

RESEARCH ARTICLE

The association of clinical phenotypes to known AD/FTD genetic risk loci and their inter-relationship

Qingqin S. Li^{1*}, Chao Tian², The 23andMe Research Team^{2¶}, David Hinds², Guy R. Seabrook³

1 Janssen Research & Development, LLC, Titusville, NJ, United States of America, **2** 23andMe, Inc., Mountain View, CA, United States of America, **3** Johnson & Johnson Innovation, South San Francisco, CA, United States of America

¶ Membership of the The 23andMe Research Team is listed in the Acknowledgments.

* qli2@its.jnj.com



OPEN ACCESS

Citation: Li QS, Tian C, The 23andMe Research Team, Hinds D, Seabrook GR (2020) The association of clinical phenotypes to known AD/FTD genetic risk loci and their inter-relationship. *PLoS ONE* 15(11): e0241552. <https://doi.org/10.1371/journal.pone.0241552>

Editor: Jordi Perez-Tur, Institut de Biomedicina de València-CSIC, SPAIN

Received: January 20, 2020

Accepted: October 18, 2020

Published: November 5, 2020

Copyright: © 2020 Li et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The full set of PheWAS results for cohort #1 is available within the manuscript and/or [Supporting Information](#) files. Individual level data cannot be shared publicly because of limitation in informed consent and ethical and legal considerations applied to maintain patient privacy. Indirect access of data are still possible by working directly with 23andMe. Please see <https://research.23andme.com/dataset-access/> or <https://research.23andme.com/collaborate/> for details. Summary data for cohort #2 is available from Open Target Genetics (<https://genetics>).

Abstract

To elucidate how variants in genetic risk loci previously implicated in Alzheimer's Disease (AD) and/or frontotemporal dementia (FTD) contribute to expression of disease phenotypes, a phenome-wide association study was performed in two waves. In the first wave, we explored clinical traits associated with thirteen genetic variants previously reported to be linked to disease risk using both the 23andMe and UKB cohorts. We tested 30 additional AD variants in UKB cohort only in the second wave. *APOE* variants defining $\epsilon 2/\epsilon 3/\epsilon 4$ alleles and rs646776 were identified to be significantly associated with metabolic/cardiovascular and longevity traits. *APOE* variants were also significantly associated with neurological traits. *ABI3* variant rs28394864 was significantly associated with cardiovascular (e.g. hypertension, ischemic heart disease, coronary atherosclerosis, angina) and immune-related trait asthma. Both *APOE* variants and *CLU* variant were significantly associated with nearsightedness. *HLA-DRB1* variant was associated with diseases with immune-related traits. Additionally, variants from 10+ AD genes (*BZRAP1-AS1*, *ADAMTS4*, *ADAM10*, *APH1B*, *SCIMP*, *ABI3*, *SPPL2A*, *ZNF232*, *GRN*, *CD2AP*, and *CD33*) were associated with hematological measurements such as white blood cell (leukocyte) count, monocyte count, neutrophil count, platelet count, and/or mean platelet (thrombocyte) volume (an autoimmune disease biomarker). Many of these genes are expressed specifically in microglia. The associations of *ABI3* variant with cardiovascular and immune-related traits are one of the novel findings from this study. Taken together, it is evidenced that at least some AD and FTD variants are associated with multiple clinical phenotypes and not just dementia. These findings were discussed in the context of causal relationship versus pleiotropy via Mendelian randomization analysis.

Introduction

Genome Wide Association Study (GWAS) is a powerful approach in identifying genetic risk loci. However, the functional elucidation on how a risk locus is related to a disease still requires

opentargets.org/) while the individual level data is available from UKBB <https://www.ukbiobank.ac.uk/researchers/> according to its term "UK Biobank is an open access resource. The Resource is open to bona fide scientists, undertaking health-related research that is in the public good. Approved scientists from the UK and overseas and from academia, government, charity and commercial companies can apply to use the Resource. To obtain access to UKB individual level data, please read UK Biobank Access Policy and Procedures before registering via the UK Biobank Access Management System (AMS). For assistance one can send a Message via AMS after signing up. For queries pre-registration you can email access@ukbiobank.ac.uk.

Funding: The study was funded by Johnson & Johnson Innovation, 23andMe, and Janssen Research & Development, LLC. The funder Johnson & Johnson Innovation provided research funding to 23andMe. The funder Janssen Research & Development, LLC provided research funding to 23andMe and support in the form of salaries for authors QSL and GRS. The funder 23andMe provided support in the form of salaries for authors CT, DAH and members of the 23andMe Research Team. All funders did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing interests: QSL and GRS are employees of Janssen Research & Development, LLC and may own stock/stock options in the company. CT, DAH and members of the 23andMe Research Team are or have been employed by 23andMe, Inc. and may own stock/stock options in the company. Due to the user agreement (available at <https://www.23andme.com/about/tos/>) where it states "Genetic and/or Self-Reported Information. Disclosure of individual-level Genetic Information and/or Self-Reported Information to third parties for Research purposes will not occur without explicit consent.", the individual level cannot be made directly available to scientific community but have to be accessed indirectly via 23andMe. A file containing full summary association statistics is available in the Supplemental Material.

much functional genomics follow up. Understanding the phenotypic spectrum associated with genetic risk variants from the GWAS signals will shed light in the understanding of disease etiology of the main disease trait of interest (i.e. dementia in this study).

Dementia refers to conditions of memory loss and other cognitive decline serious enough to interfere with daily life. Alzheimer's disease (AD) is the most common cause of dementia and accounts for 60 to 80% of dementia cases [1]. Frontotemporal dementia (FTD) is the second most common cause of dementia [2] and represents a group of brain disorders caused by degeneration of the frontal and/or temporal lobes of the brain including behavior variant frontotemporal dementia (bvFTD), primary progressive aphasia (PPA), and semantic variant primary progressive aphasia (svPPA) [3]. The pathological hallmark of AD includes the deposition of β -amyloid ($A\beta$) aggregates in the form of senile plaques (SP) and abnormally phosphorylated tau in the form of neurofibrillary tangles (NFT) [4]. A subset of FTD patients, also referred to as with Pick's disease, have abnormal accumulation of protein inside nerve cells in the damaged areas of the brain. These Pick bodies contain an abnormal form of tau, but the structure and folding of tau filaments are different between Pick's disease and AD [5].

Both AD and FTD have familial and sporadic forms of the diseases. Amyloid beta precursor protein (*APP*) [6], presenilin 1 (*PSEN1*) [7], and presenilin 2 (*PSEN2*) [8, 9] are AD familial risk genes explaining < 0.5% of AD cases with age of onset typically between 30–50 years of age, while progranulin (*GRN*) [10] and microtubule-associated protein tau (*MAPT*) [11–13] are FTD familial risk genes [2]. Apolipoprotein E (*APOE*) was suggested to play a role in late onset Alzheimer's disease (LOAD), the sporadic form of AD, long before GWAS era [14]. GWAS confirmed role of variants in *APOE* but additionally identified clusterin (*CLU*), phosphatidylinositol binding clathrin assembly protein (*PICALM*) [15], bridging integrator 1 (*BIN1*) [16], CD2 associated protein (*CD2AP*) [17, 18], complement C3b/C4b receptor 1 (*CR1*) [19], ATP binding cassette subfamily A member 7 (*ABCA7*) [20], and CD33 molecule (*CD33*) [17, 21] to be associated with LOAD. Meta-analysis of 74,046 individuals by International Genomics of Alzheimer's Project (I-GAP) revealed 19 GWAS significant AD loci including 11 novel loci [21]. These genes were implicated in cholesterol metabolism (*APOE*, *CLU*, and *ABCA7*), immune response (*CR1*, *CD33*, *CLU*, *ABCA7*), and endocytosis (*BIN1*, *PICALM*, *CD2AP*) [22].

Studies have shown that loss-of-function mutations in *GRN*, which encodes a neurotrophic factor, cause familial FTL, a progressive neurodegenerative disease affecting ~10% of early-onset dementia patients. Rs646776 was shown to be associated with plasma progranulin levels ($p = 1.7 \times 10^{-30}$) and replicated in two independent series of 508 controls ($p = 1.9 \times 10^{-19}$) and 197 FTL patients ($p = 6.4 \times 10^{-12}$) [23] with each copy of the minor C allele decreasing progranulin levels by ~15%. In addition, common variant rs5848 in the 3' UTR region of the *GRN* gene conferred risk of FTD [24] and LOAD [25], and its genotype was associated with serum progranulin level [26] as well. Rs5848 is also an eQTL variant for *GRN* transcript level in nerve tibial (q-value = 3.19×10^{-11}) and artery tibial (q-value = 1.66×10^{-9}) (data source: GTEx V6p version). In addition to the well-known AD and FTD genes, we have previously identified two genome wide significant variants associated with CSF $A\beta_{42}$ levels [27] using Alzheimer's Disease Neuroimaging Initiative (ADNI) samples. In total, 13 variants implicated in AD or FTD, or associated with CSF $A\beta_{42}$ levels were included in the wave 1 of this study for further interrogation.

To elicit additional insight on these 13 SNPs previously implicated in AD and FTD, phenotype-wide association studies (PheWAS) were performed using the database from the personal genetics company 23andMe, Inc. to identify phenotypes (both dementia and non-dementia traits) associated with these variants of interest. UK Biobank PheWAS results were also looked up via Open Target Genetics [28]. Since then newer waves of GWAS meta-analysis

including samples from UK Biobank using family history as proxy and Alzheimer's Disease Sequencing Project (ADSP) allows identification of additional novel risk loci [29–32]. In wave 2, we further include 30 variants from the more recent GWAS meta-analyses [29–32] or variants identified earlier but not prioritized in wave 1 PheWAS. PheWAS approach has been previously applied to BioVU, Vanderbilt's DNA biobank where phenotypes are defined by EMR records namely ICD codes [33], or the 23andMe research database where phenotypes are defined by self-reports [34], or both [35]. It has the potential of validating target, nominating treatment indication and/or assessing safety signal especially if the effect of a genetic variant mimics the pharmacotherapy effect [36]. Given that the genes implicated in AD/FTD were implicated in cholesterol metabolism (*APOE*, *CLU*, and *ABCA7*) and immune response (*CRI*, *CD33*, *CLU*, *ABCA7*, *TREM2*, *SPPL2A*, *SCIMP*, *HLA-DRB1*), It is foreseeable that some of the AD variants may be also associated with metabolic/cardiovascular and/or immune-related traits. Recent development of using genetic variants as an instrument variable in GWAS summary statistics based Mendelian randomization (MR) [37] provides another means to dissect the pleiotropy vs. causal relationship between related traits.

Materials and methods

Study participants

Cohort 1: 23andMe. All individuals included in the analyses were research participants of 23andMe who have provided electronic informed consent, DNA samples for genetic testing, and answered surveys online. The study was conducted according to human subject protocol, which was reviewed and approved by Ethical & Independent Review Services, a private institutional review board (<http://www.eandireview.com>). It is also consistent with the procedures involving experiments on human subjects in accord with the ethical standards of the Committee on Human Experimentation of the institution in which the experiments were done or in accord with the Helsinki Declaration of 1975. All data was completely anonymized and de-identified before access by the analyst for data analysis.

As described previously [38], DNA samples have been genotyped on one of four genotyping platforms. The v1 and v2 platforms were variants of the Illumina HumanHap550+ BeadChip (Illumina, San Diego, CA, USA), including about 25 000 custom single-nucleotide polymorphisms (SNPs) selected by 23andMe, with a total of about 560 000 SNPs. The v3 platform was based on the Illumina OmniExpress+ BeadChip, with custom content to improve the overlap with the v2 array, with a total of about 950 000 SNPs. The v4 platform is a fully custom array, including a lower redundancy subset of v2 and v3 SNPs with additional coverage of lower-frequency-coding variation, and about 570 000 SNPs. [S1 Table](#) shows which 23andMe genotype platform (v1-v4) the tested variant is genotyped on. It also shows the imputation statistics for the tested variant, including the average imputation dosages for the first (A) and second (B) alleles (freq.a and freq.b) and the average and minimum imputation quality across all batches (avg.rsqr and min.rsqr). The r^2 statistic is used to measure imputation quality, which range from 0 (worst) to 1 (best). The batch effect test is an F test from an ANOVA of the SNP dosages against a factor representing imputation batch.

