BMC Medical Genomics

RESEARCH Open Access

Downward bias in the association between *APOE* and Alzheimer's disease using prevalent and by-proxy disease sampling in the *All of Us* research program

Clayton O. Mansel[1*], Valentina Ghisays[2], Jonathan D. Mahnken[3], Russell H. Swerdlow[4], Eric M. Reiman[2,5], Jason H. Karnes[6], Joshua C. Denny[7,8] and Olivia J. Veatch[1,4,9]

Abstract

Background Recent genome-wide association studies for Alzheimer's Disease and related dementias (ADRD) have increased statistical power via larger analysis datasets from biobanks by (1) including non-age-matched controls and prevalent cases, and/or (2) including individuals who report a family history of ADRD as proxy cases. However, these methods have the potential to increase noise and distort genetic associations which are important for genomic-informed prevention and treatment of ADRD. Here, we sought to understand how the effect sizes of genetic associations in ADRD could be sensitive to these methodological choices, using *APOE* genotypes as an example.

Methods Participants in the *All of Us* Research Program over the age of 49 at enrollment ($n = 229{,}722$) were assigned one of four categories: incident ADRD (developed after enrollment in *All of Us*), prevalent ADRD (present on enrollment), proxy ADRD (participant noted a family history of ADRD), and control (no history or diagnosis of ADRD). ADRD diagnoses were determined using available electronic health records and *APOE* genotype was determined using whole-genome sequencing. Effect sizes for the associations between *APOE* risk alleles and ADRD diagnoses were compared using polychotomous logistic regression and presented as adjusted generalized ratios (AGR).

Results The mean age of the cohort was 64 ± 9 years, and it was 57% female; 65% clustered predominantly with European genetic reference populations. Among the participants, 733 (0.3%) had prevalent ADRD, 684 (0.3%) had incident ADRD, and 19,186 (8.4%) reported a family history of ADRD (proxy ADRD). The effect size for *APOE* ε4 heterozygote was similar for proxy ADRD (AGR [95% CI]: 2.10 [1.96–2.24]) but attenuated for prevalent ADRD (1.38 [1.17–1.63]) compared to incident ADRD (2.13 [1.81–2.50]). For *APOE* ε4 homozygotes, the effect sizes were significantly attenuated in both proxy (3.53 [2.93–4.26]) and prevalent (3.12 [2.20–4.45]) ADRD. Furthermore, *APOE* and ADRD association effect sizes increased when restricting the control (no ADRD) group to older age brackets.

*Correspondence:
Clayton O. Mansel
cmansel@kumc.edu

Full list of author information is available at the end of the article

Conclusions Our study highlights how genetic associations with ADRD can be sensitive to how cases are defined in biobanks like *All of Us*, with effect sizes downwardly biased when using prevalent or by-proxy cases compared to incident cases.

Keywords Alzheimer's Disease, Proxy case, *APOE*, Incidence, Prevalence, Biobank

Background

Alzheimer's Disease and related dementias (ADRD) are a devastating class of progressive neurodegenerative diseases of which over 7 million are affected in the U.S. alone and there is no cure [1, 2]. AD in particular has a long prodromal preclinical phase in which biomarkers such as glucose hypometabolism and amyloid-beta plaques appear in the brain years before noticeable cognitive decline [3]. This has led to efforts for early detection and prevention, including genomic-informed risk assessments given that ADRD is highly heritable [4]. These genomic-informed risk assessments rely on inferences drawn largely from case-control and observational study designs and are thus subject to well-documented selection and survival bias [5–7]. In an effort to increase sample size and thus statistical power, recent genome-wide association studies (GWAS) for ADRD have incorporated observational data from large biobanks such as the UK Biobank and *All of Us* Research Program [8, 9]. This has led to controversies including the use of so-called "proxy cases" in which an individual with a self-reported family history of ADRD in a first degree relative, but no evidence of cognitive decline, is included as a proxy case [10]. For example, in a recent GWAS for ADRD, 49,275 out of 111,326 total cases were proxy cases from the UK Biobank [8].

The proportion of individuals in modern ADRD GWAS that come from biobanks is quickly dwarfing ADRD-specific cohorts, despite their potential biases [5, 11, 12] and challenges with identifying clinical ADRD cases using electronic healthcare records (EHR) [13]. When an individual enrolls in a biobank such as the UK Biobank, he or she often consents to adding years of previous EHR records [14], including potential diagnoses of ADRD. Because individuals cannot have cognitive impairment when they give informed consent, this can lead to a cohort of "prevalent cases" (i.e., cases diagnosed prior to enrollment in a biobank) that are healthier or decline slower than the general population. In addition to clinical cases, biobank proxy cases can be plagued by non-participation bias in optional family health history questionnaires—which have been shown to have genetic predictors of participation, including variants associated with ADRD [15]. Previous studies have shown downward bias in the effect size estimates for ADRD risk alleles when using proxy cases in the UK Biobank [16, 17]. We note that the UK Biobank has been shown to be healthier than the general UK population [18] while participants in the *All of Us* Research Program have higher disease prevalence compared to the general U.S. population [19]. It is currently unknown if this alleviates the downward bias of prevalent and by-proxy disease sampling shown in previous studies in the UK Biobank.

Among the most consistently replicated GWAS loci for ADRD is the *APOE* genotype which is a strong predictor of disease liability, currently explaining more phenotypic variability than polygenic scores constructed from hundreds of small-effect loci derived from GWAS [4, 20, 21]. Individuals with one copy of the *APOE* ε4 allele have a 3–4 fold increases risk for AD while two copies of ε4 carries a 12–15 fold increased risk [22, 23]. As such, *APOE* genotyping is currently recommended as a part of evidence-based ADRD prevention [24] and is likely to be a part of genomic-informed ADRD care prior to polygenic scores which require much more optimization [21, 25, 26]. Given these previous studies, we sought to understand how genetic associations in ADRD could change depending on the definition of disease, using *APOE* as an example for an effect likely present in many genes. We leverage the *All of Us* Research Program, a multisite, U.S.-based biobank, including many individuals from backgrounds whose health outcomes have not been extensively studied in biomedical research [27]. Our study provides early insight into whether the biases of proxy cases and incident vs. prevalent disease sampling is present in one of the most populous and diverse biobanks in the world.

