

# Associations of Alzheimer's disease with inpatient hospital costs and with quality-adjusted life years: Evidence from conventional and Mendelian randomization analyses in the UK Biobank

Padraig Dixon<sup>1\*</sup>, Emma Anderson<sup>2</sup>

1: Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

2: Division of Psychiatry, University College London, London, UK

December 2023

## ABSTRACT

**BACKGROUND** Alzheimer's disease and other dementias are progressive neurodegenerative disorders with profound impacts on cognitive function. There is a shortage of economic evidence relating to the impact Alzheimer's disease on healthcare costs and quality-adjusted life-years (QALYs).

**METHODS.** We employed two study designs to model the association between Alzheimer's disease and healthcare costs and QALYs. We first estimated conventional multivariable models of the association between Alzheimer's disease and these core economic outcomes. However, these types of model may be confounded by diseases, processes, or traits that independently affect Alzheimer's disease and either or both of healthcare costs and QALYs. We therefore also explored a complementary approach using germline genetic variation as instrumental variables in a Mendelian randomization analysis. We used single nucleotide polymorphisms (SNPs) identified in recent genome-wide association studies of Alzheimer's disease as instruments. We studied outcome data on inpatient hospital costs and QALYs in the UK Biobank cohort.

**RESULTS** Data from up to 310,838 individuals were analyzed. N=55 cases of Alzheimer's disease were reported at or before recruitment into UK Biobank. A further N=284 incident cases were identified over follow-up. Multivariable observational analysis of the prevalent cases suggested significant impacts on costs (£1,140 in cases, 95% Confidence Interval (CI): £825 to £1,456) and QALYs (-25%, 95% CI: -28% to -21%). Mendelian randomization estimates were very imprecise for costs (£3,082, 95% CI: -£7,183 to £13,348) and QALYs (-32%, 95% CI: -149% to 85%), likely due to the small proportion of variance (0.9%) explained in Alzheimer's disease status by the most predictive set of SNPs.

**IMPLICATIONS** Conventional multivariable models suggested important impacts of Alzheimer's disease on inpatient hospital costs and QALYs, although this finding was based on very few cases which may have included instances of early-onset dementia. Mendelian randomization was very imprecise. Larger GWAS of clinical cases, improved understanding of the architecture

of the disease, and the follow-up of cohorts until old age and death will help overcome these challenges.

**Keywords:** Alzheimer’s disease, dementia, genetics, instrumental variables, healthcare costs, quality of life, QALYs, Mendelian Randomization

---

\*Corresponding author. Pdraig Dixon, Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Primary Care Building, Radcliffe Observatory Quarter, Woodstock Rd, Oxford OX2 6GG, [padraig.dixon@phc.ox.ac.uk](mailto:padraig.dixon@phc.ox.ac.uk)

# 1 Introduction

Alzheimer's disease, the most common cause of dementia, is a progressive neurodegenerative disorder that primarily affects cognitive function (1). Worldwide, some 57 million people were estimated to live with Alzheimer's disease and other dementias in 2019 (2), a number that has more than doubled since 1990 because of population growth and population ageing (2)(3). The consequences of Alzheimer's disease can be profound, especially as the disease progresses, and may include language impairment, personality changes, memory loss, and inability to manage daily tasks including self-care and rudimentary decision-making (1).

There are no cures for Alzheimer's disease. Treatment of symptoms may stabilize cognitive function in some cases (4), but non-pharmacotherapeutic care and management remains essential for most patients. A key consideration for any intervention is cost-effectiveness, which is a function of the costs, quality-of life, and mortality impacts associated with both the disease and intervention under consideration. There have been relatively few economic evaluations of interventions for Alzheimer's disease in comparison to other diseases, meaning that informed decisions are challenged by lack of evidence on how interventions affect these economic outcomes (5, 6). This type of evidence is particularly important for decision-analytic models of interventions for Alzheimer's disease.