Only participants enrolled by 2015 were included in this analysis. A similar approach using the same research database was previously described [34]. We tested the association with more than 1000 well-curated phenotypes ([S2 Table](#)), which were distributed among different phenotypic categories (e.g. cognitive, autoimmune, psychiatric etc.). GWAS were previously performed on these well-curated phenotypes and confirmed to replicate known associations and not to generate spurious false positives. For our standard PheWAS, we restricted participants to a set of individuals who have > 97% European ancestry, as determined through an analysis

of local ancestry [39]. Briefly, this algorithm first partitioned phased genomic data into short windows of about 100 SNPs. Within each window, we used a support vector machine (SVM) to classify individual haplotypes into one of 31 reference populations. The SVM classifications were then fed into a hidden Markov model (HMM) accounting for switch errors and incorrect assignments as well as generating probabilities for each reference population in each window. Finally, we used simulated admixed individuals to recalibrate the HMM probabilities so that the reported assignments were consistent with the simulated admixture proportions. The reference population data is derived from public datasets (the Human Genome Diversity Project, HapMap, and 1000 Genomes—participants provided informed consent and all data was completely anonymized and de-identified before access and analysis), as well as 23andMe customers who have reported having four grandparents from the same country. A maximal set of unrelated individuals was chosen for each phenotype using a segmental identity-by-descent (IBD) estimation algorithm [40]. Individuals were defined as related if they shared more than 700 cM IBD, including regions where the two individuals shared either one or both genomic segments identical-by-descent. This level of relatedness (roughly 20% of the genome) corresponded approximately to the minimal expected sharing between first cousins in an outbred population.

The imputed dosages rather than best-guess genotypes were used for association testing in PheWAS. Participant genotype data were imputed against the September 2013 release of 1000 Genomes Phase1 reference haplotypes, phased with ShapeIt2 [41]. Genotype data for research participants were generated from four versions of genotyping chips as described previously [38]. We phased and imputed data for each genotyping platform separately. We phased using a phasing tool Finch, which implements the Beagle haplotype graph-based phasing algorithm [42], modified to separate the haplotype graph construction and phasing steps. Finch extended the Beagle model to accommodate genotyping error and recombination, to handle cases where there were no consistent paths through the haplotype graph for the individual being phased. We constructed haplotype graphs for European and non-European samples on each 23andMe genotyping platform from a representative sample of genotyped individuals, and then performed out-of-sample phasing of all genotyped individuals against the appropriate graph.

In preparation for imputation, we split phased chromosomes into segments of no more than 10,000 genotyped SNPs, with overlaps of 200 SNPs. We excluded SNPs with Hardy-Weinberg equilibrium $p < 10^{-20}$, call rate $< 95\%$, or with large allele frequency discrepancies compared to European 1000 Genomes reference data. Frequency discrepancies were identified by computing a 2x2 table of allele counts for European 1000 Genomes samples and 2000 randomly sampled 23andMe customers with European ancestry and identifying SNPs with a chi squared $p < 10^{-15}$. We imputed each phased segment against all-ethnicity 1000 Genomes haplotypes (excluding monomorphic and singleton sites) using Minimac2 [43], using 5 rounds and 200 states for parameter estimation.

Association test results were computed using logistic regression for case control comparisons, or linear regression for quantitative traits. For survival traits, association test results using Cox proportional hazards regression were computed. We assumed additive allelic effects and included covariates for age, gender, and the top five principal components to account for residual population structure. The association test p value reported was computed using a likelihood ratio test, which was shown to be a better choice despite of its computational demands [44]. We reported raw p -values for the PheWAS association results, but interpret the results taking into account the number of variants and traits tested. An association with $p < 0.05 / (13 * 1,234) = 3.12 \times 10^{-6}$ was deemed to be significant association, other associations with $FDR < 0.05$ was deemed to be suggestive associations.

Cohort 2: UK Biobank (UKB). Pre-computed UK Biobank PheWAS results based on Neale lab UK Biobank summary statistics were looked up via Open Target Genetics [28]

(genetics.opentargets.org) for side by side comparison with PheWAS results based on the 23andMe cohort. There were three version of UKB results accessed via Open Target Genetics, Neale v1 PheWAS results were accessed in November 2019, and Neale v2 PheWAS results (<http://www.nealelab.is/uk-biobank>) were accessed in July 2020. UKB SAIGE is yet a different version of UKB PheWAS results by University of Michigan (<http://pheweb.sph.umich.edu/SAIGE-UKB/about>). There are 4593 traits in total for the Neale v2 PheWAS analysis. We report raw p-values for the PheWAS association results, but interpret the results taking into account the number of variants and traits tested. An association with $p < 0.05 / (48 * 4593) = 2.27 \times 10^{-7}$ was deemed to be significant association. No additional adjustment was made for Neale v1 PheWAS results (a few traits present in v1 were not present in v2) and/or UKB SAIGE PheWAS.

Whole-genome genetic correlations between significant PheWAS traits

For convenience of collecting whole-genome summary statistics, AD summary statistics from Jansen et al study [30] was used to calculate genetic correlation with other traits using LD Hub (v1.9.3) [45], which is a centralized database of summary-level GWAS results and a web interface for LD score regression (LDSC) [46].

Directional horizontal pleiotropy vs causal relationship

For the multiple clinical phenotypes associated with the AD/FTD variants identified in the PheWAS analysis, we attempted to untangle the relationship between trait A and trait B to determine if genetic variants impact trait A (also called exposure in the literature) and trait B (also called outcome in the literature) independently, or genetic variants' effect on trait B is mediated by trait A (or vice versa). We applied MR Egger intercept test [47, 48] to test directional horizontal pleiotropy, where the variants affect both trait A (e.g. CAD) and trait B (e.g. AD) independently. MR uses genetic variants as a proxy for an environmental exposure/trait A (e.g. CAD), assuming that: 1) the genetic variants are associated with the exposure/trait A; 2) the genetic variants are independent of confounders in the exposure-outcome association; 3) the genetic variants are associated with the outcome only via their effect on the exposure, i.e. there is no horizontal pleiotropy whereby genetic variants have an effect on an outcome (e.g. AD) independent of its influence on the exposure (e.g. CAD). If the MR Egger intercept test had a significant p-value ($p < 0.05$) (i.e. violating assumption #3 from the MR analysis), the pair of traits was excluded from the bi-directional, two-sample MR test using inverse variance weighted (IVW) method among traits identified in the PheWAS study. In this case (Egger intercept $p < 0.05$), the gene-outcome vs gene-exposure regression coefficient is estimated using MR Egger regression to correct for the bias due to directional pleiotropy, under a weaker set of assumptions than typically used in MR [49]. Both IVW and MR Egger regression however do not protect against violation of assumption #2. The MR analysis is also only feasible if there is sufficient information from MR Base [50] for analysis or if the information could be supplemented by manually adding GWAS results from publications, e.g. the recent AD meta-analysis by Jansen et al. [30] In the MR analysis, we primarily leveraged variants implicated in a trait from public summary statistics (pre-compiled as a set of instruments from NHGRI-EBI GWAS Catalog [51] in the MRInstruments R package v0.3.2 <https://github.com/MRCIEU/MRInstruments>) as an individual variant is unlikely to be powerful enough as an instrument variable unless the effect size is large. Instrument variables were constructed using the default independent genome wide significant SNPs ($p < = 5 \times 10^{-8}$) for AD and other diseases/risk factors except for FTD where a p-value threshold of $p < = 6 \times 10^{-6}$ was used because of the smaller GWAS samples size [52]. We assessed if bi-directional causal relationships exist

between AD and a number of significant PheWAS traits identified in the PheWAS analyses. For FTD, only directional MR analysis was performed using FTD as the exposure as only top GWAS hits were available publicly. All analyses were performed using the MR-Base 'TwoSampleMR' v0.5.4 package [50] in R and MR test with nominal $p < 0.05$ using inverse variance weighted and/or MR Egger method was reported. A $p < 0.05/\#$ of PheWAS traits examined is considered significant, while a $p < 0.05$ is considered suggestive.

Results

Thirteen variants were successfully imputed from the four genotyping platforms in the 23andMe cohort with the average and minimum imputation quality across all batches (avg.rsqr and min.rsqr) ranging from 0.96 to 1 for avg.rsqr (average across 13 variants was 0.995) and 0.86 to 1 for min.rsqr (average = 0.982) (S1 Table).

AD-risk variants are highly associated with neurological, longevity, metabolic, cardiovascular, eye, and immune-related traits

Selected PheWAS findings were summarized in Tables 1 and 2 for wave 1 and wave 2 PheWAS alongside the known associations reported in the NHGRI-EBI GWAS Catalog [51] for the SNPs previously associated with LOAD and/or FTD. The list of PheWAS association results using the 23andMe cohort with $FDR < 0.05$ is available as S3 Table, while the full list of PheWAS association results is available from S4 Table. An association in the 23andMe cohort with $p < 0.05 / (13 * 1,234) = 3.12 \times 10^{-6}$ was deemed to be significant association, other associations with $FDR < 0.05$ was deemed to be suggestive associations. A number of the known associations was replicated. In addition, novel associations were identified. The two SNPs, rs429358 and rs7412, defining APOE $\epsilon 2/\epsilon 3/\epsilon 4$ alleles were known to be associated with multiple neurological, longevity, metabolic and cardiovascular traits (Fig 1, Table 1, and S3 Table). Subjects carrying the minor T allele of rs7412 are APOE $\epsilon 2$ protective allele carriers and subjects carrying the minor C allele of rs429358 are APOE $\epsilon 4$ risk allele carriers. PheWAS identified significant associations with metabolic traits (high cholesterol or taking drugs to lower cholesterol, body mass index (BMI)), neurological traits (AD family history, AD, cognitive decline, mild cognitive impairment, memory problems), longevity traits (nonagenarian—at least 90 years old, healthy old—over age 60 with no cancer or disease, centenarian family), cardiovascular diseases (coronary artery disease (CAD), metabolic and heart disease), and eye problems (nearsightedness, glasses usage, myopia vs. hyperopia), and serious side effects from statins (rs429358, $p = 7.14 \times 10^{-7}$). The directionality of association is consistent with the protective vs. risk effect of two APOE SNPs in that the minor allele of rs7412 was associated with lower risk of high cholesterol, while the minor allele of rs429358 was associated with higher risk of high cholesterol ($p = 6.6 \times 10^{-295}$). Additional suggestive associations were identified ($FDR < 0.05$) for age-related macular degeneration (AMD) or blindness (rs429358 $p = 9.59 \times 10^{-5}$, $FDR = 0.004$). Interestingly, rs11136000 from *CLU* is also strongly associated with multiple eye phenotypes (nearsightedness, myopia, glasses, astigmatism) (Table 1, Fig 2, and S3 Table). Subjects carrying the minor allele of rs429358 had lower chance of nearsightedness ($p = 1.4 \times 10^{-8}$), while subjects carrying the minor allele of rs11136000 had higher chance of nearsightedness ($p = 4.5 \times 10^{-15}$). For the overlapping phenotypes, UK Biobank PheWAS results largely supported the 23andMe PheWAS findings.

HLA-DRB1 variant rs6931277 associated with AD was also associated with diseases with an immune-related etiology such as ulcerative colitis (UC, $p = 2.23 \times 10^{-10}$), self-reported rheumatoid arthritis (RA, $p = 2.30 \times 10^{-130}$), celiac disease (CeD, $p = 5.99 \times 10^{-54}$), self-reported asthma ($p = 8.57 \times 10^{-68}$), self-reported multiple sclerosis ($p = 1.22 \times 10^{-13}$), Type 1 diabetes ($p = 1.55$

Table 1. (Continued)