Methods

All of Us and inclusion criteria

The *All of Us* Research Program is a longitudinal, multisite biorepository of participants' health data including surveys, EHR, and other data linked to genomic data including whole-genome sequencing (WGS) [27, 28]. It began enrollment nationally in 2018. EHR data from dozens of healthcare providers across the U.S. are harmonized using the Observational Medical Outcomes Partnership (OMOP) Common Data Model [29]. The data were acquired under the Data Use and Registration Agreement between the University of Kansas Medical Center and *All of Us.* As of March 2025, 633,540 participants had completed initial consent and enrollment for the Curated Data Repository v8. Participants were included in the present study if they (1) had available EHR data (2) had WGS data for *APOE* genotyping (3) were assigned male or female at birth and (4) were over

the age of 49 at enrollment (enrollment date ranges from 05-31-2017 to 09-14-2023). The total sample size meeting these criteria was 258,693 participants.

Definition of phenotypes

To label ADRD cases, a previously validated computable phenotype was used which specified an individual as a case if they had five or more visits with matching ICD-9/ICD-10 codes or one visit with a matching drug prescription in their EHR record [30, 31]. Exploratory data analysis revealed that 1,321 ADRD cases had a matching drug prescription but no visit with a matching ICD-9/ICD-10 code—including many with drug prescriptions greater than 10 years prior to enrollment in *All of Us* (Supplementary Fig. 1). We therefore excluded these individuals from our analyses. This means, for our study, ADRD cases could have either 5 or more visits with matching ICD-9/ICD-10 codes or a matching drug prescription and at least 1 visit with a matching ICD-9/ICD-10 code.

To determine incident versus prevalent ADRD, we used the earliest date that the individual fully met criteria (e.g., the fifth visit with matching codes). If that date occurred prior to enrolling in *All of Us*, the individual was labelled as prevalent ADRD. If it occurred after, they were labelled as incident ADRD. All individuals with an ADRD diagnosis within 1 year of enrollment were excluded because we could not be confident in their incident vs. prevalent status given the gradual onset of ADRD and lag between symptom onset and confident diagnosis. For proxy ADRD, the Personal and Family Health History survey was used which was filled out by (26.5%) of the participants included in this study and was designed for use in *All of Us* (see supplementary file). If the individual answered "mother", "father" or "sibling" to the question "Including yourself, who in your family has had dementia (includes Alzheimer's, vascular, etc.)?" they were labelled as proxy ADRD. All remaining individuals who did not meet criteria for incident, prevalent, or proxy ADRD were labelled control.

Participants' age was defined as age at ADRD diagnosis. Sex assigned at birth was self-reported. For parental sex, father was assumed to by male and mother was assumed to be female. Global genetic ancestry reference population similarity was provided by *All of Us* for each participant and was inferred using a random forest classifier trained on samples from 1,000 Genomes and Human Genome Diversity Project samples [28]. We assign individuals to genetically defined ancestry groups based on their similarity to global reference populations, with the understanding that these categories are imperfect proxies for the continuous and admixed nature of human genetic variation. These groupings are used solely to enable analyses of genetic variation across diverse populations and are not intended to represent social or racial identities. EHR length was determined by the difference between the earliest and latest visit date across multiple OMOP Common Data Model tables.

APOE genotyping

All of Us provides qualified researchers access to WGS data that has undergone quality control procedures outlined extensively elsewhere [28]. This study specifically used the Allele Count/Allele Frequency (ACAF) unphased callset comprised of single nucleotide polymorphisms (SNPs) meeting a population-specific allele frequency threshold of > 0.01 or allele count > 100. *APOE* SNPs were extracted using *PLINK* 1.9 and *APOE* genotypes were determined based on the combined genotypes at rs429358 and rs7412 [32] and imputed using R as previously described [33]. Because multiple phenotype groups (e.g., incident ADRD) contained *APOE* genotype counts < 20, we binned *APOE* genotypes into three groups: *APOE* ε4 absent (ε2/ ε2, ε2/ ε3, ε3/ ε3), *APOE* ε4 heterozygote (ε2/ ε4, ε3/ ε4), and *APOE* ε4 homozygote (ε4/ ε4).

Statistical analysis

Summary statistics compared the mean (standard deviation) for continuous variables and n (%) for categorical variables. Group level comparisons were performed via Pearson chi-square tests for categorical variables and one-way analyses of variance for continuous variables. To determine the association between *APOE* genotype and ADRD status, we modeled the nominal response variable ADRD status using polychotomous logistic regression with a generalized logit function [34]. We included age at diagnosis, sex assigned at birth (or parental sex in the proxy cases), and global genetic ancestry as covariates considering extensive evidence of these variables interacting with the *APOE* genotype to influence risk for AD [22, 35–37] For proxy cases, parental age was unavailable in *All of Us*. The results are presented as adjusted generalized ratios (AGRs) which can be interpreted as ratio of the likelihoods of each definition of ADRD compared to control which was the referent group. Notably, this approach allows for non-equal sample sizes in different components of these ratios (e.g., incident cases vs. controls). The AGRs for *APOE* genotype and proxy ADRD were adjusted by a factor of 2 to account for the genetic relatedness between parent and child. This adjustment has been shown to be accurate for case/control odds ratios for alleles with small to medium effect sizes [10]. To ensure that this adjustment appropriate for our study, we compared adjusting the proxy AGR by a factor of 2 with another method derived directly from the allele frequencies in the cases and controls as shown in Liu et al., 2017 [10] and found that the adjusted AGR is similar (Supplementary Table 1).

Recent ADRD GWAS have stopped using age-matched controls. For example, the most recent large GWAS had controls with a mean age at assessment of 57.9 compared to 67.2 for the cases [8]. Given this, we also ran our analyses using different age cut-offs in the control group (e.g., > 49yrs, > 59yrs, > 69yrs, > 79yrs). All analyses were performed in R v4.4.0 and RStudio v2024.04.0.

