Investigating the genetic relationship between depression symptoms and Alzheimer’s Disease in clinically diagnosed and proxy cases


a Department of Basic and Clinical Neuroscience, King’s College London, London, United Kingdom
b Social, Genetic and Developmental Psychiatry, King’s College London, London, United Kingdom
c Perron Institute for Neurological and Translational Science, Perth, Australia
d Department of Biostatistics and Health Informatics, King’s College London, London, United Kingdom
e NIHR Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London, King’s College London, London, United Kingdom
f MRC Unit for Lifelong Health and Aging, University College London, London, United Kingdom
g Translational Neurogenomics Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia
h College of Medicine and Health, University of Exeter, Exeter, United Kingdom
i Old Age Psychiatry, King’s College London, London, United Kingdom
j Faculty of Medicine, Department of Brain Sciences, Imperial College London, London, United Kingdom
k Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, Australia
l Department of Medical and Molecular Genetics, King’s College London, London, United Kingdom

* Corresponding author: Dr. Petroula Proitsi (petroula.proitsi@kcl.ac.uk)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Depression is a risk factor for the later development of Alzheimer’s disease (AD), but evidence for their genetic relationship is mixed. Assessing depression symptom specific genetic associations may better clarify this relationship.

Using data from the UK Biobank, the GLAD Study and PROTECT, we performed the largest genome-wide meta-analyses (GWAS) of the nine depression symptom items, plus their sum score, on the Patient Health Questionnaire (PHQ-9) (GWAS equivalent N: 224,535–308,421). We then assessed global and local genetic correlations and statistical colocalisation between depression/depression symptoms and AD across six AD GWAS with varying proportions of clinical and proxy (family history) case ascertainment. We assessed bi-directional causal associations using Mendelian randomisation (MR) and tested the predictive utility of depression/depression symptom polygenic risk scores (PRS) for AD case/control status in three clinical AD cohorts.

Our GWAS meta-analyses identified 37 genomic risk loci across the ten depression symptom phenotypes. Of the 72 global genetic correlation tests conducted between depression/depression symptoms and AD, 20 were significant at $p_{FDR} \leq 0.05$. Only one of these was identified with AD GWAS containing clinical-only cases. Local genetic correlations were detected at 14 loci. Statistical colocalisation was not identified at these loci but was identified in the region of transmembrane protein 106B (TMEM106B) between multiple depression phenotypes and both clinical-only and clinical+proxy AD. MR and PRS analyses did not yield significant results after multiple testing correction.

Our findings do not demonstrate a causal role of depression/depression symptoms on AD and suggest that previous evidence of genetic overlap between depression and AD may be driven by the inclusion of family-history-based proxy cases/controls. However, the identification of colocalisation at TMEM106B warrants further investigation.
1. INTRODUCTION

Epidemiological studies suggest that a diagnosis of depression is a risk factor for the later development of dementia\(^1\)–\(^6\), of which Alzheimer’s disease (AD) is the most common form, accounting for approximately 80% of the over 40 million global cases\(^7\). Establishing the underlying mechanisms by which depression confers increased risk for AD offers a pathway by which new interventions might be implemented and the global dementia burden reduced\(^8\).

As twin studies have demonstrated, both depression and AD are substantially heritable – approximately 40% and 80% respectively\(^3\)–\(^10\). Further, large-scale genome-wide association studies (GWAS) have demonstrated high polygenicity, identifying over 70 genomic risk loci for AD and nearly 200 for depression\(^11\)–\(^17\). It is therefore possible that their phenotypic association is in part due to shared genetic architecture. However, results from previous investigations into the genetic overlap between the two disorders have been mixed. For example, some findings indicate non-significant genetic overlap\(^18\)–\(^19\), others a significant – if modest – genetic correlation of \(\sim 16\%\) and a risk increasing causal effect of depression on AD\(^20\)–\(^22\).

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), diagnosis of a major depressive episode requires the presence of at least five of a possible nine symptoms for \(\geq 2\) weeks, including one of the two cardinal symptoms – depressed mood or anhedonia\(^23\). Potentially hundreds of symptom combinations are possible to meet this diagnosis criteria\(^24\)–\(^26\). As such, this heterogeneity poses challenges to researchers seeking to better understand differences in the genetic contribution to depression and its subtypes\(^27\)–\(^28\). However, the decomposition of depression into individual symptoms has provided insight into unique patterns of genome-wide significant loci and cross-trait genetic associations, as demonstrated in a recent GWAS of depression symptoms on the Patient Health Questionnaire (PHQ-9) by Thorp et al.\(^29\).

A number of studies suggest that anhedonia may be a better predictor of dementia than depressed mood\(^30\)–\(^32\). Further, several depression symptoms, such as appetite changes, psychomotor dysfunction, and sleep disruption are commonly observed in non-depressed dementia patients\(^33\)–\(^35\). Taking this into account alongside the mixed nature of previous findings examining the genetic overlap between depression and AD, it is possible that leveraging depression symptom level genetic information may offer greater insight into the disorders shared genetic architecture.

