TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Baker, E., Leonenko, G., Schmidt, K. M., Hill, M., Myers, A. J., Shoai, M., de Rojas, I., Tesi, N., Holstege, H., & More Authors (2023). What does heritability of Alzheimer's disease represent? *PLoS ONE*, *18*(4 April), e0281440. Article e0281440. https://doi.org/10.1371/journal.pone.0281440

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

*This work is downloaded from Delft University of Technology.*

PLOS ONE

RESEARCH ARTICLE

# What does heritability of Alzheimer's disease represent?

**Emily Baker[1,2], Ganna Leonenko[1,2], Karl Michael Schmidt[3], Matthew Hill[2¤], Amanda J. Myers[4], Maryam Shoai[5], Itziar de Rojas[6,7], Niccoló Tesi[8,9,10], Henne Holstege[8,9,10,11], Wiesje M. van der Flier[9,11], Yolande A. L. Pijnenburg[9,11], Agustin Ruiz[6,7], John Hardy[5], Sven van der Lee[8,9,11], Valentina Escott-Price[1]***

**1** Division of Neuroscience and Mental Health, School of Medicine, Cardiff University, Cardiff, United Kingdom, **2** Dementia Research Institute, School of Medicine, Cardiff University, Cardiff, United Kingdom, **3** School of Mathematics, Cardiff University, Cardiff, United Kingdom, **4** Department of Cell Biology, Miller School of Medicine, University of Miami, Coral Gables, FL, United States of America, **5** Institute of Neurology, University College London, London, United Kingdom, **6** Research Center and Memory Clinic, ACE Alzheimer Center Barcelona, Universitat Internacional de Catalunya, Barcelona, Spain, **7** Networking Research Center on Neurodegenerative Diseases (CIBERNED), Instituto de Salud Carlos III, Madrid, Spain, **8** Genomics of Neurodegenerative Diseases and Aging, Human Genetics, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands, **9** Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands, **10** Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands, **11** Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands

¤ Current address: Vertex Pharmaceuticals, Milton, Abingdon, England
* EscottPriceV@Cardiff.ac.uk

Check for updates

OPEN ACCESS

**Citation:** Baker E, Leonenko G, Schmidt KM, Hill M, Myers AJ, Shoai M, et al. (2023) What does heritability of Alzheimer's disease represent? PLoS ONE 18(4): e0281440. https://doi.org/10.1371/journal.pone.0281440

**Editor:** Toni-Kim Clarke, Forest Research, UNITED KINGDOM

**Received:** October 11, 2022

**Accepted:** January 24, 2023

**Published:** April 28, 2023

**Copyright:** © 2023 Baker et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Since for this study the individual level genotype data were used, the data have ethical and/or legal restrictions (see below). The authors received no special privileges in accessing the data. 1. Ethical approval was granted to the UK Biobank project by the North West Multi-Centre Ethics committee. The UK Biobank data is available by submitting a research proposal to https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access. 2. The Ethics and Scientific Committees (Acta 25/2016. Ethics Committee. H. Clinic i Provincial, Barcelona, Spain) have approved the research protocol for the

## Abstract

### Introduction

Both late-onset Alzheimer's disease (AD) and ageing have a strong genetic component. In each case, many associated variants have been discovered, but how much missing heritability remains to be discovered is debated. Variability in the estimation of SNP-based heritability could explain the differences in reported heritability.

### Methods

We compute heritability in five large independent cohorts (N = 7,396, 1,566, 803, 12,528 and 3,963) to determine whether a consensus for the AD heritability estimate can be reached. These cohorts vary by sample size, age of cases and controls and phenotype definition. We compute heritability a) for all SNPs, b) excluding *APOE* region, c) excluding both *APOE* and genome-wide association study hit regions, and d) SNPs overlapping a microglia gene-set.

### Results

SNP-based heritability of late onset Alzheimer's disease is between 38 and 66% when age and genetic disease architecture are correctly accounted for. The heritability estimates decrease by 12% [SD = 8%] on average when the *APOE* region is excluded and an additional 1% [SD = 3%] when genome-wide significant regions were removed. A microglia

GR@ACE dataset. The restrictions are imposed by the data access committee, the application process to access the data is described https://ega-archive.org/studies/EGAS00001003424. 3. The KRONOS/Tgen dataset is obtained from 21 National Alzheimer's Coordinating Center (NACC) brain banks and from the Miami Brain Bank. Samples were de-identified before receipt, and the study met human studies institutional review board and HIPPA regulations. This work is declared not human-subjects research and is IRB exempt under regulation 45 CFR 46. The data is available at https://www.niagads.org/datasets/ng00028. Data use limitations: General Research Use (GRU). 4. ROSMAP/MSBB/MAYO were downloaded via the AMP-AD portal via the Synapse platform and https://www.radc.rush.edu. Data use limitations: Disease-Specific (Neurodegenerative, NPU) (DS-ND-NPU). 5. The local (VU University Medical Center) Medical Ethics Committee has approved a general protocol for biobanking and using the clinical data for research purposes for the Amsterdam Dementia cohort (ADC) data. The restrictions are imposed by the data access committee, the application process to access the data is described: https://www.neurodegenerationresearch.eu/cohort/amsterdam-dementia-cohort/.

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have read the journal's policy and have the following competing interests: MH hold shares in Vertex Pharmaceuticals. This does not alter our adherence to PLOS ONE policies on sharing data and materials. There are no patents or patents, products in development or marketed products associated with this research to declare.

gene-set explains 69–84% of our estimates of SNP-based heritability using only 3% of total SNPs in all cohorts.

## Conclusion

The heritability of neurodegenerative disorders cannot be represented as a single number, because it is dependent on the ages of cases and controls. Genome-wide association studies pick up a large proportion of total AD heritability when age and genetic architecture are correctly accounted for. Around 13% of SNP-based heritability can be explained by known genetic loci and the remaining heritability likely resides around microglial related genes.

## 1. Introduction

Autosomal dominant Alzheimer's disease accounts for only ~1% of all cases, the remaining AD cases are probably caused by a complex interplay of environmental and genetic factors. The pathological changes of aggregation of amyloid plaques and formation of intracellular neurofibrillary tangles begin in the brain long before manifestation of the first clinical symptoms due to severe neuronal loss [1]. AD can be diagnosed with certainty during life using cerebrospinal fluid (CSF) biomarkers, amyloid PET imaging and definitely at autopsy [2, 3]. However, the accuracy of clinical diagnosis, without the use of CSF or blood biomarkers or PET imaging, is relatively low and includes up to 30% of misdiagnosed patients [4–6].