Results

Of the 229,722 individuals included in the analysis dataset, 685 (0.3%) had incident ADRD, 733 (0.3%) had prevalent ADRD and 19,186 (8.4%) had proxy ADRD (Table 1). The overall mean (standard deviation) EHR length was 11.1 (9.7) years. Control, incident, and proxy ADRD majority (> 50%) assigned female at birth with the proxy ADRD subgroup at a notably higher proportion (64.7%) (Table 1). The mean age of the subgroups ranged from 63.9 to 72.7 years. The subgroup with the highest mean age was incident ADRD (72.7 yrs) followed by prevalent ADRD (70.6 yrs), proxy ADRD (65.5 yrs) and then controls (63.9 yrs) (Table 1). The proportion of participants carrying 1 *APOE* ε4 allele varied widely from 24.1% of the control group to 34.6% of the incident ADRD subgroup (Table 1). The incident ADRD subgroup also had the highest proportion of *APOE* ε4 homozygotes (7.5%) compared to the other groups (Table 1). All subgroups were mostly comprised of individuals with European ancestry. The control group had the smallest proportion of European ancestry at 63.0% followed by prevalent ADRD at 67.9%, incident ADRD at 68.3%, and proxy ADRD with 85.9% of individuals (Table 1).

Table 1 Age, Sex, *APOE* genotype, and Genetic Ancestry for 229,722 *All of Us* participants

	Control	Incident ADRD	Prevalent ADRD	Proxy ADRD	*p* value
n (%)	209,119 (91.0)	684 (0.3)	733 (0.3)	19,186 (8.4)	
Demographics					
Female Sex (%)	118,371 (56.6)	370 (54.1)	339 (46.2)	12,333 (64.3)	< 0.001
Age (mean [sd])	63.9 (9.0)	72.7 (8.9)	70.6 (10.0)	65.5 (7.9)	< 0.001
APOE Genotype					
APOE ε4 absent	154,446 (73.9)	396 (57.9)	498 (67.9)	13,025 (67.9)	< 0.001
APOE ε4 Heterozygote	50,331 (24.1)	237 (34.6)	201 (27.4)	5616 (29.3)	
APOE ε4 Homozygote	4342 (2.1)	51 (7.5)	34 (4.6)	545 (2.8)	
Global Genetic Ancestry					
African	43,131 (20.6)	113 (16.5)	112 (15.3)	1279 (6.7)	< 0.001
American Admixed/Latino	28,681 (13.7)	87 (12.7)	104 (14.2)	1082 (5.6)	
East Asian	3480 (1.7)	**	**	240 (1.3)	
European	131,801 (63.0)	467 (68.3)	498 (67.9)	16,478 (85.9)	
Middle Eastern	533 (0.3)	**	**	38 (0.2)	
South Asian	1493 (0.7)	**	**	69 (0.4)	

Total sample size is 229,722 participants in the *All of Us* Research Program who identified as male or female at birth and were over the age of 49 at enrollment. Incident and prevalent ADRD defined based on whether the participant met diagnostic criteria before or after they enrolled in *All of Us* (see methods). Proxy ADRD defined based on whether the participant reported a history of ADRD in a first-degree relative. Categorical variables are presented as n (%) and continuous variables presented as mean (standard deviation). *p* values are unadjusted and correspond to X^2 tests for categorical variables and one-way analyses of variance (ANOVA) for continuous variables

**Participant counts < 20 suppressed in accordance with *All of Us*'s Data and Statistics Dissemination Policy

The relative ratio of the proportion belonging to each ADRD status group divided by the proportion of controls among *APOE* ε4 heterozygotes (relative to ε4 absent) and homozygotes (relative to ε4 absent) is shown in Fig. 1 and Supplementary Tables 2 and is presented as adjusted generalized ratios (AGRs) (see Methods). For the *APOE* ε4 heterozygotes, the adjusted generalized ratio (AGR) (95% CI) was highest for incident ADRD at 2.13 (1.81–2.50) and was similar to proxy ADRD which had an AGR of 2.10 (1.96–2.24). In comparison, prevalent ADRD had a lower AGR of 1.38 (1.17–1.63) (Fig. 1; Supplementary Table 2). For *APOE* ε4 homozygotes, the AGR was highest for incident ADRD at 6.39 (4.74–8.62); for prevalent ADRD the AGR was 3.13 (2.20–4.45); for proxy ADRD the AGR was 3.53 (2.93–4.26) (Fig. 1; Supplementary Table 2). In summary, we note that in *APOE* ε4 heterozygotes, the prevalent ADRD AGR was significantly lower than the incident ADRD AGR (95% confidence interval limits). In *APOE* ε4 homozygotes, both the prevalent and proxy AGR was significantly lower than the incident ADRD AGR (95% confidence interval limits).

To account for the fact that some control participants may be mislabeled because of a short history of EHR visits, we restricted our dataset to only individuals with greater than one year of EHR visits ($n = 193{,}634$) and found that this did not affect the association between *APOE* and ADRD status (Supplementary Fig. 2). We also restricted our dataset to just individuals who completed the Family Health History survey ($n = 60{,}598$) to prevent misattribution of control individuals who should be labelled as proxy ADRD (Supplementary Fig. 3). We found that, again, both proxy and prevalent ADRD AGRs were significantly smaller compared to incident ADRD. Because parental/sibling age was not available in *All of Us*, we performed an additional analysis not controlling for age and observed that the the AGRs were not significantly different from our main analysis (Supplementary Fig. 4). We also included the 1,321 ADRD cases that had only a drug prescription but no matching ICD-9/10 codes in the prevalent ADRD group and found that this further depressed the *APOE* ε4 homozygote AGR compared to