SNP	Chr	BP	Freq ^{**}	Alleles ^{**}	Gene	Published associated with AD	Reported P-value	Analysis	DISEASE/TRAIT	PubMed ID	Phenotype Group	Selected 23andMe PheWAS Results				Selected UKBB PheWAS results								
												DISEASE/TRAIT	β^*	OR	P	N Cases	N Overall	DISEASE/TRAIT	β^*	OR	P	N Cases	N Overall	
rs744373	2	1,271,370,39	0.29	A/G	BINI	2.12E-16	IGAP stage 1	Alzheimer's disease (late onset)	24162737; 21390209; 21627779	29777097	Neurological	alzheimers	0.09	1.60E-01	645	157843								
rs9349407	6	47485642	0.74	C/G	CD2AP	3.92E-07	IGAP stage 1	Family history of Alzheimer's disease (late onset)	24162737; 21460841; 21460840	29777097	Neurological	iqb.alzheimers_fh	0.04	1.76E-04	25610	87378	Alzheimer's disease/dementia illnesses of mother	0.06	1.93E-10	28507	331041			
								Mean platelet volume	27863252	27863252	Autoimmune disease biomarker						Mean platelet (thrombocyte) volume	0.037	1.89E-37	350470				
								Platelet distribution width	27863252	27863252	Hematological measurement						Platelet distribution width	0.017	7.00E-37	350470				
								Platelet count	27863252	27863252	Hematological measurement						Platelet count	-1.352	2.27E-18	350474				
								High light scatter reticulocyte percentage of red cells	27863252	27863252	Hematological measurement						High light scatter reticulocyte percentage	-0.006	1.50E-10	344729				
											Anthropomorphic	height	-0.03	5.53E-06	355080	355080	Standing height	-0.131	3.50E-15	360388				
rs1136600	8	2760702	0.40	C/T	CLU	1.72E-16	IGAP stage 1	Alzheimer's disease (late onset)	24162737; 19734902	29777097	Neurological													
								Family history of Alzheimer's disease	24162737; 19734902	29777097	Neurological	iqb.alzheimers_fh	-0.03	2.31E-03	25610	87378	Alzheimer's disease/dementia illnesses of mother	-0.045	0.956	28507	331041			
											eyes	nearsightedness	0.05	4.51E-15	123722	102325	Which eye(s) affected by myopia (short sight): Left eye ^a	-0.065	1.02E-01	1377	26943			
											eyes	astigmatism	0.03	2.45E-06	77240	80566	Which eye(s) affected by astigmatism: Left eye ^a	0.058	1.66E-01	1450	8863			
rs11695792	10	93703269	0.03	G/T	FRA10AC1			Cerebrospinal fluid A β 1-42 levels	26252872	26252872	Neurological						Illnesses of siblings: Alzheimer's disease/dementia	0.201	4.43E-02	1468	25921			
rs3851179	11	86157598	0.37	C/T	PICALM	2.84E-15	IGAP stage 1	Alzheimer's disease (late onset)	24162737; 19734902	25778476	Neurological													
								Alzheimer's disease in APOE ϵ 4- carriers	24162737; 19734902	25778476	Neurological	iqb.alzheimers_fh	-0.03	1.18E-02	25610	87378	Alzheimer's disease/dementia illnesses of mother	-0.039	0.962	1.32E-05	28507	331041		
								Family history of Alzheimer's disease	29777097	29777097	Neurological						Alzheimer's disease/dementia illnesses of father	-0.048	0.954	6.69E-05	15022	312666		
											Neurological	migraine	-0.03	4.40E-05	74955	353214								
rs1503351	15	96814290	0.05	A/G	SPAT18-RN7SKP181			Cerebrospinal fluid A β 1-42 levels	26252872	26252872	Neurological						Alzheimer's disease/dementia illnesses of mother	0.046	1.047	3.59E-02	28507	331041		
rs764650	19	1046521	0.91	G/T	ABCA7	3.22E-07	IGAP stage 1	Alzheimer's disease (late onset)	24162737; 21460840	24162737; 21460840	Neurological						Alzheimer's disease/dementia illnesses of mother	-0.043	0.958	3.81E-03	28507	331041		
											Treatment/medication						Mean spheroid cell volume	-0.127	7.01E-09	344729				
rs429358	19	44908684	0.86	C/T	APOE			Brain imaging	20100581; 29860282	24162737	Neurological	alzheimers	-1.12	2.39E-56	645	157843	Alzheimer's disease ^b	-1.9	2.34E-58	404	402787			
								Alzheimer's disease (late onset)	24162737	24162737	Neurological						Delirium dementia and amnesic and other cognitive disorders ^b	-0.79	3.86E-60	1970	406353			

(Continued)

Table 1. (Continued)

SNP	Chr	BP	Freq ^{***}	Alleles ^{**}	Gene	Published associated with AD	Reported P-value	Analysis	Selected known associations ^{***} from Open Targets Genetics		Phenotype Group	Selected 23andMe PheWAS Results				Selected UKBB PheWAS results											
									DISEASE/TRAIT	PubMed ID		DISEASE/TRAIT	β^*	OR	P	N Cases	N Overall	DISEASE/TRAIT	β^*	OR	P	N Cases	N Overall				
									Cognitive decline	28078323	Neurological	cognitive_decline	-0.42			1.84E-35	3584	6928	Non-cancer illness code, self-reported: dementia/alzheimers/cognitive impairment ^a	-1.600			7.34E-14	83	337159		
									Alzheimer's disease progression score	29860282	Neurological								Alzheimer's disease/dementia illnesses of father	-0.529	0.589	1.53E-243	15022	312666			
									Family history of Alzheimer's disease	29777097	Neurological	iqb.alzheimers_fh	-0.50			3.16E-252	25610	87378	Alzheimer's disease/dementia illnesses of siblings	-0.591	0.554	3.36E-36	1609	279062			
									Cerebral amyloid deposition	26252872	Neurological	iqb.mild_cognitive_imp_fh	-0.27			4.39E-21	5497	54813	Alzheimer's disease/dementia illnesses of mother	-0.597	0.55	0.00E-400	28507	331041			
									Cerebrospinal fluid A β 1-42 levels	25027320	Neurological																
									Levy body disease	25188341, 29263008	Neurological																
									age-related macular degeneration	26691988	eyes																
									HDL cholesterol	25961943, 29403010	metabolic	Body mass index	0.09			1.95E-06	344351	344351	Body mass index (BMI)	0.112					3.54E-13		359983
									LDL cholesterol	29403010	metabolic	high_cholesterol	-0.36			6.64E-295	112221	219875	High cholesterol non-cancer illness code, self-reported	-0.261	0.77	3.66E-160	43957	361141			
									Total cholesterol levels	29403010	metabolic	iqb.low_hdl	-0.23			3.27E-80	36339	154349									
											metabolic/family illness								Diabetes illnesses of mother	0.093	1.098	4.96E-16	30772	331142			
									Coronary artery disease		cardiovascular	Coronary artery disease	-0.12			5.93E-11	13648	329205	Cholesterol lowering medication cholesterol, blood pressure, diabetes, or take exogenous hormones	-0.248	0.781	1.56E-81	24247	193148			
									Heart attack		cardiovascular	Heart attack	-0.15			1.71E-09	7573	334019	Ischemic heart disease ^b	-0.101	0.904	3.89E-16	31355	408458			
									physical activity	29899525	Physical								Leg fat percentage (left)	0.149		1.79E-18		354791			
									Parental lifespan	29030599	life-style	At least 90 years old	0.37			3.80E-37	5964	417602									
									Platelet count	27863252	lab measurement																
									Red cell distribution width	27863252	lab measurement																
									C-reactive protein levels	29403010	protein biomarker																
rs7412	19	44908822	0.08	C/T	APOE				Alzheimer's disease	28714976	Neurological	alzheimers	-0.43			1.91E-04	645	157843									

(Continued)

Table 1. (Continued)

SNP	Chr	BP	Freq ^{***}	Alleles ^{**}	Gene	Published associated with AD	Reported P-value	Analysis	Selected known associations ^{***} from Open Targets Genetics		Phenotype Group	Selected 23andMe PheWAS Results				Selected UKBB PheWAS results									
									DISEASE/TRAIT	PubMed ID		DISEASE/TRAIT	β^*	OR	P	N Cases	N Overall	DISEASE/TRAIT	β^*	OR	P	N Cases	N Overall		
									Family history of Alzheimer's disease	29777097	neurological/family illnesses	iqb.alzheimers_fm	-0.20	1.66E-23	25610	87378	Alzheimer's disease/dementia illnesses of mother	-0.217	0.805	2.45E-43	28507	331041			
								Mortality	27029810	longevity	At least 90 years old	0.18	4.54E-08	5964	417602										
								Response to statins (LDL cholesterol change)	22331829	pharmacogenomics															
								LDL cholesterol	23067351, 28371326, 28548082, 28334899	metabolic															
								HDL cholesterol	28270201	metabolic	iqb.low_hdl	-0.29	3.24E-68	36339	154349										
								Cholesterol, total	25961943, 28548082, 28270201, 28334899	metabolic	high_cholesterol	-0.62	0.00E-400	112221	219875	High cholesterol non-cancer illness code, self-reported	-0.346	0.707	2.04E-158	43957	361141				
										metabolic	Presence of any metabolic or heart-related disease	-0.23	2.06E-105	166416	318382	Heart attack/myocardial infarction non-cancer illness code, self-reported	-0.173	0.841	1.54E-09	8239	361141				
										metabolic															
										Physical	Trunk fat mass												3.88E-12	354597	
										Physical	Whole body fat mass												2.97E-11	350244	
										Physical	Hip circumference												5.83E-11	360521	
											Cholesterol lowering medication cholesterol, blood pressure, diabetes, or take exogenous hormones											5.26E-92	24247	193148	
								Lipoprotein (a) levels	28512139	lab measurement															
								Lipoprotein-associated phospholipase A2 activity change in response to danapradib	28753643	pharmacogenomics															
								Lipoprotein phospholipase A2 activity	28753643																
								Coronary artery disease	29212778, 28714975	cardiovascular	Coronary artery disease	-0.16	7.13E-10	13648	329205	Angina vascular/heart problems diagnosed by doctor	-0.19	0.827	7.52E-15	11372	360420				
								Pulse pressure	28135244	vital sign															
rs3865444	19	51224706	0.69	A/C	CD33	IGAP stage 1	5.12E-08		Alzheimer's disease	21460841, 24162737, 28714976, 21460840	Neurological											8.22E-01	83	337159	
											Neurological												1.65E-03	404	402787
								Family history of Alzheimer's disease	29777097	Neurological													1.46E-02	28507	331041
								Platelet count	27863252	Hematological measurement													7.15E-19	350474	
								White blood cell count	27863252	Hematological measurement													2.35E-16	350470	

(Continued)

Table 2. Wave 2 PheWAS results.

SNP	Chr	BP*	Freq**	Alleles**	Gene	Published associated with AD		Selected known association*** from NHGRI-EBI Catalog		Phenotype Group	Selected UKBB PheWAS results					
						Reported p-value	Analysis	DISEASE/ TRAIT	PubMed ID		DISEASE/ TRAIT	β^*	OR	p	N Cases	N Overall
<i>Additional variants associated with Alzheimer's disease</i>																
rs4575098	1	161185602	0.76	A/G	ADAMTS4	2.05E-10	Overall (phase 3) (Jansen et al. 2019)	Family history of Alzheimer's disease	30617256; 29777097	Neurological	Alzheimer's disease/dementia illnesses of mother	-0.060	0.942	2.84E-09	28507	331041
								Monocyte percentage of white cells	27863252	Hematological measurement	Monocyte percentage	-0.048		1.37E-10		349861
								Platelet distribution width	27863252	Hematological measurement	Platelet distribution width	0.012		9.62E-16		350470
								Mean platelet volume	27863252	Hematological measurement	Neutrophil count	0.022		3.85E-08		349856
								Granulocyte percentage of myeloid white cells	27863252	Hematological measurement	Mean platelet (thrombocyte) volume	0.030		6.06E-23		350470
rs10933431	2	233117202	0.234	C/G	INPP5	8.92E-10	Overall (phase 3) (Jansen et al. 2019)	Alzheimer's disease (late onset)	30617256; 24162737	Neurological	Alzheimer's disease/dementia illnesses of mother	-0.043	0.958	3.97E-05	28507	331041
rs184384746	3	57192122	0.002	C/T	HESX1	1.24E-08	AD-by-proxy (Jansen et al. 2019)	AD-by-proxy	30617256; 30617256	Neurological	Alzheimer's disease/dementia illnesses of mother	0.254	1.289	3.21E-02	28507	331041
rs6448453	4	11024404	0.747	A/G	CLNK	1.93E-09	Overall (phase 3) (Jansen et al. 2019)	Family history of Alzheimer's disease	30617256; 29777097	Neurological	Alzheimer's disease/dementia illnesses of mother	-0.049	0.952	4.68E-07	28507	331041
rs7657553	4	11721611	0.71	A/G	HS3ST1	2.16E-08	Case-control status (phase 1) (Jansen et al. 2019)	Alzheimer's disease (late onset)	30617256; 24162737	Neurological						
rs187370608	6	40974457	0.99799	A/G	TREM2	1.45E-16	Overall (phase 3) (Jansen et al. 2019)	AD-by-proxy	30617256; 30617256	Neurological	Alzheimer's disease/dementia illnesses of mother	-0.475	0.622	8.07E-09	28507	331041
rs6931277	6	32615580	0.16	A/T	HLA-DRB1	8.41E-11	Overall (phase 3) (Jansen et al. 2019)	Family history of Alzheimer's disease	30617256; 29777097	Immune system						

(Continued)

Table 2. (Continued)

