

1. Why should I share my Alzheimer's Disease (AD) research data?

The National Institute on Aging (NIA) advocates making available to the public the results and accomplishments of the activities that it funds. NIA assures that research resources developed with public funds become readily available to the broader research community in a timely manner for further research, development, application, and secondary data analysis in the expectation that this will lead to products and knowledge of benefit to the public health. Sharing research data is essential for expedited translation of Alzheimer's disease (AD) research results into new therapeutic approaches, products and procedures.

2. What is NIA's AD Genetics Sharing Policy?

NIA is committed to facilitating the broad sharing of NIH funded research resources and has had a sharing policy for AD genetics in place since 2003. The policy applies to the sharing of biological samples, phenotypic data, and genotypic data for NIA funded studies. The full text is available at:

<http://www.nia.nih.gov/research/dn/alzheimers-disease-genetics-sharing-plan>

3. What changes in the field of genetics have occurred over the last decade that impact upon the study of AD genetics?

Since the NIA AD Genetics Sharing Policy was implemented in 2003, the field of genetics has experienced monumental advances. For example, development of high throughput and massively parallel sequencing and dramatic reductions in sequencing costs have resulted in historic increases in the amount and nature of available data. Next generation and deep sequencing approaches have provided the capacity to analyze the entire exome or even the entire genome. Sequencing approaches are being extensively applied to RNA transcripts (RNA-Seq) to analyze RNA abundance and identify splice variants. Methods to map chromatin modification and protein binding to the genome by chromatin immunoprecipitation-sequencing approaches (ChIP-Seq) have also evolved. Other methodologies under development are also likely to further advance the field in the future. These advances have already accelerated AD risk factor gene detection and will dramatically promote understanding the Alzheimer's phenotype. This in turn will improve the likelihood of rapid development of novel therapeutic approaches.

4. How have changes in the field of genetics impacted upon the kind of Alzheimer's- related data that can be shared under the Sharing Policy?

New analytical approaches have recently become available to the research community to help decipher the underlying biology of AD. Advances in sequencing capability along with improved ability to analyze the data allow a more integrated analytical approach to discern

the underlying genetic architecture of AD. Broad sharing of diverse, but related, data provides a unique opportunity to discover the underlying causes of AD and to discover new therapeutic approaches.

5. What are some examples of the types of data related to genetics that may be shared under the Sharing Policy in order to capture the full spectrum of genetic and genomic variation in Alzheimer's disease?

There are a myriad of small and large-scale differences in the human genome. Complete characterization of the genetics of AD will require identification of the full spectrum of genomic variation in large and diverse sample sets. A partial list of the types of data that NIA includes in the NIA AD Genetics Sharing Policy is:

- Single nucleotide variants (SNPs). Platforms include polymerase-chain reaction (PCR)-based assays, SNP genotyping arrays including Genome Wide Association Studies (GWAS) and exome chip analysis, or high-throughput sequencing technologies;
- Whole exome sequences;
- Whole genome sequences;
- Sequences of targeted genomic regions;
- Structural variants such as chromosomal rearrangements, insertions and deletions; copy number variants; structural variation within functional genetic elements; structural variation in non-coding regions of the genome;
- Expression data including genome-wide expression analysis;
- Epigenetic data such as DNA methylation, histone modification, and chromatin conformation changes;
- Transcriptome data (e.g. data on gene products such as RNAs);
- Proteomic data (e.g. gene products such as proteins);
- Metabolomic data (e.g. indirect products of genes, such as metabolites);
- Sequence changes in mitochondrial DNA;
- Biomarker data;
- Neuroimaging data that are associated with genetics findings.

6. What is NIAGADS?

NIAGADS is the **NIA Genetics of Alzheimer's Disease Data Storage Site**. NIAGADS is a national data repository which is designed to facilitate access by qualified investigators to data for the study of various aspects of the AD phenotype. NIAGADS serves the AD research community at large as a site to enhance the capacity to find new therapeutic approaches and drug therapies for AD. Data generated from NIA funded genetics studies

are stored at NIAGADS, and include a broad range of data from genome wide association studies, next generation sequencing, targeted genome sequencing, and related primary and secondary data. NIAGADS also provides links to existing data sets, such as those for the Alzheimer's Disease Neuroimaging Initiative (<http://www.adni-info.org/>), the Alzheimer's Disease Cooperative Study (<http://www.adcs.org/>) and a variety of epidemiological cohorts. NIAGADS is continually updated and provides an opportunity for AD researchers to analyze multiple layers of AD-related data in a single work-space. It is anticipated that in the future NIAGADS will become a nexus for AD researchers to discern the architecture of the disease. The URL to the NIAGADS web site is <http://www.niagads.org/>.