However, any association between depression and AD must also consider the potential influence of differences in case/control ascertainment in AD GWAS. A review by Escott-Price et al.\(^36\) notes that recent large-scale AD GWAS contain a relatively small proportion of clinically ascertained cases/controls, with a large percentage of
cases ascertained by proxy – that is, cases and controls are defined as individuals with and without a self-reported parental history of AD/dementia, respectively. The combination of clinical and proxy samples in AD GWAS meta-analyses has proved an effective way of boosting sample size and variant discovery\textsuperscript{15–17,37}. However, evidence suggests that this has come at the expense of specificity in regard to genomic risk loci and an apparent stagnation in the percentage of variance explained by common variants\textsuperscript{36}. Most importantly for cross-trait analysis, recent studies indicate that the direction of Mendelian Randomisation (MR) causal estimates for AD risk factors on AD can be in the opposite direction depending on whether the AD outcome GWAS contains clinical and proxy cases/controls or is more strictly clinically ascertained\textsuperscript{38–40}.

To address these points, we perform the first and largest genome-wide meta-analyses of PHQ-9 depression symptom items using data from the Genetic Links to Anxiety and Depression (GLAD) Study\textsuperscript{41}, the PROTECT Study\textsuperscript{42,43} and two questionnaires from UK Biobank (UKB)\textsuperscript{44}. We obtained summary statistics from previous large-scale GWAS for clinical\textsuperscript{11} and broad\textsuperscript{12} depression, and six AD GWAS. Specifically, three with clinical+proxy case/control ascertainment\textsuperscript{15–17}, one with proxy-only\textsuperscript{37}, and two with clinical-only\textsuperscript{14,45}. We used these GWAS to assess the presence, strength and differences in genetic overlap between depression, depression symptoms and AD, with the additional aim of better understanding the influence of different AD case ascertainment strategies on associations.
2. METHODS

For an analysis flowchart, see Figure 1.
**Figure 1:** Analysis flowchart for the present study describing the analyses undertaken following genetic and phenotypic quality control (see: Methods) for each of the depression symptom items of the Patient Health Questionnaire (PHQ-9) in each of the four samples – UK Biobank (Mental Health Questionnaire and Experience of Pain Questionnaire), the Genetic Links in Anxiety and Depression (GLAD) Study and the PROTECT Study. GWAS = Genome-wide association analysis; AD = Alzheimer’s Disease; MTAG = Multi-trait Analysis of GWAS; FUMA = Functional Mapping and Annotation; LAVA = Local Analysis of [co]Variant Association; ADNI = Alzheimer’s Disease Neuroimaging Initiative; PC = genetic principal components.

### 2.1 GWAS in the UK Biobank, GLAD and PROTECT

#### 2.1.1 Patient Health Questionnaire-9 (PHQ-9) phenotypes

The PHQ-9 is a well-validated clinical screening questionnaire used to assess depression symptom severity on nine individual symptoms in the DSM-IV\(^{46,47}\). The severity of each symptom is measured by the self-reported persistence of that symptom over the preceding two weeks on a scale of 0 to 3. Scores of 3 indicate an individual experienced that symptom nearly every day, 2 indicates an individual experienced that symptom more than half the days, 1 indicates an individual experienced that symptom for several days and 0 indicates no experience of that symptom at all. The sum of an individual’s scores over all nine items (sum-score) ranges from 0-27. For an overview of the PHQ-9 items and response distribution for each sample, see Table 1. Sum-score distributions can be seen in Supplementary Table 1.
Table 1: Patient Health Questionnaire (PHQ-9) derived depression symptom phenotypes, their corresponding question on the PHQ-9, and the distribution and percentage of responses for each of the four samples on which genome-wide association analysis was performed.
2.1.2 Study population

In each GWAS sample, individuals were only retained if they had reported European ancestry and had provided a valid response to all PHQ-9 items. Individuals were excluded if they had reported a previous professional diagnosis of schizophrenia, psychosis, mania, hypomania, bipolar, or manic depression (UKB Field ID: 20544) or a previous prescription of medication for a psychotic experience (UKB Field ID: 20466).

2.1.3 GWAS software

All GWAS were conducted using REGENIE v3.1.3. In step one of REGENIE, ridge regression is applied to a subset of quality controlled (QC’d) variants to fit, combine and decompose a set of leave-one-chromosome-out (LOCO) predictions. QC for step one was undertaken using PLINK v1.9. In step two, imputed variants are tested for association with the phenotype. LOCO predictions from step one are included as covariates to control for proximal contamination. For all GWAS, genotyping batch, sex, age, and age-squared were included as covariates, as were the maximum available genetic principal components (PCs) for GLAD (10 PCs) and PROTECT (20 PCs) to control for population stratification. For the UK Biobank analyses, 16 PCs were included as recommended by Privé et al. Assessment centre was also included as a covariate for UK Biobank analyses.

A total of 40 GWAS were conducted for the meta-analyses – one for each of the nine PHQ-9 depression symptom phenotypes as well as the sum-score across all nine items in each of the four samples. To maximise statistical power, PHQ-9 phenotypes were treated as continuous (ranging 0-3 for individual items and 0-27 for the sum score) and analysed using linear regression. GWAS were restricted to the autosomes.

2.1.4 GWAS: UK Biobank (UKB)

UKB is a large-scale biomedical database and research resource consisting of ~500,000 individuals with data across a broad range of phenotypes, including mental health outcomes. Individuals in UKB have been genotyped on the custom UK Biobank Axiom or UKBiLEVE arrays, with imputed data available for ~90 million variants imputed with IMPUTE2 using the Haplotype Reference Consortium (HRC) and combined UK10K + 1000 Genomes Phase 3 reference panels.