In this paper, we addressed this evidence gap by estimating associations of Alzheimer's disease with healthcare costs and with quality-adjusted life years (QALYs) in the UK Biobank (7) cohort using two types of study design. We first explored the associations of Alzheimer's disease with costs and QALYs in conventional multivariable models, which we henceforth simply refer to as multivariable models, using linear adjusted linear regressions that adjust for basic demographic information. Multivariable models are widely used as means of assessing associations between disease status and cost and QALY outcomes but may be confounded by omitted variables that influence both disease status and these outcomes. These confounding variables may include any processes or traits that independently influence Alzheimer's disease and either or both of healthcare costs and QALYs.

We therefore also used germline genetic variation associated with the risk of Alzheimer's disease as instrumental variables in a Mendelian Randomization. Mendelian Randomization exploits the natural experiment of quasi-random genetic variation allocation from parents to children during conception. Some of this allocation includes genetic variants known to influence the susceptibility to Alzheimer's disease. This approach therefore avoids, in principle, confounding influences that might otherwise bias conventional multivariable models.

## **2 Data**

Our principal data source was the prospective UK Biobank cohort (research ethics approval reference 11/NW/0382), which recruited over 500,000 people between 2006 and 2010 from 22 centres across England, Scotland, and Wales (6)(7–9).

### **2.1 Construction of cost and QALY data**

We have previously described our creation of healthcare cost and quality-adjusted life years data (8-10). Briefly, the UK Biobank data was linked to medical records containing information on inpatient hospital episodes of care associated with diagnoses and procedures. From these data, we created per-patient, per year costs of inpatient hospital care from the linked Hospital Episode Statistics resource. We calculated costs from recruitment into Biobank up to the date of participant death or the censoring date for our Hospital Episode Statistics data of 31 March 2015.

QALY data were constructed by assigning quality of life decrements associated with 240 different health conditions (11) to the Biobank cohort. These specific health conditions were reported by individuals or recorded within their health records. This, in conjunction with data on mortality, facilitated the development of individual quality-adjusted life years (QALYs) from the time of recruitment until March 31, 2017. The individual, person-specific QALY estimates represent absolute percentage changes in QALYs over an average year of follow-up.

### **2.2 Identification of cases**

Alzheimer's disease cases in UK Biobank were defined as follows. Cases were identified as occurring at or before the baseline UK Biobank recruitment appointment if they were self-reported at that appointment, or if there was a prior record of inpatient care before the baseline appointment coded with ICD9 codes beginning with 331 or with ICD10 codes beginning with F00 ("Dementia in Alzheimer disease") or G30 ("Alzheimer disease"). The self-report question was not specific to Alzheimer's disease but instead queried history of "Dementia/Alzheimers/Cognitive Impairment" Only cases identified at or before the baseline recruitment appointment were used in the analysis to ensure that the exposure was known to have happened before outcomes were observed. In addition to these cases, we also identified incident cases from either hospital records (using the ICD codings above) or from self-report at a post-baseline interview to which some participants were invited. (23). For the Mendelian randomization analysis, we considered both prevalent (known at or before baseline) and incident cases (occurring after baseline recruitment) since the exposure variable was PRSs instrumenting lifelong genetic liability to Alzheimer's disease. The multivariable analysis considered only Alzheimer's disease reported at or before baseline to ensure the cost and QALY outcomes were observed after exposure was recorded.

### **2.3 Polygenic risk scores for Alzheimer's disease**

We created weighted polygenic risk scores (PRSs) from genome-wide association studies (GWASs) for the two-sample Mendelian randomization analysis. We used Alzheimer's disease GWASs that did not use data from the UK Biobank cohort (our outcome data source) in order to avoid bias from sample overlap (12). We examined two recent GWASs in individuals of European ancestry with clinically confirmed Alzheimer's disease or dementia: Kunkle et al (13) and Bellenguez et al (14). The Kunkle et al GWAS was published in 2019 and analysed data from 63,296 individuals in a discovery sample with clinically diagnosed late-onset Alzheimer's disease. Bellenguez et al was published in 2022 and used a two-stage design with up 487,511 individuals included in discovery sample analysis comprising both diagnosed and proxy cases (based on whether respondents reported their parents having dementia), although we only used results from the GWAS of clinically diagnosed cases in our analyses.

