Project 1 Notes 9/18/17

* not all variables from data dictionary are in the dataset
* not planning on including comorbidities unless there is a reason (example of comorbidity = hbp)
  + hbp:
    - not sure what the difference is between yes and yes from data trajectory or no and no from data trajectory

Notes 9/20/17

Info About the Interim Analysis Presentation:

* Might ask – which of these variables do you really want me to look at
* might notice that there are 10 levels for this categorical variable (how can we re-do this)
* comment on any weird things you’ve noticed in data that you have questions about
* not graded on quality of presentation
* opportunity to get teachers to get a peek at what we’re trying to do so we can get feedback
* initially – do I understand the data (can I get enough questions on the slide to help understand questions)
* may includes some descriptives, tables or graphs
* final (third slide) should say what

My rough plan for now: 3 slides + title slide

Slide 1: Overview – describing what they want to answer and make sure you understand it

Slide 2: Things you’ve noticed in the data any far and questions

Slide 3: My plan for analysis

Notes from last project:

How do you know you’ve ruled out clinically meaningful? Have investigator look at confidence intervals

If any clinically meaningful values are outside your CI then you’ve ruled them out and you don’t have an underpowered subject

If there is some meaningful range included in the CI, then

The overall treatment effect partial F-test will be the same regardless of the reference group you choose. The partial F test looks at all possible treatment group comparisons to see if there is a significant overall effect of treatment. Changing the reference group changes the subset of pairwise comparisons that you are doing. Only look at pairwise comparisons if get a significant partial F test (overall effect of treatment). Once you have that then you need to think about what types of pairwise comparisons you’d like to do.

Potentially joint modeling (modeling more than one outcome at once):

* Really have to think through assumptions
* Covariance structures can get more complex
* May not be more beneficial
* DO NOT NEED TO CONSIDER THIS TYPE OF MODELLING IN 6623

Comments when talking with Camille and asking questions:

Q: Are there any covariates of particular interest to you or that you think are especially important?

A: It is a cohort study so it is important to control for general demographic stuff like age, BMI, and race/ethnicity

Think that there might be relationships with ethnicity and drug use.

Think we will also need to control socioeconomic status in some way (income or education) because it may affect how people were able to get access to their medication.

Comment on adherence: People lie about how often they took their medicine, but we can’t track them so we have to go with what we have. Want to use it because it is a cohort study. Usually if people take 95% or more of there medication they have a good treatment response (my note: consider breaking this into a categorical variable)

Email Camille to ask her to look up list of covariates controlled for and typically include in similar studies

Make coding for impossible values and missing values all be missing (NAs for R)

Q: Do you definitely want to look at all 4 outcome variables?

Yes!