UKB participants have completed the PHQ-9 in two online surveys. In total, 157,345 individuals provided responses as part of the Mental Health Questionnaire (UKB-MHQ) (Category: 136) between 2016 and 2017, and 167,199 individuals had provided responses as part of the Experience of Pain Questionnaire (UKB-EoP) (Category: 154) between 2019 and 2020.
After filtering for self-reported European ancestry, valid PHQ-9 responses, and previous diagnosis/prescription exclusions, 144,630 (UKB-MHQ) and 155,027 (UKB-EoP) individuals remained prior to genetic QC for REGENIE. In step one, single-nucleotide polymorphisms (SNPs) with a call rate > 98%, minor allele frequency (MAF) > 1%, and Hardy-Weinberg equilibrium test \( p > 1 \times 10^{-6} \) were retained, as were individuals with variant missingness < 2%, no unusual levels of heterozygosity, and not mismatched on sex. Individuals were retained if they were determined to be of European ancestry based on 4-means clustering on the first two principal components.

For the final GWAS analyses, 143,171 (mean age [SD] = 63.70 [7.68]; %female = 56.38%) and 152,932 (mean age [SD] = 65.95 [7.63]; %female = 56.57%) individuals proceeded from the MHQ and EoP questionnaires respectively. Of these, 108,601 individuals had provided responses on both questionnaires. In step two, a total of 9,746,698 imputed variants were retained with MAF ≥ 0.01 and imputation INFO score ≥ 0.7.

### 2.1.5 GWAS: Genetic Links to Anxiety and Depression (GLAD) Study

The GLAD study has the specific goal of recruiting a large cohort of re-contactable individuals with anxiety or depression into the National Institute for Health and Care Research (NIHR) Mental Health BioResource with genetic, environmental and phenotypic data collected\(^4\). Genotyping for GLAD was conducted using the UK Biobank v2 Axiom array and imputed using the TopMed imputation pipeline\(^5\).

After filtering for self-reported European ancestry, valid PHQ-9 responses, and previous diagnosis/prescription exclusions, 15,472 individuals remained prior to genetic QC for REGENIE step one. Genotype data was provided by the study team and had been filtered to retain SNPs with a genotype call rate > 95%, MAF > 1%, Hardy-Weinberg equilibrium test \( p > 1 \times 10^{-10} \), and individuals with genotype missingness < 5%. Individuals were additionally excluded if they had unusual levels of heterozygosity, mismatched on sex and of non-European ancestry based on 4-means clustering. A total of 15,171 individuals (mean age [SD] = 39.27 [14.61]; %female = 78.30%) were retained for the final analysis. In step two, a total of 13,979,187 imputed variants with MAF ≥ 0.001 and INFO ≥ 0.7 were analysed.

### 2.1.6 GWAS: PROTECT Study

PROTECT is an online registry of ~25,000 UK-based individuals that aims to track cognitive health in older adults. Individuals were only considered eligible for inclusion in PROTECT if they were older than 50, had no previous dementia diagnosis and had internet access. Genetic data are available alongside phenotypic data for ~10,000 of the participants. These individuals were genotyped on the Illumina Infinium Global
Screening Array and imputed on the 1000 Genomes reference panel using the Michigan imputation server and genotype phasing using Eagle.

After filtering for self-reported European ancestry, valid PHQ-9 responses, and previous diagnosis/prescription exclusions, 7,589 individuals remained for genetic QC for step one of REGENIE. Genetic data in PROTECT had been previously QC’d prior to imputation to only retain individuals and variants with a call rate > 98%, Hardy-Weinberg equilibrium test $p > 0.00001$ and excluding unusual heterozygosity. Variants used in step one were down-sampled from the imputed data using a snplist from the Illumina Infinium Global Screening Array provided by the PROTECT investigators. Variants were retained if they had MAF > 1%. After mismatched sex and 4-means clustering ancestry exclusions, a total of 7,589 individuals (mean age [SD] = 61.96 [7.07]; %female = 75.13%) proceeded to step two. In step two, 9,388,534 imputed variants with MAF $\geq 0.001$ and imputation INFO score $\geq 0.7$ were analysed.
2.2 GWAS summary statistics

An overview of additional summary statistics obtained for this study can be seen in Table 2.

Table 2: An overview of previously conducted genome-wide association studies for depression and Alzheimer’s disease used in this study. Heritability estimates were calculated naively on the liability scale from these standardised summary statistics using Linkage Disequilibrium Score Regression (LDSC), taking a population prevalence of 15% for depression and 5% for Alzheimer’s disease.

