To submit data, please fill out the Data Registration Template. A member from NIAGADS will work with you on data transfer. If you have any questions regarding the template forms or uploading issue, please contact [NIAGADS@pennmedicine.upenn.edu](mailto:NIAGADS@pennmedicine.upenn.edu).

NOTE: All documents related to the application should be provided in English. For institutions where English is not the primary language, please provide translations of documents along with the original document. Translated documents should be signed by the institutional signing official.

**Required Policy Documents**

The following documents are required in order to deposit and share your data through NIAGADS:

1. [Institutional Certification for ADRD Studies](http://dss.niagads.org/wp-content/uploads/docs/Institutional-Certification-for-Alzheimers-Disease-and-Related-Dementias-ADRD-Studies.docx) that covers all subjects in your study. Multiple certifications may be required.
2. Signed copy of the [NIA AD Genomics Sharing Plan](https://www.nia.nih.gov/sites/default/files/2017-06/revised-ADSP-sharing-plan-6-13-17.docx).
3. Dataset Registration Template completed: **01\_DSS\_Dataset\_Registration\_Template.docx**

# Subject and Sample ID Re-mapping

All sequenced subjects and any non-sequenced connecting family members submitted to be harmonized with ADSP data will be renamed to fit the ADSP UID schema. A separate protocol is provided within this packet; please see the document named **02\_ADSPID\_Assignment\_Instructions.docx** for further instructions. NIAGADS will remap all data you submit with the ADSP UID.

# Phenotypes

Phenotypes should be provided according to the ADSP format. Use the data dictionary, **03\_ADSP\_Phenotypes\_Augmentation\_DD.docx**, as a reference to reformat all phenotypes to the ADSP format. A template is provided for use with the data dictionary above; **03\_ADSP\_Phenotypes\_Augmentation\_DS.xlsx**.

# GWAS

If available, please provide GWAS or exome chip data for all sequenced subjects and any connecting family members. Genotype data should be in PLINK binary format, preferably on build hg38, and formatted to the forward strand. This will help with calling structural variants. Please make sure that we are able to map the IDs in the GWAS to the IDs used in the sequencing.

# Sequencing Data

## **CRAM, BAM or FASTQ**

Sequencing read data can be submitted in any of formats:

* + - 1. **FASTQ**: please save all reads, including those that could not be mapped to the reference genome.
      2. **BAM**: please save all reads, including those that could not be mapped to the reference genome.
      3. **CRAM**: please save all reads, including those that could not be mapped to the reference genome.

No matter what your input data file format is, please include information about how the sequencing was performed (see “Sequencing Protocol/Pipeline, Kit Info, WES target regions” section below for more details). Also please send us the list of samples that will be transferred (sample manifest), as well as sequencing quality control metrics that have been collected (see “Sequencing Quality Control Metrics” section below for more details).

## **Seq Protocol/Pipeline, Kit Info, WES target regions**

Provide any relevant sequencing information, including the following:

* Sequencing Center
* Sequencer Machine
* Read Length
* PCR Free or PCR Amplified?
* Kit Name/Version
* Copy of the WES target regions if applicable
* Sequencing Quality Control Metrics

NOTE: Please provide md5 checksum for every submitted data file to ensure submission completion.