7. Where are AD Genetics and related data stored?

Depending upon the type of data, there are at least two sites where AD genetics and related data can be shared. NIA established the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS) (<http://www.niageneticsdata.org/>), at the University of Pennsylvania. To facilitate broad and consistent access to NIH-supported GWAS datasets, the NIH has developed a central NIH GWAS data repository, currently named the NIH database of Genotypes and Phenotypes (dbGaP) <http://www.ncbi.nlm.nih.gov/gap>. NIA's policy <http://www.nia.nih.gov/research/dn/alzheimers-disease-genetics-sharing-plan> is that all AD genetics data, including secondary analysis data, derived from NIA funded studies be deposited in NIAGADS (or another NIA approved site) and dbGaP or both.

8. How have changes in the field of bioinformatics impacted upon the kind of Alzheimer's- related data that can be shared at NIAGADS under the Sharing Policy?

Resources such as the National Center for Biotechnology Information (NCBI), the University of California, Santa Cruz (UCSC) Genome Browser website (<http://genome.ucsc.edu/>); Ensemble (<http://uswest.ensembl.org/index.html>); and the KEGG Pathway data base (<http://www.genome.jp/kegg/pathway.html>) are now commonly used by scientists in a variety of fields to help analyze their data. Recent advances in information technology make it possible to simultaneously examine sequence, expression, and epigenomic data at the level of a single gene in a common workspace such as NIAGADS. Combining the power of existing research resources and the wealth of data now available to AD researchers shared under the Sharing Policy will support the Alzheimer's community to help meet the challenge of rapid discovery of novel therapeutic approaches to the disease.

9. What does NIAGADS do?

NIAGADS provides a flexible web-based data entry system using standardized common data elements for AD research studies that can be accessed by qualified investigators for a variety of basic science and clinical research studies. NIAGADS is able to accept and link to associated genetic and phenotypic data that are available in other AD-related databases. NIAGADS also provides a large database of publicly available sequence and annotation data along with an integrated tool set for examining and comparing the genomes of affected and unaffected individuals, aligning sequence to genomes, and displaying and sharing users' annotated data. AD sequence, expression, epigenomic, and related data are available at the level of specific genes. The database archives, processes and distributes genetic data, and publicly displays results. This unique research resource interfaces with existing NIA funded research resources such as the National Cell Repository for Alzheimer's Disease (<http://ncrad.iu.edu/>) and the National Alzheimer's Coordinating Center (<http://www.alz.washington.edu/>). The latest information technology optimizes the accessibility and usefulness of the available information. NIAGADS is a key facet of the NIA's policy to effectively leverage investments already made in AD research.

10. I did not request support for sharing data in my application, which was funded by NIA. Can I charge requestors for the costs associated with sharing the data?

In many instances, AD related data generated by NIA funded studies can be deposited at NIAGADS. Data that are deposited at NIAGADS are shared cost free. In instances where data do not appear to be appropriate for NIAGADS, investigators should work with their Program Officers to find the best way to make your data widely accessible. A list of NIA's Division of Neuroscience Program Officers is available at <http://www.nia.nih.gov/about/offices/division-neuroscience-dn/#staff>.

11. How do I use NIAGADS?

In the spring of 2012, NIAGADS <http://www.niagads.org/> underwent an extensive upgrade. It will now support the following new features.

- An enhanced web-based user interface to browse and search for datasets deposited at NIAGADS, and links to download datasets (some datasets are subject to approval);

- A web-based integrated genomics database with a genome browser for genome-wide association study results and interlinks with gene/pathway annotation;
- Computer programs, guidelines, and workflows maintained by NIAGADS to facilitate the analysis of genetic data. These resources can either be downloaded and installed on the user's own computers, or can be invoked on Amazon Elastic Compute Cloud (EC2). Please note that users need to cover expenses to use their own local computing resource or Amazon (EC2);
- An online user forum for questions and answers, discussion, and feature requests.