The heritability (the proportion of phenotypic variance explained by genetics [7]) of late onset Alzheimer's Disease liability is generally agreed to be between 60–80% from twin studies [8, 9]. The largest contributor to genetic risk is the *APOE* gene and genome-wide association studies (GWAS) have been successful in identifying over 80 common and rare loci significantly associated with AD [10–17]. *APOE* and these other variants do not explain all genetic liability for AD. The hope is that with larger GWAS sample sizes, not only more risk loci will be identified, but also a larger proportion of total heritability will be explained. The amount of heritability still remaining to be found is under debate.

Heritability analyses were largely designed for the analysis of disorders of children and early adulthood in which both case and control designations have some certainty due to early in life onset and therefore were not influenced by age. Unfortunately, in AD these characteristics do not apply. The clinical diagnosis of AD is not particularly accurate, and the age dependence of the disease causes both obvious and subtle problems with analysis. For example, Beach et al. [4] show that 23.4% of people did not have frequent neuritic plaque density, despite their positive clinical diagnoses. Escott-Price et al. [5] estimate misdiagnosis rate up to 36% in controls unscreened for the APOE genotype, and up to 29% when E3 homozygous subjects are used as controls in clinical studies. The most important problem in estimating heritability is that an individual's genetic loading for disease remains the same at any age, but the prevalence of AD is dependent on age. AD risk increases with age, people at the highest risk develop disease earlier, and therefore risk allele frequency decreases with age [18, 19]. Thus, heritability estimates are age dependent [20] and for reliable assessment at any individual age, it is necessary for cases and controls to be age matched. It is also possible that there will be some differences in the heritability of disease between populations, related to different haplotype length and to the presence/absence of rare mutations in the population e.g. the presenilin mutation (E280A) in Antioquia, Colombia [20]. All the above is reflected in widely different SNP-based heritability estimates across different datasets in AD, from as high as 53% [21] to as low as 3% [16]. The latter is unlikely to be true as the *APOE* gene alone explains 4% of the

**Table 1.**

| Publication | Sample Size | Narrow sense heritability on 5% liability scale |
|---|---|---|
| Harold et al. 2009 | 11,025 | 17% |
| Corneveaux et al. 2010 | 1,594 | 42% |
| Naj et al. 2011 | 21,165 | 25% |
| Lambert et al. 2013 | 54,162 | 9% |
| Marioni et al. 2018 | 368,440 | 3% |
| Kunkle et al. 2019 | 63,926 | 7% |
| Jansen et al. 2019 | 455,258 | 6% (without UK Biobank), 2% (with UK Biobank) |
| Rojas et al. 2021 | 409,435 | 3% |
| Wightman et al. 2021 | 1,126,563 | 3% |
| Bellenguez et al. 2022 | 487,551 | 3% |

https://doi.org/10.1371/journal.pone.0281440.t001

variance when studying incident AD [22]. Estimates are likely to be low due to LDSC underestimating heritability from summary [23]. Table 1 shows the heritability estimates for a number of different large scale studies [24].

The variability of the reported heritability estimates arise from various sources, related to the populations studied and technical issues. The differences in heritability estimates may either be on the observed scale i.e. for the proportion of cases and controls as in the sample, or on the liability scale, i.e. assuming a disease prevalence in a particular population, which varies depending on the age group and population where the prevalence has been reported. For example, 2% lifetime prevalence was reported in the US in 2019 [25], 3% in 2020 in individuals aged 65–74 in the US [26], 5% lifetime prevalence in Europeans from a meta-analysis of multiple studies [27], 17% in 2020 in individuals aged 75–84 in the US [26], 32% in 2020 in individuals aged 85+ in the US [26]. Prevalence is not well-suited to late-onset disorders since this varies with age, with the risk of AD increasing substantially when reaching oldest ages [28]. For heritability analyses, the issue of prevalence is mostly resolved when using an age-matched cohort with relatively narrow age ranges in cases and controls. Age-related prevalence needs to be accounted for when computing heritability estimates.

The main aim of this study is to determine late onset AD heritability in a variety of AD data cohorts to understand the variability introduced by the liability model and age and evaluate whether consistent estimates can be determined for AD SNP-based heritability. Current methods do not allow age-related prevalence to be accounted for in one analysis, therefore, we computed heritability on the liability scale using three thresholds; 2, 5 and 15%. The convergence of the results across the thresholds will demonstrate that existing methods are appropriate when prevalence is correctly specified for the age of study participants. Next, we sought to utilise heritability estimates to give insights regarding where in the genome we should search for missing heritability, by investigating a gene-set specific to microglia which are known to be important in AD pathology. For this purpose, we investigate the proportion of heritability which can be explained using SNPs overlapping a specific gene-set related to microglia. We assess the proportion of heritability explained by this gene-set in comparison to the total heritability in the sample and compare this to the proportion of SNPs which explain this heritability.

## 2. Results

### 2.1 Cohort heritability estimates

First we present results for the heritability estimates calculated on the liability threshold with AD prevalence of 2%, 5% and 15% in all datasets; for A) ADC with amyloid confirmed AD

