



Article

The Effect of Genetic Predisposition to Alzheimer's Disease and Related Traits on Recruitment Bias in a Study of Cognitive Aging

Lina M. Gomez¹ , Brittany L. Mitchell¹, Kerrie McAloney¹, Jessica Adsett¹, Natalie Garden¹, Madeline Wood¹, Santiago Diaz-Torres^{1,2}, Luis M. Garcia-Marin^{1,2}, Michael Breakspear³, Nicholas G. Martin¹  and Michelle K. Lupton^{1,2,4}

¹QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia, ²School of Biomedical Sciences, Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia, ³School of Psychological Sciences, The University of Newcastle, Newcastle, New South Wales, Australia and ⁴School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Queensland, Australia

Abstract

The recruitment of participants for research studies may be subject to bias. The Prospective Imaging Study of Ageing (PISA) aims to characterize the phenotype and natural history of healthy adult Australians at high future risk of Alzheimer's disease (AD). Participants approached to take part in PISA were selected from existing cohort studies with available genomewide genetic data for both successfully and unsuccessfully recruited participants, allowing us to investigate the genetic contribution to voluntary recruitment, including the genetic predisposition to AD. We use a polygenic risk score (PRS) approach to test to what extent the genetic risk for AD, and related risk factors predict participation in PISA. We did not identify a significant association of genetic risk for AD with study participation, but we did identify significant associations with PRS for key causal risk factors for AD, IQ, household income and years of education. We also found that older and female participants were more likely to take part in the study. Our findings highlight the importance of considering bias in key risk factors for AD in the recruitment of individuals for cohort studies.

Keywords: Alzheimer's disease; recruitment; bias; genetics; polygenic; PRS

(Received 22 May 2023; accepted 30 May 2023; First Published online 21 July 2023)

The recruitment process for participants in research studies can be biased due to factors affecting their likelihood to volunteer, resulting in a sample that is not necessarily representative of the general population (Oswald et al., 2013). Recruitment bias can be exacerbated when the target population is composed of individuals with a specific disorder or related history, which may affect their willingness to participate (Patten, 2000). For instance, individuals with lower cognitive function may be more willing to enrol in research, as they may perceive themselves as being in greater need of interventions (Li et al., 2022). Similarly, people with dementia or a family history of dementia may be motivated to participate in research as a way to contribute to the development of treatment strategies (Milani et al., 2021). Several factors, such as lack of confidence, depressive symptoms (Ghanbari et al., 2023), unemployment and low educational attainment (Haring et al., 2009) may lead to a decline in participation, while willingness to participate has been shown to be associated with favorable profiles, such as younger age and higher education attainment (Enzenbach et al., 2019; Ganguli et al., 2015).

Recruitment bias can impact prevalence estimates. In addition, it can affect association estimates, caused by collider bias, where

two variables independently influence a third collider variable. Conditioning on the collider (in this case, study participation) can contribute to a spurious association between the two variables (Zheng et al., 2012).

Alzheimer's disease (AD) is a weakening neurological condition, characterized by progressive neurodegeneration and decline of cognitive function leading to dementia (Andrews et al., 2021). The symptomatic burden of dementia typically occurs late in life, but it is preceded by a long preclinical phase, characterized by neuropathological impairment. Therefore, much of the ongoing dementia research is focused on elucidating early pathological changes (Jack et al., 2018). However, no study has analyzed the effect of genetic predisposition to AD and causal traits on study participation in middle-aged and older individuals.

The Prospective Imaging Study of Ageing: Genes, Brain and Behaviour (PISA) aims to characterize the phenotype and natural history of healthy adult Australians at high future risk of AD (Lupton et al. 2021). Potential PISA participants were selected from previous genetic cohort studies on the condition that they had genomewide genetic data already available. This offers a unique opportunity to investigate the genetic contribution to voluntary recruitment, enabling us to test for differences between those successfully recruited into PISA and those who were not.

