Supplementary Table 2 – All curated and harmonized/derived data fields in ADVP

Column index	Column names	How the information was obtained	Description of the column
1	Pubmed ID	Extract	Pubmed PMID of the publication
2	Record Type	Infer	SNP-based or Gene-based association
3	SNP	Extract	rsID shown in the publication
4	Coordinates	Extract	Chromosome and base-pair informaton shown in publication (note can vary across genome builds)
5	Locus	Extract	Reported gene in the association record in the results table in the publication (not the eQTL or gene-based gene)
6	Reported gene	Extract	Gene name for gene-based test. NA if that's SNP-based test.
7	Interactions	Extract	Reported interaction results
8	Population	Infer	Choose from the following, derived from "Population" column, population mapped into these categories: African American, Arab, Asian, Carribbean Hispanic, Caucasian, Hispanic, Non-Hispanic Caucasian, Non-Hispanic White
9	Population (detailed)	Extract	Population description from the publication
10	Cohort	Infer	Cohorts derived from Cohort (detailed) column: if consortium name was available (e.g. ADGC, CHARGE, these will be used)
11	Cohort (detailed)	Extract	Cohort description from the publication
12	Sample size	Extract	Total sample size
13	Subset Analyzed	Extract	What samples were used for analyses. This is designed for distinguishing between same snps with different p-values in the same data table in a publication.
14	Phenotype	Infer	Derived from Phenotype (detailed) column. Choose from the following: AD, ADRD, Cognitive, Expression, Fluid biomarker, Imaging, Neuropathology, Non-ADRD, Other
15	Phenotype (detailed)	Extract	Defined as outcome of the regression analyses, shown is what is described in text
16	Association type	Infer	Choose from the following: eQTL, Disease risk, Endophenotype, AAO/Survival, Pleiotropy, Cross phenotype
17	RA1	Extract	Reported Allele 1 - the first allele reported in the association record in the publication, if any
18	RA2	Extract	Reported Allele 2 - the second allele reported in the association record in the publication, if any
19	AF	Extract	Reported allele frequency across all samples (for this associationr record) in the publication
20	P-value	Extract	p-value reported; show corrected p-value if available
21	Effect Size	Extract	Effect size type (OR, Beta, FDR etc) and value of the effect size
22	Confidence Interval	Extract	95% confidence interval of the effect size (if any)
23	Stage	Infer	Stage of the analysis reported in the publication, e.g., “Stage n” (n=1,2,3). If nothing is reported, choose from the followings: “Discovery”, “Validation”, “Meta-analysis", "Joint-analysis"
24	Model	Extract	Description of model: what kind of statistical model, and if the analyses were adjusted for anything

Column index	Column names	How the information was obtained	Description of the column
25	Imputation	Extract	How the data is imputed. Choose from the following: 1000G, HapMap, HRC
26	View in GenomicsDB	Compute / cross-reference	URL link for viewing the record in NIAGADS genomicsDB
27	Nearest gene	Compute	Distance from the SNP to the Locus (basepair information included)
28	Most severe consequence	Compute / cross-reference	Functional information (VEP provided by NIAGADS genomicDB)