<table>
<thead>
<tr>
<th>Phenotype (Original GWAS)</th>
<th>Excluded Samples</th>
<th>Cases (% Clinical)</th>
<th>Controls (% Clinical)</th>
<th>Total</th>
<th>SNP h2 (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Depression GWAS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broad Depression (Howard et al. (2019))</td>
<td>23andMe</td>
<td>170,756 (25.3%)</td>
<td>329,443 (29.0%)</td>
<td>500,199</td>
<td>0.0798 (0.003)</td>
</tr>
<tr>
<td>Clinical Depression (Wray et al. (2018))</td>
<td>UK Biobank; 23andMe</td>
<td>45,591 (100%)</td>
<td>97674 (100%)</td>
<td>143,261</td>
<td>0.1012 (0.007)</td>
</tr>
<tr>
<td><strong>Alzheimer’s Disease GWAS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical + Proxy AD (Bellenguez et al. (2022) (Stage 1))</td>
<td>N/A</td>
<td>85,934 (45.5%)</td>
<td>401,577 (14.0%)</td>
<td>487,511</td>
<td>0.0306 (0.003)</td>
</tr>
<tr>
<td>Clinical + Proxy AD (Wightman et al. (2021))</td>
<td>23andMe</td>
<td>86,531 (46.1%)</td>
<td>676,386 (26.1%)</td>
<td>762,917</td>
<td>0.0237 (0.004)</td>
</tr>
<tr>
<td>Clinical + Proxy AD (Jansen et al. (2019))</td>
<td>N/A</td>
<td>71,880 (33.5%)</td>
<td>383,378 (14.4%)</td>
<td>455,258</td>
<td>0.0234 (0.003)</td>
</tr>
<tr>
<td>Proxy-only AD (Marioni et al. (2018))</td>
<td>IGAP</td>
<td>42,035 (0%)</td>
<td>272,243 (0%)</td>
<td>314,278</td>
<td>0.0165 (0.003)</td>
</tr>
<tr>
<td>Clinical-only AD (Wightman et al. (no UKB) (2021))</td>
<td>UK Biobank; 23andMe</td>
<td>39,918 (100%)</td>
<td>358,140 (29.1)</td>
<td>398,058</td>
<td>0.0431 (0.008)</td>
</tr>
<tr>
<td>Clinical-only AD (Kunkle et al. (2019) (Stage 1))</td>
<td>N/A</td>
<td>21,982 (100%)</td>
<td>41,944 (86.2%)</td>
<td>63,926</td>
<td>0.068 (0.011)</td>
</tr>
</tbody>
</table>

2.2.1 Clinical and broad depression

To examine potential differences in genetic overlap with AD between depression as a disorder compared to individual depression symptoms, summary statistics for two previously conducted GWAS of clinical and broad depression were obtained from the Psychiatric Genomics Consortium (PGC) (https://pgc.unc.edu/for-researchers/download-results/). For clinical depression, we used a subsample of the Major Depressive Disorder (MDD) GWAS by Wray et al.\textsuperscript{11} that excluded samples from the UKB and 23andMe, and contained only individuals for whom case ascertainment was defined through structured diagnostic interview or electronic health records. For the broad definition depression GWAS, we used a subsample of the depression GWAS by Howard et al.\textsuperscript{12}, which also excluded samples from 23andMe. In addition
to clinical cases and controls used by Wray et al., this broad depression GWAS included individuals in the UKB for whom case-control ascertainment was based on self-reported responses to the questions “Have you ever seen a general practitioner for nerves, anxiety tension or depression?” and “Have you ever seen a psychiatrist for nerves anxiety, tension or depression?”.

2.2.2 Alzheimer’s disease

Summary statistics were obtained from six previously conducted AD GWAS: three with proxy+clinical, one with proxy-only and two with clinical-only case ascertainment. All three of the proxy+clinical AD GWAS – Bellenguez et al., Wightman et al. and Jansen et al. – and the proxy-only AD GWAS – Marioni et al. used data from the UKB for proxy AD samples.

There are some key differences in the way these AD GWAS define proxy cases and controls. Bellenguez et al. define proxy cases/controls as a binary phenotype, whereby individuals reporting a parent with AD or dementia are considered cases and those reporting no parental history are considered controls. Wightman et al. and Jansen et al. instead define proxy cases/control as a continuous phenotype, summing the number of parents an individual has reported with dementia and down-weighting unaffected parents by their age (or age of death).

For the proxy-only Marioni et al. GWAS, summary statistics were obtained from a meta-analysis of paternal and maternal AD. Here, proxy phenotyping was based on the self-report of either maternal or paternal AD, including the parent’s age at time of reporting/age of death as a covariate.

Summary statistics from a clinical-only subsample of the GWAS by Wightman et al. that excluded proxy cases/controls from the UKB were obtained from the authors. Summary statistics for a final clinical-only AD GWAS were obtained from Stage 1 of the GWAS by Kunkle et al.

2.3 Summary statistic standardisation

Summary statistics from all 40 depression symptom GWAS, the two depression GWAS and the six AD GWAS, were standardised using the MungeSumstats in R version 4.2.1. Using dbSNP 141 and the BSgenome.Hsapiens.1000genomes.hs37d5 reference genome, missing rsIDs were corrected, duplicates and multi-allelic variants removed, effect alleles and the direction of their effects aligned to the reference genome, and variants filtered at INFO score ≥ 0.7 and MAF ≥ 0.01. The GLAD Study and Bellenguez et al. summary statistics were lifted over from GRCh38 to GRCh37.