In the present study, we aimed to investigate whether the likelihood of participating in the PISA study was influenced by factors associated with a higher risk of AD, including genetic risk of

Corresponding author: Lina Gomez; Email: lina.gomez@qimrberghofer.edu.au

Cite this article: Gomez LM, Mitchell BL, McAloney K, Adsett J, Garden N, Wood M, Diaz-Torres S, Garcia-Marin LM, Breakspear M, Martin NG, Lupton MK. (2023) The Effect of Genetic Predisposition to Alzheimer's Disease and Related Traits on Recruitment Bias in a Study of Cognitive Aging. *Twin Research and Human Genetics* 26: 209–214, <https://doi.org/10.1017/thg.2023.26>



AD and causally associated risk factors. Using Mendelian randomization, we and others have previously shown that out of several potentially modifiable AD risk factors, educational attainment (EA), intelligence (measured by intelligence quotient; IQ) and socioeconomic status (measured by household income; HI) have a causal association with AD (Thorpe et al., 2022). Therefore, we used polygenic risk scores (PRS) to assess the relationship between the genetic risk for AD and educational attainment, intelligence, and household income. By examining the association between these factors and the likelihood of participating in the PISA study, we aimed to assess whether there is a heritable component that is mediating agreement to participate in PISA, and whether this could create bias in the investigation of risk factors and prodromal disease markers for AD.

Materials and Methods

Sample Recruitment

PISA is a prospective cohort study of healthy Australians in mid to late adulthood. The sample recruitment pool consisted of 15,531 participants with available genomewide genotyping, derived from extensive in-house cohorts, including studies of Australian twins and their families, and genomewide association studies (GWAS) for physical or psychiatric conditions conducted over four decades at QIMR Berghofer Medical Research Institute (Heath et al., 2011; Medland et al., 2009). We attempted to re-contact all research participants between 40 to 80 years old (years of birth: 1946–1986). Those who were known to be deceased, were living overseas, had requested not be included in further studies or had no available valid contact details were excluded, resulting in a final sample size of 13,432. Over 9000 participant records ($N=9685$, 72%) were successfully updated. The majority (61.7%) of the participants in the study were women, with a mean age of 58.8 and standard deviation of 7.4 years. All recontacted participants were invited to complete an online survey that aimed to update aspects related to lifestyle and cognitive and behavioral function (full details are given in Lupton et al., 2021). The survey began with an information and consent page followed by a core module that captured the central dataset. Once the core module was complete, the survey included 10 additional modules and the option to consent to linking medical records (described in full in Lupton et al., 2021). No incentive or remuneration was offered to participants to take part in this stage of the study. Participants were offered the option of completing a paper version of the survey, which was mailed to them upon request. Subjects were informed that they might be invited to take part in further research. Overall, 4801 participants completed the core module of the online PISA survey, with a mean age of 60 years and standard deviation of 7 years.

For the purposes of this study, those who took part by completing the core survey module of the PISA questionnaire were considered as participating. All participants provided online informed consent prior to participating in the study, which included permission to link to previously collected data. The PISA study protocol (P2210) was approved by the Human Research Ethics Committees (HREC) of QIMR Berghofer Medical Research Institute.

Single Nucleotide Polymorphism (SNP)-Based Heritability

Genomewide genotyping of participants within our recruitment pool was performed using a range of genotyping arrays as previously detailed (Cuellar-Partida et al., 2015; Medland et al.,

2009). Datasets were combined with strict quality control procedures and imputed to the Haplotype Reference Consortium (HRC) Release 1 reference panel (Loh et al., 2016).

We employed genome-based restricted maximum likelihood (GREML) as implemented in GCTAv.1.91.3 beta (Yang et al., 2011) to estimate the proportion of variance in the PISA study (on the observed scale) explained by measured genetic differences (SNP-based heritability). This approach leverages a genetic-relatedness matrix (GRM) and restricted maximum likelihood to partition the variance of a phenotype into genetic and environmental components (Yang et al., 2010). This method is frequently used to assess whether genetic covariation explains a significant proportion of the covariation of a trait. For this study, a GRM based on a subset of unrelated individuals of European ancestry was leveraged to identify evidence for a genetic component to taking part in PISA.