We used an in-house protocol (15) to quality-control genetic data from UK Biobank. We applied an  $R^2$  threshold of 0.001 within a 10,000-kilobase window to clump genome-wide significant SNPs to minimize linkage disequilibrium. For SNPs not present in the UK Biobank outcome dataset, we conducted a search for proxy SNPs using the European subsample of the 1,000 Genomes Project as a reference panel, setting a lower  $R^2$  limit of 0.6 (11). We restricted the main analysis to unrelated individuals of white British ancestry (to avoid confounding by ancestral background) living in England or Wales at recruitment, and after removal for sex mismatches, sex chromosomal abnormalities, and withdrawals of consent.

Using the set of SNPs remaining after this process, PRSs were calculated as the weighted sum of the risk-increasing effect alleles of genome-wide significant SNPs, with each SNP's weight determined by the log-odds ratio coefficient obtained from the corresponding GWAS. We proxied variation explained in Alzheimer's disease by the PRS using Nagelkerke's pseudo  $R^2$  calculated from a logistic regression of Alzheimer's disease case status on each respective PRS, as well as the  $R^2$  adjusted for the effective sample size in these logistic regressions, which is a function of the proportion of observations at each distinct value of the Alzheimer's disease case status variable.

We also examined the impact of including two variants from the APOE (Apolipoprotein E) region, encoded by the rs7412 and rs429358 SNPs. The APOE  $\epsilon 2$  allele (encoded by the rs7412 SNP) is protective for incident Alzheimer's disease, the APOE  $\epsilon 4$  allele (encoded by rs429358) increases risk, and together the interplay between these two SNPs exerts an important influence on genetic susceptibility to Alzheimer's disease (16, 17). We therefore examined the impact of excluding this pair of SNPs from each analysis and assessed how well these two SNPs alone explained disease-case status.

## **2.4 Data and code availability**

The derived cost and QALY variables will be archived with UK Biobank as returned variables and will be made available to individuals who obtain the necessary permissions from the study's

data access committee. Analysis code is available here: <https://github.com/pdixon-econ/alzheimer-cost-qaly>.

### 3 Methods

We estimated conventional multivariable linear regressions for QALYs and healthcare costs in relation to Alzheimer's disease status, with age, sex, 40 genetic principal components, Townsend Deprivation Index, and recruitment centre as controls. We also re-ran these regressions separately for men and women. All multivariable analyses were restricted to the same group of individuals for which we could perform the Mendelian Randomization analysis, which comprised individuals who gave consent to be genotyped, and who passed the quality control and processing steps described below. The multivariable analysis differed from the Mendelian Randomization analyses by studying only prevalent cases (reported before or at recruitment into UK Biobank) to ensure exposure to Alzheimer's disease occurred before the outcomes were observed.

We then undertook the Mendelian Randomization analysis (18-21). Under three core assumptions, Mendelian Randomization may be used to estimate the causal effect of an exposure on an outcome. The first of these assumptions is that the instrument be associated with the exposure of interest. The genetic variants used as instruments in MR are typically single nucleotide polymorphisms (SNPs), which represent single nucleotide changes in the genetic code. SNPs are a type of allele; alleles are the alternative forms of a gene at a particular location in the genome. The associations of specific SNPs with disease risk are typically tested using genome-wide association studies. The second assumption, sometimes referred to as the independence requirement, is that there is no confounding of the instrument/outcome relationship. The third assumption, often known as the exclusion restriction, is that the instrument affects the outcome only via the exposure. These may include other diseases, traits or behaviours, or may involve processes such population structure (since ancestral background influences both the distribution of genetic variants and also disease risk not necessarily associated with those variants) and geographic structure (since individuals with the same or

similar ancestral background living in different types of location may experience different environmental influences that affect the risk of Alzheimer's disease).

Two other important influences may also violate these assumptions. First, horizontal pleiotropy – the influence of a single variant on multiple phenotypes – may violate the exclusion restriction assumption if a variant influences the outcome by paths that are independent of the exposure. Second, linkage disequilibrium, which refers to the non-random association of alleles at different genomic loci to be inherited together. This could lead to the association of a variant included in analysis with excluded variants, potentially violating the third assumption.