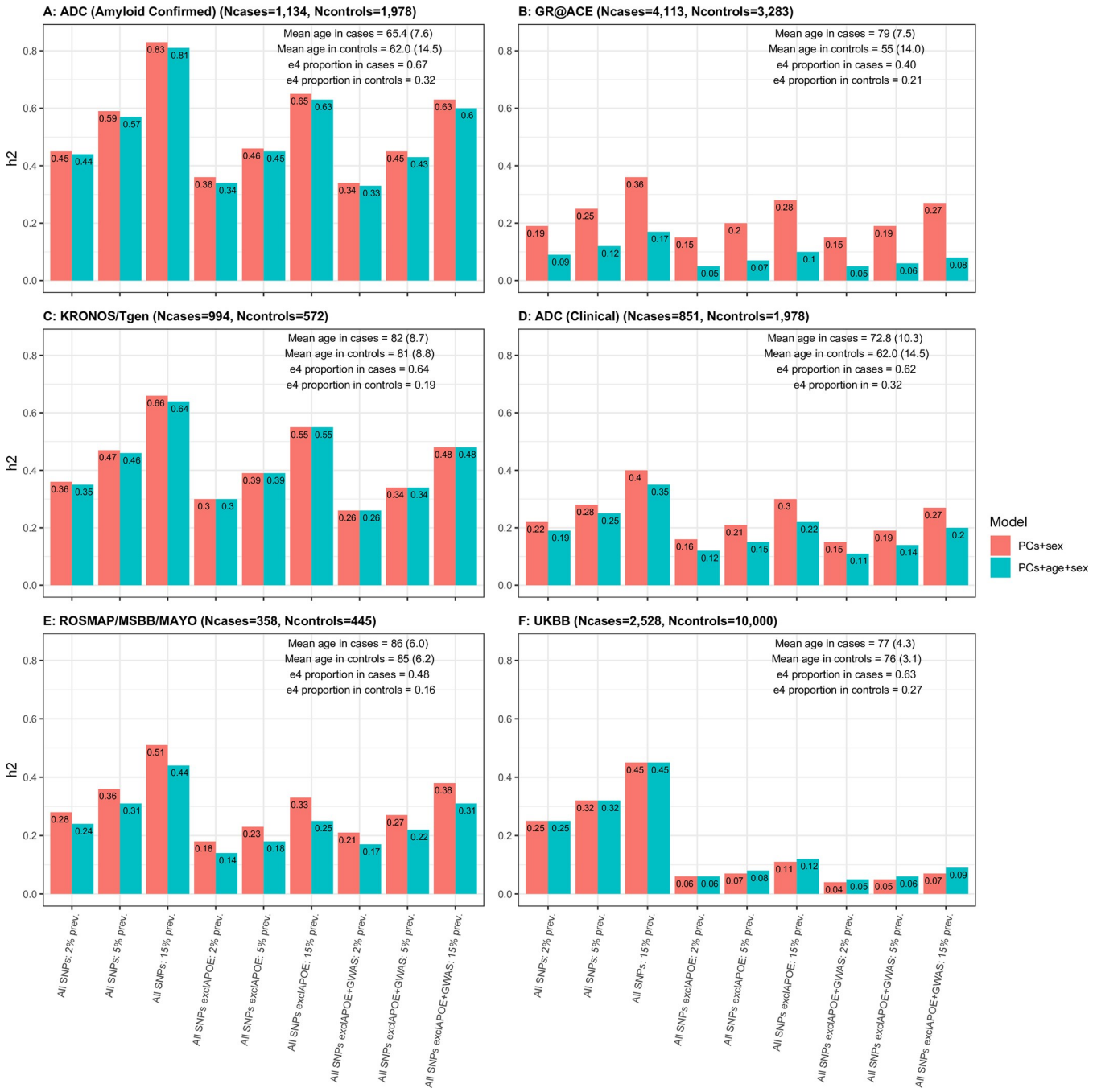


**Fig 1.** Heritability Estimates for AD prevalence of 2%, 5%, 15% in A) ADC with amyloid confirmed AD cases, B) GR@ACE, C) KRONOS/Tgen, D) ADC with clinical AD cases, E) ROSMAP/MSBB/MAYO, F) UKBB AD cases with controls aged 70+. Two models are considered: estimates adjusted for PCs and sex and PCs, sex and age.

https://doi.org/10.1371/journal.pone.0281440.g001

cases, B) GR@ACE, C) KRONOS/Tgen, D) ADC with clinical AD cases, E) ROSMAP/MSBB/MAYO and F) UKBB with controls aged 70+, see Fig 1.

The results presented in Fig 1 show great variability in the heritability estimates even within the same liability threshold analyses (all estimates from all analyses can be observed in S1-S6 Tables in S1 File). When age is added as a covariate to an age mis-matched study (see e.g.

(Fig 1B)), the estimates of heritability drop substantially, whereas in age-matched, pathologically confirmed cohorts of cases and controls, the heritability remains almost unchanged (see e.g. (Fig 1A and 1C)). Since age is a proxy of AD, adjusting for age in age mis-matched cohorts is biasing analyses towards the null hypothesis. Age mis-matched cohorts also cause issues in results because the prevalence of age-related disorders change with age, and therefore, the adjustment to the liability scale will not be appropriate for ages of cases or controls.

The heritability estimates decrease by 12% on average when the *APOE* region is removed and decrease ~1% further when the 0.5MB regions around GWAS index SNPs are additionally excluded. The largest decrease of more than 25% is observed in the UK Biobank cohort (Fig 1F) after removal of the *APOE* region.

In GR@ACE, the analysis was restricted to AD cases diagnosed with probable AD at both first and second diagnoses (N = 1,851). The heritability estimates increased for all prevalences by 10% [SD = 3%] on average to 0.27, 0.35 and 0.49 for 2, 5 and 15% prevalences respectively, when adjusted for PCs and sex. Thus making estimates in this sample more comparable to the other cohorts.

We investigate the impact of the age of controls in UKBB by using four age bins for the control subset ($\leq$60, 60–70, 70–80 and 80+ years old). It is seen from S1 Fig in S1 File that heritability estimates are fairly consistent for controls at all ages, with estimates being slightly increased for the group with the youngest controls ($\leq$60 years old). The model adjusted for PCs, sex and age did not converge in the two youngest control groups since there was little overlap in age distributions between cases and controls.

The p-value of the heritability estimates were directly linked to the size of the cohorts (see S2 Fig and S1, S4-S6 Tables in S1 File). In the KRONOS/Tgen dataset (N = 1,566) the significance reaches $p = 3.22\times10^{-3}$ when all SNPs were included and p = 0.02 after exclusion of *APOE* and GWAS regions. In ADC (clinical) and ROSMAP/MSBB/MAYO all heritability estimates are non-significant for all models, see S2, S3 Tables in S1 File.

The heritability estimates in cohorts with pathologically/amyloid confirmed diagnosis (Fig 1, left plots) are higher (0.36–0.59) compared to cohorts with a clinical diagnosis only (Fig 1, right plots) ($h^2 = 0.25$–$0.34$). This is expected as a pathologically/amyloid confirmed diagnosis is more accurate than a clinical diagnosis of AD which may contain up to 30% of misdiagnosed individuals [4, 5]. Heritability estimates adjusting for PCs only are very similar to those adjusting for PCs and sex, see S3 Fig in S1 File.

As noted above, the additional adjustment for age has little impact on heritability estimates in the pathologically/amyloid confirmed data but reduces the estimates in the GR@ACE data by more than 13%. This result suggests that the decrease in heritability estimate could be mainly attributed to the difference in age distribution between cases and controls. Although it is tempting to adjust for age by including it as a covariate, it is difficult to do this effectively. If there is a systematic age difference between cases and controls, the age covariate largely absorbs the disease status effect, and the analysis is biased towards the null hypothesis. This suggests that the observed heritability should be estimated without adjustment for age but accounted for when transforming to the liability scale. For example, in GR@ACE data, the mean cases' age (79 [SD = 7.5]) is above the average onset of e44 and e4 carriers (which is 68 and 76, respectively [29], whereas the controls are below this age 54.5 [SD = 14.0]. Therefore, if they live until their 80s, more than 15% of controls could develop AD, indicating that they have genetic liability to the disease.