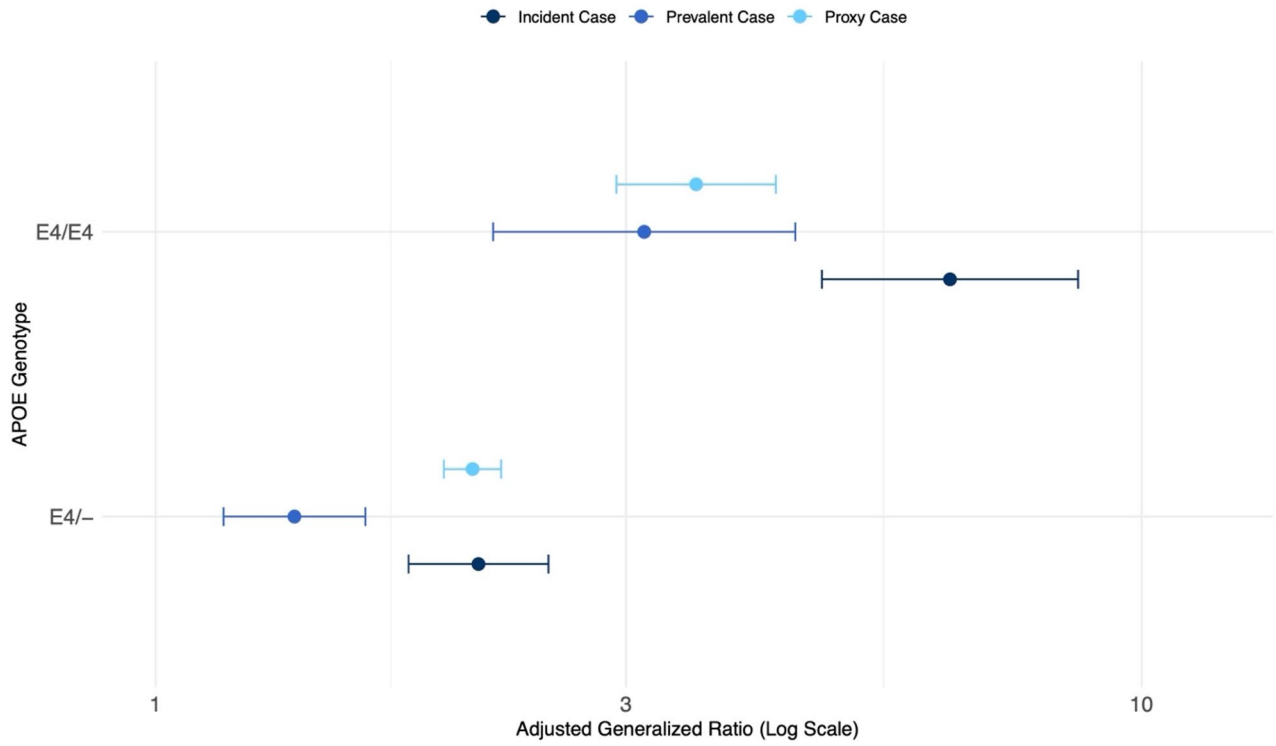

Fig. 1 Adjusted generalized ratios for the association between *APOE* ε4 genotype and ADRD status in *All of Us*. Total sample size is 229,722 participants in the *All of Us* Research Program who identified as male or female at birth and were over the age of 49 at enrollment. Incident and prevalent ADRD defined based on whether the participant met diagnostic criteria before or after they enrolled in *All of Us* (see Methods). Proxy ADRD defined based on whether the participant reported a history of ADRD in a first-degree relative. Adjusted generalized ratio's plotted (log scale) and derived from a polychotomous logistic regression with ADRD status as the outcome (outcome referent = control) adjusted for age, sex, and global genetic ancestry. Proxy ADRD AGRs were adjusted by a factor of 2 to account for the genetic relatedness between the parent and child

incident ADRD and did not significantly change the ε4 heterozygote prevalent AGR (Supplementary Fig. 5).

Because ADRD has a long prodromal period, it can be hard to delineate between true incident vs. prevalent diagnosis. To address this, we excluded all ADRD cases diagnosed within one year of *All of Us* enrollment (see methods) in our main analysis. In a sensitivity analysis, we included these individuals in a fourth "uncertain timing" category and note that for both the *APOE* ε4 heterozygote and homozygote, the AGR falls in between the incident and prevalent ADRD AGR (Supplementary Fig. 6). Even though we adjusted our analyses for global genetic ancestry, we also performed our analyses in the European subgroup (since it was the largest) and found that the difference between incident and prevalent AGRs were slightly attenuated (Supplementary Fig. 7). One of the purported benefits of using proxy disease sampling is to help alleviate misclassification bias from controls who are too young to receive a diagnosis of ADRD. To address this, we reanalyzed the association between *APOE* and ADRD status using different age cut-offs in the control group (e.g., > 49yrs, > 59yrs, > 69yrs, > 79yrs) and showed that, even for proxy cases, the older control groups numerically increased the AGRs (Fig. 2).

Discussion

Our study demonstrates that using proxy and prevalent case definitions of ADRD attenuates the observed association between *APOE* genotypes and disease risk when compared to incident cases. To our knowledge, this is the first time bias in genetic association studies has been studied in *All of Us*. With over 20,000 researchers using the *All of Us* researcher workbench, it is important to understand how the definitions of cases and controls can affect downstream analyses. Proxy case definitions based on a family history of ADRD is a common approach in modern ADRD GWAS and other case/control study designs [8, 38–40]. We showed that the association between being an *APOE* heterozygote and ADRD was similar between proxy and incident cases, indicating that this method may be appropriate for small to medium genetic effect sizes, as has been argued previously [10]. However, for *APOE* ε4 homozygotes, the AGR for proxy cases was attenuated even after multiple correction methods. We note that these correction methods [10] assume a log-additive (per-allele) model and thus are likely inappropriate for large effect-size genotypes like ε4/ε4. Moreover, reporting/survival bias of parental history can further depress proxy effect sizes. It has been argued

