Research Use Statement for Application for Genomic Data from NIAGADS Please limit to 2,200 characters max. The statement should include the following components:

- Objectives of the proposed research;
- Study design;
- Analysis plan, including the phenotypic characteristics that will be evaluated in association with genetic variants

Research Use Statement:

Alzheimer's Disease (AD) has a complex genetic component. While important loci have been discovered using genome-wide association studies (GWAS) for AD under the assumption of linear additive allele effects on gene, much AD genetic contribution remains unknown. These efforts in AD research have discovered genetically heritable traits associated with AD such as APOE, the greatest known genetic indicator of AD development to date. This assumption is not always suitable, however, and can cause important traits to be left undiscovered. To address this, we aim to identify novel genetic factors of AD development through biologically-informed, non-additive genetic models in our GWAS.

Study Design:
Previously, we assembled 18 GWAS datasets (ADNI, GenADA, and the 16 of the GWAS sets requested in this present application), a genome wide association study (GWAS) was performed in a case-control setting, analyzing each dataset independently and combining the results by meta-analysis. Currently, we intend to improve on this by adding 5 new GWAS datasets added in this update, TARCC, WashU2, ROSMAP2, WHICAP, and MTC GWAS. After adding these datasets to our meta-analyses, novel results that demonstrate the effects of rare (minor allele frequency < 5%) variants in increasing risk of developing AD or protecting those with increased genetic risk will be annotated to identify SNP locations through databases such as 1,000 Genomes. These results will be further validated through testing association with CSF markers (beta-amyloid, phospho-tau and total-tau) in the ADNI data set and through ROSMAP proteomic data.

Analysis Plan:
For each dataset, we will apply a stringent quality-control pipeline to the data before and after the imputation using the University of Michigan's imputation server. The association testing will be carried out by logistic regression (or linear regression using CSF biomarker levels as a quantitative trait), adjusting for age, sex, APOE4 status, and population structure using PLINK(v1.9). The analysis will be conducted using biologically-informed models of genetics in humans, as opposed to the usual methodology commonly observed in GWAS studies, such as the additive allele model or the dominant allele model. The results will be combined in a meta-analysis using the method implemented in the GWAMA software tool, after applying genomic control within each individual study. Finally, we wish to further validate any SNP findings with gene expression data, which will help discern the functional importance of our results.

Non-Technical Summary for Application for Genomic Data from NIAGADS
Investigators will provide a non-technical summary of their proposed research. If

the project is approved, this statement will be publicly available for lay audiences to read the purpose and objectives of the research. Please limit to 1,100 characters.

We aim to identify additional genetic factors contributing to the development of AD through the use of biologically-informed, non-additive genetic models in our GWAS to test for significant association. The use of novel models to understand and characterize AD may illuminate novel gene candidates that function via underexplored mechanisms that are closer in accuracy to the underlying molecular biology.