REVIEW ARTICLE

Genome-wide association studies for Alzheimer’s disease: bigger is not always better

Valentina Escott-Price1,2 and John Hardy3,4

As the size of genome-wide association studies increase, the number of associated trait loci identified inevitably increase. One welcomes this if it allows the better delineation of the pathways to disease and increases the accuracy of genetic prediction of disease risk through polygenic risk score analysis. However, there are several problems in the continuing increase in the genome-wide analysis of ‘Alzheimer’s disease’. In this review, we have systematically assessed the history of Alzheimer’s disease genome-wide association studies, including their sample sizes, age and selection/assessment criteria of cases and controls and heritability explained by these disease genome-wide association studies. We observe that nearly all earlier disease genome-wide association studies are now part of all current disease genome-wide association studies. In addition, the latest disease genome-wide association studies include (i) only a small fraction (~10%) of clinically screened controls, substituting for them population-based samples which are systematically younger than cases, and (ii) around 50% of Alzheimer’s disease cases are in fact ‘proxy dementia cases’. As a consequence, the more genes the field finds, the less the heritability they explain. We highlight potential caveats this situation creates and discuss some of the consequences occurring when translating the newest Alzheimer’s disease genome-wide association study results into basic research and/or clinical practice.

1 Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK
2 Dementia Research Institute at Cardiff, Cardiff University, Cardiff, UK
3 UCL Institute of Neurology, Queen Square, London, UK
4 UCL Dementia Research Institute, UCL, London, UK

Correspondence to: Valentina Escott-Price
Cardiff University, Hadyn Ellis Building
Maindy Road, Cardiff CF24 4HQ, UK
E-mail: escottpricev@cardiff.ac.uk

Correspondence may also be addressed to: John Hardy
UCL Institute of Neurology, Queen Square
London WC1N 3BG, UK
E-mail: j.hardy@ucl.ac.uk

Keywords: Alzheimer’s disease; genome-wide association study; heritability

Abbreviations: FTD = frontotemporal dementia; GWAS = genome-wide association study; LDSC = linkage disequilibrium score; MDD = major depressive disorder; PRS = polygenic risk score.

Received December 17, 2021. Revised March 15, 2022. Accepted May 13, 2022. Advance access publication May 17, 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of the Guarantors of Brain.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Introduction

As the size of genome-wide association studies (GWASs) increase, the number of associated trait loci identified inevitably increase. One welcomes this if it allows the better delineation of the pathways to disease and increases the accuracy of genetic prediction of disease risk through polygenic risk score (PRS) analysis. However, there are several problems in the continuing increase in the genome-wide analysis of ‘Alzheimer’s disease’. The first is that the diagnostic accuracy for Alzheimer’s disease is poor: of the order of 80% in clinic-based series based both on neuropathological criteria and on genetic analysis and this is certain to be worse in the case of the proxy cases used in the larger and more recent GWAS. The second is that, while for many rare diseases, age matching of controls makes little difference to the results obtained, because Alzheimer’s disease is a common cause of mortality, the risk gene APOE also has the greatest effect on longevity. This makes age-matching essential for accurate risk assessment. In addition, a simple inclusion of age as a covariate in the GWAS creates a robust but biologically spurious association between Alzheimer’s disease and age-associated variants, similar to the association identified between sex- and height-associated variants. Thus, in case of Alzheimer’s disease, the appropriate use of age-matched controls is important. A final major problem in the published GWAS is that for most of them, only summary statistics are made available. These problems are systemic in nearly all the ‘Alzheimer’ GWASs, including ones in which we have been co-authors, except those using neuropathologically defined disease samples and as data from different studies are meta-analysed together, these systematic errors get baked into the updated analyses. An indicator of diluting the true Alzheimer’s disease associations is the reported heritability estimates. If in a small clinically assessed GWAS (N = 11 789 with 3 genome-wide significant loci identified), the heritability was estimated as $h^2 = 17%$ (SE = 3%), the latest GWAS with a sample size of more than 1.1M people with 38 independent genome-wide significant loci, accounts only for 3% (SE = 0.6%) of heritability. These errors then get incorporated into PRS analyses and also, perhaps, incorrectly contribute to the suggestion that neurodegenerative diseases share disease mechanisms. In this regard, for example, the designation of TMEM106B and GRN loci as Alzheimer’s disease loci (both are known frontotemporal dementia loci) is of particular concern, even though they appear in both clinic-based and proxy GWAS data sets. A related problem is likely to be the reported evidence of APOE association with clinical frontotemporal dementia (FTD). What is needed is larger GWAS of Alzheimer’s cases based on either neuropathological or on good biomarker data as, at present, such studies are underpowered. Neuropathological GWAS should give definitive risk loci for disease, whereas GWAS based on biomarker data perhaps give information on disease progression. The danger is that as larger and larger studies of cases with unsatisfactory diagnoses are analysed, the statistical weight behind unwarranted conclusions will become stronger.