## 2.2 Gene-set heritability estimates

Table 2 demonstrates the proportion of heritability and number of SNPs in the microglia gene-set compared to those including all SNPs for ADC with amyloid confirmed AD cases,

**Table 2. Proportion of heritability and SNPs explained by a microglia gene-set in all data cohorts across all disease prevalences.**

| Data Cohort | Sample Size | Microglia | |
|---|---|---|---|
| | | Proportion of h2 | Proportion of SNPs |
| KRONOS/Tgen | 1,566 | 0.64 | 0.028 |
| ROSMAP/MSBB/MAYO | 803 | 0.68–0.69 | 0.030 |
| GR@ACE | 7,396 | 0.50–0.53 | 0.032 |
| UKBB (70+ controls) | 12,528 | 0.80–0.82 | 0.032 |
| ADC (amyloid confirmed) | 3,112 | 0.67–0.69 | 0.032 |
| ADC (clinical) | 2,829 | 0.91–0.93 | 0.032 |

Heritability estimates adjusted for PCs and sex.

https://doi.org/10.1371/journal.pone.0281440.t002

GR@ACE, KRONOS/Tgen, ADC with clinical AD cases, ROSMAP/MSBB/MAYO and UKBB with controls aged 70+. The absolute heritability estimates adjusted for PCs and sex for each cohort can be seen in S4 Fig in S1 File.

It can be seen that by selecting cell-type specific SNPs, a substantial proportion of heritability is explained using fewer SNPs (approximately 3% of SNPs in the microglia gene-set). The proportion of heritability explained for the microglia gene-set was 68–69% in ROSMAP/MSBB/MAYO, 80–82% in UKBB, 64% in KRONOS/Tgen, 67–69% for amyloid confirmed ADC and 91–93% for clinical cases ADC. The range of values represent the proportions across all AD disease prevalences.

In general, the microglia gene-set has lower heritability estimates compared to all SNPs, however, the reduction is not proportional to the reduction in the number of SNPs, see Table 2. It can be seen in S4 Fig in S1 File that the microglia gene-set produces comparable heritability estimates with the model excluding the *APOE* region. We also present heritability estimates for the microglia gene-set with the same parameters as in S4 Fig in S1 File but adjusted for PCs, sex and age in S5 Fig and S1-S6 Tables in S1 File. Thus, despite this gene-set utilising a much-reduced number of SNPs, it is able to explain a substantial proportion of AD heritability.

We also investigate the impact of the age of controls in UKBB by using four age bins for the control subset ($\leq$60, 60–70, 70–80 and 80+ years old) on the proportion of heritability explained by microglia SNPs. The results are seen in S6 Fig in S1 File, the microglia gene-set explains a consistent proportion of heritability across all control age groups (75–87%).

## 3. Discussion

To date, reported SNP-based heritability estimates in AD have been very varied across different datasets and methodologies. We studied five different cohorts and harmonized analytical methods to estimate SNP-based heritability. We estimate that the SNP-based heritability is between 36% and 59% in pathologically or CSF confirmed AD and 25% to 32% in clinically assessed cohorts when assuming, for the purposes of the heritability model, AD prevalence of 5%. The regions related to microglial genes (only 3% of SNPs) explain between 50% and 93% of the SNP based heritability. This shows the importance of further development of biologically relevant AD gene-sets/pathways that could reduce the signal to noise ratio by highlighting the most influential SNPs/genes in AD. Novel loci are most likely to be expected in these regions.

These heritability estimates have implications for understanding AD risk. For example, Zhang et al. [30] proposed that AD is oligogenic (based on <100 SNPs) based on simulations

assuming 9% narrow-sense heritability excluding *APOE*. The larger heritability estimates reported here will change conclusions about the number of loci contributing to AD risk.

We studied the effects of age and *APOE* on heritability estimates. The results show that heritability estimates are systematically reduced when the *APOE* region is excluded. The reduction varies across cohorts with the largest decrease in UK Biobank, likely due to the age of the UK Biobank cases which is ~76–77 which is the age at onset for e4 carriers [29]. When GWAS hits are additionally excluded, the heritability estimates reduce further but only by a small amount.

The inclusion of age as a covariate clearly has a large impact on the heritability estimates for data cohorts where the mean age of cases and controls differs substantially. Where there is little difference in age between cases and controls, heritability estimates do not change. Based on these observations we recommend that age should not be used as covariate, since a difference in age distribution between cases and controls will lead to adjustment for 'caseness' by biasing the analysis towards the null, and therefore reducing the heritability estimates significantly. Instead, we suggest that the genetic architecture of AD is different depending on age at clinical onset. Indeed, it is known that very early AD cases (aged 30–50) are mostly attributed to rare highly penetrant mutations in *APP* and *PSEN* genes. The disease prevalence at this age in the population is then close to the frequencies of these risk alleles (<1%). *APOE* e44 carriers have age at onset of about 68, and the disease prevalence at this age is likely to be around or slightly larger than e44 frequency (~2–3%), due to the variation in the age at onset of e4 heterozygotes and non-carriers. The mean age of clinical onset of e4 non-carriers is ~84 years of age [29]. The disease prevalence at this age is reported as something between 17–32% [26]. The disease at this age is likely to be attributed to a large number of common SNPs associated with a variety of disease development mechanisms, including comorbid disorders. It is worth noting that the density of AD pathology required for an AD diagnosis is less as age increases [31]. Furthermore, several studies have shown age dependent association of AD polygenic risk score (PRS) with Alzheimer's disease and cognitive function, with almost no association in those with age below 50 years [32], with GWAS significant SNP-based PRS association in samples with mean age 60–65 [17, 33], and with genome-wide PRS association in samples aged 65+ [32, 34–36]. Mars et al. [19] showed 4–9 years earlier age at onset with high polygenic risk in other age-related diseases (type 2 diabetes, atrial fibrillation, breast cancer and prostate cancer), but they do not explain the increasing prevalence of these diseases with age in the population, as earlier onset implies earlier mortality. A simulation study [18] showed that the AD risk increases with age, and therefore, that risk allele frequency decreased with age due to people at the highest risk developing disease earlier [18]. In this circumstance it is perhaps not surprising that the architecture of genetic risk is different at different ages. Therefore, we suggest that for neurodegenerative disorders, the heritability estimates on the liability scale should be adjusted for the *age-related* prevalence of cases. If the controls are not screened for the disease, the proportion of cases in the sample needs to be uplifted to account for the genetic liability for the disease of individuals who do not yet show symptoms, and the observed heritability adjusted accordingly before transforming it to the liability scale. For example, the observed heritability in the GR@ACE data was estimated $h_o^2 = 0.30$ (see S1 Table in S1 File) with the proportion of cases $P = 0.56$ with mean age 84 years. Assuming that 15% of controls (who are on average 54 years old) will develop the disease given time, the actual proportion of cases is $P_{\text{actual}} = 0.62$, and therefore $h_l^2 = 0.38$, (see equation (23) in [37]), which is 2% higher than shown in Fig 1B ("All SNPs: 15% prev"). In contrast, in the ADC—amyloid confirmed sample (mean age in cases 65.4), the observed heritability does not need to be adjusted (as ages of cases and controls are similar), and the SNP-based heritability on the liability scale should be reported as $h_l^2 = 0.45$ (Fig 1B ("All SNPs: 2% prev")).