Polygenic Risk Scores (PRS)

A PRS is a way to estimate an individual's genetic susceptibility for a particular trait or disease based on their genetic makeup. It is calculated using the sum of an individual's genomewide or location-specific genotypes, weighted by effect size estimates for SNPs associated with the trait or disease of interest (Clark et al., 2022).

We computed PRS, using the SBayesR approach, to investigate the extent to which the genetic risk for AD, EA, IQ, and HI predicted participation in PISA. We estimated PRS using GWAS summary statistics for AD (Bellenguez et al., 2022), EA (Lee et al., 2018), IQ (Savage et al., 2018) and HI (Hill et al., 2019). For the AD PRS, SNPs within 500 kb distance at either side of the *APOE* locus were excluded to ensure the exclusion of the *APOE*-associated signal. The *APOE* genotype was considered separately from the AD PRS (derived from SNPs rs429358 and rs7412) and was either obtained from genomewide SNP chip data (where either the *APOE* SNPs were directly genotyped on the array or imputed with a high degree of certainty) or directly genotyped using TaqMan SNP genotyping assays, as described in Lupton et al. (2018). Furthermore, as both the GWAS for EA and IQ included potentially overlapping samples with the PISA cohort, we computed PRS using summary data from GWAS meta-analyses that exclude QIMR Berghofer cohorts, including any PISA participants and their family members (referred to as 'leave one out' (LOO) QIMRB summary statistics). Prior to estimating SBayesR PRS, we excluded low imputation quality ($r^2 < 0.6$), allele frequency (MAF < 0.01), nonautosomal, and strand-ambiguous variants. Imputed genotype dosage data were used to calculate PRS by multiplying the variant effect size by the dosage of the effect allele. Finally, the total sum was calculated across all variants. The SBayesR PRS was generated using Plink 2.0 (Purcell et al., 2007; Chang et al., 2015).

Statistical Analyses

We examined the association of PRS for AD, EA, IQ and HI PRS and also *APOE* $\epsilon 4$ carrier status with recruitment bias using a logistic regression model in GCTA (1.91.3 beta), which accounted for sex, age, and the first five genetic ancestry principal components (PCs) as fixed effects. Relatedness among individuals was accounted for as a random effect with a genetic relatedness matrix. The AD PRS excludes the *APOE* region and is therefore independent of *APOE* $\epsilon 4$. A Nagelkerke's R^2 was used to estimate the variance explained by the PRS or *APOE* status.

Table 1. Demographics of participants to take part in PISA

	Whole sample (<i>N</i> = 13,432)	Participated in PISA (<i>n</i> = 4801)	Not participated in PISA (<i>n</i> = 8631)	Participated in PISA OR (95% CI)
Age in years				
Mean (SD)	59.1 (7.4)	60.08 (7.15)	58.8 (7.4)	1.03 (1.03, 1.04)
Female	8548 (63.6%)	3219 (67%)	5329 (61.7%)	1.31 (1.23, 1.39)
Male	4884 (36.4%)	1582 (33%)	3302 (38.3%)	
APOE ε4 +ve	3935 (29.3%)	1411 (32.2%)	2524 (31.35%)	1.02 (0.94, 1.09)
APOE ε4 -ve	9206 (70.7%)	3281 (67.8%)	5925 (68.6%)	

Note: 291 Missing data proportion of APOE.

Results

Demographic Factors

Participants who took part in PISA were, on average, older than those who did not ($OR = 1.007$; 95% CI [1.006, 1.008] per year of age) p value = $7.33E-30$ (Table 1). The participation rate stratified by decade showed that the largest number of participants were in the 50–69 age range, representing 78.6% of the whole sample (Table 2). Female participants were more likely to participate in PISA than males (Table 1) ($OR = 1.063$; 95% CI [1.04, 1.08]) p value = $1.18E-12$.

Genetic Factors

The GREML analysis suggested the presence of genetic contribution to the likelihood of participating in PISA. The SNP-based heritability on the observed scale was 0.14 ($SE = 0.013$, $p < .0001$, phenotypic variance of the observed scale ~ 0.22).

The results of the association of genetic variables (PRS and APOE), sex and age with participation in the PISA study are shown in Table 3 and Figure 1. Neither AD PRS nor presence of APOE ε4 were associated with participation in PISA. The EA, IQ and HI PRS were positively associated with participating in PISA.