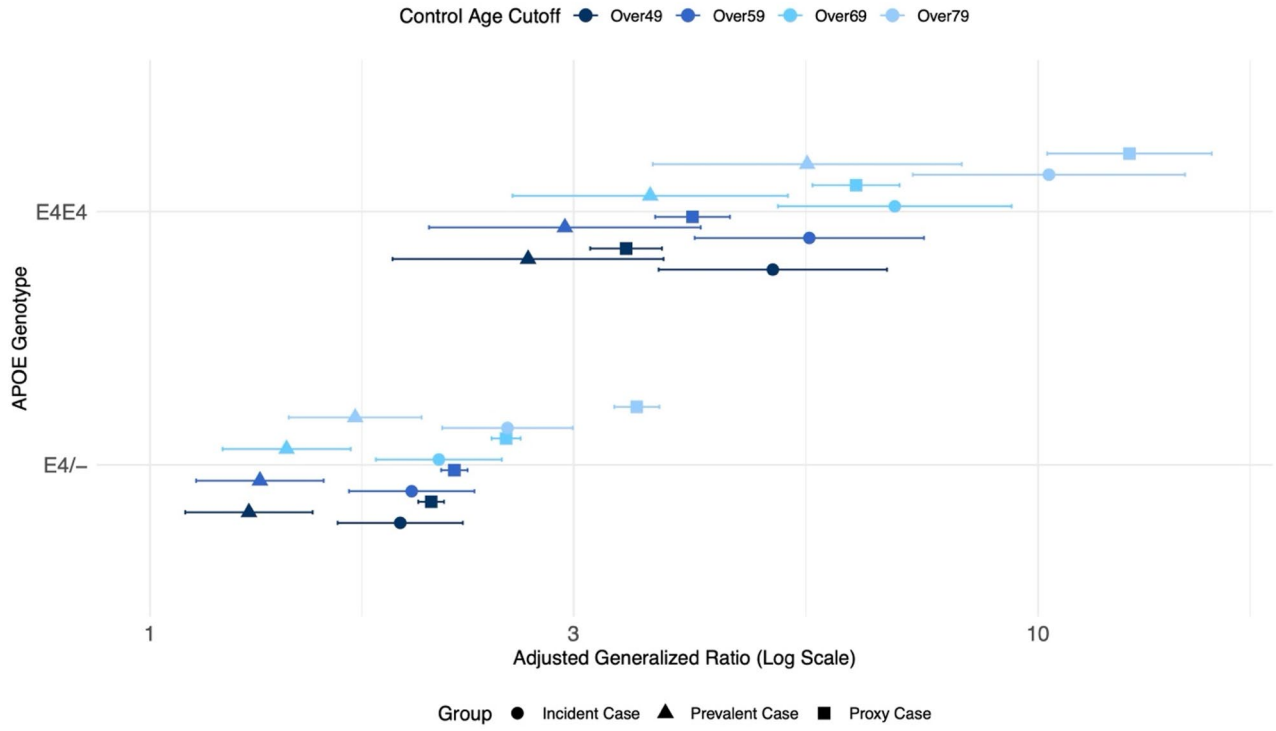

Fig. 2 Adjusted generalized ratios for the association between *APOE* ε4 genotype and ADRD status in *All of Us*. Total sample size is 229,722 participants in the *All of Us* Research Program who identified as male or female at birth and were over the age of 49 at enrollment. Incident and prevalent ADRD defined based on whether the participant met diagnostic criteria before or after they enrolled in *All of Us* (see methods). Proxy ADRD defined based on whether the participant reported a history of ADRD in a first-degree relative. Adjusted generalized ratio's plotted (log scale) and derived from a polychotomous logistic regression with ADRD status as the outcome (outcome referent = control) adjusted for age, sex, and global genetic ancestry. Data are stratified to include only non-demented participants over the age of 49 yrs ($n = 229{,}722$), 59 yrs ($n = 154{,}138$), 69 yrs ($n = 77{,}854$), and 79 yrs ($n = 31{,}519$)

previously that proxy disease sampling underestimates genetic associations in ADRD in the UK Biobank, a finding corroborated by our study [16, 17, 41].

Strikingly, for both *APOE* ε4 homozygotes and heterozygotes, the association between *APOE* and ADRD was significantly lower using prevalent cases compared to incident cases. This could be because the prevalent ADRD cases are healthier than the incident cases, a bias that our study shows is present in *All of Us* despite it being a sicker cohort than the UK Biobank and the U.S. general population [19]. While biobanks like *All of Us* can take advantage of decades of EHR data on enrolled participants, our study highlights how this can underestimate the association of genetic risk factors like *APOE* in ADRD, especially when evaluating prevalent conditions whose presence may be negatively correlated with the probability of enrolling.

Adjusted generalized ratios (AGRs), odds ratios (ORs), and relative risk (RR) are not directly comparable. In the case of ORs and RR, heterogeneous cases and controls are each collapsed into two binary groups. Our study highlights how heterogeneity in case definitions can change the effect size of genetic associations studies. Nonetheless, the effect size of our association between *APOE* and incident ADRD cases (AGR: 2.13 for ε4/- and 6.39 for ε4/ε4) appears closest to epidemiologic evidence estimating 3–4 fold increased risk for *APOE* ε4 heterozygotes and 12–15 fold increased risk for *APOE* ε4 homozygotes [22, 23]. Attenuation of our AGR compared to previous reported ORs could also be explained by our use of EHR-defined ADRD which captures both AD and AD-related dementias, the latter of which has a weaker association with *APOE* genotype [42]. Furthermore, we showed that all effect sizes are greater when filtering the control group to older cohorts. This highlights a major challenge for ADRD case-control study designs in which controls can be incorrectly labelled given that the disease starts long before cognitive symptoms appear. Modern blood-based biomarkers could help remedy this issue by at least excluding controls who are positive for beta amyloid in the brain [43].

For the purposes of discovering novel biological pathways to prioritize for further functional validation, the use of biobanks and proxy cases in GWAS may be sensible [10]. But for genomic-informed risk assessments, in which the magnitude of the association between genotype and phenotype affects the potential absolute risk reduction of costly interventions [44], we argue that more

precision is needed. Proponents of proxy case genetic association studies (GWAX) argue that the effect sizes are similar between GWAS and GWAX for large-effect loci [8, 10] and that there is not significant heterogeneity of individual single nucleotide polymorphism (SNP) effect sizes [17]. However, using mendelian randomization, others have found significant genetic heterogeneity when using clinical ADRD vs. proxy ADRD [41]. A recent GWAS found that only 0.6% of loci identified using proxy cases from *All of Us* were present when using clinical AD cases—including many with the opposite direction of effect [9]. Our study demonstrated that prevalent cases in *All of Us* have a downward bias in genetic associations with ADRD to a similar magnitude as proxy cases. Thus, both sampling strategies may be inappropriate for building genomic-informed risk assessments.

Given the decreasing cost of genotyping and genome sequencing, genomic-informed medicine presents an enticing opportunity to improve risk stratification and tailor treatments for many diseases. For instance, previous work in atherosclerotic cardiovascular disease has proven the utility of genomic-informed risk assessments which have been shown to successfully motivate behavior-change in almost half of a recent prospective cohort [45]. The same potential exists for ADRD prevention [46]. However, ADRD faces unique challenges in genetic association studies given that it is a disease of late life, has a long prodromal phase, and has no validated biomarker that is widely used for accurate diagnosis outside of specialist memory centers. GWAS including proxy cases and prevalent biobank cases may be distorting important genetic associations needed to build effective genomic-informed prediction models and thus may be doing more harm than good. The inferences drawn from GWAS matter a lot to this effort as evidenced by a recent benchmarking study which found that the choice of GWAS had the greatest effect of any methodological choice in genomic-informed ADRD prediction modelling [47].