The limitations of this study are that some of the datasets have small sample sizes, therefore have broad confidence intervals which impacts the statistical significance of the heritability estimates. Another limitation is that all cohorts investigated are Caucasian, but the GR@ACE cohort has a different genetic architecture compared to the other cohorts (Spain compared to North Caucasian) [38]. AD diagnostic criteria in the cohorts of this study are different, varying from pathologically confirmed, amyloid confirmed to clinical diagnosis which introduces noise in the heritability estimates.

In conclusion, for late onset diseases such as AD, the heritability cannot be represented as a single number, but in fact depends upon the age of the cases and controls in the sample where the heritability is to be determined. We also determined that a major fraction of AD heritability independent of cohort age is attributed to the microglia gene-set, highlighting the importance of biologically relevant gene-sets in AD development.

## 4. Methods

The cohorts which were investigated are 1) Genome Research at Fundacio ACE (GR@ACE) [39], 2) KRONOS/Tgen [40–43], 3) Religious Orders Study and the Rush Memory and Aging Project (ROSMAP) data [44–46], The Mount Sinai Brain Bank (MSBB), MAYO Clinic Brain Bank (MAYO), 4) UK Biobank (UKBB) data [47] and 5) the Amsterdam Dementia Cohort (ADC) [48]. These data vary in terms of sample size, age, the definition of late onset AD and control phenotypes (e.g. pathologically confirmed or clinically defined AD cases; age-matched or population cohort controls). All samples are anonymised and therefore it is not possible to identify any individuals.

Heritability was computed in each series independently a) for all available SNPs in each data cohort, b) for all SNPs excluding the *APOE* region (chr19: 44.4–46.5Mb), and c) for all SNPs but with both *APOE* SNPs and SNPs within 0.5Mb of previously reported genome-wide association study (GWAS) hits excluded. For comparability with other studies (e.g. [30]), the estimates were adjusted to the liability scale based on AD disease prevalence in the population (5% [27]). We however present and discuss the results for 2%, 5% and 15% prevalence.

### 4.1 Population description

The GR@ACE data [39] consists of 4,113 cases and 3,283 controls. AD cases are classified as individuals with dementia who were diagnosed with either possible or probable AD at any time. Written informed consent was obtained from all participants. The Ethics and Scientific Committees have approved this research protocol (Acta 25/2016. Ethics Committee. H. Clinic i Provincial, Barcelona, Spain).

The KRONOS/Tgen dataset is obtained from 21 National Alzheimer's Coordinating Center (NACC) brain banks and from the Miami Brain Bank as previously described [40–43]. The cohort consists of 994 AD cases and 572 controls of European descent. This work is declared not human-subjects research and is IRB exempt under regulation 45 CFR 46.

ROSMAP [44–46], MSBB (The Mount Sinai Brain Bank) and The Mayo Clinic Brain Bank (MAYO) have been whole-genome sequenced, harmonised and analysed together. This sample contains 803 individuals; 358 AD cases and 445 controls. The parent studies and sub-studies were all approved by an Institutional Review Board of Rush University Medical Center and all participants signed an informed consent, Anatomical Gift Act, and a repository consent to share data and biospecimens.

The UKBB is a large prospective cohort of individuals from the UK recruited between 2006–2010 [47]. Inclusion criteria was for cases -all individuals who were diagnosed with AD based on ICD-10 code F00 or G30, N = 2,528 and for controls -a subset of 10,000 individuals

**Table 3. Summary of demographics for all cohorts.**

| Data | Demographics | Cases | Controls |
|---|---|---|---|
| GR@ACE | N | 4113 | 3283 |
| | Age at onset/interview [SD] | 79 [7.5] | 54.5 [14.0] |
| | Sex [M/F/NA] | 1256/2856/1 | 1676/1605/2 |
| KRONOS/Tgen | N | 994 | 572 |
| | Age of death [SD] | 81.9 [8.7] | 81 [8.8] |
| | Sex [M/F] | 361/633 | 355/217 |
| ROSMAP/MSBB/MAYO | N | 358 | 445 |
| | Age of death [SD] | 85.9 [6.0] | 84.5 [6.2] |
| | Sex [M/F] | 100/258 | 167/278 |
| UKBB | N | 2528 | 10,000 |
| | Age at interview [SD] | 76.8 [4.3] | 75.2 [3.3] |
| | Sex [M/F] | 1227/1301 | 4706/5294 |
| ADC | N | 1985 (1134 CSF, 851 clinical) | 1978 |
| | Age at interview [SD] | 65.4 [7.6] (CSF) 72.8 [10.3] (clinical) | 62.0 [14.5] |
| | Sex [M/F] | 540/594 (CSF) 373/478 (clinical) | 1031/947 |

https://doi.org/10.1371/journal.pone.0281440.t003

with no AD or dementia diagnosis who were aged over 70 (UKBB (controls 70+)). All participants provided informed written consent to participate in UK Biobank. Ethical approval was granted to the UK Biobank project by the North West Multi-Centre Ethics committee. Data were released to us after application project reference 13310.

A secondary analysis to investigate the impact of the age of controls was carried out using four different control subsets; 1) aged ≤60 years old, 2) aged 60–70 years old, 3) aged 70–80 years old and 4) aged 80+ years old.

The Amsterdam Dementia cohort (ADC) data [48, 49] is a cohort of AD cases and controls, consisting of 1,985 cases (1,134 CSF confirmed and 851 clinically diagnosed) and 1,978 controls. All patients gave informed consent. The local Medical Ethics Committee has approved a general protocol for biobanking and using the clinical data for research purposes.