Discussion

We investigated systematic differences between individuals who decided to participate in the PISA study (i.e., complete the core survey module) with those who, despite being contacted and participating in previous studies, did not participate. Investigating variables that associate with study participation is invariably challenging, since data are required to be collected on individuals who are not willing to participate in a study. The PISA study offers a unique opportunity in this research area, as the recruitment pool consists of participants from previous studies, where genetic data has already been collected. PISA consists of middle-aged and older individuals aged ≥ 40 years with an aim of discovering markers of early AD pathology to identify modifiable risk factors and establish the very earliest phenotypic and neuronal signs of disease onset. Therefore, it is crucial to identify factors that are associated with the likelihood of recruitment into the study and have the potential to introduce bias in association estimates in downstream analyses.

Analysis of the relationship between sex, age and study participation revealed that older female individuals were more likely to participate in PISA (Tables 1 and 2, Figure 1). It is well known that women are more likely than men to participate in clinical research. Research into the factors that influence women's decision to participate has found that decisions may be impacted

Table 2. Participation rate by decade

Age group	<i>N</i>	% participated in PISA in each age group
<40	1	0.01
40–49	1567	11.67
50–59	5907	43.97
60–69	4789	35.65
70–80	1151	8.57
>80	17	0.13

Table 3. The association of genetic variables, sex and age with participation in the PISA study, $p < .0083$ (Bonferroni corrected threshold)

Variable	Standardized beta	<i>SE</i>	<i>t</i> statistic	<i>p</i> value	<i>R</i> ²
EA PRS	0.047	0.0043	10.98	6.50E-28	.545
IQ PRS	0.036	0.0044	8.142	4.21E-16	.316
AD PRS	-0.004	0.0042	-1.07	.28	8.90E-05
HI PRS	0.023	0.0043	5.447	5.22E-08	.003
APOE ε4 +ve	0.003	0.0092	0.32	.75	7.62E-04
Sex (female)	0.061	0.009	7.11	1.18E-12	.004
Age	0.007	0.0006	11.74	7.33E-30	.010

Note: EA, educational attainment; AD, Alzheimer's disease; HI, household income.

not only by individual characteristics, but also by interpersonal relationships and social norms within the community (Baker et al., 2005; Liu & Dipietro Mager, 2016). We found in our recruitment pool of middle-aged and older individuals that participants aged 50 and older were more likely to participate than those in their 40s, which may reflect a lack of spare time in the younger age group who are more likely to be of working age and with younger families. It is also possible that older participants perceived more personal risk of cognitive decline and dementia. The recruitment rate reduces once participants reach 70. These finding echoes those in studies of age-related diseases, where the odds of participation are greater in older participants (aged around 55–75 years) and reduced in younger age groups, and once participants reach over 75 years (Beard et al., 1994; Benfante et al., 1989).

Using a PRS approach we tested whether the genetic liabilities to AD and causally associated traits are associated with the decision to participate in the PISA study. We tested for an association of EA, IQ and HI, which are highly related to each other with complex multifaceted relationships. We have previously shown all three to