Neither the second nor the third assumption is fully testable, but potential violations can be identified. Below, we describe the modelling approaches and data processing steps we took to account for possible failures of the core Mendelian Randomization assumptions. If the assumptions are satisfied, the analysis will identify groups that have varying levels of exposure to the risk of Alzheimer's disease and for which outcomes – costs and QALYs – may be compared.

### **3.1 Polygenic risk scores**

For the Mendelian randomization models, we implemented two-stage least square (2SLS) instrumental variable models. We first regressed the Alzheimer's disease case status variable on the PRS, adjusting for age at baseline recruitment in UK Biobank, the first forty genetic principal components (as a control for population structure), and the UK Biobank recruitment location (as a control for geographic structure). Predicted values from this first stage regression were then used in a second stage regression on the cost and QALY outcomes respectively. 2SLS models were estimated using the AER package (22) in R. We inspected F-statistics from the first-stage regression to evaluate instrument strength (23).

The effect estimates from these PRS models reflect genetic liability to disease, and may be interpreted as the average effect (on either costs or QALYs) per unit increase in the log-odds of having Alzheimer's disease (23, 24). This assumes either a constant effect of genetic liability to



Alzheimer's disease on the outcomes, or a monotonic effect of the SNPs on Alzheimer's disease risk under a local average treatment effect interpretation.

### **3.2 Pleiotropy-robust sensitivity analysis**

We also examined potential violations of the exclusion restriction assumption due to horizontal pleiotropy. We examined heterogeneity in effect estimates across individual SNPs using Cochran's Q statistic (25). Heterogeneity may be due to sampling error, or to the presence of SNPs with a large influence (relative to the mean effect of all SNPs) on outcome, which could suggest a channel or channels of influence that go from the SNP to the outcome other than via Alzheimer's disease. This latter form of heterogeneity would indicate a likely violation of the exclusion restriction.

In our main analysis, we calculated IVW estimates for individual SNP instrumental variables, combining them via random-effects meta-analysis with weights based on SNP-outcome precision. This assumes no exclusion restriction violations, or restrictions that affect only the variance of estimates but not the overall effect size. We then implemented three estimators robust to more general forms of pleiotropy, each with their own assumptions: the MR Egger estimator, which is consistent under assumptions related to instrument strength and pleiotropic effects (14, 15); a penalized weighted median estimator, consistent if  $\geq 50\%$  SNPs are valid instruments (16); and a weighted mode estimator, consistent if the largest homogenous SNP cluster is valid, even if more than 50% SNPs are invalid instruments (17).

## **4 Results**

We analysed data from data from up to 310,838 unrelated individuals in multivariable and Mendelian randomization analyses, of whom 53.5% were female. Mean age at recruitment in this sample was 56.9 years. Cost data were available for a median of 6.1 years of follow-up and QALY data for a mean of 8.1 years. Mean healthcare costs per person, per year of follow-up were £88 in 2019 prices and mean per-person QALYs (averaged over all years of follow-up) were 0.91.

There were 55 cases of Alzheimer’s disease self-reported at or before baseline, and 339 cases including these cases as well as incident diagnoses recorded in inpatient care records over the follow-up. Thus just 0.1% of the sample had Alzheimer’s disease before or during UK Biobank recruitment and follow-up, although 12.6% (n=39,195) of the sample reported that at least one of their parents had Alzheimer’s disease.

A PRS comprised of 22 SNPs based on Kunkle et al (13) (including the rs7412 and rs429358 SNPs in the APOE locus) was the most predictive of the PRSs tested, with a Nagelkerke pseudo-R<sup>2</sup> of 0.9% and an R<sup>2</sup> adjusted for effective sample size of 3.9%. Results for other PRSs tested are given in Supplementary material (Table S1) and range from <0.0% to 0.9% for Nagelkerke’s R<sup>2</sup> and <0.0% to 3.9% for R<sup>2</sup> adjusted for effective sample size. This 22-SNP PRS was associated with prevalent and incident Alzheimer’s disease (log odds-ratio 0.57, 95% CI: 0.41 to 0.74) and with self-reported parental history of Alzheimer’s disease (log odds-ratio 0.25, 95% CI: 0.23 to 0.27), but there was no evidence of an association with the self-reported baseline cases (log odds-ratio 0.16, 95% CI:-0.30 to 0.62). This latter finding could be a consequence of the non-specific self-reported question asked at baseline in UK Biobank, instances of early-onset dementia assessed against a PRS for late-onset disease, or low statistical power given the small number of cases.