2.4 SNP heritability

SNP heritability (\(h^2_{SNP}\)) estimates were calculated for all GWAS used in this study with Linkage Disequilibrium Score Regression (LDSC). Briefly, LDSC calculates \(h^2_{SNP}\) by regressing the effect sizes from GWAS summary statistics on their
LD score as computed in a reference panel – in this case HapMap3 variants contained within the European sample of 1000 Genomes Phase 3. Liability scale $h^2_{SNP}$ was calculated naively from the standardised depression GWAS and AD GWAS using a 15% and 5% population prevalence respectively\textsuperscript{11,16}. Heritability $Z$-scores were calculated for all phenotypes by dividing the $h^2_{SNP}$ estimates by their standard error.

### 2.5 GWAS meta-analysis of depression symptoms

To leverage the maximum genetic information available controlling for the sample overlap between the UKB-MHQ and UKB-EoP samples, REGENIE output for each PHQ-9 phenotype from the UKB-EoP, GLAD Study and PROTECT were first subject to Inverse Variance Weighted (IVW) meta-analysis using METAL\textsuperscript{58}. All available variants were included, for a total of 8,425,618 ($N =$ 175,692). Multi-trait Analysis of GWAS (MTAG)\textsuperscript{59} v1.0.8 was then used to meta-analyse the METAL output with the UKB-MHQ sample. While MTAG is commonly used for the joint genetic analysis of multiple traits or multiple measurements of the same trait, by assuming the heritability of included phenotypes are equal (\texttt{--equal-h2}) and their genetic correlation is one (\texttt{--perfect-gencov}), MTAG performs an IVW meta-analysis of the same measures of the same trait, accounting for sample overlap using the cross-trait intercept from LDSC\textsuperscript{56,57}. Heritability estimates for all samples, plus the METAL meta-analysis, are in Supplementary Table 2. Genetic correlations between the UKB-MHQ and METAL GWAS are in Supplementary Table 3. For greater detail on the IVW function of MTAG, see the online methods of the original MTAG paper\textsuperscript{59}. A total of 8,196,874 SNPs with MAF > 0.01 were present for MTAG analysis.

This MTAG function provides one set of summary statistics and two GWAS-equivalent sample sizes – one for each original sample included. A single, weighted GWAS-equivalent $N$ was obtained for each PHQ-9-MTAG GWAS, using the following formula;

$$N(\text{final}) = \frac{(N_1(\text{pre})\times N_1(\text{post})) + (N_2(\text{pre})\times N_2(\text{post}))}{N_1(\text{pre}) + N_2(\text{pre})}$$

where $N_1$ and $N_2$ represents the UKB-MHQ and METAL GWAS respectively, $\text{pre}$ is the mean sample size prior to inclusion in MTAG and $\text{post}$ is the GWAS-equivalent sample estimated by MTAG following analysis.

### 2.6 Genomic risk loci and gene annotation

GWAS meta-analysis results were annotated using FUMA GWAS\textsuperscript{60} v3.1.6a. Genomewide significance was set at $p \leq 5e-8$. Lead variants at genomic risk loci were defined by clumping all variants correlated at $r^2 > 0.1$ 250kb either side. Clumping was performed using the European sample of the 1000 Genomes Phase 3 reference
panel. Lead variants were mapped to genes within 10kb using positional mapping and eQTLs from four brain (BrainSeq\textsuperscript{61}, PsychENCODE\textsuperscript{62}, CommonMind\textsuperscript{63}, and BRAINEAC\textsuperscript{64}) and five blood (Blood-eQTL\textsuperscript{65}, BIOS\textsuperscript{66}, eQTLGen cis and trans\textsuperscript{67}, Twins UK\textsuperscript{68} and xQTLServer\textsuperscript{69}) eQTL datasets, alongside all 54 tissue type eQTLs from GTEx v8 (https://gtexportal.org/home/tissueSummaryPage).

2.7 Genetic correlations

Genetic correlation can be understood as the genome-wide correlation of genetic effects between two phenotypes, and as such can be viewed as an estimate of pleiotropy\textsuperscript{70}.

Genetic correlations were calculated between each depression phenotype and the six AD GWAS using High Definition Likelihood v1.4.1 (HDL)\textsuperscript{71}. HDL extends the LDSC framework by leveraging LD information from across the whole LD reference panel through eigen decomposition, thus shrinking standard errors and improving precision. A pre-computed eigenvector/value LD reference panel calculated from 335,265 individuals of European ancestry in the UKB was obtained via the HDL GitHub (https://github.com/zhenin/HDL/wiki/Reference-panels). This reference panel was calculated using 1,029,876, imputed, autosomal HapMap3 SNPs, with bi-allelic SNPs outside the MHC region, MAF > 5%, a call rate > 95% and INFO > 0.9 retained.

2.8 Local genetic correlations

Local genetic correlation assesses the correlation of genetic effects between two phenotypes in specific region of the genome. It provides a more refined examination of the genetic overlap between two phenotypes, allowing for the identification of key regions driving shared genetic architecture.