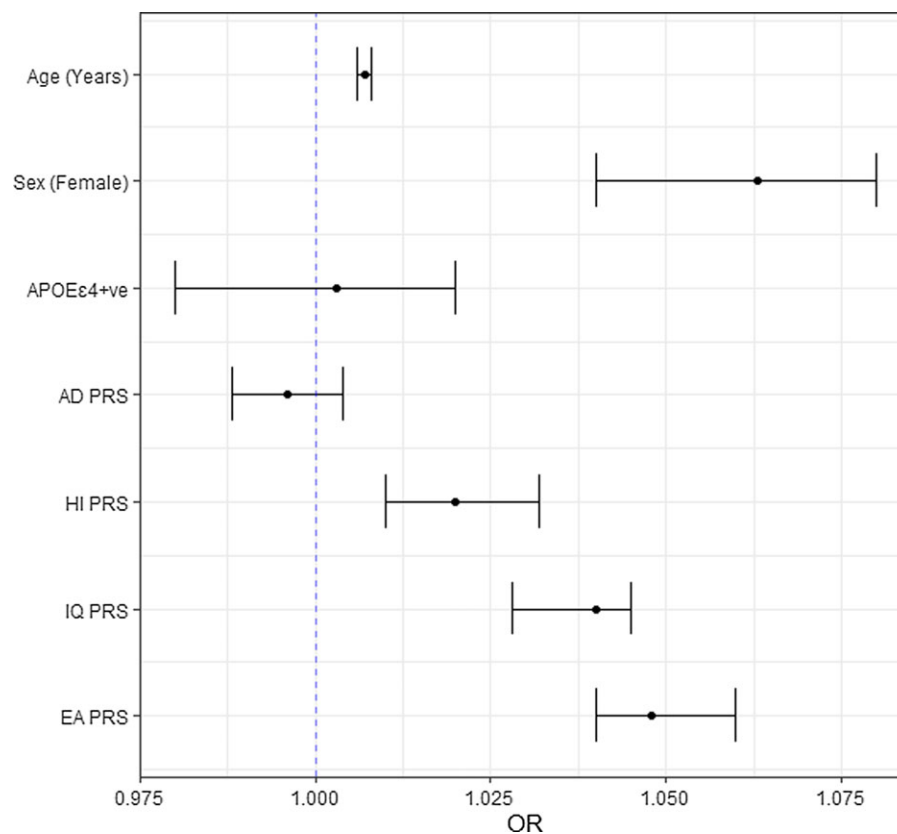


Figure 1. Forest plots depict odds ratios (OR) and 95% confidence intervals for the association between age, sex, PRS and presence of APOE ϵ 4 with participation in the PISA study.

be causally associated with risk of AD at a genetic level using Mendelian randomization analysis, and that this causal association is led by the cognitive component of EA (intelligence). (Thorp et al., 2022).

We did not find evidence suggesting that the genetic risk of AD for both AD PRS (with APOE region removed) or APOE genotype are associated with recruitment bias in our sample of middle-aged and older research participants. However, we identified an association of PRS for key causal risk factors for AD, EA, IQ and HI. Previous studies have highlighted the strong relationship between education level and voluntary participation (Enzenbach et al., 2019; Lissner et al., 2003; van Loon et al., 2003). The over-representation of more highly educated individuals in epidemiological studies is likely due to greater health awareness and interest in science (Galea & Tracy, 2007). In addition, the association of low socioeconomic status with reduced participation in cohort studies and clinical trials has been well documented, and is thought to be due to several factors, including psychosocial and practical barriers to participation. (Howe et al., 2013; Stuber et al., 2020).

A recent investigation of genomic predictors of participation in optional components of the large scale UK Biobank study (up to 451,306 participants) also identified an association of genetic predictors of educational duration, intelligence and lower levels of deprivation (using four measures) with participation, as well as a negative correlation with AD in this well powered sample. Causal associations were confirmed using Mendelian randomization (Tyrrell et al., 2021). We did not replicate these previous findings of an association of recruitment with genetic risk for AD, including an association with APOE ϵ 4 carrier status (Tyrrell et al. 2021). This may be due to decreased power in our dataset, or potentially

the influence of an opposing bias where those with a strong family history of AD due to inheritance of APOE 4 are more likely to be compelled to volunteer for a study researching ageing and dementia.

Limitations of the present study must be acknowledged. The participants in the recruitment pool had all previously been successfully recruited into genetic studies, and are therefore more likely to volunteer again. These baseline studies are not representative of the current Australian population, comprising predominantly those of white European ancestry, an over-representation of females, and individuals recruited based on a phenotypes of interest, including twin pairs and their family members, as well as physical and psychiatric conditions. In this study, a lack of response was interpreted as a passive refusal to participate, but a subset of the recruitment pool may have out-of-date contact details. This was mitigated as much as possible by updating contact details using electoral roll information, contacting participants through several modalities including email, letter and phone when available, and using available death records to exclude those who had died. As much as possible we aimed to prevent accessibility issues for any participants without internet access or computing capabilities by making paper copies of the survey available upon request.