#### 4.1 Results of main analysis

The multivariable analysis indicated material increases in cost and decreases QALYs associated with self-reported cases of Alzheimer’s disease at baseline. These associations were similar for men and women (Table 1) with overlapping confidence intervals, although point estimates for men indicated more severe outcomes for both costs and QALYs.

**Table 1** Adjusted multivariable estimates of association between self-reported Alzheimer’s disease and costs and QALYs

	N	Effect estimate	95% confidence interval
<b>Healthcare costs</b>			
All	310,496	£1,140	£825 to £1,456

Men	143,841	£1,281	£851 to £1,710
Women	166,655	£906	£432 to £1,380
<b>QALYs</b>			
All	310,496	-25%	-28% to -21%
Men	143,841	-27%	-32% to -22%
Women	166,655	-21%	-26% to -15%

These models adjust for age, sex, the Townsend deprivation index, and UK Biobank recruitment centre. Note that Townsend deprivation index was not available at baseline for all Biobank participants and the N of these regressions is therefore smaller than the total sample size.

The Mendelian randomization estimates were very imprecise, being consistent with both large and small impacts of Alzheimer’s disease on both costs and QALYs (Table 2). The uncertainty around point estimates means that conclusions about the potential magnitudes of these impacts is not possible. Given this uncertainty, we did not disaggregate results by sex.

**Table 2** 2SLS PRS estimates

	Effect estimate	95% confidence interval
<b>Outcome</b>		
Healthcare costs	£3,082	-£7,183 to £13,348
QALYs	-32%	-149% to 85%

The F-statistic was 47.8 for both costs and QALY outcomes (since the first stage regression in each case is identical). While this is above the conventional  $F > 10$  rule of thumb sometimes used in this context, weak instrument bias relates to instrument strength in a continuous rather than binary manner and we cannot rule out the possibility of bias. If present, weak instruments would bias these Mendelian randomization estimates in the direction of the conventional multivariable estimates.

## 4.2 Sensitivity analysis

Cochran’s Q test for heterogeneity was consistent with the null for the cost outcome ( $Q=13.3$ ,  $p=0.82$ ) but there was evidence of heterogeneity in the QALY estimates ( $Q=39.8$ ,  $p<0.01$ ). Inspection of the IVW and various pleiotropy robust estimators (MR Egger, penalized weighted median, and weighted mode) revealed similar point estimates for the QALY outcome but these

were consistent with the null in all cases and associated with considerable uncertainty (See Supplementary Material Table S3).

## 5 Discussion

We examined the association between Alzheimer's disease and the core health economic outcomes of healthcare cost and QALYs using two types of study design in the large UK Biobank cohort. Our analysis was limited to data from individuals of European ancestry. Conventional multivariable estimates indicated large impacts on each outcome when assessed using cases identified at or before recruitment into the UK Biobank cohort. However, the number of such cases was very small relative to the size of the cohort. Moreover, these cases may have included instances of early onset dementia, although we cannot establish if this the case.

We also estimated Mendelian randomization models using genome-wide significant genetic variants from Kunkle et al. While this type of analysis may be less biased by confounding and reverse causation than the multivariable models, the results of the Mendelian randomization analysis (both the 2SLS models and the pleiotropy robust models) were too imprecise to draw meaningful conclusions about the magnitude of Alzheimer's disease impacts on these outcomes. This uncertainty was likely driven by the small number of cases in the UK Biobank cohort, and the modest explanatory power of the best performing polygenic risk score analysed.