There are potential remedies to the biases demonstrated in this study. Future genetic association studies could incorporate survival data using time-matched sampling or weighting by the time observed between diagnosis and sampling [48, 49]. Second, if the probability of being enrolled in the biobank could be estimated, then inverse probability weighting could be used to generate less biased estimates [50]. Third, each passing year provides more and more incident cases in biobanks, remedying this issue over time. However, any statistical correction of bias will pale in comparison to resolving the issue at the GWAS itself. Modern ADRD GWAS have larger and larger sample sizes but smaller and smaller single nucleotide polymorphism (SNP) heritability estimates [51]. We argue that GWAS could be tailored to the specific goal in mind. For genomic medicine inferences, GWAS should not use proxy cases, use age-matched controls, use incident cases to minimize selection biases where possible, and, given the availability of accurate biomarkers now for ADRD [43], recommend use of biomarkers where possible to minimize misclassification of case-control status. At the very least, the summary statistics for associations excluding proxy cases should be made available to the research community, something that was lacking in the most recent large ADRD GWAS [8].

Our study was limited to the use of EHRs to infer ADRD case/control status, which is not as accurate as diagnoses at specialist memory centers and prevented us from distinguishing between AD and other closely related dementias which could lead to an underestimation of *APOE*'s association with disease. We were also not able to evaluate *APOE* ε2 allele-specific effects due to small sample sizes. In addition, we did not have a large enough sample size to stratify by race/ethnicity and thus could not determine whether the prevalent and proxy bias differed in non-White or Hispanic individuals. Notably, we had a much higher proportion of individuals of European ancestry in the proxy case cohort. We hypothesize that this could be because individuals identifying as White were more likely to fill out the Family Health History survey. Another limitation of our study was that *All of Us* did not have parental/sibling age available to mitigage survivial bias in the proxy ADRD cases as has been done in the UK Biobank [38]. Finally, it may be difficult to distinguish prevalent from incident ADRD cases given its long prodromal period. However, we still show a difference in effect size despite the overlap between groups potentially biasing the result toward the null. It remains possible that the effects observed in this work are unique to *APOE* given it's relatively high prevalence in the population. Future work should attempt to replicate our findings using other genes associated with ADRD and other biobanks or datasets.

Conclusions

Despite these limitations, our study highlights a fundamental question in genomic-informed medicine: how do we infer an association between a gene and a disease for use in the clinic? Our study showed how these associations are sensitive to changes in methodology. We showed that using proxy and prevalent case definitions of ADRD attenuates the observed association between *APOE* genotypes and disease risk when compared to incident cases. We also showed that using different age cut-offs in the controls affects the association between *APOE* genotypes and ADRD risk. Early personalized interventions for ADRD are likely to be expensive and laborious [24], thus there exists a need for greater precision to determine who is likely to benefit the most.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-026-02362-1.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

We acknowledge All of Us participants for their contributions, without whom this research would not have been possible. We also thank the National Institutes of Health's All of Us Research Program for making available the participant and genetic data examined in this study. The All of Us Research Program is supported by the National Institutes of Health, Office of the Director: Regional Medical Centers: 1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1OT2 OD026556; 1 OT2 OD026550; 1 OT2 OD 026552; 1 OT2 OD026553; 1 OT2 OD026548; 1OT2 OD026551; 1 OT2 OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085U; Data and Research Center: 5 U2C OD023196; Biobank: 1 U24OD023121; The Participant Center: U24 OD023176; Participant Technology Systems Center: 1U24 OD023163; Communications and Engagement: 3 OT2 OD023205; 3 OT2 OD023206; and Community Partners: 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2OD025276. In addition, the *All of Us* Research Program would not be possible without the partnership of its participants.

Authors' contributions

COM, JCD, and OJV were involved in the initial conception, design, and execution of the study. VG, EMR, JHK, and SHR assisted COM in the analysis and interpretation of the results. JDM provided expertise in statistical analysis. CM and OJV drafted the initial manuscript. All authors read and approved the final manuscript.

Funding

This work was performed with funding from NIH P30 AG072973 (CM, JDM) and T32 AG078114 (CM); NIH P20 GM130423 (OJV), 5UL1TR002366 (OJV), and P30 AG035982 (OJV, JDM & RHS). JHK is funded by the NIH's Office of the Director under awards 1OT2OD026549 and OT2OD036485 and the National Heart, Lung, and Blood Institute (NHLBI) under awards R21HL172036, R01 HL156993, and R01 HL158686. This work was also supported by the NHGRI Intramural Research Program ZIA HG200417 (JCD).

Data availability

All data used in this study can be accessed by qualified researchers at allofus.nih.gov. Source code for these analyses is available within the *All of Us* ResearcherWorkbenchat (https://workbench.researchallofus.org/workspaces/aou-rw-55fcf7f4/alzheimersdiseaseapoeprevalencevsincidencev8/data).

Declarations

Ethics approval and consent to participate

Informed consent for all participants is conducted in person or through an eConsent platform that includes primary consent, HIPAA Authorization for Research use of EHRs and other external health data, and Consent for Return of Genomic Results. The protocol was reviewed by the Institutional Review Board (IRB) of the All of Us Research Program. The All of Us IRB follows the regulations and guidance of the NIH Office for Human Research Protections for all studies, ensuring that the rights and welfare of research participants are overseen and protected uniformly.