Materials and methods

We have reviewed the GWAS for Alzheimer’s disease derived from analysis of populations of historical European ancestry.
and assessed their sample sizes, diagnosis and age distributions of cases and controls where possible, as well as the number of genome-wide associated loci they report. The numbers of clinically assessed cases and controls were calculated from the numbers of cases and controls reported in the publication, excluding cases with family-history-based diagnosis (proxy) and controls from the population cohorts in all previous studies contributed to the publication via meta-analysis.

We have extracted the single-nucleotide polymorphisms (SNP)-based heritability estimates for the GWAS from the publications where available and calculated the heritability ourselves if the authors did not provide them in the paper. For the latter, we have downloaded the corresponding summary statistics and used the Linkage Disequilibrium Score (LDSC) regression approach. We estimated heritability ourselves for fix studies using the default settings of the LDSC regression software and pre-calculated LDSCs from the 1000 Genomes European reference population, supplied with the LDSC software. Although Jansen et al. provided heritability estimate for Phase 1 in their Supplemental Note, we have also downloaded the study’s summary statistics, which included the UK Biobank (UKBB; combining Phases 1 and 2). Wightman et al. provided their own heritability estimate, with the same approach, reference population and software options. For the pathology confirmed sample of 1011 cases and 583 controls, we used the summary statistics as reported in Escott-Price et al. Due to the relatively small sample size, the LDSC heritability estimates were negative for these summary statistics when default LDSC parameters were used. Since in the pathology confirmed sample, there were no confounders (such as age mismatching or misdiagnosis), we estimated the heritability for this sample by constraining the intercept using theno-intercept flag. All heritability estimates were recalculated on a liability scale assuming a population prevalence of 5%.

Data availability
Data sharing is not applicable to this article as no new data were created or analysed in this study.

Results
Apart from four early GWASs (2009–2011), none of the current GWASs are independent (see Fig. 1 and Table 1). The latest GWASs (2019 onwards) include a large proportion of ‘cases’ are based upon the reported impression of offspring that their parent had dementia (usually referred to as ‘proxy Alzheimer’s disease cases’). The accuracy of these impressions is suspect, but, even assuming that 80% of parents have dementia, only 60% of them are likely to have had Alzheimer’s disease. This will introduce significant noise into the data set resulting in about 50% of parental cases having a different form of dementia or no dementia at all. This and any other diagnostic imprecision may specifically limit the detection of variants of small effect, which are the basis of the polygenic architecture of Alzheimer’s disease.

The number of clinically assessed controls drops down to ~10% as the majority of them are population based, and consequently not age matched. If in the pathology assessed GWAS and (mostly) clinically assessed GWAS, the average age difference was about 1 year, in the latest GWAS, it is about 10 years or simply impossible to trace (Table 1).

Counterintuitively, the exponential increase in sample size provides only marginal increases in the identification of novel GWAS significant loci: 2 in the samples of ~10 000 people, and 7 in the sample of ~1 126 563 people. Remarkably, the heritability estimates drop from ~40% to 2–3% (see Table 1, Fig. 2) as the sample size increases, despite the fact that all earlier GWAS are included to the latest ones (see Fig. 1).

Discussion
Why the heritability estimates are not accurate?
In the context of Alzheimer’s disease, heritability itself is a complex concept since it is possible that everyone would develop Alzheimer’s disease if they lived long enough (but see Morris); and genetic risk appears to determine when this occurs, not if it will occur: thus heritability estimates are exquisitely age dependent. Twin studies report heritability of Alzheimer’s disease 59–78% usually referred as broad-sense heritability. The SNP-based (narrow-sense) heritability estimates are varied across different data sets between 31 and 31%.

Different approaches are used for heritability estimates [genome-wide complex trait analysis (GCTA) and LDSC] with the latter gaining more popularity as it requires only summary statistics. However, the two approaches disagree in their estimates even for the same Alzheimer’s disease data sets, while for neurodevelopmental
Table 1 History of AD GWAS and their SNP-based heritability assuming 5% disease prevalence estimated with LDSC regression\(^6\)