SNP	Chr	BP*	Freq**	Alleles**	Gene	Published associated with AD		Selected known association*** from NHGRI-EBI Catalog	Phenotype Group	Selected UKBB PheWAS results					
						Reported p-value	Analysis			DISEASE/ TRAIT	PubMed ID	DISEASE/ TRAIT	β^*	OR	p
								Ulcerative colitis	Immune system	Ulcerative colitis ^b	-0.206	0.8138331	2.23E-10	3195	337978
									Immune system	Rheumatoid arthritis non-cancer illness code, self-reported	0.691	1.997	2.30E-130	4017	361141
									Immune system	Celiac disease ^b	-0.702	0.496	5.99E-54	1855	336638
									Respiratory system	Asthma non-cancer illness code, self-reported	0.162	1.176	8.57E-68	41934	361141
									Neurological	Multiple sclerosis non-cancer illness code, self-reported	-0.366	0.693	1.22E-13	1326	361141
								Inflammatory bowel disease	Immune system	Inflammatory bowel disease and other gastroenteritis and colitis	-0.140	0.869	2.87E-07	4528	339311
								Type 2 diabetes	Endocrine system	Type 1 diabetes	0.617	1.853	1.55E-60	2660	391416
								White blood cell count	Endocrine system	Type 2 diabetes ^b	0.119	1.126	4.63E-17	18945	407701
								Neutrophil count	Hematological measurement	White blood cell (leukocyte) count	0.099	1.29E-58	1.29E-58		350470
								Monocyte percentage of white cells	Hematological measurement	Neutrophil count	0.074	4.43E-67	4.43E-67		349856
								Family history of Alzheimer's disease	Hematological measurement	Monocyte percentage	-0.044	3.49E-08	3.49E-08		349861
rs1859788	7	100374211	0.697	A/G	ZCWPW1	2.22E-15	Overall (phase 3) (Jansen et al, 2019)	30617256; 29777097	Neurological	Alzheimer's disease/dementia illnesses of mother	0.044	1.045	1.80E-06	28507	331041
									Hematological measurement	Haemoglobin concentration	0.016	4.94E-10	4.94E-10		350474
										Pulse rate, automated reading	-0.169	6.58E-09	6.58E-09		340162
rs114360492	7	143411065	0.50	C/T	EPHA1	3.59E-11	Overall (phase 3) (Jansen et al, 2019)	30617256; 24162737; 21460840	Neurological	Alzheimer's disease/dementia illnesses of father	-0.042322	0.9585607	2.63E-04	15022	312666
						2.10E-09	Overall (phase 3) (Jansen et al, 2019)	30617256							

(Continued)

Table 2. (Continued)

SNP	Chr	BP*	Freq**	Alleles**	Gene	Published associated with AD		Selected known association*** from NHGRI-EBI Catalog	Phenotype Group	Selected UKBB PheWAS results							
						Reported p-value	Analysis			DISEASE/TRAIT	PubMed ID	DISEASE/TRAIT	β^*	OR	p	N Cases	N Overall
rs11257238	10	11675398	0.62	C/T	ECHDC3	1.26E-08	Overall (phase 3) (Jansen et al. 2019)	Alzheimer's disease (late onset)	Neurological								
rs2081545	11	60190907	0.617	A/C	MS4A6A	1.55E-15	Overall (phase 3) (Jansen et al. 2019)	Alzheimer's disease (late onset)	Neurological								
rs11218343	11	121564878	0.96	C/T	SORL1	1.09E-11	Overall (phase 3) (Jansen et al. 2019)	Heel bone mineral density	Bone measurement								
rs12590654	14	92472511	0.657	A/G	SLC24A4	1.65E-10	Overall (phase 3) (Jansen et al. 2019)	Alzheimer's disease (late onset)	Neurological								
rs442495	15	58730416	0.68	C/T	ADAM10	1.31E-09	Overall (phase 3) (Jansen et al. 2019)	Family history of Alzheimer's disease	Neurological								
									Hematological measurement								
rs117618017	15	63277703	0.124	C/T	APH1B	3.35E-08	Overall (phase 3) (Jansen et al. 2019)	Family history of Alzheimer's disease	Neurological								
rs59735493	16	31121779	0.70	A/G	KAT8	3.98E-08	Overall (phase 3) (Jansen et al. 2019)	Platelet count	Hematological measurement								

(Continued)

Table 2. (Continued)

SNP	Chr	BP*	Freq**	Alleles**	Gene	Published associated with AD		Selected known association*** from NHGRI-EBI Catalog		Phenotype Group	Selected UKBB PheWAS results					
						Reported p-value	Analysis	DISEASE/ TRAIT	PubMed ID		DISEASE/ TRAIT	β^*	OR	p	N Cases	N Overall
rs113260531	17	5235685	0.881	A/G	SC1MP	9.16E-10	Overall (phase 3) (Jansen et al., 2019)	Family history of Alzheimer's disease	30617256; 29777097; 28092683	Neurological	Alzheimer's disease/dementia illnesses of father	-0.078	0.925	1.20E-05	15022	312666
rs28394864	17	49373413	0.53	A/G	ABI3	1.87E-08	Overall (phase 3) (Jansen et al., 2019)	White blood cell count	27863252	Hematological measurement	White blood cell (leukocyte) count	-0.041		5.71E-08		350470
										Hematological measurement	Neutrophil count	-0.028		7.12E-08		349856
										Cardiovascular	Hypertension ^b	-0.0476		1.60E-13	77977	408343
										Cardiovascular	Ischemic heart disease ^b	-0.0529		4.58E-09	31355	408458
								Coronary artery disease	29212778	Cardiovascular	Coronary atherosclerosis ^b	-0.0682		6.47E-10	20023	397126
										Cardiovascular	Angina vascular/heart problems diagnosed by doctor	-0.077758		6.51E-09	11372	360420
										Respiratory system	Asthma blood clot, dvt, bronchitis, emphysema, asthma, rhinitis, eczema, allergy diagnosed by doctor	-0.054635		1.77E-13	41633	360527
											Impedance of whole body	1.79524		1.16E-29		354795
								Eosinophil counts	27863252	Hematological measurement	Eosinophil count	-0.015258		6.74E-18		349856
										Anthropometric measurement	Standing height	0.147437		7.32E-23		360388

(Continued)

Table 2. (Continued)

SNP	Chr	BP*	Freq**	Alleles**	Gene	Published associated with AD		Selected known association*** from NHGRI-EBI Catalog	Phenotype Group	Selected UKBB PheWAS results						
						Reported p-value	Analysis			DISEASE/ TRAIT	PubMed ID	β**	OR	p	N Cases	N Overall
rs2632516	17	58331728	0.548	C/G	BZRAP1-AS1	1.42E-09	Case-control status (phase 1) (Jansen et al. 2019)	Alzheimer's disease (late onset)	Neurological							
								Cognitive performance	Neurological							
								Monocyte count	Hematological measurement				1.84E-31			349856
								Plateletcrit	Hematological measurement				7.32E-16			350471
								Platelet count	Hematological measurement				1.02E-11			350474
rs8093731	18	31508995	0.01	C/T	SUZ12P1	4.63E-08	Case-control status (phase 1) (Jansen et al. 2019)	Alzheimer's disease (late onset)	Neurological							
rs76726049	18	58522227	0.9863	C/T	ALPK2	3.30E-08	Overall (phase 3) (Jansen et al. 2019)	Family history of Alzheimer's disease	Neurological				0.873		2.17E-04	28507
								Alzheimer's disease (late onset)	Neurological				1.8964809		5.96E-04	404
rs76320948	19	45738583	0.96	C/T	AC074212.3	4.64E-08	Overall (phase 3) (Jansen et al. 2019)	Alzheimer's disease (late onset)	Neurological				0.64			402787
rs6014724	20	56423488	0.0905	A/G	CASS4	6.56E-10	Overall (phase 3) (Jansen et al. 2019)	Alzheimer's disease (late onset)	Neurological				-0.054		3.82E-04	28507
								Alzheimer's disease (late onset) and AD-by-proxy	Neurological				0.947			331041
rs7185636	16	19796841	0.84	C/T	IQCK	5.30E-08	Overall (Kunkle et al. 2019)	Alzheimer's disease (late onset) and AD-by-proxy	Neurological							
								Impedance of whole body					-1.84821		5.69E-19	354795
								Comparative body size at age 10					0.0179134		2.67E-17	354996
								Body mass index (bmi)	Anthropometric measurement				0.0766973		1.77E-07	359983
rs2830500	21	26784537	0.686	A/C	ADAMTS1	2.60E-08	Overall (Kunkle et al. 2019)	Alzheimer's disease (late onset) and AD-by-proxy	Neurological							

(Continued)

Table 2. (Continued)

SNP	Chr	BP*	Freq ^{***}	Alleles**	Gene	Published associated with AD		Selected known association ^{***} from NHGRI-EBI Catalog		Phenotype Group	Selected UKBB PheWAS results					
						Reported p-value	Analysis	DISEASE/ TRAIT	PubMed ID		DISEASE/TRAIT	β^*	OR	p	N Cases	N Overall
rs3845261	17	5104941	0.347	C/T	ZNF232	4.00E-08	UK Biobank paternal and maternal meta-analysis	AD-by-proxy	29777097	Neurological	Alzheimer's disease/dementia illnesses of mother	0.029	1.030	1.30E-03	28507	331041
rs72824905	16	81908423	0.01	C/G	PLCG2	5.38E-10	Sims et al., 2017	White blood cell count	27863252	Hematological measurement	White blood cell (leukocyte) count	0.021		4.24E-05		350470

P-values are in bold font if passing the PheWAS multiple testing correction threshold of 2.32×10^{-7} in UKB cohort.

*chromosomal position based on genome build GRCh38 coordinate.

**The Alleles column describes the two possible alleles at the variant location, listed in alphabetical order. In this study, the first allele will be called "A allele" and the second allele will be called the "B allele"; effect (β): The effect size, ln(Odds Ratio [OR]) for binary traits, defined per copy of the B allele.

*** Highlighted associations do not necessarily pass genome wide significance threshold in all PubMed ID citations. Freq_B is frequency of B allele in Non-Finnish European from gnomAD.

^aAnalysis result was based on UKB Neale analysis accessed November 2019

^bUKB SAIGE, while the rest of UKB results reported are based on UKB Neale v2 analysis accessed July 3, 2020.

<https://doi.org/10.1371/journal.pone.0241552.t002>

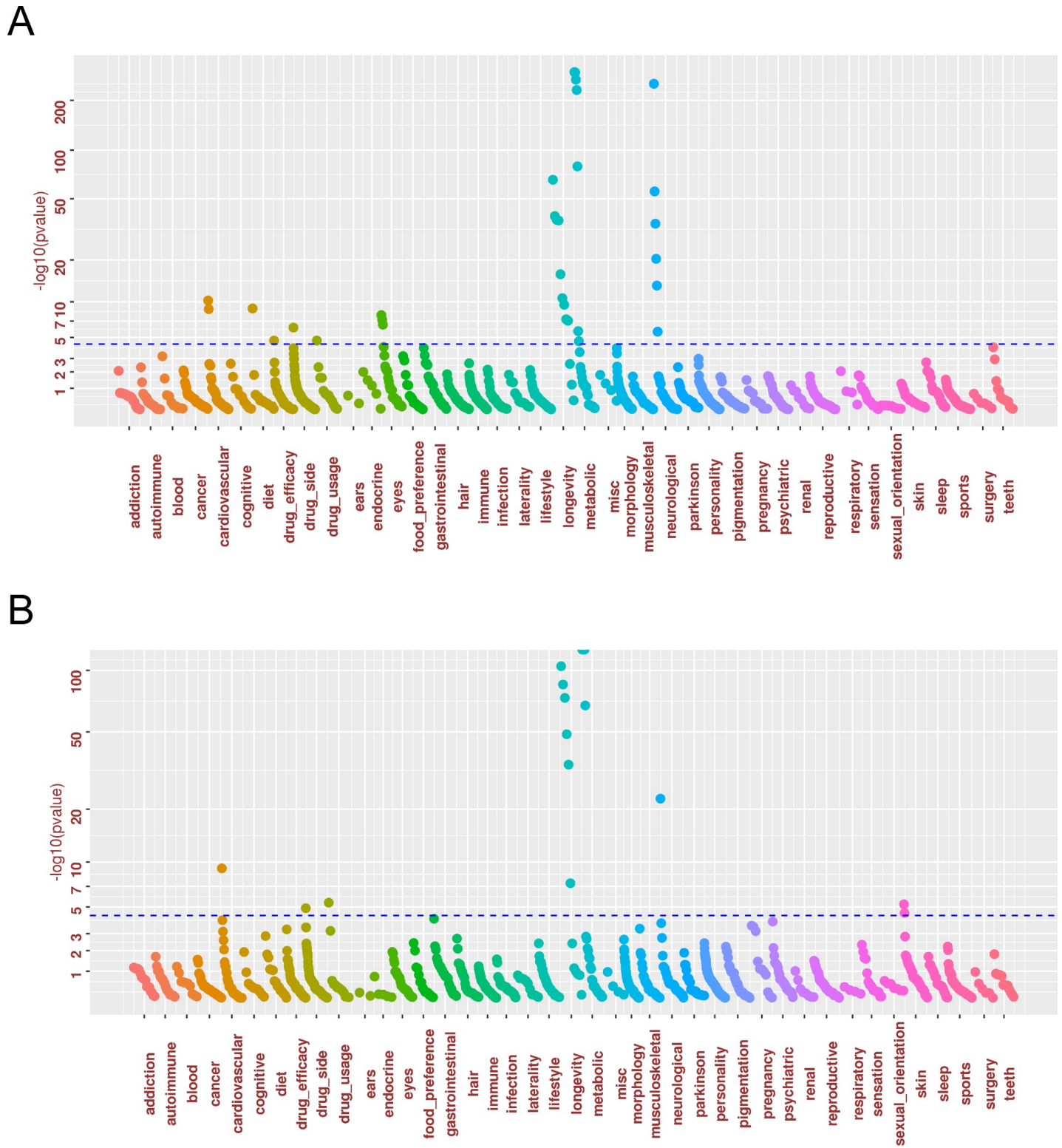


Fig 1. PheWAS association plot for APOE variants. A grey line is drawn at Positions $p = 5 \times 10^{-5}$ (a score of about 4.3), which is a threshold for significance after controlling for the Family-Wise-Error-Rate (FWER) using Bonferroni correction. (a) rs429358; (b) rs7412.