This is the first Mendelian randomization analysis of the causal effect of Alzheimer's disease on QALYs and any form of healthcare cost. The literature assessing and evaluating findings from more conventional study designs against these outcomes typically reports very large cost impacts e.g. (26-31). The 1999 work by Neumann et al (32) remains a heavily used source of evidence on quality-of-life associated with this condition, finding that disease status was a significant predictor of health-related quality-of-life scores amongst patients, with negative impacts increasing in the severity of the condition. A recent systematic review and meta-analysis (32) of health-related quality-of-life also found this dose-dependent relationship over

disease severity. Point estimates were broadly comparable to the point effects obtained in our multivariable and Mendelian Randomization estimates.

Our cost data was limited to inpatient costs incurred in NHS hospitals linked to the Biobank cohort. We did not model outpatient or other healthcare costs; for example, we could not account for care provided in outpatient neurology clinics. Moreover, inpatient costs are likely to be much higher in the advanced stages of disease and towards the end of life; thus, we may not fully capture the extent of those in a relatively young cohort like the UK Biobank. Total healthcare costs associated Alzheimer's disease are very likely to be higher than those we report here, although in principle the Mendelian randomization estimates may offer an unbiased estimate of exposure to the condition on inpatient hospital costs for this age range. We also did not have access to carer costs, which are a very significant issue for most Alzheimer's disease patients and their families (33). While research continues on the most appropriate way to account for carer burden in economic evaluation (34, 35), further evidence on the scale of potential impacts in this group is needed (36-39).

We generated QALYs by assigning quality of life decrements to 240 specific health conditions reported in episodes of inpatient hospital care, amongst which conditions were Alzheimer's disease and other dementias. This created a modest degree of circularity, since case status therefore appeared in both the exposure and outcome. However, the overall contribution of case status to variability in the QALY outcome was very small given the low number of such cases recorded in UK Biobank. Our indirect calculation of quality-of-life data avoided complexities associated with declining ability to self-complete quality-of-life questionnaires as the disease progresses (40). The QALY calculation was based on conditions reported in episodes of hospital care and this may have overstated quality-of-life at baseline by considering only inpatient care.

Two types of selection bias may have affected both types of analysis. The first is that UK Biobank is a volunteer cohort, which relied on participants being sufficiently healthy and motivated to participate. While Alzheimer's disease primarily manifests in later life, very unwell individuals, including with early or prodromal disease, would have been less likely to participate

(41). UK Biobank participants are known to be healthier (on average) compared to the wider population from which the cohort was drawn (31). Even if this type of selection related to morbidity and survival was not present or not material (given the late onset of most cases of Alzheimer's disease), the implicit conditioning on participation may also amount to survival or collider bias (42). This would arise if factors related to Alzheimer's disease and the outcomes (such as high healthcare costs or low quality-of-life driven by significant morbidity) themselves influenced participation, which could confound the associations that we present in both multivariable and Mendelian randomization models.

The Mendelian randomization estimates reflect the impact of lifelong exposure to elevated genetic liability to incident Alzheimer's disease, whereas the multivariable estimates reflect both this lifelong exposure as well as possible confounding influences from other variables. Neither the multivariable nor the Mendelian randomization estimate necessarily correspond to how outcomes would change under modification of Alzheimer's disease risk, including via hypothetical manipulation of disease risk factors. Nevertheless, Mendelian randomization estimates likely reflect many of the common germline influences on genetic risk for this condition.

Our analysis has highlighted several areas for future work. There remains a need for further, larger GWASs of clinically diagnosed Alzheimer's disease (43, 44). Improved understanding of the genetic basis of the disease will support aetiology more generally, as well as downstream causal analysis utilizing Mendelian randomization. The Mendelian randomization analysis focused on relatively common SNPs; future analysis could explore the impact of rare variants with large effects on disease risk (43). UK Biobank aimed to recruit individuals aged between 40 and 69, and linked care records enable longitudinal aspects of disease incidence and progression to be evaluated as the cohort ages. However, recruiting older individuals to future cohorts, as will continued study of cohorts such as the various international longitudinal studies of ageing, will comprise an important element of future research in this area.