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Sample size (Stage 1)</th>
<th>Mean age at assessment*</th>
<th>Clinical/pathology assessment (%)</th>
<th>SNP-based heritability on liability scale (5% prevalence)</th>
<th>Number of GWAS significant loci(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Cases</td>
<td>Controls</td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td>2010</td>
<td>Corneveaux et al.(^8)</td>
<td>1594</td>
<td>1011</td>
<td>583</td>
<td>81.9</td>
<td>80.8</td>
</tr>
<tr>
<td>2009</td>
<td>Harold et al.(^10)</td>
<td>11 025</td>
<td>31 77</td>
<td>7848</td>
<td>78.6</td>
<td>51</td>
</tr>
<tr>
<td>2009</td>
<td>Lambert et al.(^*)</td>
<td>8260</td>
<td>22 43</td>
<td>6017</td>
<td>68.5</td>
<td>74</td>
</tr>
<tr>
<td>2010</td>
<td>Seshadri et al.(^24)</td>
<td>14 283</td>
<td>13 15</td>
<td>12 968</td>
<td>82.7</td>
<td>72.8</td>
</tr>
<tr>
<td>2011</td>
<td>Naj et al.(^25)</td>
<td>21 165</td>
<td>10 273</td>
<td>10 892</td>
<td>74.7</td>
<td>76.3</td>
</tr>
<tr>
<td>2013</td>
<td>Lambert et al.(^26)</td>
<td>54 162</td>
<td>17 008</td>
<td>37 154</td>
<td>76.6</td>
<td>70.5</td>
</tr>
<tr>
<td>2019</td>
<td>Kunkle et al.(^17)</td>
<td>63 926</td>
<td>21 982</td>
<td>41 944</td>
<td>72.9</td>
<td>72.4</td>
</tr>
<tr>
<td>2018</td>
<td>Marioni et al.(^18)</td>
<td>368 440</td>
<td>70 306</td>
<td>298 134</td>
<td>Not known</td>
<td>67.3</td>
</tr>
<tr>
<td>2019</td>
<td>Jansen et al.(^19)</td>
<td>455 258</td>
<td>71 880</td>
<td>383 378</td>
<td>Not known</td>
<td>67.3</td>
</tr>
<tr>
<td>2021</td>
<td>Rojas et al.(^17)</td>
<td>409 435</td>
<td>81 611</td>
<td>308 979</td>
<td>Not known</td>
<td>67.3</td>
</tr>
<tr>
<td>2022</td>
<td>Bellenguez et al.(^21)</td>
<td>487 511</td>
<td>85 934</td>
<td>401 577</td>
<td>67.2</td>
<td>57.9</td>
</tr>
<tr>
<td>2021</td>
<td>Wightman et al.(^1)</td>
<td>1 126 563</td>
<td>90 338</td>
<td>1 036 225</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

\(^a\) Mean age at assessment (if not reported) was estimated as weighted (by the sample sizes) average of the ages at assessments reported in the contributing studies.

\(^b\) Pathology confirmed.

\(^c\) Heritability is estimated using summary statistics of imputed GWAS\(^8\).

\(^d\) Transformation to liability scale with 5% prevalence is reported by Zhang et al.\(^11\).

\(^e\) Estimated with GCTA software.\(^17\)

\(^f\) Without/with UK Biobank data.

\(^g\) Reported in Moreno-Grau et al.\(^28\).

\(^h\) With UK Biobank data.

\(^i\) Without UK Biobank data.

ожно, что наследственность альдоме не зависит от возраста на один год.

В таблице приведены данные по истории GWA и их SNP-генных наследственности предполагая 5% заболеваемость рассчитанной с помощью LDSC регрессии. В таблице указаны следующие данные:

- Год
- Автор
- Свойство (этап 1): Общее, Пациенты, Участники
- Возраст при оценке
- Клиническое/патологическое оценение (%)
- SNP-генетическая наследственность на либеральном шкале (5% заболеваемости)
- Количество значимых GWA генов

Например, на протяжении нескольких лет, в одном и том же наборе данных, наследственность составляет 31% с GCTA и 17% с LDSC. Как LDSC использует только упрощённые статистики, оно не может определить степень связи между участниками, специфические для нейродегенеративных расстройств. В частности, могло быть несколько генетических архитектоник APOE-ε4, когда сравниваются с носителями. В противном случае, он познал, что альдома-ε4 частота генов уменьшается с возрастом,\(^5,37\) в то время как альдома-ε4 частота генов уменьшается с возрастом. В нейродегенеративных расстройствах (где методологии согласуются), диагноз вероятно быть более значимым для возраста, чем для возрастного сбора. В нейродегенеративных расстройствах (где методологии согласуются), диагноз вероятно быть более значимым для возраста, чем для возрастной группы. В нейродегенеративных расстройствах (где методологии согласуются), диагноз вероятно быть более значимым для возраста, чем для возрастной группы.

Однако, например, Parkinson’s disease and major depressive disorder (MDD) have incorporated data sets from both UKBB and 23andMe and have not observed a corresponding decrease in heritability or the discovery of few GWS loci than expected. The reason for Parkinson’s disease is likely due to the clinical diagnosis being more precise, than for Alzheimer’s disease. In addition, Parkinson’s disease has lower prevalence in the population, so the addition of unscreened controls does not add much noise. While in MDD the prevalence it is similar to Alzheimer’s disease, it is an earlier onset disorder. Finally, for both disorders, there is no known genetic factor that modifies the age at onset and the rate of mortality (the latter changes the allele frequencies in an age-dependent way).

