**EPIDEMIOLOGY 340.600: STATA PROGRAMMING**

**Lab 3**

This lab is optional; you are NOT required to complete these questions. Please use this lab as an opportunity to review the course material and prepare yourself for the homework questions. Sample responses to the lab questions will be provided separately.

1. Start Stata, open your do-file editor, write the header, and load transplants.dta.
2. ctr\_id indicates the ID of the transplant center where the patient received the transplant. Count the number of recipients at each center, and store in a new variable volume.
3. List ctr\_id and volume to see how many patients each center has. Maybe let’s try this: list ctr\_id volume
4. Whelp. This is not what we wanted. Generate a variable ctr\_tag that “tags” one observation per center.
5. Now list ctr\_id and volume, but just for one record per center.
6. Calculate the mean age of the patients at each center, and store in a new variable mean\_age.
7. For each primary diagnosis subgroup (use variable dx), run a regression with age as the predictor and peak PRA (peak\_pra) as the outcome.

1. Now let’s make the output cleaner. Count the number of cases within each diagnosis group. If there are more than 500 cases, run the regression and display the output. If not, display “There are fewer than 500 cases.”
2. Define a program called reg\_pra. This program will perform the same tasks as described in Question 8, but the regression will take one or more variables specified by the user as the predictor.
3. You have all your commands in your do file, right? Run your do file from the beginning and make sure your do file does exactly the same thing.