## 6 Conclusion

Alzheimer's disease is associated with significant impacts on the core health economic outcomes of (inpatient hospital) costs and QALYs. However, these associations were estimated with considerable uncertainty in causal analyses. Further interrogation and understanding of the genetic architecture of Alzheimer's disease, and on the wider costs and quality-of-life impacts (including at the more advanced stages of disease) of the condition, will shed further light on these impacts.

## Declarations

**Funding statement:** P.D. received support from NIHR ARCs Dementia Capacity Building Post-Doctoral Training Scheme (DEM-COMM). E.A. is supported by a UKRI Future Leaders Fellowship (MR/W011581/1)

**Conflict of interest statement:** The authors declare no conflicts of interest.

**Acknowledgments:** This research was conducted using the UK Biobank Resource under Application Number 29294.



## References

1. Scheltens P, De Strooper B, Kivipelto M, Holstege H, Chételat G, Teunissen CE, et al. Alzheimer's disease. *The Lancet*. 2021;397(10284):1577-90.
2. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Health*. 2022;7(2):e105-e25.
3. Nichols E, Szeke CEI, Vollset SE, Abbasi N, Abd-Allah F, Abdela J, et al. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Neurology*. 2019;18(1):88-106.
4. Soria Lopez JA, González HM, Léger GC. Chapter 13 - Alzheimer's disease. In: Dekosky ST, Asthana S, editors. *Handb Clin Neurol*. 167: Elsevier; 2019. p. 231-55.
5. Knapp M, Wong G. Economics and dementia: Challenges and responses. *Dementia*. 2023;14713012231193141.
6. Landeiro F, Morton J, Gustavsson A, Potashman M, Lecomte P, Belger M, et al. Health economic modeling for Alzheimer's disease: Expert perspectives. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*. 2022;8(1):e12360.
7. 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.
8. Dixon P, Davey Smith G, Hollingworth W. The Association Between Adiposity and Inpatient Hospital Costs in the UK Biobank Cohort. *Appl Health Econ Health Policy*. 2018.
9. Dixon P, Hollingworth W, Harrison S, Davies NM, Davey Smith G. Mendelian Randomization analysis of the causal effect of adiposity on hospital costs. *J Health Econ*. 2020;70:102300.
10. Harrison S, Dixon P, Jones HE, Davies AR, Howe LD, Davies NM. Long-term cost-effectiveness of interventions for obesity: A mendelian randomisation study. *PLoS Med*. 2021;18(8):e1003725.
11. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. *Med Decis Making*. 2011;31(6):800-4.
12. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol*. 2016;40(7):597-608.
13. 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 A $\beta$ , tau, immunity and lipid processing. *Nat Genet*. 2019;51(3):414-30.
14. Bellenguez C, Küçükali F, Jansen IE, Kleindam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. 2022;54(4):412-36.
15. Mitchell R, Hemani G, Dudding T, Corbin L, S. H, Paternoster L. UK Biobank Genetic Data: MRC-IEU Quality Control, version 2. 2019.