Detailed information and demographics for all the cohorts can be found in Table 3 and S1 File.

## 4.2 Heritability estimates

Heritability estimates are computed using the Genome-wide Complex Trait Analysis (GCTA) [50, 51] software to estimate the proportion of phenotypic variance explained by SNPs. GCTA software was chosen as the primary approach for calculation of heritability estimates since a) individual genotypes were available to us, and b) when a large proportion of the SNP-based heritability is explained by a single variant, the genome-based restricted maximum likelihood, implemented in GCTA, is unbiased whereas the alternative approach (LDScore regression [52]) in this case provides systematically lower estimates [53]).

The restricted maximum likelihood (GREML-LDMS) analysis was used to estimate SNP-based heritability whilst correcting for LD bias, by splitting data into LD quartiles and stratifying SNPs based on the segment-based LD score and MAF = 0.05. For this analysis, a region of 200kb was used to compute the segment-based LD score. The heritability was estimated in two scenarios 1) adjusting for principal components (PCs) and sex, and 2) for PCs, age and sex. The GR@ACE and KRONOS/Tgen data were adjusted for 5 PCs; the ROSMAP/MSBB/MAYO dataset is adjusted for 8 PCs, UKBB is adjusted for 15 PCs and the ADC is adjusted for 10 PCs, determined from PC plots.

The GCTA software was applied to the five datasets separately, using a) all available SNPs, b) excluding the *APOE* region (chr19:44.4–46.5Mb), and c) excluding SNPs in the *APOE* region and those within 0.5Mb of known GWAS hits [54]. Observed heritability estimates were re-scaled to the liability threshold based on 2%, 5% and 15% prevalences which represent a range of prevalences previously published [25–27].

### 4.3 Gene-sets

A number of biological gene-sets have been defined which may enable the AD genetic signal to be focused to specific biological functions. We investigated the proportion of heritability explained by SNPs in genes related to microglia. [55] defined microglia regions based on GWAS signatures and epigenetic/gene regulatory data. [56] have redefined the list of SNPs to include established regulatory regions of the genes. We have used SNPs within these regions and heritability based on these SNPs was computed to compare heritability in each data cohort.

1. Research Center and Memory clinic. ACE Alzheimer Center Barcelona, Universitat Internacional de Catalunya, Spain.
2. CIBERNED, Center for Networked Biomedical Research on Neurodegenerative Diseases, National Institute of Health Carlos III, Ministry of Economy and Competitiveness, Spain,
3. Dep. of Surgery, Biochemistry and Molecular Biology, School of Medicine. University of Málaga. Málaga, Spain,
4. Grupo de Medicina Xenómica, Centro Nacional de Genotipado (CEGEN-PRB3-ISCIII). Universidad de Santiago de Compostela, Santiago de Compostela, Spain.
5. Fundación Pública Galega de Medicina Xenómica- CIBERER-IDIS, Santiago de Compostela, Spain.
6. Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas, CIBERDEM, Spain, Hospital Clínico San Carlos, Madrid, Spain,
7. CAEBI. Centro Andaluz de Estudios Bioinformáticos, Sevilla, Spain
8. Unidad Clínica de Enfermedades Infecciosas y Microbiología. Hospital Universitario de Valme, Sevilla, Spain.

## Supporting information

**S1 File.**
(DOCX)

**S1 Data.**
(CSV)

## Acknowledgments

We thank the Dementia Research Institute, Welsh Government, Joint Programming for Neurodegeneration, Dementia Platforms UK and MRC Centre for Neuropsychiatric Genetics and Genomics.

Research of the Alzheimer center Amsterdam is part of the neurodegeneration research program of Amsterdam Neuroscience. The Alzheimer Center Amsterdam is supported by Stichting Alzheimer Nederland, Stichting VUmc fonds, Gieskes-Strijbisfonds, Hersensticht-ing, Stichting Dioraphte and Edwin Bouw Fonds.

We would like to thank patients and controls who participated in this project. We are indebted to Trinitat Port-Carbó legacy and her family for their support of Ace Alzheimer Center Barcelona research programs. Ace Alzheimer Center Barcelona is one of the participating centers of the Dementia Genetics Spanish Consortium (DEGESCO).

Part of the data used in this project have been generated by GR@ACE/DEGESCO consortium. Funds obtained to generate this dataset have been detailed elsewhere (de Rojas et al. 2021).

The GR@ACE study group

Aguilera Nuria[1], Alarcon Emilio[1], Alegret Montserrat[1,2], Boada Mercè[1,2], Buendia Mar[1], Cano Amanda[1], Cañabate Pilar[1,2], Carracedo Angel[4,5], Corbatón-Anchuelo A[6], de Rojas Itziar[1,2], Diego Susana[1], Espinosa Ana[1,2], Gailhajenet Anna[1], García-González Pablo[1,2], Guitart Marina[1], González-Pérez Antonio[7], Ibarria Marta[1], Lafuente Asunción[1], Macias Juan[8], Maroñas Olalla[4], Martín Elvira[1], Martínez Maria Teresa[6], Marquié Marta[1,2], Montrreal Laura[1], Moreno-Grau Sonia[1,2], Moreno Mariona[1], Raúl Nuñez-Llaves[1], Olivé Clàudia[1], Orellana Adelina[1], Ortega Gemma[1,2], Pancho Ana[1], Pelejà Ester[1], Pérez-Cordon Alba[1], Pineda Juan A[8], Puerta Raquel[1], Preckler Silvia[1], Quintela Inés[3], Real Luis Miguel[3,8], Rosende-Roca Maitee[1], Ruiz Agustín[1,2], Sáez Maria Eugenia[7], Sanabria Angela[1,2], Serrano-Rios Manuel[6], Sotolongo-Grau Oscar[1], Tárraga Luís[1,2], Valero Sergi[1,2], Vargas Liliana[1].

## Author Contributions

**Conceptualization:** Valentina Escott-Price.

**Data curation:** Matthew Hill, Amanda J. Myers, Maryam Shoai, Itziar de Rojas, Niccoló Tesi, Henne Holstege, Wiesje M. van der Flier, Yolande A. L. Pijnenburg, Agustin Ruiz, John Hardy, Sven van der Lee.