<https://doi.org/10.1371/journal.pone.0241552.g001>

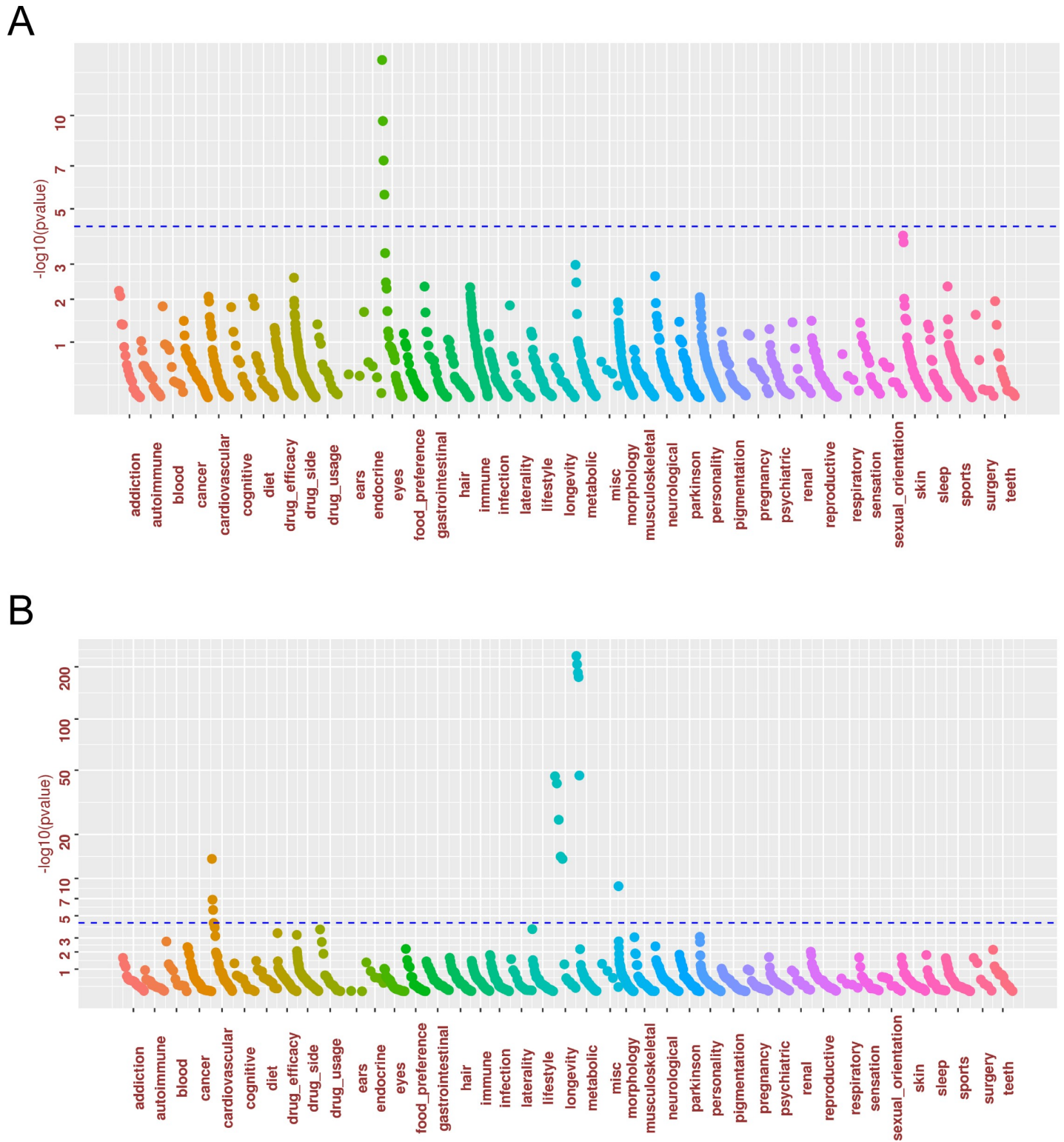


Fig 2. PheWAS association plot. A grey line is drawn at Positions $p = 5 \times 10^{-5}$ (a score of about 4.3), which is a threshold for significance after controlling for the FWER using Bonferroni correction. (a) rs11136000 (*CLU*); (b) rs646776.

<https://doi.org/10.1371/journal.pone.0241552.g002>

$\times 10^{-60}$) in the UKB cohort. In addition, *SPPL2A*, *CD33*, *SCIMP*, *ADAMTS4*, *APH1B*, *BZRA-PI-AS1*, *ZNF232*, *GRN*, and *CD2AP* variants were associated with mean platelet (thrombocyte) volume (an autoimmune disease biomarker), neutrophil count, monocyte count/percentage, and/or white blood cell (leukocyte) count. The significance of these associations is unknown given the large study sample size and the small effect size, but the directionality in most cases (if not all) is consistent with reported in the literature in a study with a large sample size ($>160,000$ subjects) [53]. *ABI3* variant was also associated with asthma in the UKB cohort ($p = 1.77 \times 10^{-13}$). The association results from the 23andMe cohort for selected immune-related conditions are listed in S5 Table. In the UKB cohort, *CD33* variant rs3865444 had nominal association with asthma ($p = 2.84 \times 10^{-4}$), while *CD2AP* variant rs9349407 had nominal association with UC ($p = 0.0004$) and *PICALM* variant rs3851179 had nominal association with hay fever or allergic rhinitis ($p = 4.42 \times 10^{-4}$).

FTD variant rs646776 is known to be associated with LDL-cholesterol levels, cardiovascular diseases, and plasma progranulin levels. Our PheWAS analysis (Fig 2) replicated the associations of rs646776 with metabolic traits (high cholesterol or taking drugs to lower cholesterol ($p = 2.74 \times 10^{-207}$; high cholesterol, $p = 1.35 \times 10^{-185}$), low high-density lipoproteins (HDL, $p = 3.37 \times 10^{-47}$), cardiovascular traits (heart attack ($p = 1.37 \times 10^{-7}$), CAD ($p = 1.79 \times 10^{-14}$), angina ($p = 2.34 \times 10^{-6}$)). Rs646776 was also significantly associated with longevity traits (heart metabolic disease in old people, healthy old) and height ($p = 1.80 \times 10^{-9}$) and suggestively associated with coronary bypass surgery ($p = 5.24 \times 10^{-5}$, FDR = 0.004), aortic stenosis ($p = 0.0001$, FDR = 0.01) and angioplasty ($p = 0.0007$, FDR = 0.04). rs5848 had a suggestive association with Parkinson's disease in both the 23andMe cohort ($p = 1.38 \times 10^{-5}$) and the UKB cohort ($p = 3.89 \times 10^{-3}$).

In addition to *APOE* variants and rs646776, *ABI3* variant rs28394864 was also associated with cardiovascular traits, such as hypertension ($p = 1.60 \times 10^{-13}$), ischemic heart disease ($p = 4.58 \times 10^{-9}$), coronary atherosclerosis ($p = 6.47 \times 10^{-10}$), and angina ($p = 6.51 \times 10^{-9}$). Despite *CLU* and *ABCA7* are both implicated in cholesterol metabolism [22], neither the *CLU* variant nor the *ABCA7* variant was strongly associated with metabolic/cardiovascular traits in the PheWAS analyses in either the 23andMe cohort or the UKB cohort despite the fairly substantial sample size for those traits in both cohorts.

The PheWAS results for the UKB cohort are available in S6 Table (Neale v1) and S7 Table (Neale v2 and UKB SAIGE).

No significant genetic correlation between PheWAS traits and AD

Despite that multiple traits were associated with the same individual variants in PheWAS analysis, there was no significant genetic correlation among these traits (e.g. LDL/HDL cholesterol, Type 2 Diabetes, CAD, CeD, RA, UC, multiple sclerosis, and BMI) and AD at the genome level (S8 Table). As positive controls, IGAP AD [21] and UKB trait (Neale v1) Illnesses of mother: AD/dementia showed significant genetic correlation with Jansen et al AD results [30] ($r_g = 0.901$, $p = 3.10 \times 10^{-13}$; $r_g = 0.63$, $p = 1.09 \times 10^{-6}$, respectively).

A causal role of cholesterol on AD revealed by MR analysis

Among the significant PheWAS traits associated with the AD/FTD disease variants, a set of genetic variant instruments from MR Base for BMI [54, 55], T2DM [56–58], lipid traits including HDL cholesterol, LDL cholesterol, and total cholesterol [59], CAD [60, 61], extreme height [55], parental attained age [62], traits defined from UK Biobank such as “reason for glasses/contact lenses: For short-sightedness i.e. only or mainly for distance viewing such as driving, cinema, etc. (called ‘myopia’)”, and “non-cancer illness code self-reported: angina” analyses by

the Neale Lab v1), RA [63], Inflammatory bowel disease (UC and Crohn's disease (CD)) [64, 65], CeD [66], multiple sclerosis (MS) [67, 68] were obtained. When treating AD as outcome and using $p < 5 \times 10^{-8}$ to select variants as instrument variables, the MR Egger intercept test suggested a directional horizontal pleiotropy for extreme height (Egger intercept $p = 0.009$), total cholesterol ($p = 0.02$), RA ($p = 0.02$) and parents' age at death (Egger intercept $p = 0.02$). MR Egger analysis suggested that metabolic traits (e.g. LDL cholesterol ($p = 4.7 \times 10^{-4}$) and total cholesterol ($p = 9.8 \times 10^{-5}$)) and RA had protective effect on AD with higher level of LDL or total cholesterol increasing the risk of AD and having RA reduce the risk of AD (S9 Table).

Conversely, genetic variant instrument for AD [30] suggested that AD possibly had causal effect on MS ($p = 0.0001$ using inverse variance weighted method) and coronary heart disease ($p = 0.003$ using MR Egger method, S9 Table). For FTD, the results may be inconclusive due to few SNPs were used in the instrument variable and the SNPs chosen were suggestively significant from the GWAS with smaller sample size. These MR tests would be still significant after correcting for the number of traits tested ($n = 15$, $p < 0.05/15 \sim 0.003$). A full list of MR results is listed in S9 Table.

Discussion

The PheWAS study showed that both *APOE* variants defining $\epsilon 2/\epsilon 3/\epsilon 4$ alleles, *ABI3* variant rs28394864, and rs646776 had significant associations with metabolic/cardiovascular and/or longevity traits. *APOE* variants were additionally significantly associated with neurological traits. *HLA-DRB1* variant was associated with immune-related traits. Both *APOE* variants and *CLU* variant were significantly associated with eye phenotypes. The associations of *ABI3* variant rs28394864 with cardiovascular traits (hypertension, ischemic heart disease, coronary atherosclerosis, angina), and asthma are novel findings from this study.

The novel finding of PheWAS associations of *ABI3* variant is of most interest. Rare variant (rs616338, p.Ser209Phe, $p = 4.56 \times 10^{-10}$, OR = 1.43, $MAF_{cases} = 0.011$, $MAF_{controls} = 0.008$) in *ABI3* was previously reported, and *ABI3* is specifically expressed in microglia (S2 Fig, similar expression pattern in human compared to other AD genes implicated by human genetics including *TREM2*, *HLA-DRB1*, *PLCG2*, *SORL1*, *SCIMP*, and *MS4A6A*) and thought to play a role in microglia-mediated innate immunity in AD [69]. Given its role in immune response, the PheWAS association with asthma is not completely unexpected, and its association with cardiovascular traits might reflect the role of immune dysregulation on those disease processes.