### Longevity

Потенциальный биас в оценках ефекта GWA и значимости локуса (с учётом наследственности) может быть введён, как SNPs связаны с обоими заболеваниями Alzheimer’s disease and age. The APOE ε4 is the prime suspect as it is associated with a shorter lifespan\(^41\) and with other ‘killers’ in the population such as heart disease and stroke.\(^32\) The frequency of APOE-ε4 reduces with age from ~0.18 in the general population to 0.1 in the age group 85+.\(^37\) Despite this reduced APOE-ε4 frequency in the very old (85+), Alzheimer’s disease prevalence is higher in this latter age group.

### Lack of study independence

Мы полагаем, что русский-двойник GWA, где крупные исследования включают все меньшие доли, несущие имперfection и маскировки, не дает ясности в понимании Alzheimer’s disease genetic architecture. Это GWA установили с использованием только суммарных статистик доступных для исследователей (i) не разрешает исследование дальнейшей гипотезы в подгруппах, e.g. Alzheimer’s disease predictability by the hypothesis-driven-specific (gene-network) PRS, и (ii) маскирует понимание Alzheimer’s disease heritability estimates.

### Consequences

Недавно все ‘Alzheimer’ GWA страдают от всех этих недостатков, мы делаем в них, в частности, недостаток возраста соответствия, низкую диагностическую точность и недостаток прозрачности данных.
This is leading to potentially serious issues (for example drug trials targeted at FTD genes in Alzheimer’s disease cases\(^45\)). This problem relates not only to the primary ‘new’ studies, but also the ones in which they are meta-analysed. If earlier GWAS studies have shown that genetics of Alzheimer’s disease and Parkinson’s disease is distinct,\(^46\) now papers appear discussing genetic overlap between ‘Alzheimer’s disease’ and Parkinson’s disease. However, ‘Alzheimer’s disease’ cohorts certainly include dementia with Lewy body (DLB) cases and overlap between Parkinson’s disease and DLB is well established.\(^47\) Thus, in many ways, this genetic sloppiness is having consequences both in terms of the loci associated with disease and therefore passed on to cell biologists and for the utility of PRS analyses for clinical prediction of disease. For example, one of the consequences of the reported low SNP-based heritability is the conclusion that late onset Alzheimer’s disease is oligogenic (~100 genes),\(^11\) where the authors assumed 9% heritability in their simulation study, whereas earlier publications suggest that the disease is polygenic (thousands of genes).\(^5\), \(^48\)

What is needed?

The GWASs have clearly made an enormous contribution to our understanding of Alzheimer’s disease, chiefly through the identification of microglial and brain lipid metabolism\(^49\) as important risk components, and have focussed attention on the way the brain responds to amyloid deposition.\(^50\) Larger and larger GWASs now display the law of diminishing returns. A clear distinction needs to be introduced between Alzheimer’s disease GWAS and GWAS for dementia to avoid sending the misleading messages to molecular biologists: the latest big GWAS needs to be labelled as dementia GWAS, not Alzheimer’s disease GWAS. In these dementia GWAS, the Russian doll needs to be unpacked so that the summary statistics for each of them can be made available without an application process.

The consensus on the heritability of Alzheimer’s disease captured by the SNPs needs to be reached. If there is extensive missing heritability, as is widely believed, then epistatic interactions are likely candidates for this missing heritability where risks at unlinked loci act multiplicatively rather than additively. The possibility to detect epistatic loci is widely
debated (28). However, this possibility is impaired if the case/control definition is inaccurate, and is forever lost if all that is available are summary statistics.

We need to understand more subtle phenotypic variability within the disease and the genetic factors which influence the rate of decline in disease. In this context, more genotyping of deeply phenotyped sample series and of cases with pathological confirmation are needed. In both cases, consents and protocols are required which permit academic access to individual level data to allow post hoc informed cleaning of these data. This would be preferable to ever larger GWAS of poorly characterized individuals. In parallel, we certainly need to understand the architecture of disease in non-European populations, and, within the genes we have already found, the identification of variability which would help disease modelling.

One way forward would be to develop a framework where the ever larger dementia GWAS hits were systematically evaluated in GWAS derived solely from Alzheimer’s disease pathologically confirmed samples, independent from the dementia GWAS. The current research trajectory will lead to ever more confusion, especially amongst those who are not aware of the problems we outline.

Funding
The authors thank the Dementia Research Institute [UKDRI supported by the Medical Research Council (UKDRI-3003), Alzheimer’s Research UK and Alzheimer’s Society].

Competing interests
The authors report no competing interests.

References