Overall, this work shows that in the PISA study key causal risk factors for AD related to cognition are associated with recruitment. Therefore, participants are not accurately representative of the population. Although this does not necessarily invalidate study findings in the investigation of risk factors for age-related cognitive decline and dementia, it must be considered when interpreting findings, specifically for informing analysis strategies, using the

appropriate control variables and interpreting the generalizability of findings to the general population (Elwood, 2013; Rothman et al., 2013).

Acknowledgments. We would like to acknowledge the contribution of the research participants. The population-based sample for PISA comprises of research participants who have taken part in genetic epidemiology studies led by NM (with colleagues and collaborators) at QIMR Berghofer and elsewhere over the past 40 years. Demographic and contact information as well as genomewide SNP chip data were utilized from these previous studies for participant selection and recruitment. Principal sources of funding for these early studies were from grants to NGM from Australian NHMRC and to Andrew Heath and Pam Madden (Washington University, St Louis) from NIH (mainly NIAAA and NIDA) and we gratefully acknowledge these contributions.

Financial support. The Prospective Imaging Study of Ageing (PISA) is funded by the National Health and Medical Research Council (Grant ID: APP1095227). LMGM is supported by a UQ Research Training Scholarship from The University of Queensland (UQ)

Competing interests. None.

Ethical standard. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Andrews, S. J., Fulton-Howard, B., O'Reilly, P., Marcora, E., Goate, A. M., & Collaborators of the Alzheimer's Disease Genetics Consortium. (2021). Causal associations between modifiable risk factors and the Alzheimer's phenotype. *Annals of Neurology*, 89, 54–65. <https://doi.org/10.1002/ana.25918>
- Baker, L., Lavender, T., & Tincello, D. (2005). Factors that influence women's decisions about whether to participate in research: An exploratory study. *Birth*, 32, 60–66. <https://doi.org/10.1111/j.0730-7659.2005.00346.x>
- Beard, C. M., Lane, A. W., O'Fallon, W. M., Riggs, B. L., & Melton, L. J., 3rd. (1994). Comparison of respondents and nonrespondents in an osteoporosis study. *Annals of Epidemiology*, 4, 398–403. [https://doi.org/10.1016/1047-2797\(94\)90075-2](https://doi.org/10.1016/1047-2797(94)90075-2)
- Bellenguez, C., Küçükali, F., Jansen, I. E., Kleindam, L., Moreno-Grau, S., Amin, N., Naj, A. C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., Holmans, P. A., Boland, A., Damotte, V., van der Lee, S. J., Costa, M. R., Kuulasmaa, T., Yang, Q., de Rojas, I., Bis, J. C., ... Lambert, J.-C. (2022). New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nature Genetics*, 54, 412–436. <https://doi.org/10.1038/s41588-022-01024-z>
- Benfante, R., Reed, D., MacLean, C., & Kagan, A. (1989). Response bias in the Honolulu Heart Program. *American Journal of Epidemiology*, 130, 1088–1100. <https://doi.org/10.1093/oxfordjournals.aje.a115436>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Clark, K., Leung, Y. Y., Lee, W.-P., Voight, B., & Wang, L.-S. (2022). Polygenic Risk scores in Alzheimer's disease genetics: Methodology, applications, inclusion, and diversity. *Journal of Alzheimer's Disease*, 89, 1–12. <https://doi.org/10.3233/JAD-220025>
- Elwood, J. M. (2013). Commentary: On representativeness. *International Journal of Epidemiology*, 42, 1014–1015. <https://doi.org/10.1093/ije/dyt101>
- Enzenbach, C., Wicklein, B., Wirkner, K., & Loeffler, M. (2019). Evaluating selection bias in a population-based cohort study with low baseline participation: the LIFE-Adult-Study. *BMC Medical Research Methodology*, 19, 135. <https://doi.org/10.1186/s12874-019-0779-8>
- Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of Epidemiology*, 17, 643–653. <https://doi.org/10.1016/j.annepidem.2007.03.013>