The observed PheWAS associations of *APOE* variants with metabolic/cardiovascular traits are not surprising. While vascular and metabolic risk factors such as hypertension, hyperlipidemia /hypercholesterolemia, hyperinsulinemia, and obesity at midlife, diabetes mellitus (DM), and cardiovascular and cerebrovascular diseases (including stroke, clinically silent brain infarcts and cerebral microvascular lesions) are generally thought to increase the risk of dementia and AD [70–73], the directional impact of a factor could be age-dependent, for example, hypertension, obesity and hypercholesterolemia are risk factors at middle age (<65 years) for late-life dementia and AD, but protective late in life (age >75 years) [74]. It seems to be odd that AD patient had a lower risk of developing CAD [73], but it is consistent with a meta-analysis [72] and this meta-analysis also reported that metabolic syndrome decreases the risk of AD. In the MR analysis from this study, AD increased the risk of CAD (S9 Table), but this result was supported by MR Egger method only. Taking age into consideration may help better delineate the relationship. Furthermore, several cardiovascular risk factors demonstrated associations with more rapid cognitive decline as expected, however it was also reported that recent or active hypertension and hypercholesterolemia were associated with

slower cognitive decline for AD patients [75]. These epidemiology studies suggested that it appears to be a complex interplay between AD and metabolic/cardiovascular risk factors and conditions, and the occasionally contradictory findings may be due to age of the population, sampling biases and/or other confounding factors. Nevertheless, the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) study demonstrated that multidomain intervention (diet, exercise, cognitive training, vascular risk monitoring) had beneficiary effect on the primary outcome, i.e. change in cognition as measured through comprehensive neuropsychological test battery (NTB) in an at-risk elderly population (aged 60–77) with CAIDE (Cardiovascular Risk Factors, Aging and Dementia) Dementia Risk Score of at least 6 points and cognition at mean level or slightly lower than expected for his/her age group, suggesting targeting modifiable vascular and lifestyle-related risk factors could improve or maintain cognitive functioning [76]. The PheWAS analysis suggested the minor allele of rs7412 (defining $\epsilon 2$ allele), a known protective allele for AD (OR = 0.74), was also a protective allele for having high cholesterol, low HDL, having heart metabolic disease or CAD. Similarly, the minor allele of rs429358 (defining $\epsilon 4$ allele), a known risk allele for AD (OR = 2.17), was also a risk allele for having high cholesterol, low HDL, having heart metabolic disease or CAD. Our MR analysis demonstrated that LDL and total cholesterol had a causal relationship to the development of AD using MR Egger. This MR result is however sensitive to the MR methods used as other methods such as weighted mode, weighted median, or simple mode (not pre-specified analyses) did not provide evidence or only provide suggestive evidence for the causal effect of LDL on AD. A recent MR analysis on 24 potentially modifiable risk factors [77] concluded that genetically predicted cardiometabolic factors were not associated with AD as there was no evidence of causal relationship after excluding one pleiotropic genetic variant (not disclosed in the publication) near the *APOE* gene (also near *APOC1* and *TOMM40* genes). The evidence we obtained was far weaker than that reported by Larsson et al., for all variants [77]. Despite there were few SNPs driving the causal evidence in single variant analysis, leave-one-analysis did not differ substantially from the analysis including all variants for LDL trait except rs7412 (S1 Fig). This study opted to report the findings using the inverse variance weighted method (when Egger intercept is not significant) as also adopted by Howard et al. [78], where a minimal of 30 SNPs used in instrument variable was also imposed, or MR Egger regression results (when the intercept is significant). We did not filter out analysis with less than 30 variants. Both compromises are limitations in this study thus those results shall be interpreted with caution. In addition, both IVW and MR Egger methods do not protect against the violation of the MR assumption when the pleiotropic effects act via a confounder of the exposure-outcome association [49].

The observed PheWAS associations of rs646776 variant with metabolic/cardiovascular traits are also not surprising. SNP rs646776 was reported to be robustly associated with low-density lipoprotein cholesterol (LDL-C, $p = 3 \times 10^{-29}$) with each copy of the minor allele decreasing LDL cholesterol concentrations by ~5–8 mg/dl [79]. The association was strengthened in a meta-analysis of ~100,000 individuals of European descent for LDL-C ($p = 5 \times 10^{-169}$) and was also detected for total cholesterol ($p = 7 \times 10^{-130}$) [80]. However, which gene is the causative gene for rs646776 effect is less clear despite it was selected to be included in the PheWAS analysis based on the association with plasma progranulin levels. Rs646776 at the 1p13 locus was also strongly associated with transcript levels of three neighboring genes: sortilin (*SORT1*) ($p = 3 \times 10^{-26}$), cadherin EGF LAG seven-pass G-type receptor 2 (*CELSR2*) ($p = 2 \times 10^{-12}$) and proline and serine rich coiled-coil 1 (*PSRC1*) ($p = 3 \times 10^{-12}$) [79]. The conditional analysis suggested that *SORT1* eQTL effect might be the dominant effect [79]. Rs599839, a SNP in LD with rs646776, was also reported to be associated with CAD [81]. The minor allele conferring lower level of LDL cholesterol also conferred lower risk of CAD. Rs646776 was also identified in a

bivariate analysis to be associated with circulating IGF-I and IGF-binding protein-3 (IGFBP-3) ($p = 6.87 \times 10^{-9}$) in a meta-analysis of 21 studies including 30,884 adults of European ancestry [82]. The growth hormone/insulin-like growth factor (IGF) axis can be manipulated in animal models to promote longevity. IGF related proteins including IGF-I and IGFBP-3 have also been implicated in risk of human diseases including cardiovascular diseases and diabetes. This is particularly interesting given the observation that rs646776 is associated with longevity in the PheWAS analysis.

It is surprising and puzzling to see the effect of AD variants on multiple eye phenotypes especially myopia that have onset in early childhood or teens. The association with age related macular degeneration (AMD) was reported previously [83]. The AMD association is interesting because the histopathological hallmark of AMD is amyloid- β (A β) in optic never drusen [84]. Drusen of the macula are very small yellow and white spots that appear in one of the layers of the retina named Bruch's membrane and are remnant nondegradable proteins and lipids (lipofuscin), which is the earliest visible sign of dry macular degeneration. In addition to the amyloid phenotype, AMD and AD also share other common histologic feature such as vitronectin accumulation and immunologic features such as increased oxidative stress, and apolipoprotein and complement activation pathways [85]. The common etiopathogenetic and morphological manifestations of AD and age-related eye diseases in amyloid genesis may have a broader implication in understanding the disease mechanism, identifying new biomarkers and treatment [86]. A recent study showed that the soft drusen area in amyloid-positive patients was significantly larger than that in amyloid-negative patients [87]. Ocular and visual information processing deficit were other possible biomarkers for AD [88]. Recently it was also reported that thinner retinal nerve fiber layer is associated with an increased risk of dementia including AD, suggesting that retinal neurodegeneration may serve as a preclinical biomarker for dementia [89]. Risk variant for AD rs429358 in our PheWAS results had a protective effect for AMD and blindness ($p = 9.6 \times 10^{-5}$, FDR = 0.004), perhaps reflecting the equilibrium of A β in brain vs. retina like the situation between brain vs. CSF. A variety of other visual problems reported in patients with AD have been reviewed in details [90] including loss of visual acuity (VA), color vision and visual fields; changes in pupillary response to mydriatics, defects in fixation and in smooth and saccadic eye movements; changes in contrast sensitivity and in visual evoked potentials (VEP); and disturbances of complex visual functions, though they have not been studied as a risk factor of AD or outcome of having AD. In the MR analysis, we cannot directly test the causal relationship between AMD despite the GWAS with large sample size is available due to standard error of odds ratio was not reported in the paper [83].

Subjects with AD risk variant rs1113600 in *CLU* gene had a higher chance of being nearsighted, while subjects with AD risk variant rs429358 in *APOE* gene had a lower chance of being nearsighted. It was reported that wearing reading glasses correlated significantly with high mini-mental state examination for the visually impaired (MMSE-blind) after adjustment for sex and age (OR = 2.14, 95% CI = 1.16–3.97, $p = 0.016$), but reached borderline significance after adjustment for education [91]. There was a trend toward correlation between myopia and better MMSE-blind ($r = -0.123$, $p = 0.09$, Pearson correlation) [91]. On the other hand, myopia may be a surrogate phenotype for intelligence (or education), as a genetic correlation between myopia and intelligence was shown in a small cohort of 1500 subject ($p < 0.01$) [92]. Larsson et al., suggested that genetically predicted educational attainment was significantly associated with AD per year of education completed (OR = 0.89, $p = 2.4 \times 10^{-6}$) and per unit increase in log odds of having completed college/university (OR = 0.74, $p = 8.0 \times 10^{-5}$), while intelligence had a suggestive association with AD (OR = 0.73, $p = 0.01$) [77]. Our MR analysis did not provide evidence on the causal relationship of myopia phenotype on AD.

Furthermore, although genetic variants (rs429358 and rs1113600) are associated with multiple phenotypes, the associations are not necessarily independent of each other. In fact, MR Egger intercept test did not support the independent relationship except for height and a few other traits. Overall, the relationship and interpretation between traits seem to be complex and require further examination.

Other limitations of our study design also merit comment. The sample size for PheWAS varies from trait to trait depending on the prevalence rate of the trait and availability of data. For example, the cohort size for AD in the 23andMe database was not large in 2015 (~640 cases and ~158K controls) when the PheWAS analysis for the 23andMe cohort was performed, which is a limitation for this study especially for replicating the association with AD. This may explain why some of the known SNP associations with AD were not replicated or only had nominally significant association in the 23andMe cohort. Furthermore, FTD is not a self-reported question collected in the 23andMe database and therefore could not be tested in the PheWAS analysis. Even if this was included, the sample size would have been smaller than that for AD based on population prevalence rate. Similarly, the cohorts for CD (n~3,600), UC (n~6,200), bipolar (~9,700), and schizophrenia (~700) were limited in size. However, the cohort sizes for other psychiatric disorders (e.g. depression, anxiety and panic) were sufficiently large (>250K cases). The PheWAS 23andMe cohort size for AD used in this study was limited, and therefore only *APOE* variants, the loci with the largest effect size, were confirmed to be associated with neurological traits. The sample size for a specific trait shall be taken into consideration when interpreting the PheWAS results. PheWAS typically uses “light” phenotyping (based on self-reported as in the 23andMe and surveys deployed by UKB or based on diagnostic ICD codes or medication / procedure usage pattern), the stringency of phenotype is certainly not as good as clinical ascertained phenotype, but the tradeoff is the power to survey a large number of diverse phenotypes within a single study.

The nominal causal effect between immune-related traits (except multiple sclerosis and rheumatoid arthritis) and AD/FTD would have been insignificant if correcting for the expanded list of diseases and risk factors from MR Base tested. Some of the instrument variable used consisted of small number of SNPs and may have weakened the real causal effect if exist. The observation does not seem to be purely by chance especially in light of the report on immune related enrichment of FTD where they found up to 270-fold genetic enrichment between FTD and RA, up to 160-fold genetic enrichment between FTD and UC, and up to 175-fold genetic enrichment between FTD and CeD. Overall, the immune overlap seems to be common to both FTD and AD at the genome level (represented by genome wide significant SNPs used as an instrument variable for AD and other diseases/risk factors except FTD where a p-value threshold of 5×10^{-6} was used because of the smaller GWAS samples size), while there could still be specificity of neuroinflammation for risk variants in *CRI*, *CD33*, *CLU*, *ABCA7*, *TREM2*, *SORL1*, *MS4A6A*, *SPPL2A*, *SCIMP*, *PLCG2*, *ABI3*, and *HLA-DRB1*. Different MR analysis methods have different assumptions (which in reality do not always hold or even rarely hold) and power, the inference of the causal effect may be inconclusive or only suggestive unless the causal effect size is so huge that most of methods give unequivocal concordant results. Both IVW and MR Egger methods used in this study are vulnerable to false positives when the exposure and outcome traits are both affected by a heritable confounder [49]. Different exposure or outcome GWAS may also vary by study sample size, and number of variants with summary association statistics available (for outcome GWAS as this may limit the ability to leverage proxy SNPs in LD (default $r^2 > = 0.8$) with the set of SNPs in the instrument variable) and impact the strength of instrument variable and the power of MR analysis. Future re-analysis when studies with larger sample size and more complete summary association statistics will be warranted to interrogate the causal relationship.

Supporting information

S1 Table. Genotype platform representation and imputed variant statistics in PheWAS.
(XLSX)

S2 Table. A list of 1234 phenotypes from the 23andMe cohort used for PheWAS analyses and sample size for each trait using the rs3865444 PheWAS results as an example.
(XLSX)

S3 Table. Detailed PheWAS results (FDR < 0.05) from the 23andMe cohort.
(XLSX)

S4 Table. A full list of detailed PheWAS results from the 23andMe cohort.
(XLSX)

S5 Table. Detailed results of PheWAS for immune-related traits.
(XLSX)

S6 Table. A full list of detailed results of PheWAS from the UKB cohort from Open Target Genetics (Neale v1) accessed in May 2019.
(XLSX)

S7 Table. A full list of detailed results of PheWAS from the UKB cohort from Open Target Genetics (Neale v2 and UKB SAIGE) accessed in July 2020.
(XLSX)

S8 Table. MR analysis for PheWAS traits (full results) and selected immune-related traits ($p < 0.1$).
(XLSX)

S9 Table. MR analysis for exposure = LDL and outcome = AD sensitivity analysis.
(XLSX)

S1 Fig. MR analysis on the LDL exposure and AD outcome.
(DOCX)

S2 Fig. Cell type specificity of *ABI3*, *SORL1*, *HLD-DRB1*, *MS4A6A*, *TREM2*, *PLCG2*, and *SCIMP*.
(DOCX)

Acknowledgments

We thank the 23andMe research participants for their contributions to this study. We gratefully acknowledge all the studies and databases that made GWAS summary data available: **C4D** (Coronary Artery Disease Genetics Consortium), **CARDIoGRAM** (Coronary ARtery Disease Genome wide Replication and Meta-analysis), **DIAGRAM** (DIAbetes Genetics Replication And Meta-analysis), **GIANT** (Genetic Investigation of ANthropometric Traits), **GLGC** (Global Lipids Genetics Consortium), **IMSGC** (International Multiple Sclerosis Genetic Consortium), **UK Biobank**.

