RESEARCHERS DATA DICTIONARY

Genetic Data (RDD-Gen)

Introduction

The *Researcher's Data Dictionary-Genetic Data* (RDD-Gen) is intended to be the primary resource for researchers interested in identifying UDS and/or MDS subjects for whom genetic data are available. The RDD-Gen describes variables that contain either genetic data (APOE genotype) or information about the availability of genetic data that can be obtained, by request, from the Alzheimer's Disease Genetics Consortium (ADGC), or the National Institute of Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS).

Note that updates to these data and inclusion of genetic data for additional UDS subjects are obtained per the discretion of ADGC and the individual Alzheimer's Disease Centers (ADCs).

ABBREVIATIONS

NACC	**National Alzheimer's Coordinating Center**
ADGC	**Alzheimer's Disease Genetics Consortium**
NCRAD	**National Cell Repository for Alzheimer's Disease**
NIAGADS	**National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site**
ADC	**NIA/NIH Alzheimer's Disease Center**

Definitions

- Variables with source type **ADGC** are coded exactly as they are provided by ADGC.

- **Derived variables** are developed by NACC. These variables provide information that is collected indirectly from data in the Uniform Data Set (UDS), Neuropathology (NP) Data Set, the individual ADCs, the National Cell Repository for Alzheimer's Disease (NCRAD) and ADGC — for example, **NACCAPOE** provides the APOE genotype for each subject, when available. This information is obtained from ADGC, the individual ADCs, NCRAD, and the NACC Neuropathology Data Set. This derived variable combines the four sources of information so that one variable captures all available data.

Requesting genotype and sequencing data available at ADGC or NIAGADS

Genetic data may be obtained from ADGC or NIAGADS; however, data that is marked as available through ADGC can only be requested by an ADRC. Data available in NIAGADS can be requested by any qualified investigator. Available data types include GWAS, exome chip genotyping, whole exome sequencing, and whole genome sequencing. Please contact these groups directly for access to the genotype and sequencing data.

Alzheimer's Disease Genetics Consortium . . . . http://www.adgenetics.org/
NIA Genetics of Alzheimer's Disease Data Storage Site . . . . https://www.niagads.org

Table of Variables

	Variable name	Short descriptor	Data type	Data source
1	**ADGCGWAS**	GWAS available from ADGC (y/n)	Numeric cross-sectional	ADGC
2	**ADGCEXOM**	ExomeChip available at ADGC (y/n)	Numeric cross sectional	ADGC
3	**ADGCRND**	ADGC data-selection round	Character cross-sectional	ADGC
4	**ADGCEXR**	ExomeChip genotyping round	Character cross-sectional	ADGC
5	**NGDSGWAS**	GWAS available at NIAGADS (y/n)	Numeric cross-sectional	ADGC
6	**NGDSEXOM**	ExomeChip available at NIAGADS (y/n)	Numeric cross-sectional	ADGC
7	**NGDSWGS**	Whole genome sequencing available at NIAGADS (y/n)	Numeric cross-sectional	ADGC
8	**NGDSWES**	Whole exome sequencing available at NIAGADS (y/n)	Numeric cross-sectional	ADGC
9	**NGDSGWAC**	NIAGADS GWAS accession number	Character cross-sectional	ADGC
10	**NGDSEXAC**	NIAGADS ExomeChip accession number	Character cross-sectional	ADGC
11	**NGDSWGAC**	NIAGADS whole genome sequencing accession number	Character cross-sectional	ADGC
12	**NGDSWEAC**	NIAGADS whole exome sequencing accession number	Character cross-sectional	ADGC
13	**NACCNCRD**	Samples are available from NCRAD (y/n)	Numeric cross-sectional	NCRAD
14	**NACCAPOE**	APOE genotype	Numeric cross-sectional	NACC derived
15	**NACCNE4S**	Number of APOE e4 alleles	Numeric cross-sectional	NACC derived

Variable Definitions

1	Variable name	**ADGCGWAS**
	Short descriptor	GWAS available from ADGC (y/n)
	Data type	Numeric cross-sectional
	Data source	ADGC
	Allowable codes	0 = No 1 = Yes
	Description / derivation	Indicator of whether GWAS data are available from ADGC. Note that genetic data at ADGC is only available for request by ADGC or ADRC investigators.
2	Variable name	**ADGCEXOM**
	Short descriptor	ExomeChip available at ADGC (y/n)
	Data type	Numeric cross-sectional
	Data source	ADGC
	Allowable Codes	0 = No 1 = Yes
	Description / derivation	Indicator of whether ExomeChip data are available from ADGC. Note that genetic data at ADGC is only available for request by ADGC or ADRC investigators.
3	Variable name	**ADGCRND**
	Short descriptor	ADGC data-selection round
	Data type	Character cross-sectional
	Data source	ADGC
	Allowable codes	ADC 1 = Round 1 ADC 2 = Round 2 ADC 3 = Round 3 ADC 4 = Round 4 ADC 5 = Round 5 ADC 6 = Round 6 ADC 7 = Round 7 ADC 8 = Round 8 ADC 9 = Round 9 ADC 10 = Round 10 AA = African American round 88 = Not applicable/no genotype data available 99 = Missing/ could not be determined
	Description / derivation	This variable indicates the GWAS round in which this participants's sample was analyzed. Note that a participant's sample may be run in multiple rounds.

