**Wenru Zhou**

**Data Analysis section**

We will use linear regression to conduct the analysis. The assumption of outcome is normal distribution. If the outcome is skewed, appropriate transformation should be performed. Covariates that will be adjusted in model are age, sex, severity level (CDR of 0 and 0.5), present or absent of E4 allele, BMI, history of hypercholesterolemia, NSAID use; immune-related health conditions. To correct the inflated type I error caused by multiple comparison of six predictor, we plan to use Bonferroni adjustment. The alpha level is set as 0.008. Power is set as 0.8. We used best subset method to perform model selection: we will choose the models with highest R-square and lowest Cp and with main predictors in them. Model with less covariates would be preferred. Then interactions will be tested with the remaining covariates.

Aim1:

Hypothesis a: We are going to fit 12 sets of models: there are six kinds of models corresponding to six markers of cytokine and chemokine predictors: IL-6, TNF-alpha, MCP-1, Eotaxin-1, Beta-2 microglobulin, ACT at baseline. For each model, the outcomes are the change in cortical thickness and the change in memory over one year. Additional covariates that should be adjusted are Baseline cortical thickness and baseline memory separately corresponding to change in cortical thickness as the outcome and change in memory as the outcome.

Hypothesis b: All are the same as hypothesis a expect that we will use change in markers as main predictors instead of markers at baseline.

Aim 2:

Hypothesis a: We are going to fit 12 sets of models: there are six kinds of models corresponding to six markers of cytokine and chemokine predictors: IL-6, TNF-alpha, MCP-1, Eotaxin-1, Beta-2 microglobulin, ACT at baseline. For each model, the outcomes are the amyloid deposition and cortical thickness at baseline. Confounders are age and sex.

Hypothesis b: We are going to fit 12 sets of models: there are six kinds of models corresponding to six markers of cytokine and chemokine predictors: IL-6, TNF-alpha, MCP-1, Eotaxin-1, Beta-2 microglobulin, ACT at baseline. And for each model we will add amyloid and cortical thickness at baseline separately. Interaction will be considered. The outcome for each model is change in memory.

**Sample size calculation:**

Given the fact that the correlation between all outcomes and main predictors are from 0.2 to 0.6. We are going to use correlation=0.2 to estimate the maximum sample size. The initial sample size=298. Considering about 10% loss follow up. The final sample size should be 332 to meet all our need. We cannot conclude if this calculation is conservative or not because we don’t know how many predictors precision predictors are, and how many predictors confounders are.

The markers should be dichotomized into two groups: High and low. The detectable difference of slope in interaction between each marker and amyloid, or each marker and cortical thickness is 0.407. Assuming the standard deviation of two groups and outcome are the same, the detectable correlation is also 0.407.

Sample size was computed in G\*Power 3.1