- Ganguli, M., Lee, C.-W., Hughes, T., Snitz, B. E., Jakubcak, J., Duara, R., & Chang, C.-C. H. (2015). Who wants a free brain scan? Assessing and correcting for recruitment biases in a population-based sMRI pilot study. *Brain Imaging and Behavior*, 9, 204–212. <https://doi.org/10.1007/s11682-014-9297-9>
- Ghanbari, R., Lovasi, G. S., & Bader, M. D. M. (2023). Exploring potential for selection bias in using survey data to estimate the association between institutional trust and depression. *Annals of Epidemiology*, 77, 61–66. <https://doi.org/10.1016/j.annepidem.2022.11.004>
- Haring, R., Alte, D., Völzke, H., Sauer, S., Wallaschofski, H., John, U., & Schmidt, C. O. (2009). Extended recruitment efforts minimize attrition but not necessarily bias. *Journal of Clinical Epidemiology*, 62, 252–260. <https://doi.org/10.1016/j.jclinepi.2008.06.010>
- Heath, A. C., Whitfield, J. B., Martin, N. G., Pergadia, M. L., Goate, A. M., Lind, P. A., McEvoy, B. P., Schrage, A. J., Grant, J. D., Chou, Y.-L., Zhu, R., Henders, A. K., Medland, S. E., Gordon, S. D., Nelson, E. C., Agrawal, A., Nyholt, D. R., Bucholz, K. K., Madden, P. A. F., & Montgomery, G. W. (2011). A quantitative-trait genome-wide association study of alcoholism risk in the community: Findings and implications. *Biological Psychiatry*, 70, 513–518. <https://doi.org/10.1016/j.biopsych.2011.02.028>
- Hill, W. D., Davies, N. M., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., Di Angelantonio, E., Roberts, D. J., Xueyi, S., Davies, G., Liewald, D. C. M., Porteous, D. J., Hayward, C., Butterworth, A. S., McIntosh, A. M., Gale, C. R., & Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, 10, 5741. <https://doi.org/10.1038/s41467-019-13585-5>
- Howe, L. D., Tilling, K., Galobardes, B., & Lawlor, D. A. (2013). Loss to follow-up in cohort studies: Bias in estimates of socioeconomic inequalities. *Epidemiology*, 24, 1–9. <https://doi.org/10.1097/EDE.0b013e31827623b1>
- Jack, C. R., Jr, Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeblerlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., ... Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14, 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>
- Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghziyan, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., ... Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, 50, 1112–1121. <https://doi.org/10.1038/s41588-018-0147-3>
- Li, M., Kim, H., Sereika, S. M., Nissley, T. J., & Lingler, J. H. (2022). Willingness to participate in clinical research among individuals with cognitive impairment. *Research in Gerontological Nursing*, 15, 76–84. <https://doi.org/10.3928/19404921-20220131-01>
- Lissner, L., Skoog, I., Andersson, K., Beckman, N., Sundh, V., Waern, M., Edin Zylberstein, D., Bengtsson, C., & Björkelund, C. (2003). Participation bias in longitudinal studies: Experience from the population study of women in Gothenburg, Sweden. *Scandinavian Journal of Primary Health Care*, 21, 242–247. <https://doi.org/10.1080/02813430310003309-1693>
- Liu, K. A., & Dipietro Mager, N. A. (2016). Women's involvement in clinical trials: Historical perspective and future implications. *Pharmacy Practice*, 14. <https://doi.org/10.18549/PharmPract.2016.01.708>
- Loh, P.-R., Danecek, P., Palamara, P. F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G. R., Durbin, R., Price A. L. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nature Genetics*, 48, 1443–1448. <https://doi.org/10.1038/ng.3679>
- Lupton, M. K., Medland, S. E., Gordon, S. D., Goncalves, T., MacGregor, S., Mackey, D. A., Young, T. L., Duffy, D. L., Visscher, P. M., Wray, N. R., Nyholt, D. R., Bain, L., Ferreira, M. A., Henders, A. K., Wallace, L., Montgomery, G. W., Wright, M. J., & Martin, N. G. (2018). Accuracy of Inferred APOE genotypes for a range of genotyping arrays and imputation reference panels. *Journal of Alzheimer's Disease*, 64, 49–54. <https://doi.org/10.3233/JAD-171104>