This study used de-identified data from the All of Us Research Program, accessed via the Researcher Workbench. According to U.S. federal regulations (45 CFR 46), research using de-identified data is not considered human subjects research and is exempt from IRB review; this determination was confirmed by the University of Kansas Medical Center Institutional Review Board. All participants in the All of Us Research Program provided written informed consent at the time of enrollment. The All of Us Research Program protocol was approved by its designated ethics oversight committees and conducted in accordance with the principles of the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Cell Biology and Physiology, University of Kansas Medical Center, G011 Wahl Hall East, Mail Stop 3043, 3901 Rainbow Boulevard, Kansas City, KS 66160, USA
[2]Banner Alzheimer's Institute, Phoenix, AZ, USA
[3]Department of Biostatistics & Data Science, University of Kansas Medical Center, Kansas City, KS, USA
[4]Alzheimer's Disease Research Center, University of Kansas Medical Center, Kansas City, KS, USA
[5]Early Detection and Prevention Division, Translational Genomics Research Institute, Phoenix, AZ, USA
[6]Department of Pharmacy Practice and Science, R. Ken Coit College of Pharmacy, University of Arizona, Tucson, AZ 85721, USA
[7]Center for Precision Health Research, National Human Genome Research Institute, National Institute of Health, Bethesda, MD, USA
[8]All of Us Research Program, Office of the Director, National Institutes of Health, Bethesda, MD, USA
[9]Department of Psychiatry and Behavioral Sciences & Division of Medical Informatics, University of Kansas Medical Center, Kansas City, KS, USA

Received: 5 August 2025 / Accepted: 27 March 2026
Published online: 11 April 2026

References

1. Alzheimer's A. 2024 Alzheimer's Association Facts and Figures. Published online 2024.
2. Huang LK, Kuan YC, Lin HW, Hu CJ. Clinical trials of new drugs for Alzheimer disease: a 2020–2023 update. J Biomed Sci. 2023;30(1):83. https://doi.org/10.1186/s12929-023-00976-6.
3. Jack CR, Wiste HJ, Weigand SD, et al. Amyloid-first and neurodegeneration-first profiles characterize incident amyloid PET positivity. Neurology. 2013;81(20):1732–40. https://doi.org/10.1212/01.wnl.0000435556.21319.e4.
4. Karlsson IK, Escott-Price V, Gatz M, et al. Measuring heritable contributions to Alzheimer's disease: polygenic risk score analysis with twins. Brain Commun. 2022;4(1):fcab308. https://doi.org/10.1093/braincomms/fcab308.
5. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47(1):226–35. https://doi.org/10.1093/ije/dyx206.
6. Helgason A, Yngvadóttir B, Hrafnkelsson B, Gulcher J, Stefánsson K. An Icelandic example of the impact of population structure on association studies. Nat Genet. 2005;37(1):90–5. https://doi.org/10.1038/ng1492.
7. Haworth S, Mitchell R, Corbin L, et al. Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. Nat Commun. 2019;10(1):333. https://doi.org/10.1038/s41467-018-08219-1.
8. Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet. 2022;54(4):412–36. https://doi.org/10.1038/s41588-022-01024-z.
9. Willett JDS, Waqas M, Choi Y, et al. Identification of 16 novel Alzheimer's disease loci using multi-ancestry meta-analyses. Alzheimers Dement J Alzheimers Assoc. 2025;21(2):e14592. https://doi.org/10.1002/alz.14592.
10. Liu JZ, Erlich Y, Pickrell JK. Case–control association mapping by proxy using family history of disease. Nat Genet. 2017;49(3):325–31. https://doi.org/10.1038/ng.3766.
11. Shapland CY, Gkatzionis A, Hemani G, Tilling K. Use of genetic correlations to examine selection bias. Genet Epidemiol. 2025;49(1):e22584. https://doi.org/10.1002/gepi.22584.
12. Schoeler T, Speed D, Porcu E, Pirastu N, Pingault JB, Kutalik Z. Participation bias in the UK Biobank distorts genetic associations and downstream analyses. Nat Hum Behav. 2023;7(7):1216–27. https://doi.org/10.1038/s41562-023-01579-9.
13. Maserejian N, Krzywy H, Eaton S, Galvin JE. Cognitive measures lacking in EHR prior to dementia or Alzheimer's disease diagnosis. Alzheimers Dement. 2021;17(7):1231–43. https://doi.org/10.1002/alz.12280.
14. Ollier W, Sprosen T, Peakman T. UK Biobank: From Concept to Reality. Pharmacogenomics. 2005;6(6):639–46. https://doi.org/10.2217/14622416.6.6.639.