The AD research community is encouraged to apply for a free user account, which is required for restricted areas such as individual-level genotype data by accessing the NIAGADS website and applying online. Information is provided on how to apply for data, how applications are reviewed, and the types of data that are available

12. What is the Alzheimer's Disease Sequencing Project (ADSP)?

On February 7, 2012, a new Presidential Initiative was announced to fight Alzheimer's Disease (AD). As part of this effort, the National Human Genome Research Institute (NHGRI) was asked by the Director of the National Institutes of Health (NIH) to use \$25M already committed to its Large-Scale Sequencing Program for genomic studies in AD. The NIH director asked the National Institute on Aging (NIA) and the NHGRI to work together to develop and execute a Large Scale Sequencing Project (LSSP) to analyze the genomes of a large number of well characterized individuals in order to identify a broad range of AD risk and protective gene variants, with the ultimate goal of facilitating the identification of new pathways for therapeutic approaches and prevention. The analysis will also provide insight as to why individuals with known risk factor genes escape from developing AD. The project developed jointly by NIA and NHGRI is called the Alzheimer's Disease Sequencing Project (ADSP). The ADSP will conduct and facilitate analysis of whole exome and whole genome sequencing data to extend previous discoveries that may ultimately result in new directions for AD therapeutics.

13. How will data from the Alzheimer's Disease Sequencing Project (ADSP) be shared?

The Bermuda Principles (<http://www.gene.ucl.ac.uk/hugo/bermuda.htm>) in 1996 and the Ft. Lauderdale Large Scale Biological Sequencing Projects accord in 2003 (<http://www.genome.gov/Pages/Research/WellcomeReport0303.pdf>) were developed by the scientists engaged in the International Human Genome Sequencing Consortium and their funding agencies. These documents observe that pre-publication data release might conflict with a fundamental scientific incentive: publishing the first analysis of one's own

data. It is not possible to absolutely guarantee this incentive without applying restrictions that would undermine the rationale for rapid, unrestricted release of data from community resources. Nonetheless, it is essential that excellent scientists continue to be attracted to these projects. To encourage this, the scientific community should understand that pre-publication data release needs active community-wide support if it is to continue to receive widespread support from the producers. The ability of the producers to analyze and publish their own data should be respected by the research community and the contributions and interests of the data producers should be recognized and respected by the users of the data. As an extension of the Bermuda Principles and the Ft. Lauderdale Accord, the following obtains with regard to ADSP data:

In a Memorandum of Understanding signed by the members of the ADSP, it was agreed that ADSP sequence and phenotypic data would be made available rapidly after generation and that all partners in the ADSP would have immediate access to sequence data through an NIH approved data base. In keeping with the Bermuda Principles, the Ft. Lauderdale accord, and the ADSP MOU, **data generated by the ADSP will be made available to the research community at large immediately after quality control checks and variant calls are completed.** Data can be accessed by application either through dbGaP: <http://www.ncbi.nlm.nih.gov/gap> or the NIA Genetics of Alzheimer's Disease Data Storage Site <http://www.niagads.org/>.

In the spirit of the clear benefit that ensues from converting such data sets into community resources as rapidly as possible, it is expected that **users of the data generated by the ADSP will withhold publication until the producers of the data have published their findings.** ADSP participants will publish their data in an expeditious fashion at least one major paper reporting the results of the ADSP to be jointly submitted by all of the members.

14. Will NIAGADS accept applications for genetic data from foreign institutions?

Yes, NIAGADS accepts applications from investigators in countries outside of the US. In keeping with NIH policy, NIAGADS requests additional information for applications from investigators at institutions outside the US.

- First, NIAGADS requests information to ensure applicants fit the NIH definition of a qualified investigator. Investigators must be permanent employees of their institution at a level equivalent to a tenure-track professor or senior scientist. Laboratory staff and trainees such as graduate students and postdoctoral fellows are not permitted to submit project requests. NIAGADS expects that the official

Institution Signing Official will verify the PI's status.

- NIAGADS requests an ethics board, institutional review board or institutional equivalent approval document.
- NIAGADS requests a signature on the application of an Information Technology (IT) Director, who is an institutional staff member with the authority to vouch for the institution's IT capacities.
- NIAGADS requests a data security plan that states that all of the data are stored at the university under security policies and procedures. This includes the necessary assurances that data will be encrypted and not accessible to other users.