

Research Use Statement for Application for Genomic Data from NIAGADS

We know now that genetic dissection of complex diseases (like Alzheimer's disease) require very large sample sizes. This in turn requires the careful analysis of as many samples as possible. We have requested access to as many NIAGADS Alzheimer's datasets as possible. We also have access to 11 Alzheimer-related datasets from dbGaP. Furthermore, smaller samples from other sources will be included (e.g., Norway and Sweden). By combining all these datasets on the individual level, we aim to identify novel genetic risk factors for Alzheimer's disease. The main phenotypic focus of these analyses is Alzheimer's disease. However, where applicable we additionally focus on mild cognitive impairment, dementia and endophenotypes, such as CSF biomarkers and MRI imaging.

We will combine results from NIAGADS Alzheimer's datasets with those from other repositories. This is now a standard approach in human genetics and is done routinely for many disorders. We are unaware of how this might practically/realistically increase risk to participants. The uses proposed here are consistent with the Ethical Committee permissions and Informed Consents from the original studies.

The proposed uses (non-profit, focus on Alzheimer's disease) are consistent with the data use limitations for all of the requested datasets. We have obtained IRB approvals as required.

Non-Technical Summary for Application for Genomic Data from NIAGADS

Alzheimer's disease is a major public health problem worldwide. Considerable genomic analysis has been done, but the data have not been optimally analyzed. The extant consortia used a meta-analysis approach that entailed site-based quality control, imputation, and analysis. Experience has shown that this is less powerful and centralized analysis of all of these datasets would be the next step. We thus propose to apply the Psychiatric Genomic Consortium's centralized "mega-analysis" pipeline to all available individual genetic and phenotypes datasets that studied Alzheimer's disease.

The key scientific question is whether our "mega-analysis" procedures can improve upon the "meta-analysis" that was done by the original investigators.

The data will be processed through an instantiation of the "ricopili" GWAS analysis pipeline. This 50,000 line program processes all of the data consistently, and includes careful quality control, imputation, evaluation/control of biases like ancestry effects, and statistical analysis.