Data Release Date: October 30, 2018, **Dataset Version:** 2018.09.17

This release includes the ADSP quality control checked GATK joint called VCF containing all 4789 whole genomes released as part of 2018.07.30 (described below). Available file types include project level VCFs and quality control companion files.

File Manifest: https://dss.niagads.org/dss_file-manifest_adspumbrella.xlsx

Phenotype Updates:
Documentation about phenotypes released with the ADSP data mapped to hg37 in dbGaP are available on the dbGaP website in the ADSP entry Release Notes for phs000572.v7. Several updates have been made to the phenotypes since the last release in dbGaP, April 2017, and are described below:

1. Twelve sex mismatches found in the Discovery WES dataset; 8 samples updated, 4 dropped from study (this wes dataset has not been released through NIAGADS yet, it is currently only available through dbGaP, phs000572.v7):
   - C-ASPS-30075-BL-ASPS-801050- updated to male
   - C-ASPS-50003-BL-ASPS-199000- updated to male
   - C-ASPS-51466-BL-ASPS-36800- updated to male
   - C-RS-40001-BL-ERA-5663001- updated to male
   - C-ASPS-51642-BL-ASPS-63800- updated to female
   - A-RAS-RA000011-BL-UPN-27627- updated to female
   - C-ASPS-52021-BL-ASPS-127401- updated to female
   - C-ASPS-51379-BL-ASPS-38900- updated to female
   - A-LOAD-LD012112-BR-NCR-10AD24166- incorrect sex, sample dropped
   - C-RS-30149-BL-ERA-8643003- incorrect sex, sample dropped
   - C-RS-50723-BL-ERA-4525001- incorrect sex, sample dropped
   - A-MAP-MA000771-BR-RUS-003- incorrect sex, sample dropped

2. Several pedigree structural inconsistencies were found:
   - A-CUHS-CU000723 (Family CU0015F) is unrelated to the rest of the pedigree and was removed.
   - A-CUHS-CU000970 (Family CU0022F) was found to be unrelated to A-CUHS-CU000978 and a half-sib to A-CUHS-CU000971, A-CUHS-CU000973, and A-CUHS-CU000972. A dummy parent was created to replace A-CUHS-CU000978. Dummy parent ID = A-CUHS-CU009813.
   - A-CUHS-CU001246 (Family CU0029F) is a spouse control and does not add any additional information. This subject along with all members within the branch have been removed from the pedigree.
   - A-CUHS-CU003128 (Family CU0082F) was sequenced, but appears unrelated to the rest of the family. Sample has been dropped from the study.
   - Many subjects in the CU sample set were reported as unaffected, but should have been marked as unknown AD status. These have been corrected within the phenotype file.

- A-CUHS-CU001552 (Family CU0036F) had the incorrect parents reported. The correct parents are A-CUHS-CU001556 and A-CUHS-CU001558.
3. Two samples are dropped from the hg38 version of the data as the BAM file was found to be discordant with the GWAS data. The two samples are:
   - A-MIA-UM001976-BL-MIA-20010205
   - A-MIA-UM000315-BL-MIA-19961724

4. In the 2018.07.30 release, it was noted that 23 ADNI IDs contained a lowercase 's' instead of an uppercase 'S'. All ADNI subjects should have an uppercase 'S', however due to time constraints we decided to change the phenotypes to lowercase 's' to match the sequencing IDs. In addition, all ADNI subjects in the phenotype files were missing the leading "ADNI_" in the ID and we have updated the IDs to contain this. These changes have been updated in this release, October 30, 2018. The updated files are labeled '2018.07.30.v2' in the filename.

Data Release Date: July 30, 2018, **Dataset Version:** 2018.07.30

This release includes whole-genome sequencing (WGS) data from the ADSP and ADNI studies. Available file types include CRAMs, GATK-called gVCFs, sequencing metrics, phenotypes, and pedigree structures for family based subjects.

These data were processed using the Genome Center for Alzheimer's Disease (GCAD) pipeline, VCPA1.0. All samples were mapped to Genome Reference Consortium Human Build 38 (GRCh38) and variant called using GATK.

Sample Set	Accession	Number of Subjects	Number of Samples
ADSP Discovery	snd10000	n = 574	n = 580
ADSP Extension	snd10001	n = 3367	n = 3400
ADNI-WGS-1	snd10002	n = 809	n = 809

The subjects belong to the following consent levels as indicated by the submitting study IRBs:

Consent Level*	# Samples
DS-ADRDAGE-IRB-PUB	214
DS-ADRD-IRB-PUB	98
DS-ADRDMEM-IRB-PUB-NPU	20
DS-AGEADLT-IRB-PUB	173
DS-AGEADLT-IRB-PUB-NPU	77
DS-AGEBRMEM-IRB-PUB-GSO	7
DS-DEMND-IRB-PUB	186
DS-DEMND-IRB-PUB-NPU	91

DS-ND-IRB-PUB	61
DS-ND-IRB-PUB-MDS	4
DS-ND-IRB-PUB-NPU	64
DS-NEURO-IRB-PUB	173
DS-NEURO-IRB-PUB-NPU	1
GRU-IRB-PUB	3110
GRU-IRB-PUB-NPU	36
HMB-IRB-PUB	250
HMB-IRB-PUB-GSO	102
HMB-IRB-PUB-NPU	122

*Consent level definitions can be found on the Data Use Limitations page.