23andMe Research Team: Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, Nicholas A. Furlotte, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Joyce Y. Tung, Vladimir Vacic, Catherine H. Wilson and Amir S. Zare.

Author Contributions

Conceptualization: Qingqin S. Li.

Formal analysis: Qingqin S. Li, Chao Tian.

Funding acquisition: Guy R. Seabrook.

Investigation: Qingqin S. Li, Chao Tian, David Hinds.

Methodology: Qingqin S. Li, Chao Tian, David Hinds.

Project administration: Qingqin S. Li, Chao Tian.

Resources: Qingqin S. Li.

Supervision: David Hinds.

Visualization: Qingqin S. Li, Chao Tian.

Writing – original draft: Qingqin S. Li.

Writing – review & editing: Qingqin S. Li, Chao Tian, David Hinds, Guy R. Seabrook.

References

1. Lane CA, Hardy J, Schott JM. Alzheimer's disease. *European journal of neurology*. 2018; 25(1):59–70. <https://doi.org/10.1111/ene.13439> PMID: 28872215.
2. Olszewska DA, Lonergan R, Fallon EM, Lynch T. Genetics of Frontotemporal Dementia. *Curr Neurol Neurosci Rep*. 2016; 16(12):107. <https://doi.org/10.1007/s11910-016-0707-9> PMID: 27878525.
3. Olney NT, Spina S, Miller BL. Frontotemporal Dementia. *Neurologic clinics*. 2017; 35(2):339–74. <https://doi.org/10.1016/j.ncl.2017.01.008> PMID: 28410663; PubMed Central PMCID: PMC5472209.
4. Scheltens P, Blennow K, Breteler MM, de Strooper B, Frisoni GB, Salloway S, et al. Alzheimer's disease. *Lancet*. 2016; 388(10043):505–17. [https://doi.org/10.1016/S0140-6736\(15\)01124-1](https://doi.org/10.1016/S0140-6736(15)01124-1) PMID: 26921134.
5. Falcon B, Zhang W, Murzin AG, Murshudov G, Garringer HJ, Vidal R, et al. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature*. 2018; 561(7721):137–40. <https://doi.org/10.1038/s41586-018-0454-y> PMID: 30158706; PubMed Central PMCID: PMC6204212.
6. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature*. 1991; 349(6311):704–6. <https://doi.org/10.1038/349704a0> PMID: 1671712.
7. Schellenberg GD, Bird TD, Wijsman EM, Orr HT, Anderson L, Nemens E, et al. Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. *Science*. 1992; 258(5082):668–71. <https://doi.org/10.1126/science.1411576> PMID: 1411576.
8. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science*. 1995; 269(5226):973–7. <https://doi.org/10.1126/science.7638622> PMID: 7638622.
9. Rogaev EI, Sherrington R, Rogaeva EA, Levesque G, Ikeda M, Liang Y, et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature*. 1995; 376(6543):775–8. <https://doi.org/10.1038/376775a0> PMID: 7651536.
10. Cruts M, Gijselink I, van der Zee J, Engelborghs S, Wils H, Pirici D, et al. Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature*. 2006; 442(7105):920–4. <https://doi.org/10.1038/nature05017> PMID: 16862115.
11. Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, et al. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*. 1998; 393(6686):702–5. <https://doi.org/10.1038/31508> PMID: 9641683.
12. Poorkaj P, Bird TD, Wijsman E, Nemens E, Garruto RM, Anderson L, et al. Tau is a candidate gene for chromosome 17 frontotemporal dementia. *Ann Neurol*. 1998; 43(6):815–25. <https://doi.org/10.1002/ana.410430617> PMID: 9629852.
13. Spillantini MG, Murrell JR, Goedert M, Farlow MR, Klug A, Ghetti B. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proceedings of the National Academy of Sciences*

- of the United States of America. 1998; 95(13):7737–41. <https://doi.org/10.1073/pnas.95.13.7737> PMID: 9636220; PubMed Central PMCID: PMC22742.
14. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America*. 1993; 90(5):1977–81. <https://doi.org/10.1073/pnas.90.5.1977> PMID: 8446617; PubMed Central PMCID: PMC46003.
 15. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nature genetics*. 2009; 41(10):1088–93. <https://doi.org/10.1038/ng.440> PMID: 19734902; PubMed Central PMCID: PMC2845877.
 16. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. *Jama*. 2010; 303(18):1832–40. <https://doi.org/10.1001/jama.2010.574> PMID: 20460622; PubMed Central PMCID: PMC2989531.
 17. Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buross J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nature genetics*. 2011; 43(5):436–41. <https://doi.org/10.1038/ng.801> PMID: 21460841; PubMed Central PMCID: PMC3090745.
 18. Chen H, Wu G, Jiang Y, Feng R, Liao M, Zhang L, et al. Analyzing 54,936 Samples Supports the Association Between CD2AP rs9349407 Polymorphism and Alzheimer's Disease Susceptibility. *Molecular neurobiology*. 2015; 52(1):1–7. <https://doi.org/10.1007/s12035-014-8834-2> PMID: 25092125.
 19. Jun G, Naj AC, Beecham GW, Wang LS, Buross J, Gallins PJ, et al. Meta-analysis confirms CR1, CLU, and PICALM as Alzheimer disease risk loci and reveals interactions with APOE genotypes. *Archives of neurology*. 2010; 67(12):1473–84. <https://doi.org/10.1001/archneurol.2010.201> PMID: 20697030; PubMed Central PMCID: PMC3048805.
 20. Liu G, Li F, Zhang S, Jiang Y, Ma G, Shang H, et al. Analyzing large-scale samples confirms the association between the ABCA7 rs3764650 polymorphism and Alzheimer's disease susceptibility. *Molecular neurobiology*. 2014; 50(3):757–64. <https://doi.org/10.1007/s12035-014-8670-4> PMID: 24643655.
 21. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature genetics*. 2013; 45(12):1452–8. <https://doi.org/10.1038/ng.2802> PMID: 24162737; PubMed Central PMCID: PMC3896259.
 22. Karch CM, Goate AM. Alzheimer's disease risk genes and mechanisms of disease pathogenesis. *Biological psychiatry*. 2015; 77(1):43–51. <https://doi.org/10.1016/j.biopsych.2014.05.006> PMID: 24951455; PubMed Central PMCID: PMC4234692.
 23. Carrasquillo MM, Nicholson AM, Finch N, Gibbs JR, Baker M, Rutherford NJ, et al. Genome-wide screen identifies rs646776 near sortilin as a regulator of progranulin levels in human plasma. *American journal of human genetics*. 2010; 87(6):890–7. <https://doi.org/10.1016/j.ajhg.2010.11.002> PMID: 21087763; PubMed Central PMCID: PMC2997361.
 24. Rademakers R, Eriksen JL, Baker M, Robinson T, Ahmed Z, Lincoln SJ, et al. Common variation in the miR-659 binding-site of GRN is a major risk factor for TDP43-positive frontotemporal dementia. *Human molecular genetics*. 2008; 17(23):3631–42. <https://doi.org/10.1093/hmg/ddn257> PMID: 18723524; PubMed Central PMCID: PMC2581433.
 25. Xu HM, Tan L, Wan Y, Tan MS, Zhang W, Zheng ZJ, et al. PGRN Is Associated with Late-Onset Alzheimer's Disease: a Case-Control Replication Study and Meta-analysis. *Molecular neurobiology*. 2017; 54(2):1187–95. <https://doi.org/10.1007/s12035-016-9698-4> PMID: 26820675.
 26. Hsiung GY, Fok A, Feldman HH, Rademakers R, Mackenzie IR. rs5848 polymorphism and serum progranulin level. *Journal of the neurological sciences*. 2011; 300(1–2):28–32. <https://doi.org/10.1016/j.jns.2010.10.009> PMID: 21047645; PubMed Central PMCID: PMC3085023.
 27. Li QS, Parrado AR, Samtani MN, Narayan VA, Alzheimer's Disease Neuroimaging I. Variations in the FRA10AC1 Fragile Site and 15q21 Are Associated with Cerebrospinal Fluid Abeta1-42 Level. *PloS one*. 2015; 10(8):e0134000. <https://doi.org/10.1371/journal.pone.0134000> PMID: 26252872; PubMed Central PMCID: PMC4529186.
 28. Carvalho-Silva D, Pierleoni A, Pignatelli M, Ong C, Fumis L, Karamanis N, et al. Open Targets Platform: new developments and updates two years on. *Nucleic acids research*. 2019; 47(D1):D1056–D65. <https://doi.org/10.1093/nar/gky1133> PMID: 30462303; PubMed Central PMCID: PMC6324073.
 29. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nature genetics*. 2019; 51(3):414–30. <https://doi.org/10.1038/s41588-019-0358-2> PMID: 30820047; PubMed Central PMCID: PMC6463297.

30. Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature genetics*. 2019; 51(3):404–13. <https://doi.org/10.1038/s41588-018-0311-9> PMID: 30617256; PubMed Central PMCID: PMC6836675.
31. Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, et al. GWAS on family history of Alzheimer's disease. *Translational psychiatry*. 2018; 8(1):99. <https://doi.org/10.1038/s41398-018-0150-6> PMID: 29777097; PubMed Central PMCID: PMC5959890.
32. Liu JZ, Erlich Y, Pickrell JK. Case-control association mapping by proxy using family history of disease. *Nature genetics*. 2017; 49(3):325–31. <https://doi.org/10.1038/ng.3766> PMID: 28092683.
33. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*. 2010; 26(9):1205–10. <https://doi.org/10.1093/bioinformatics/btq126> PMID: 20335276; PubMed Central PMCID: PMC2859132.
34. Ehm MG, Aponte JL, Chiano MN, Yerges-Armstrong LM, Johnson T, Barker JN, et al. Phenome-wide association study using research participants' self-reported data provides insight into the Th17 and IL-17 pathway. *PloS one*. 2017; 12(11):e0186405. <https://doi.org/10.1371/journal.pone.0186405> PMID: 29091937; PubMed Central PMCID: PMC5665418.
35. Diogo D, Tian C, Franklin CS, Alanne-Kinnunen M, March M, Spencer CCA, et al. Phenome-wide association studies across large population cohorts support drug target validation. *Nature communications*. 2018; 9(1):4285. <https://doi.org/10.1038/s41467-018-06540-3> PMID: 30327483; PubMed Central PMCID: PMC6191429.
36. Scott RA, Freitag DF, Li L, Chu AY, Surendran P, Young R, et al. A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease. *Sci Transl Med*. 2016; 8(341):341ra76. <https://doi.org/10.1126/scitranslmed.aad3744> PMID: 27252175; PubMed Central PMCID: PMC5219001.
37. Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. *Jama*. 2017; 318(19):1925–6. <https://doi.org/10.1001/jama.2017.17219> PMID: 29164242.
38. Li QS, Tian C, Seabrook GR, Drevets WC, Narayan VA. Analysis of 23andMe antidepressant efficacy survey data: implication of circadian rhythm and neuroplasticity in bupropion response. *Translational psychiatry*. 2016; 6(9):e889. <https://doi.org/10.1038/tp.2016.171> PMID: 27622933.
39. Durand EY, Do CB, Mountain JL, Macpherson JM. Ancestry composition: a novel, efficient pipeline for ancestry deconvolution.
40. Henn BM, Hon L, Macpherson JM, Eriksson N, Saxonov S, Pe'er I, et al. Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PloS one*. 2012; 7(4):e34267. <https://doi.org/10.1371/journal.pone.0034267> PMID: 22509285; PubMed Central PMCID: PMC3317976.
41. Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010; 467(7319):1061–73. <https://doi.org/10.1038/nature09534> PMID: 20981092; PubMed Central PMCID: PMC3042601.
42. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American journal of human genetics*. 2007; 81(5):1084–97. <https://doi.org/10.1086/521987> PMID: 17924348; PubMed Central PMCID: PMC2265661.
43. Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. *Bioinformatics*. 2015; 31(5):782–4. <https://doi.org/10.1093/bioinformatics/btu704> PMID: 25338720; PubMed Central PMCID: PMC4341061.
44. Xing G, Lin CY, Wooding SP, Xing C. Blindly using Wald's test can miss rare disease-causal variants in case-control association studies. *Ann Hum Genet*. 2012; 76(2):168–77. <https://doi.org/10.1111/j.1469-1809.2011.00700.x> PMID: 22256951.
45. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*. 2017; 33(2):272–9. Epub 2016/11/03. <https://doi.org/10.1093/bioinformatics/btw613> PMID: 27663502; PubMed Central PMCID: PMC5542030.
46. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*. 2015; 47(3):291–5. Epub 2015/02/03. <https://doi.org/10.1038/ng.3211> PMID: 25642630; PubMed Central PMCID: PMC4495769.
47. Hagenaars SP, Gale CR, Deary IJ, Harris SE. Cognitive ability and physical health: a Mendelian randomization study. *Scientific reports*. 2017; 7(1):2651. <https://doi.org/10.1038/s41598-017-02837-3> PMID: 28572633; PubMed Central PMCID: PMC5453939.

48. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. 2017; 32(5):377–89. <https://doi.org/10.1007/s10654-017-0255-x> PMID: 28527048; PubMed Central PMCID: PMC5506233.
49. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015; 44(2):512–25. <https://doi.org/10.1093/ije/dyv080> PMID: 26050253; PubMed Central PMCID: PMC4469799.
50. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018; 7. <https://doi.org/10.7554/eLife.34408> PMID: 29846171; PubMed Central PMCID: PMC5976434.
51. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic acids research*. 2014; 42(Database issue):D1001–6. <https://doi.org/10.1093/nar/gkt1229> PMID: 24316577; PubMed Central PMCID: PMC3965119.
52. Pottier C, Zhou X, Perkerson RB 3rd, Baker M, Jenkins GD, Serie DJ, et al. Potential genetic modifiers of disease risk and age at onset in patients with frontotemporal lobar degeneration and GRN mutations: a genome-wide association study. *Lancet Neurol*. 2018; 17(6):548–58. [https://doi.org/10.1016/S1474-4422\(18\)30126-1](https://doi.org/10.1016/S1474-4422(18)30126-1) PMID: 29724592; PubMed Central PMCID: PMC6237181.
53. Astle WJ, Elding H, Jiang T, Allen D, Ruklika D, Mann AL, et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*. 2016; 167(5):1415–29 e19. Epub 2016/11/20. <https://doi.org/10.1016/j.cell.2016.10.042> PMID: 27863252; PubMed Central PMCID: PMC5300907.
54. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015; 518(7538):197–206. <https://doi.org/10.1038/nature14177> PMID: 25673413; PubMed Central PMCID: PMC4382211.
55. Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nature genetics*. 2013; 45(5):501–12. <https://doi.org/10.1038/ng.2606> PMID: 23563607; PubMed Central PMCID: PMC3973018.
56. Replication DIG, Meta-analysis C, Asian Genetic Epidemiology Network Type 2 Diabetes C, South Asian Type 2 Diabetes C, Mexican American Type 2 Diabetes C, Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in multi-Ethnic Samples C, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics*. 2014; 46(3):234–44. <https://doi.org/10.1038/ng.2897> PMID: 24509480; PubMed Central PMCID: PMC3969612.
57. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics*. 2012; 44(9):981–90. <https://doi.org/10.1038/ng.2383> PMID: 22885922; PubMed Central PMCID: PMC3442244.
58. Wood AR, Tyrrell J, Beaumont R, Jones SE, Tuke MA, Ruth KS, et al. Variants in the FTO and CDKAL1 loci have recessive effects on risk of obesity and type 2 diabetes, respectively. *Diabetologia*. 2016; 59(6):1214–21. <https://doi.org/10.1007/s00125-016-3908-5> PMID: 26961502; PubMed Central PMCID: PMC4869698.
59. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nature genetics*. 2013; 45(11):1274–83. <https://doi.org/10.1038/ng.2797> PMID: 24097068; PubMed Central PMCID: PMC3838666.
60. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature genetics*. 2015; 47(10):1121–30. <https://doi.org/10.1038/ng.3396> PMID: 26343387; PubMed Central PMCID: PMC4589895.
61. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nature genetics*. 2011; 43(4):333–8. <https://doi.org/10.1038/ng.784> PMID: 21378990; PubMed Central PMCID: PMC3119261.
62. Pilling LC, Kuo CL, Sicinski K, Tamosauskaite J, Kuchel GA, Harries LW, et al. Human longevity: 25 genetic loci associated in 389,166 UK biobank participants. *Aging (Albany NY)*. 2017; 9(12):2504–20. <https://doi.org/10.18632/aging.101334> PMID: 29227965; PubMed Central PMCID: PMC5764389.
63. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*. 2014; 506(7488):376–81. Epub 2014/01/07. <https://doi.org/10.1038/nature12873> PMID: 24390342; PubMed Central PMCID: PMC3944098.
64. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. 2012; 491(7422):119–24.

- Epub 2012/11/07. <https://doi.org/10.1038/nature11582> PMID: 23128233; PubMed Central PMCID: PMC3491803.
65. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature genetics*. 2015; 47(9):979–86. Epub 2015/07/21. <https://doi.org/10.1038/ng.3359> PMID: 26192919; PubMed Central PMCID: PMC4881818.
 66. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nature genetics*. 2011; 43(12):1193–201. Epub 2011/11/08. <https://doi.org/10.1038/ng.998> PMID: 22057235; PubMed Central PMCID: PMC3242065.
 67. International Multiple Sclerosis Genetics C, Wellcome Trust Case Control C, Sawcer S, Hellenthal G, Pirinen M, Spencer CC, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*. 2011; 476(7359):214–9. Epub 2011/08/13. <https://doi.org/10.1038/nature10251> PMID: 21833088; PubMed Central PMCID: PMC3182531.
 68. International Multiple Sclerosis Genetics C, Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kempainen A, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature genetics*. 2013; 45(11):1353–60. Epub 2013/10/01. <https://doi.org/10.1038/ng.2770> PMID: 24076602; PubMed Central PMCID: PMC3832895.
 69. Sims R, van der Lee SJ, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, et al. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nature genetics*. 2017; 49(9):1373–84. Epub 2017/07/18. <https://doi.org/10.1038/ng.3916> PMID: 28714976; PubMed Central PMCID: PMC5669039.
 70. Kivipelto M, Mangialasche F, Ngandu T. Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. *Nat Rev Neurol*. 2018; 14(11):653–66. <https://doi.org/10.1038/s41582-018-0070-3> PMID: 30291317.
 71. Eid A, Mhatre I, Richardson JR. Gene-environment interactions in Alzheimer's disease: A potential path to precision medicine. *Pharmacol Ther*. 2019. <https://doi.org/10.1016/j.pharmthera.2019.03.005> PMID: 30877021.
 72. Xu W, Tan L, Wang HF, Jiang T, Tan MS, Tan L, et al. Meta-analysis of modifiable risk factors for Alzheimer's disease. *Journal of neurology, neurosurgery, and psychiatry*. 2015; 86(12):1299–306. <https://doi.org/10.1136/jnnp-2015-310548> PMID: 26294005.
 73. Cheng CK, Tsao YC, Su YC, Sung FC, Tai HC, Kung WM. Metabolic Risk Factors of Alzheimer's Disease, Dementia with Lewy Bodies, and Normal Elderly: A Population-Based Study. *Behavioural neurology*. 2018; 2018:8312346. <https://doi.org/10.1155/2018/8312346> PMID: 29971140; PubMed Central PMCID: PMC6008802.
 74. Solomon A, Mangialasche F, Richard E, Andrieu S, Bennett DA, Breteler M, et al. Advances in the prevention of Alzheimer's disease and dementia. *J Intern Med*. 2014; 275(3):229–50. <https://doi.org/10.1111/joim.12178> PMID: 24605807; PubMed Central PMCID: PMC4390027.
 75. Irimata KE, Dugger BN, Wilson JR. Impact of the Presence of Select Cardiovascular Risk Factors on Cognitive Changes among Dementia Subtypes. *Current Alzheimer research*. 2018. <https://doi.org/10.2174/1567205015666180702105119> PMID: 29962344.
 76. Ngandu T, Lehtisalo J, Solomon A, Levalahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet*. 2015; 385(9984):2255–63. [https://doi.org/10.1016/S0140-6736\(15\)60461-5](https://doi.org/10.1016/S0140-6736(15)60461-5) PMID: 25771249.
 77. Larsson SC, Traylor M, Malik R, Dichgans M, Burgess S, Markus HS, et al. Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ*. 2017; 359:j5375. <https://doi.org/10.1136/bmj.j5375> PMID: 29212772; PubMed Central PMCID: PMC5717765.
 78. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shiri M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience*. 2019; 22(3):343–52. <https://doi.org/10.1038/s41593-018-0326-7> PMID: 30718901.
 79. Kathiresan S, Melander O, Guiducci C, Surti A, Burt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature genetics*. 2008; 40(2):189–97. <https://doi.org/10.1038/ng.75> PMID: 18193044; PubMed Central PMCID: PMC2682493.
 80. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010; 466(7307):707–13. <https://doi.org/10.1038/nature09270> PMID: 20686565; PubMed Central PMCID: PMC3039276.

81. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. *The New England journal of medicine*. 2007; 357(5):443–53. <https://doi.org/10.1056/NEJMoa072366> PMID: 17634449; PubMed Central PMCID: PMC2719290.
82. Teumer A, Qi Q, Nethander M, Aschard H, Bandinelli S, Beekman M, et al. Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. *Aging cell*. 2016; 15(5):811–24. <https://doi.org/10.1111/acer.12490> PMID: 27329260; PubMed Central PMCID: PMC5013013.
83. Fritsche LG, Igl W, Bailey JN, Grassmann F, Sengupta S, Bragg-Gresham JL, et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nature genetics*. 2016; 48(2):134–43. <https://doi.org/10.1038/ng.3448> PMID: 26691988; PubMed Central PMCID: PMC4745342.
84. Ratnayaka JA, Serpell LC, Lotery AJ. Dementia of the eye: the role of amyloid beta in retinal degeneration. *Eye*. 2015; 29(8):1013–26. <https://doi.org/10.1038/eye.2015.100> PMID: 26088679; PubMed Central PMCID: PMC4541342.
85. Cerman E, Eraslan M, Cekic O. Age-related macular degeneration and Alzheimer disease. *Turkish journal of medical sciences*. 2015; 45(5):1004–9. <https://doi.org/10.3906/sag-1406-146> PMID: 26738339.
86. Ermilov VV, Tiurenkov IN, Nesterova AA, Zagrebin VA. [Alzheimer's disease and geriatric eye diseases in the aspect of amyloid genesis]. *Arkhiv patologii*. 2013; 75(2):37–42. PMID: 24006764.
87. Shoda C, Kitagawa Y, Shimada H, Yuzawa M, Tateno A, Okubo Y. Relationship of Area of Soft Drusen in Retina with Cerebral Amyloid-beta Accumulation and Blood Amyloid-beta Level in the Elderly. *Journal of Alzheimer's disease: JAD*. 2018; 62(1):239–45. <https://doi.org/10.3233/JAD-170956> PMID: 29439351.
88. Chang LY, Lowe J, Ardiles A, Lim J, Grey AC, Robertson K, et al. Alzheimer's disease in the human eye. Clinical tests that identify ocular and visual information processing deficit as biomarkers. *Alzheimer's & dementia: the journal of the Alzheimer's Association*. 2014; 10(2):251–61. <https://doi.org/10.1016/j.jalz.2013.06.004> PMID: 24011928.
89. Mutlu U, Colijn JM, Ikram MA, Bonnemaier PWM, Licher S, Wolters FJ, et al. Association of Retinal Neurodegeneration on Optical Coherence Tomography With Dementia: A Population-Based Study. *JAMA neurology*. 2018. <https://doi.org/10.1001/jamaneurol.2018.1563> PMID: 29946702.
90. Armstrong RA. Alzheimer's Disease and the Eye. *J Optom*. 2009;(2):103–11.
91. Spierer O, Fischer N, Barak A, Belkin M. Correlation Between Vision and Cognitive Function in the Elderly: A Cross-Sectional Study. *Medicine*. 2016; 95(3):e2423. <https://doi.org/10.1097/MD.0000000000002423> PMID: 26817872; PubMed Central PMCID: PMC4998246.
92. Williams KM, Hysi PG, Yonova-Doing E, Mahroo OA, Snieder H, Hammond CJ. Phenotypic and genotypic correlation between myopia and intelligence. *Scientific reports*. 2017; 7:45977. <https://doi.org/10.1038/srep45977> PMID: 28383074; PubMed Central PMCID: PMC5382686.