in these models, the influenza vaccine effect is significantly different from zero for a given random effect, while in the GEE it is not. Regardless of whether the interaction is included, the reason the influenza vaccine effects in the GEEs are lower in magnitude than the GLMMs is because the effect in the GLMMs must always be considered in conjunction with the random effects (the random intercept and coefficients components). This is to say, the GLMM parameter, considered with a given random effect, gives the *subject-specific average effect*, while the GEE parameter, accounting for a given correlation structure, gives the *population-average effect.*