- Lupton, M. K., Robinson, G. A., Adam, R. J., Rose, S., Byrne, G. J., Salvado, O., Pachana, N. A., Almeida, O. P., McAloney, K., Gordon, S. D., Raniga, P., Fazlollahi, A., Xia, Y., Ceslis, A., Sonkusare, S., Zhang, Q., Kholghi, M., Karunanithi, M., Mosley, P. E., . . . Breakspear, M. (2021). A prospective cohort study of prodromal Alzheimer's disease: Prospective Imaging Study of Ageing: Genes, Brain and Behaviour (PISA). *NeuroImage Clinical*, 29, 102527. <https://doi.org/10.1016/j.nicl.2020.102527>
- Medland, S. E., Nyholt, D. R., Painter, J. N., McEvoy, B. P., McRae, A. F., Zhu, G., Gordon, S. D., Ferreira, M. A. R., Wright, M. J., Henders, A. K., Campbell, M. J., Duffy, D. L., Hansell, N. K., Macgregor, S., Slutske, W. S., Heath, A. C., Montgomery, G. W., & Martin, N. G. (2009). Common variants in the trichohyalin gene are associated with straight hair in Europeans. *American Journal of Human Genetics*, 85, 750–755. <https://doi.org/10.1016/j.ajhg.2009.10.009>
- Milani, S. A., Cottler, L. B., & Striley, C. W. (2021). Perceptions of research participation among a sample of Florida residents aged 50 and over reporting dementia. *Ageing International*. <https://doi.org/10.1007/s12126-021-09441-x>
- Oswald, L. M., Wand, G. S., Zhu, S., & Selby, V. (2013). Volunteerism and self-selection bias in human positron emission tomography neuroimaging research. *Brain Imaging and Behavior*, 7, 163–176. <https://doi.org/10.1007/s11682-012-9210-3>
- Patten, S. B. (2000). Selection bias in studies of major depression using clinical subjects. *Journal of Clinical Epidemiology*, 53, 351–357. [https://doi.org/10.1016/s0895-4356\(99\)00215-2](https://doi.org/10.1016/s0895-4356(99)00215-2)
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575. <https://doi.org/10.1086/519795>
- Rothman, K. J., Gallacher, J. E. J., & Hatch, E. E. (2013). Why representativeness should be avoided. *International Journal of Epidemiology*, 42, 1012–1014. <https://doi.org/10.1093/ije/dys223>
- Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., Nagel, M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., . . . Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, 50, 912–919. <https://doi.org/10.1038/s41588-018-0152-6>
- Stuber, J. M., Middel, C. N. H., Mackenbach, J. D., Beulens, J. W. J., & Lakerveld, J. (2020). Successfully recruiting adults with a low socioeconomic position into community-based lifestyle programs: A qualitative study on expert opinions. *International Journal of Environmental Research and Public Health*, 17, 2764. <https://doi.org/10.3390/ijerph17082764>
- Thorp, J. G., Mitchell, B. L., Gerring, Z. F., Ong, J.-S., Gharahkhani, P., Derks, E. M., & Lupton, M. K. (2022). Genetic evidence that the causal association of educational attainment with reduced risk of Alzheimer's disease is driven by intelligence. *Neurobiology of Aging*, 119, 127–135. <https://doi.org/10.1016/j.neurobiolaging.2022.07.011>
- Tyrrell, J., Zheng, J., Beaumont, R., Hinton, K., Richardson, T. G., Wood, A. R., Davey Smith, G., Frayling, T. M., & Tilling, K. (2021). Genetic predictors of participation in optional components of UK Biobank. *Nature Communications*, 12, 1–13. <https://doi.org/10.1038/s41467-021-21073-y>
- Van Loon, A. J. M., Tijhuis, M., Picavet, H. S. J., Surtees, P. G., & Ormel, J. (2003). Survey non-response in the Netherlands: Effects on prevalence estimates and associations. *Annals of Epidemiology*, 13, 105–110. [https://doi.org/10.1016/s1047-2797\(02\)00257-0](https://doi.org/10.1016/s1047-2797(02)00257-0)
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42, 565–569. <https://doi.org/10.1038/ng.608>
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 88A, 76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28, 3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>