LAVA\textsuperscript{72} was used to assess regions of local genetic correlation between each depression phenotype and the six AD GWAS across 2495 semi-independent, pre-defined LD-blocks of at least 2500 base pairs (https://github.com/josefin-werme/LAVA). Loci-specific heritability estimates were calculated for each phenotype for each block. If both a depression phenotype and an AD GWAS showed significant local heritability at a specific locus (Bonferroni corrected \( p \)-value \( \leq 2\text{e-5} \) (0.05/2495)), bivariate local genetic correlation was tested. Bivariate results were considered significant at \( \rho_{DR} \leq 0.05 \) correcting for the total number of bivariate tests. Sample overlap was accounted for using an LDSC intercept matrix. Analysis was restricted to the 5,531,969 non-strand-ambiguous variants shared across all GWAS summary statistics and on the European ancestry 1000 Genomes Phase 3 reference panel.
2.9 Colocalisation

Colocalisation using COLOC is a Bayesian statistical method to assess the probability of two phenotypes sharing a causal variant in a pre-defined genomic region.

Within the colocalisation framework, posterior probability is assessed for five hypotheses:

- $H_0$: There is no causal variant in the region for either phenotype
- $H_1$: There is a causal variant in the region for the first phenotype
- $H_2$: There is a causal variant in the region for the second phenotype
- $H_3$: Distinct causal variants for each trait in the region
- $H_4$: Both traits share a causal variant

Colocalisation analysis was conducted using the COLOC-reporter pipeline\(^\text{73}\) (https://github.com/ThomasPSpargo/COLOC-reporter). COLOC-reporter extracts variants in user-defined genomic regions and calculates the LD matrix for this region from a user-defined reference panel. It then harmonises the summary statistics to match the allele order of the reference panel, flipping effect directions accordingly. Observed versus expected Z-scores are assessed using the diagnostic tools provided in the susieR R package\(^\text{74}\). Z-score outliers are omitted. Following this QC, Sum of Single Effects (SuSI\(^\text{74}\)) fine-mapping is conducted to identify 95% credible sets in these regions for each phenotype. Identification of credible sets in both phenotype allows for the relaxation of the single causal variant assumption. All possible credible sets are then assessed pairwise for shared signal between phenotypes, improving resolution for colocalisation inference in regions containing multiple signals\(^\text{75}\). Should no 95% credible set be identified or only identified for one phenotype, colocalisation under the single causal variant assumption is performed using the coloc.abf\(^\text{76}\). A posterior probability $\geq 80\%$ for $H_4$ was considered evidence of colocalisation between two phenotypes ($\text{PP.}H_4 \geq 0.8$). The SuSI\(^\text{74}\) model assumed at most 10 causal variants ($L = 10$) per credible set. We used default priors.

Trait pairs with LAVA correlations significant at $p_{\text{FDR}} \leq 0.05$ were passed to COLOC-reporter. Where the same depression phenotype showed nominally significant local genetic correlation at the same locus but with different AD GWAS, these phenotype pairs were included as a sensitivity analysis. Regions +/-250kb ($r^2 \geq 0.1$) from lead variants at genome-wide significant loci from the MTAG-PHQ-9, broad and clinical depression GWAS were also examined for evidence of colocalisation with the 6 AD GWAS.
2.10 Mendelian randomisation (MR)

MR is a statistical method that uses genetic variants associated with an exposure as instrumental variables to assess the causal effect of that exposure on an outcome of interest\textsuperscript{77}. Two Sample MR framework estimates the causal relationships between an exposure and outcome using GWAS summary statistics. Valid MR instruments are defined by three key assumptions: (1) relevance – IVs are strongly associated with the exposure of interest; (2) independence – there are no confounders in the association between IVs and the outcome of interest; and (3) exclusion restriction – instruments are not associated to the outcome other than via the exposure, for example through horizontal pleiotropy\textsuperscript{77}.

Sample overlap is a known source of bias in Two Sample MR\textsuperscript{78}. Given the likelihood of sample overlap between the depression phenotypes and the AD GWAS containing proxy cases/controls due to participants from the UKB, MR analysis was primarily conducted using Causal Analysis Using Summary Effect estimates (CAUSE)\textsuperscript{79} v1.2.0. CAUSE is a Bayesian MR method robust to sample overlap, correlated and uncorrelated pleiotropy, also avoiding exclusion restriction assumption violations\textsuperscript{80}.

In CAUSE, the presence of a significant causal effect, $\gamma$, is derived by comparing the expected log pointwise posterior density (ELPD) of a sharing model with $\gamma$ fixed to zero to a causal model where $\gamma$ is a free parameter for a one-tailed p-value test. The posterior median of $\gamma$ describes the causal effect estimate and is reported alongside its 95% credible intervals. CAUSE uses a larger set of instruments that traditional MR methods. As such, clumping was performed with the default setting at $r^2 \geq 0.01$ and a p-value $\leq 0.001$ within a 10,000kb window.

Causal estimates were also calculated using the traditional IVW method. IVW estimates are biased by the presence of horizontal pleiotropy\textsuperscript{81}. Sensitivity analyses were therefore conducted using MR-Egger, weighted-median and weighted-mode MR. These methods allow for varying degrees of pleiotropy while providing unbiased causal estimates\textsuperscript{82-84}. MR-PRESSO\textsuperscript{85} was also implemented. MR-PRESSO identifies and excludes outlying instruments based on their contribution to heterogeneity and provides a corrected causal estimate.