16. Kulminski AM, Shu L, Loika Y, He L, Nazarian A, Arbeev K, et al. Genetic and regulatory architecture of Alzheimer's disease in the APOE region. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*. 2020;12(1):e12008.
17. Yamazaki Y, Zhao N, Caulfield TR, Liu C-C, Bu G. Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. *Nature Reviews Neurology*. 2019;15(9):501-18.
18. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. *Nature Reviews Methods Primers*. 2022;2(1):6.
19. Pingault J-B, O'Reilly PF, Schoeler T, Ploubidis GB, Rijdsdijk F, Dudbridge F. Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics*. 2018;19(9):566-80.
20. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1-22.
21. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89-98.
22. Kleiber C, Zeileis A, Zeileis MA. Package 'aer'. R package version. 2020;12(4).
23. Clarke PS, Windmeijer F. Instrumental variable estimators for binary outcomes. *Journal of the American Statistical Association*. 2012;107(500):1638-52.
24. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res*. 2015.
25. Burgess S, Thompson S. *Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation*. Boca Raton, Florida: CRC Press; 2015.
26. Colucci L, Bosco M, Fasanaro AM, Gaeta GL, Ricci G, Amenta F. Alzheimer's disease costs: what we know and what we should take into account. *J Alzheimers Dis*. 2014;42(4):1311-24.
27. El-Hayek YH, Wiley RE, Khoury CP, Daya RP, Ballard C, Evans AR, et al. Tip of the Iceberg: Assessing the Global Socioeconomic Costs of Alzheimer's Disease and Related Dementias and Strategic Implications for Stakeholders. *J Alzheimers Dis*. 2019;70(2):323-41.
28. Alzheimer's Association. 2023 Alzheimer's disease facts and figures. Chicago; 2023.
29. Aranda MP, Kremer IN, Hinton L, Zissimopoulos J, Whitmer RA, Hummel CH, et al. Impact of dementia: Health disparities, population trends, care interventions, and economic costs. *J Am Geriatr Soc*. 2021;69(7):1774-83.
30. Kelley AS, McGarry K, Gorges R, Skinner JS. The burden of health care costs for patients with dementia in the last 5 years of life. *Ann Intern Med*. 2015;163(10):729-36.
31. Cantarero-Prieto D, Leon PL, Blazquez-Fernandez C, Juan PS, Cobo CS. The economic cost of dementia: A systematic review. *Dementia*. 2019;19(8):2637-57.
32. Neumann PJ, Kuntz KM, Leon J, Araki SS, Hermann RC, Hsu MA, et al. Health utilities in Alzheimer's disease: a cross-sectional study of patients and caregivers. *Med Care*. 1999;37(1):27-32.
33. Kließ MK, Martins R, Connolly MP. Major Cost Drivers in Assessing the Economic Burden of Alzheimer's Disease: A Structured, Rapid Review. *The Journal of Prevention of Alzheimer's Disease*. 2021;8(3):362-70.

34. Dixon P, Round J. Caring for Carers: Positive and Normative Challenges for Future Research on Carer Spillover Effects in Economic Evaluation. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2019;22(5):549-54.
35. Al-Janabi H, Van Exel J, Brouwer W, Coast J. A Framework for Including Family Health Spillovers in Economic Evaluation. *Med Decis Making*. 2015.
36. Nickel F, Barth J, Kolominsky-Rabas PL. Health economic evaluations of non-pharmacological interventions for persons with dementia and their informal caregivers: a systematic review. *BMC Geriatr*. 2018;18(1):69.
37. Pelone F, Jacklin P, Francis JM, Purchase B. Health economic evaluations of interventions for supporting adult carers in the UK: a systematic review from the NICE Guideline. *Int Psychogeriatr*. 2022;34(9):839-52.
38. Lin P-J, D’Cruz B, Leech AA, Neumann PJ, Sanon Aigbogun M, Oberdhan D, et al. Family and Caregiver Spillover Effects in Cost-Utility Analyses of Alzheimer’s Disease Interventions. *Pharmacoeconomics*. 2019;37(4):597-608.
39. Lin P-J, Neumann PJ. Valuing Alzheimer Disease Therapies—Considering Costs and Benefits Beyond the Patient. *JAMA Network Open*. 2021;4(10):e2131913-e.
40. Keetharuth AD, Hussain H, Rowen D, Wailoo A. Assessing the psychometric performance of EQ-5D-5L in dementia: a systematic review. *Health and Quality of Life Outcomes*. 2022;20(1):139.
41. Tyrrell J, Zheng J, Beaumont R, Hinton K, Richardson TG, Wood AR, et al. Genetic predictors of participation in optional components of UK Biobank. *Nature Communications*. 2021;12(1):886.
42. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol*. 2017;47(1):226-35.
43. Andrews SJ, Renton AE, Fulton-Howard B, Podlesny-Drabiniok A, Marcora E, Goate AM. The complex genetic architecture of Alzheimer's disease: novel insights and future directions. *eBioMedicine*. 2023;90.
44. Escott-Price V, Hardy J. Genome-wide association studies for Alzheimer’s disease: bigger is not always better. *Brain Communications*. 2022;4(3):fcac125.