**Formal analysis:** Emily Baker, Ganna Leonenko, Sven van der Lee.

**Investigation:** Emily Baker, John Hardy, Sven van der Lee, Valentina Escott-Price.

**Methodology:** Karl Michael Schmidt, Valentina Escott-Price.

**Supervision:** Valentina Escott-Price.

**Writing – original draft:** Emily Baker, Valentina Escott-Price.

**Writing – review & editing:** Emily Baker, Ganna Leonenko, Karl Michael Schmidt, Matthew Hill, Amanda J. Myers, Maryam Shoai, Itziar de Rojas, Niccoló Tesi, Henne Holstege, Wiesje M. van der Flier, Agustin Ruiz, John Hardy, Sven van der Lee, Valentina Escott-Price.

## References

1. Hubbard BM, Fenton GW, Anderson JM. A quantitative histological study of early clinical and preclinical Alzheimer's disease. Neuropathology and applied neurobiology. 1990; 16(2):111–21. https://doi.org/10.1111/j.1365-2990.1990.tb00940.x PMID: 2345598
2. Jack CR Jr., Holtzman DM. Biomarker modeling of Alzheimer's disease. Neuron. 2013; 80(6):1347–58.
3. Ramirez-Bermudez J. Alzheimer's disease: critical notes on the history of a medical concept. Arch Med Res. 2012; 43(8):595–9. https://doi.org/10.1016/j.arcmed.2012.11.008 PMID: 23178566
4. Beach TG, Monsell SE, Phillips LE, Kukull W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005–2010. J Neuropathol Exp Neurol. 2012; 71(4):266–73. https://doi.org/10.1097/NEN.0b013e31824b211b PMID: 22437338
5. Escott-Price V, Baker E, Shoai M, Leonenko G, Myers AJ, Huentelman M, et al. Genetic analysis suggests high misassignment rates in clinical Alzheimer's cases and controls. Neurobiology of aging. 2019; 77:178–82. https://doi.org/10.1016/j.neurobiolaging.2018.12.002 PMID: 30851568

6. Shim YS, Roe CM, Buckles VD, Morris JC. Clinicopathologic study of Alzheimer's disease: Alzheimer mimics. J Alzheimers Dis. 2013; 35(4):799–811. https://doi.org/10.3233/JAD-121594 PMID: 23481687
7. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. Nature Reviews Genetics. 2013; 14(2):139–49. https://doi.org/10.1038/nrg3377 PMID: 23329114
8. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining alzheimer disease. Archives of general psychiatry. 2006; 63(2):168–74.
9. Bergem AL, Engedal K, Kringlen E. The role of heredity in late-onset Alzheimer disease and vascular dementia. A twin study. Archives of general psychiatry. 1997; 54(3):264–70. https://doi.org/10.1001/archpsyc.1997.01830150090013 PMID: 9075467
10. Bellenguez C, Kucukali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nature genetics. 2022; 54(4):412–36. https://doi.org/10.1038/s41588-022-01024-z PMID: 35379992
11. 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
12. 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.
13. Marioni RE, Harris SE, Zhang Q, McRae AF, Hagenaars SP, Hill WD, 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 PMID: 29777097
14. 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; 49:1373. https://doi.org/10.1038/ng.3916 PMID: 28714976
15. van der Lee SJ, Knol MJ, Chauhan G, Satizabal CL, Smith AV, Hofer E, et al. A genome-wide association study identifies genetic loci associated with specific lobar brain volumes. Commun Biol. 2019; 2:285. https://doi.org/10.1038/s42003-019-0537-9 PMID: 31396565
16. Wightman DP, Jansen IE, Savage JE, Shadrin AA, Bahrami S, Holland D, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. Nature genetics. 2021; 53(9):1276–82. https://doi.org/10.1038/s41588-021-00921-z PMID: 34493870
17. de Rojas I, Moreno-Grau S, Tesi N, Grenier-Boley B, Andrade V, Jansen IE, et al. Common variants in Alzheimer's disease and risk stratification by polygenic risk scores. Nature communications. 2021; 12 (1):3417. https://doi.org/10.1038/s41467-021-22491-8 PMID: 34099642
18. Oliynyk RT. Age-related late-onset disease heritability patterns and implications for genome-wide association studies. PeerJ. 2019; 7:e7168. https://doi.org/10.7717/peerj.7168 PMID: 31231601
19. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, et al. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med. 2020; 26(4):549–57. https://doi.org/10.1038/s41591-020-0800-0 PMID: 32273609
20. Fuller JT, Cronin-Golomb A, Gatchel JR, Norton DJ, Guzman-Velez E, Jacobs HIL, et al. Biological and Cognitive Markers of Presenilin1 E280A Autosomal Dominant Alzheimer's Disease: A Comprehensive Review of the Colombian Kindred. J Prev Alzheimers Dis. 2019; 6(2):112–20. https://doi.org/10.14283/jpad.2019.6 PMID: 30756118
21. Ridge PG, Hoyt KB, Boehme K, Mukherjee S, Crane PK, Haines JL, et al. Assessment of the genetic variance of late-onset Alzheimer's disease. Neurobiology of aging. 2016; 41:200 e13- e20. https://doi.org/10.1016/j.neurobiolaging.2016.02.024 PMID: 27036079
22. Slooter AJ, Cruts M, Kalmijn S, Hofman A, Breteler MM, Van Broeckhoven C, et al. Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. Arch Neurol. 1998; 55(7):964–8. https://doi.org/10.1001/archneur.55.7.964 PMID: 9678314
23. Ni G, Moser G, Schizophrenia Working Group of the Psychiatric Genomics C, Wray NR, Lee SH. Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood. American journal of human genetics. 2018; 102(6):1185–94. https://doi.org/10.1016/j.ajhg.2018.03.021 PMID: 29754766
24. Escott-Price V, Hardy J. Genome wide association studies for Alzheimer's disease: bigger is not always better. Brain Communications. 2022.
25. Alzheimer's Association. 2019 Alzheimer's disease facts and figures. Alzheimer's & Dementia. 2019; 15(3):321–87.
26. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. Neurology. 2013; 80(19):1778–83. https://doi.org/10.1212/WNL.0b013e31828726f5 PMID: 23390181

27. Niu H, Álvarez-Álvarez I, Guillén-Grima F, Aguinaga-Ontoso I. Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. Neurologia (Barcelona, Spain). 2017; 32(8):523–32.

28. Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. American journal of public health. 1998; 88(9):1337–42. https://doi.org/10.2105/ajph.88.9.1337 PMID: 9736873

29. Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol. 2013; 9(2):106–18. https://doi.org/10.1038/nrneurol.2012.263 PMID: 23296339

30. Zhang Q, Sidorenko J, Couvy-Duchesne B, Marioni RE, Wright MJ, Goate AM, et al. Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. Nature communications. 2020; 11(1). https://doi.org/10.1038/s41467-020-18534-1 PMID: 32968074

31. Mann DM, Yates PO, Marcyniuk B. Alzheimer's presenile dementia, senile dementia of Alzheimer type and Down's syndrome in middle age form an age related continuum of pathological changes. Neuropathology and applied neurobiology. 1984; 10(3):185–207. https://doi.org/10.1111/j.1365-2990.1984.tb00351.x PMID: 6234474

32. Paranjpe MD, Chaffin M, Zahid S, Ritchie S, Rotter JI, Rich SS, et al. Neurocognitive trajectory and proteomic signature of inherited risk for Alzheimer's disease. PLoS genetics. 2022; 18(9):e1010294. https://doi.org/10.1371/journal.pgen.1010294 PMID: 36048760

33. Fulton-Howard B, Goate AM, Adelson RP, Koppel J, Gordon ML, Alzheimer's Disease Genetics C, et al. Greater effect of polygenic risk score for Alzheimer's disease among younger cases who are apolipoprotein E-epsilon4 carriers. Neurobiology of aging. 2021; 99:101 e1- e9.

34. Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain. 2015; 138(12):3673–84. https://doi.org/10.1093/brain/awv268 PMID: 26490334

35. Bellou E, Baker E, Leonenko G, Bracher-Smith M, Daunt P, Menzies G, et al. Age-dependent effect of APOE and polygenic component on Alzheimer's disease. Neurobiology of aging. 2020; 93:69–77. https://doi.org/10.1016/j.neurobiolaging.2020.04.024 PMID: 32464432

36. Leonenko G, Baker E, Stevenson-Hoare J, Sierksma A, Fiers M, Williams J, et al. Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores. Nature communications. 2021. https://doi.org/10.1038/s41467-021-24082-z PMID: 34301930

37. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. American journal of human genetics. 2011; 88(3):294–305. https://doi.org/10.1016/j.ajhg.2011.02.002 PMID: 21376301

38. Gayan J, Galan JJ, Gonzalez-Perez A, Saez ME, Martinez-Larrad MT, Zabena C, et al. Genetic structure of the Spanish population. BMC Genomics. 2010; 11:326. https://doi.org/10.1186/1471-2164-11-326 PMID: 20500880

39. Moreno-Grau S, de Rojas I, Hernández I, Quintela I, Montrreal L, Alegret M, et al. Genome-wide association analysis of dementia and its clinical endophenotypes reveal novel loci associated with Alzheimer's disease and three causality networks: The GR@ACE project. Alzheimer's & Dementia. 2019; 15(10):1333–47. https://doi.org/10.1016/j.jalz.2019.06.4950 PMID: 31473137

40. Corneveaux JJ, Myers AJ, Allen AN, Pruzin JJ, Ramirez M, Engel A, et al. Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. Human molecular genetics. 2010; 19(16):3295–301. https://doi.org/10.1093/hmg/ddq221 PMID: 20534741

41. Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, et al. A survey of genetic human cortical gene expression. Nat Genet. 2007; 39(12):1494–9. https://doi.org/10.1038/ng.2007.16 PMID: 17982457

42. Petyuk VA, Chang R, Ramirez-Restrepo M, Beckmann ND, Henrion MYR, Piehowski PD, et al. The human brainome: network analysis identifies HSPA2 as a novel Alzheimer's disease target. Brain. 2018.

43. Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, Holmans P, et al. Genetic control of human brain transcript expression in Alzheimer disease. Am J Hum Genet. 2009; 84(4):445–58. https://doi.org/10.1016/j.ajhg.2009.03.011 PMID: 19361613

44. Bennett D, Schneider J, Arvanitakis Z, Wilson R. Overview and Findings from the Religious Orders Study. Current Alzheimer research. 2011; 9:628–45.

45. Bennett DA, Schneider JA, Buchman AS, Barnes LL, Boyle PA, Wilson RS. Overview and findings from the rush Memory and Aging Project. Curr Alzheimer Res. 2012; 9(6):646–63. https://doi.org/10.2174/156720512801322663 PMID: 22471867

46. Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA. Religious Orders Study and Rush Memory and Aging Project. Journal of Alzheimer's Disease. 2018; 64:S161–S89. https://doi.org/10.3233/JAD-179939 PMID: 29865057

47. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015; 12(3):e1001779. https://doi.org/10.1371/journal.pmed.1001779 PMID: 25826379

48. van der Flier WM, Scheltens P. Amsterdam Dementia Cohort: Performing Research to Optimize Care. J Alzheimers Dis. 2018; 62(3):1091–111. https://doi.org/10.3233/JAD-170850 PMID: 29562540

49. Tesi N, van der Lee SJ, Hulsman M, Jansen IE, Stringa N, van Schoor NM, et al. Immune response and endocytosis pathways are associated with the resilience against Alzheimer's disease. Transl Psychiatry. 2020; 10(1):332. https://doi.org/10.1038/s41398-020-01018-7 PMID: 32994401

50. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A Tool for Genome-wide Complex Trait Analysis. The American Journal of Human Genetics. 2011; 88(1):76–82. https://doi.org/10.1016/j.ajhg.2010.11.011 PMID: 21167468

51. Yang J, Bakshi A, Zhu Z, Hemani G, Vinkhuyzen AAE, Lee SH, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. Nature genetics. 2015; 47(10):1114–20. https://doi.org/10.1038/ng.3390 PMID: 26323059

52. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nature genetics. 2015; 47(11):1236–41. https://doi.org/10.1038/ng.3406 PMID: 26414676

53. Zhang Q, Prive F, Vilhjalmsson B, Speed D. Improved genetic prediction of complex traits from individual-level data or summary statistics. Nature communications. 2021; 12(1):4192. https://doi.org/10.1038/s41467-021-24485-y PMID: 34234142

54. Andrews SJ, Fulton-Howard B, Goate A. Interpretation of risk loci from genome-wide association studies of Alzheimer's disease. The Lancet Neurology. 2020; 19(4):326–35. https://doi.org/10.1016/S1474-4422(19)30435-1 PMID: 31986256

55. Gosselin D, Skola D, Coufal NG, Holtman IR, Schlachetzki JCM, Sajti E, et al. An environment-dependent transcriptional network specifies human microglia identity. Science (New York, NY). 2017; 356 (6344):eaal3222. https://doi.org/10.1126/science.aal3222 PMID: 28546318

56. Tansey KE, Cameron D, Hill MJ. Genetic risk for Alzheimer's disease is concentrated in specific macrophage and microglial transcriptional networks. Genome Med. 2018; 10(1):14. https://doi.org/10.1186/s13073-018-0523-8 PMID: 29482603