Pleiotropy was assessed using the MR-Egger intercept test\textsuperscript{86} (significant pleiotropy: $p \leq 0.05$). Heterogeneity tests were also conducted for IVW and MR-Egger estimates using their respective Cochran’s Q test\textsuperscript{87} (significant heterogeneity: $p$-value $\leq 0.05$). Instrument strength was calculated via the minimum and mean F-statistic (recommended F-statistic $\geq 10$)\textsuperscript{88} ($\beta^2/SE^2$). $I^2$ was calculated to ensure measurement error was sufficiently low so as to ensure validity of results from MR-Egger (recommended $I^2 \geq 0.90$)\textsuperscript{89}.
For IVW-MR, MR-Egger, weighted-median and weighted-mode MR, instruments clumped at a $r^2 \geq 0.001$ and a $p$-value $\leq 5e$-8 within 10,000kb. Where no instruments were available at $p \leq 5e$-8 or where < 5 instruments were available, a $p$ $\leq 5e$-6 threshold was used.

For all analyses, instruments were clumped using data from individuals of European ancestry in 1000 Genomes Phase 3. IVW, MR-Egger, weighted-median and weighted-mode analyses were conducted using the TwoSampleMR package v0.5.6.

The APOE gene is known to be associated with non-AD phenotypes such as cardiovascular disease and type 2 diabetes (T2D). Both have been linked to depression in previous MR analyses. As such, the inclusion of APOE would violate MR’s independence assumption. All MR analyses were therefore conducted excluding variants in the APOE region (chr19:45,020,859–45,844,508 (GRCh37)) as per Lord et al.94.

2.11 Polygenic risk scores (PRS)

Polygenic risk scores (PRS) describe the sum of an individual’s risk alleles, weighted by their effect size. The ten PHQ-9-MTAG, clinical depression and broad definition depression GWAS were processed for PRS using the BayesR-SS function contained within MegaPRS and implemented in LDAK v5.2.1. BayesR-SS assumes the BLD-LDAK heritability model, which incorporates 65 genome annotations pertaining to genomic features such as whether the variants are in coding regions or highly conserved. Annotation files were obtained from the LDAK website (http://dougspeed.com/bldldak/). PRS calculation was restricted to the 1,217,311 HapMap3 SNPs, with strand ambiguous SNPs excluded.

PRS predictive utility for AD case/control status was assessed using logistic regression in three clinically ascertained AD cohorts: AddNeuroMed and Dementia Case Register Studies (ANM) (N\textsubscript{cases} = 564; N\textsubscript{controls} = 345), the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (https://adni.loni.usc.edu) (N\textsubscript{cases} = 356; N\textsubscript{controls} = 360), and the Genetic and Environmental Risk in Alzheimer’s Disease (GERAD1) Consortium (https://portal.dementiasplatform.uk/CohortDirectory/Item?fingerPrintID=GERAD) (N\textsubscript{cases} = 2,661; N\textsubscript{controls} = 1,124) (Supplementary Table 4). All analyses controlled for age, sex, and 10 PCs, and were restricted to individuals aged $\geq 65$ to provide clean controls. As a sensitivity analysis, PRS were also calculated in these cohorts excluding the APOE region. Genetic QC and imputation steps for these cohorts can be viewed in detail in the study by Lord et al.96.
3. RESULTS

3.1 PHQ-9 genome-wide meta-analyses

The MTAG meta-analyses identified a total of 40 genomic risk loci between the ten PHQ-9 phenotypes (GWAS equivalent \( N \) range: 224,535 – 308,421). Only one depression symptom – suicidal thoughts – identified no genome-wide significant variants. Three lead SNPs were shared with more than one PHQ-9 phenotype, leaving a total of 37 unique genomic risk loci (Table 3). The significance of each of the lead variants in each of the samples contributing to the meta-analysis can be seen in Supplementary Table 5. eQTL mapping in FUMA mapped lead variants at genomic risk loci to 76 genes (Supplementary Table 6). \( h^2_{SNP} \) for the MTAG-PHQ-9 GWAS ranged from 1.12% for suicidal thoughts to 6.78% for the PHQ-9 sum-score. \( h^2_{SNP} Z \)-scores were all > 4 (range: 6.59 – 18.50) (Supplementary Table 7), indicating sufficient heritability to obtain reliable genetic correlation estimates in downstream analyses\(^57\). Genomic inflation factors (\( \lambda_{GC} \)) ranged from 1.0638 to 1.2156, with LDSC intercepts ranging from 0.9997 to 1.0007, indicating inflation was due to polygenic signal as opposed to confounding due to population stratification\(^56,99\). Manhattan and QQ plots can be viewed in Figure 2.
Table 3: Genomic risk loci from the 10 MTAG-PHQ-9 genome-wide association meta-analyses. Nearest gene is based on positional mapping and ANNOVAR annotation. Expression quantitative trait loci (eQTL) mapping was limited to \( p_{eQTL} \leq 0.05 \) in the selected eQTL datasets (see: Methods). Specific tissues with significant \( p_{eQTL} \) can be seen in Supplementary Table 2.
Figure 2: Manhattan and QQ-plots for each of the 10 MTAG-PHQ-9 GWAS meta-analyses. The red lines on the Manhattan plots indicate genome-wide significance ($p \leq 5 \times 10^{-8}$), while the blue lines indicate suggestive significance ($p \leq 1e^{-5}$). Top variants at each genomic risk loci are annotated.