4	Variable name	**ADGCEXR**
	Short descriptor	ExomeChip genotyping round
	Data type	Character cross-sectional
	Data source	ADGC
	Allowable codes	ADC 7 = Round 7
		ADC 8 = Round 8
		Exome1 = Exome round 1
		Exome2 = Exome round 2
		Exome3 = Exome round 3
	Allowable Codes	88 = Not applicable/no genotype data available
		99 = Missing/ could not be determined
	Description / derivation	This variable indicates the ExomeChip round in which this participants's sample was analyzed. Note that a participant's sample may be run in multiple rounds or on a chip that contained both exonic and non-exonic markers.
5	Variable name	**NGDSGWAS**
	Short descriptor	GWAS available at NIAGADS (y/n)
	Data type	Numeric cross-sectional
	Data source	ADGC
	Allowable Codes	0 = No
		1 = Yes
	Description / derivation	Indicator of whether GWAS data are available from NIAGADS.
6	Variable name	**NGDSEXOM**
	Short descriptor	ExomeChip available at NIAGADS (y/n)
	Data type	Numeric cross-sectional
	Data source	ADGC
	Allowable Codes	0 = No
		1 = Yes
	Description / derivation	Indicator of whether ExomeChip data are available from NIAGADS.
7	Variable name	**NGDSWGS**
	Short descriptor	Whole genome sequencing available at NIAGADS (y/n)
	Data type	Numeric cross-sectional
	Data source	ADGC
	Allowable Codes	0 = No
		1 = Yes
	Description / derivation	Indicator of whether whole genome sequencing data are available from NIAGADS.

8	Variable name	**NGDSWES**
	Short descriptor	Whole exome sequencing available at NIAGADS (y/n)
	Data type	Numeric cross-sectional
	Data source	ADGC
	Allowable Codes	0 = No
		1 = Yes
	Description / derivation	Indicator of whether whole exome sequencing data are available from NIAGADS.

9	Variable name	**NGDSGWAC**
	Short descriptor	NIAGADS GWAS accession number
	Data type	Character cross-sectional
	Data source	ADGC
	Allowable Codes	Prefix 'NG" followed by 5 numerals
		88 = Not applicable/no genotype data available
	Description / derivation	This variable indicates the accession number for this participant's GWAS data. Visit the following URL https://www.niagads.org/datasets/ and enter the accession number in the search bar.

10	Variable name	**NGDSEXAC**
	Short descriptor	NIAGADS ExomeChip accession number
	Data type	Character cross-sectional
	Data source	ADGC
	Allowable Codes	Prefix 'NG" followed by 5 numerals
		88 = Not applicable/no genotype data available
	Description / derivation	This variable indicates the accession number for this participant's ExomeChip data. Visit the following URL https://www.niagads.org/datasets/ and enter the accession number in the search bar.

11	Variable name	**NGDSWGAC**
	Short descriptor	NIAGADS whole genome sequencing accession number
	Data type	Character cross-sectional
	Data source	ADGC
	Allowable Codes	Prefix 'NG" followed by 5 numerals
		88 = Not applicable/no genotype data available
	Description / derivation	This variable indicates the accession number for this participant's whole genome sequencing data. Visit the following URL https://www.niagads.org/datasets/ and enter the accession number in the search bar.

12	Variable name	**NGDSWEAC**
	Short descriptor	NIAGADS whole exome sequencing accession number
	Data type	Character cross-sectional
	Data source	ADGC
	Allowable Codes	Prefix 'NG" followed by 5 numerals
		88 = Not applicable/no genotype data available
	Description / derivation	This variable indicates the accession number for this participant's whole exome sequencing data. Visit the following URL https://www.niagads.org/datasets/ and enter the accession number in the search bar.

13	Variable name	**NACCNCRD**
	Short descriptor	Samples are available from NCRAD (y/n)
	Data type	Numeric cross-sectional
	Data source	NCRAD
	Missing Codes	0 = No
		1 = Yes
	Description / derivation	Indicator of whether samples are available from NCRAD.

14	Variable name	**NACCAPOE**
	Short descriptor	APOE genotype
	Data type	Numeric cross-sectional
	Data source	NACC derived
	Allowable Codes	1 = e3,e3
		2 = e3,e4
		3 = e3,e2
		4 = e4,e4
		5 = e4,e2
		6 = e2,e2
		9 = Missing/ unknown/ not assessed
	Description / derivation	APOE genotype is run independently by the ADC and reported to NACC on the NACC Neuropathology Form. APOE genotype is also reported by ADGC and NCRAD. In the rare case that the ADC-reported genotype and the genotype reported by ADGC are not the same, the genotype is set to 9 = Missing for that subject.

15	Variable name	**NACCNE4S**
	Short descriptor	Number of APOE e4 alleles
	Data type	Numeric cross-sectional
	Data source	NACC derived
	Allowable Codes	0 = No e4 allele
		1 = 1 copy of e4 allele
		2 = 2 copies of e4 allele
		9 = Missing/ unknown/ not assessed
	Description / derivation	APOE genotype is run independently by the ADC and reported to NACC on the NACC Neuropathology Form. APOE genotype is also reported by ADGC and NCRAD. In the rare case that the ADC-reported genotype and the genotype reported by ADGC are not the same, the genotype is set to 9 = Missing for that subject.