15. Tyrrell J, Zheng J, Beaumont R, et al. Genetic predictors of participation in optional components of UK Biobank. Nat Commun. 2021;12(1):886. https://doi.org/10.1038/s41467-021-21073-y.
16. Wu Y, Sun Z, Zheng Q, et al. Pervasive biases in proxy genome-wide association studies based on parental history of Alzheimer's disease. Nat Genet. 2024;56(12):2696–703. https://doi.org/10.1038/s41588-024-01963-9.
17. De La Fuente J, Grotzinger AD, Marioni RE, Nivard MG, Tucker-Drob EM. Integrated analysis of direct and proxy genome wide association studies highlights polygenicity of Alzheimer's disease outside of the APOE region. PLOS Genet. 2022;18(6):e1010208. https://doi.org/10.1371/journal.pgen.1010208. Epstein MP, ed.
18. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol. 2017;186(9):1026–34. https://doi.org/10.1093/aje/kwx246.
19. Zeng C, Schlueter DJ, Tran TC, et al. Comparison of phenomic profiles in the *All of Us* Research Program against the US general population and the UK Biobank. J Am Med Inf Assoc. 2024;31(4):846–54. https://doi.org/10.1093/jamia/ocad260.
20. Stocker H, Möllers T, Perna L, Brenner H. The genetic risk of Alzheimer's disease beyond APOE ε4: systematic review of Alzheimer's genetic risk scores. Transl Psychiatry. 2018;8(1):166. https://doi.org/10.1038/s41398-018-0221-8.
21. Ramanan VK, Heckman MG, Przybelski SA, et al. Polygenic Scores of Alzheimer's Disease Risk Genes Add Only Modestly to APOE in Explaining Variation in Amyloid PET Burden. J Alzheimers Dis JAD. 2022;88(4):1615–25. https://doi.org/10.3233/JAD-220164.
22. Belloy ME, Napolioni V, Greicius MD. A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward. Neuron. 2019;101(5):820–38. https://doi.org/10.1016/j.neuron.2019.01.056.
23. Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. JAMA. 1997;278(16):1349–56.
24. Frisoni GB, Altomare D, Ribaldi F, et al. Dementia prevention in memory clinics: recommendations from the European task force for brain health services. Lancet Reg Health - Eur. 2023;26:100576. https://doi.org/10.1016/j.lanepe.2022.100576.
25. Lennon NJ, Kottyan LC, Kachulis C, et al. Selection, optimization and validation of ten chronic disease polygenic risk scores for clinical implementation in diverse US populations. Nat Med. 2024;30(2):480–7. https://doi.org/10.1038/s41591-024-02796-z.
26. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51(4):584–91. https://doi.org/10.1038/s41588-019-0379-x.
27. The All of Us Research Program Investigators. The All of Us Research Program. N Engl J Med. 2019;381(7):668–76. https://doi.org/10.1056/NEJMsr1809937.
28. The All of Us Research Program Genomics Investigators, Manuscript Writing Group, Bick AG, et al. Genomic data in the All of Us Research Program. Nature. 2024;627(8003):340–6. https://doi.org/10.1038/s41586-023-06957-x.
29. Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. Ann Intern Med. 2010;153(9):600–6. https://doi.org/10.7326/0003-4819-153-9-201011020-00010.
30. Ritchie MD, Denny JC, Crawford DC, et al. Robust Replication of Genotype-Phenotype Associations across Multiple Diseases in an Electronic Medical Record. Am J Hum Genet. 2010;86(4):560–72. https://doi.org/10.1016/j.ajhg.2010.03.003.
31. Carlson C. Group Health Cooperative. Dementia. PheKB. https://phekb.org/phenotype/10.
32. Zhong L, Xie YZ, Cao TT, et al. A rapid and cost-effective method for genotyping apolipoprotein E gene polymorphism. Mol Neurodegener. 2016;11:2. https://doi.org/10.1186/s13024-016-0069-4.
33. Khajouei E, Ghisays V, Piras IS et al. Phenome-Wide Association of APOE Alleles in the All of Us Research Program. MedRxiv Prepr Serv Health Sci. Published online September 4, 2024:2024.09.04.24313010. https://doi.org/10.1101/2024.09.04.24313010.
34. Agresti A. Categorical Data Analysis. Third edition. Wiley-Interscience; 2013.
35. Altmann A, Tian L, Henderson VW, Greicius MD, Alzheimer's Disease Neuroimaging Initiative Investigators. Sex modifies the APOE-related risk of developing Alzheimer disease. Ann Neurol. 2014;75(4):563–73. https://doi.org/10.1002/ana.24135.
36. Sebastiani P, Gurinovich A, Nygaard M, et al. APOE Alleles and Extreme Human Longevity. J Gerontol Biol Sci Med Sci. 2019;74(1):44–51. https://doi.org/10.1093/gerona/gly174.
37. Naslavsky MS, Suemoto CK, Brito LA, et al. Global and local ancestry modulate APOE association with Alzheimer's neuropathology and cognitive outcomes in an admixed sample. Mol Psychiatry. 2022;27(11):4800–8. https://doi.org/10.1038/s41380-022-01729-x.
38. Marioni RE, Harris SE, Zhang Q, et al. GWAS on family history of Alzheimer's disease. Transl Psychiatry. 2018;8(1):99. https://doi.org/10.1038/s41398-018-0150-6.
39. Valero S, Marquié M, De Rojas I, et al. Interaction of neuropsychiatric symptoms with APOE ε4 and conversion to dementia in MCI patients in a Memory Clinic. Sci Rep. 2020;10(1):20058. https://doi.org/10.1038/s41598-020-77023-z.
40. Gibson J, Russ TC, Adams MJ, et al. Assessing the presence of shared genetic architecture between Alzheimer's disease and major depressive disorder using genome-wide association data. Transl Psychiatry. 2017;7(4):e1094. https://doi.org/10.1038/tp.2017.49.
41. Liu H, Hu Y, Zhang Y, et al. Mendelian randomization highlights significant difference and genetic heterogeneity in clinically diagnosed Alzheimer's disease GWAS and self-report proxy phenotype GWAX. Alzheimers Res Ther. 2022;14(1):17. https://doi.org/10.1186/s13195-022-00963-3.
42. Williams DM, Heikkinen S, Hiltunen M, FinnGen, Davies NM, Anderson EL. The proportion of Alzheimer's disease attributable to apolipoprotein E. Npj Dement. 2026;2(1). https://doi.org/10.1101/2023.11.16.23298475.
43. Ashton NJ, Brum WS, Di Molfetta G, et al. Diagnostic Accuracy of a Plasma Phosphorylated Tau 217 Immunoassay for Alzheimer Disease Pathology. JAMA Neurol. 2024;81(3):255. https://doi.org/10.1001/jamaneurol.2023.5319.
44. Yaffe K, Vittinghoff E, Dublin S, et al. Effect of Personalized Risk-Reduction Strategies on Cognition and Dementia Risk Profile Among Older Adults: The SMARRT Randomized Clinical Trial. JAMA Intern Med. 2024;184(1):54. https://doi.org/10.1001/jamainternmed.2023.6279.
45. Widén E, Junna N, Ruotsalainen S, et al. How Communicating Polygenic and Clinical Risk for Atherosclerotic Cardiovascular Disease Impacts Health Behavior: an Observational Follow-up Study. Circ Genomic Precis Med. 2022;15(2). https://doi.org/10.1161/CIRCGEN.121.003459.
46. Andrews SJ, Jonson C, Fulton-Howard B, et al. The Role of Genomic-Informed Risk Assessments in Predicting Dementia Outcomes. Genetic and Genomic Medicine. Preprint posted online April. 2024;30. https://doi.org/10.1101/2024.04.27.24306488.
47. Bellou E, Kim W, Leonenko G, et al. Benchmarking Alzheimer's disease prediction: personalised risk assessment using polygenic risk scores across various methodologies and genome-wide studies. Alzheimers Res Ther. 2025;17(1):6. https://doi.org/10.1186/s13195-024-01664-9.
48. Mandal S, Qin J, Pfeiffer RM. Incorporating survival data into case-control studies with incident and prevalent cases. Stat Med. 2021;40(28):6295–308. https://doi.org/10.1002/sim.9183.
49. Waltoft BL, Pedersen CB, Nyegaard M, Hobolth A. The importance of distinguishing between the odds ratio and the incidence rate ratio in GWAS. BMC Med Genet. 2015;16:71. https://doi.org/10.1186/s12881-015-0210-1.
50. Mansournia MA, Altman DG. Inverse probability weighting. *BMJ*. Published online January 15, 2016:i189. https://doi.org/10.1136/bmj.i189.
51. Escott-Price V, Hardy J. Genome-wide association studies for Alzheimer's disease: bigger is not always better. Brain Commun. 2022;4(3):fcac125. https://doi.org/10.1093/braincomms/fcac125.

Publisher's Note