# Project 4 Final Report

Rachel Weber
December 11,2018

G*Power version 3.1.92 was used for power and sample size calculations. Minimum sample size for recruitment of aMCI patients and healthy controls combined was 300 which provided a power of .8 and detectable difference in correlation of .2. Patients are to be recruited in equal proportions and with 10% extra to accommodate some patients loss to follow-up. Target recruitment is set to 330 patients. Models proposed will be run separately for each biomarker (IL-6, TNF-alpha, MCP-1, Eotaxin-1, Beta-2 microglobulin, and ACT). Alpha levels for all tests are fixed at 0.008 by the Bonferroni adjustment dividing $\alpha = 0.05$ by 6.

Analysis using multivariable regression will determine the contribution of demographic and biological measures to the outcomes of interest for Aims 1 and 2. Regression will predict each outcome using age, sex, APOE, NSAID, BMI, baseline outcome values, and the primary predictor of interest. GBM methods will determine relative information provided by each variable. The top variables whose information sums to .75 or higher will be included as confounding variables in the final model. Regardless of how much information the predictor of interest provides, it will be included in the final model for the purpose of specifically addressing the stated hypothesis.

Ancillary analysis will use all biomarkers as predictors of memory decline and cortical thickness in a single model for which PCA will determine which markers contribute the most to each outcome. This will inform us of potential targeted studies in the future and the amount of shared information the markers present.

Aim 1: For hypothesis (a), multivariable regression will determine the association between memory decline or cortical thickness and baseline inflammatory markers. Memory decline and cortical thickness

are assumed to be distributed normally. Confounding variables found to be significant predictors through GBM will be used in each of the 6 final models. This allows for variance in confounder affects between inflammatory markers. For hypothesis (b), regression will predict the same outcomes but with the changes in cytokine or chemokine levels between baseline and 1 year as predictors in place of baseline levels. For both hypotheses, overall F tests will determine model significance after variable selection from GBM. Model and covariate significance will be determined by the 0.008 threshold.

Aim 2: For hypothesis (a), the primary predictor for amyloid deposition will be change in inflammatory markers over 1 year. Regression like that in Aim 1 will be used to appropriately adjust for significant confounding variables. An overall F-test will determine model significance after variable selection. This will confirm or refute the notion that amyloid deposition is associated with inflammatory markers.

For hypothesis (b), disease progression is defined as either memory decline or change in cortical thickness between baseline and year 1. This study is powered to detect an effect size of .41 (Cohen's q) in a model with the interaction term—meaning we can identify a difference in correlation of .39. The interaction between amyloid deposition and biomarker levels along with all previous confounding variables will be used to predict disease progression. This analysis will necessitate the dichotomization of the biomarker variable into 'low' and 'high'. It is assumed that the $\sigma_{low}$ and $\sigma_{high}$ are equal. This enables change in correlation to equate to change in slope. If the interaction term is significant at the 0.008 level, it will be included in the final models predicting memory decline and change in cortical thickness. Significance of the interaction will be assessed at the individual cytokine/chemokine level. GBM methods will again be used to select variables contributing 75% of the relative information to the model.