

## Assignment #3

Due November 14<sup>th</sup>, 2013

### CHL 5225: Advanced Statistical Methods in Clinical Trials

Data for the third assignment is in the file HIVtrialdata.xls. These data are from a randomized placebo-controlled trial of four treatment arms:

- 1 = zidovudine alternating monthly with 400mg didanosine,
- 2 = zidovudine plus 2.25mg of zalcitabine,
- 3 = zidovudine plus 400mg of didanosine, and
- 4 = zidovudine plus 400mg of didanosine plus 400mg of nevirapine.

The primary outcome variable is the CD4 cell count.

CD4 counts were measured at baseline (week 0) and weeks 8, 16, 24, 32 and 40 after randomization.

The variables in the data set are:

- ID: patient identifier
- Treatment (1, 2, 3, or 4)
- Age: (age in years)
- Gender: (1 = Male, 0 = Female)
- CD4 count
- Baseline CD4 count (CD4 count at week 0)
- Day: days from randomization
- Week: nearest 8 week visit from randomization

1. Plot the mean CD4 count by treatment group and week of follow-up. Show the numbers of patients contributing data at each time point. Comment on the trends that you see. Notice that several patients contribute more than one CD4 count to a single visit. Select the CD4 count closest in time to the 8 week visit.
2. Compare rates of dropout by treatment arm. Calculate mean CD4 counts and changes in CD4 counts by week of follow-up for (a) patients with follow-up at week 40 and (b) patients without follow-up at week 40. Are patients with complete follow-up similar or different from patients with incomplete follow-up? How does this finding affect your interpretation of the results of the trial?
3. Fit a GEE model to estimate the effects of the four treatment arms on CD4 counts. Include CD4 counts from weeks 8 to 40 as part of the outcome. Consider the CD4 count at week 0 (basecd4) as a covariate.
  - a. Describe how you chose whether to model time as a linear variable or with dummy variables for each follow-up visit.

- b. Fit a model with just treatment effect. Examine the correlation matrices for exchangeable, unstructured, mdep(4) and AR(1) and compare the QIC for each correlation structure. Choose the best correlation structure for this model and describe how you chose the correlation matrix.
    - c. Determine whether age and gender are important predictors of CD4 counts
4. Fit another GEE model in which the CD4 count at week 0 is part of the outcome vector. Include only participants from treatment arms 3 and 4. Use categorical variables to model week and assume an exchangeable correlation structure.
  - a. Include covariates for treatment and week of follow-up in the model. Interpret the meaning of the estimated coefficients.
  - b. Include covariates for treatment, week of follow-up and the interaction of treatment with week of follow-up. Interpret the meaning of the estimated coefficients. How does this differ from the interpretation of coefficients in question 3?