3.2 Genetic correlations

Of the 72 bivariate genetic correlations calculated between the 12 depression phenotypes and the six AD GWAS, 24 were nominally significant and 20 remained significant at $\rho_{\text{FDR}} \leq 0.05$ ($r_g$ range: -0.25 – 0.35; $p$-value range: $1.25 \times 10^{-2}$ – $4.01 \times 10^{-5}$; $\rho_{\text{FDR}}$ range: $4.5 \times 10^{-2}$ – $1.9 \times 10^{-3}$). Of these, 19 were identified when the AD GWAS in the pair contained either clinical+proxy cases and controls, or proxy-only cases and controls (Figure 3; Supplementary Table 8). Only one $\rho_{\text{FDR}}$ significant association was found when using a clinical AD GWAS – between suicidal thoughts and Wightman et al. ($r_g = -0.25$, $p = 6.78 \times 10^{-3}$, $p$-value = $6.78 \times 10^{-3}$; $\rho_{\text{FDR}} = 3.48 \times 10^{-3}$).

All depression phenotypes were significantly genetically correlated with each other ($r_g$ range: 0.57– 0.98; $p$-values $\leq 3.71 \times 10^{-23}$) (Supplementary Table 9). Only one PHQ-9 symptom pair – concentration problems and psychomotor changes – showed a genetic correlation that was not statistically different from one (95% CI included one), indicating genetic heterogeneity across depression symptoms.
Figure 3: Heatmap of High Definition Likelihood (HDL) calculated genetic correlations between depression items and the six Alzheimer’s disease genome-wide association studies of varying proxy/clinical case/control ascertainment. The colour of each square indicates the strength of the correlation on a scale of -1 to 1. Fill size of the square indicates uncorrected p-value significance. Genetic correlations significant at $p_{FDR} \leq 0.05$ are circled.

### 3.3 Local genetic correlations

After univariate testing, a total of 4271 bivariate local genetic correlation tests were conducted across 324 genomic loci. Of these tests, 716 were nominally significant and 15 remained significant at $p_{FDR} \leq 0.05$ across 14 unique genomic loci (local $r_g$ range: -0.81 – 0.82; $p$-value range: $1.48 \times 10^{-4} – 4.2 \times 10^{-6}$; $p_{FDR}$ range = $4.22 \times 10^{-2} – 1.38 \times 10^{-3}$) (Supplementary Table 10). Of the 15 statistically significant tests, ten were identified when using clinical+proxy/proxy-only AD GWAS. No depression phenotype showed a statistically significant association at the same genomic locus with more than one AD GWAS. However, for ten of the 15 statistically significant tests, nominally significant local genetic correlation was observed between the depression phenotype and at least one additional AD GWAS at the same loci (Supplementary Table 11). Only locus 1790 (chr12:51769420– 53039987) showed $p_{FDR}$ significant local genetic correlation with more than one depression phenotype – concentration problems and sleep problems, both with the clinical-only Wightman et al. GWAS. The number of positively and negatively correlated loci identified between each phenotype pair can be viewed in Supplementary Table 12.
3.4 Colocalisation

Following LAVA, 14 \( \rho_{\text{FDR}} \) significant regions of local genetic correlation were passed to the COLOC-reporter pipeline across 15 depression-AD phenotype pairs. A further 14 colocalisation tests were conducted where nominally significant local genetic correlation was observed at a \( \rho_{\text{FDR}} \) significant loci between the same depression phenotype and a different AD GWAS. As such, a total of 29 statistical colocalisation tests were conducted to follow up LAVA results. No 95% credible sets were identified by SuSiE for any phenotype pairs in these regions. As such, all analyses were conducted under the single causal variant assumption of \textit{coloc.abf}. There was no evidence of colocalisation in any of these loci (mean \( \text{PP.H}_4 = 0.59\% \)) (Supplementary Table 11). All but two of these tests indicated no causal variant present in either phenotype (\( \text{PP.H}_0 > 0.8 \)). The two tests in locus 319 (chr2:126754028-127895644), indicated strong probability of a causal variant for the Kunkle et al.\(^{14} \) and Wightman et al.\(^{45} \) clinical AD GWAS (\( \text{PP.H}_2 > 0.9 \)). This locus contains \textit{BIN1}, a known risk gene for AD involved in tau regulation\(^{100,101} \).

An additional 762 colocalisation tests were conducted with the six AD GWAS using regions +/-250kb (\( r^2 > 0.1 \)) lead variants from the MTAG-PHQ-9, broad and clinical depression GWAS. SuSiE identified evidence of colocalisation in regions +/-250kb of lead variants at genomic risk loci 14 (depressed mood), 15 (appetite change), 16 (PHQ-9 sum score) and for broad depression at chr7:12000402-12500402 (\( \text{PP.H}_1 \) range: 0.79 – 0.85), all with the same three AD GWAS – Bellenguez et al.\(^{17} \), Wightman et al.\(^{16} \), and Wightman et al. (excluding the UKB)\(^{45} \) (Supplementary Table 13). These colocalisations were all in the region of the transmembrane protein 106B gene (\textit{TMEM106B}) – visualised using LocusZoom\(^{102} \) in Figure 4. Colocalisation was also identified for the same phenotype pairs at the same loci under the single causal variant assumption of \textit{coloc.abf} (Supplementary Table 14). The same depression phenotypes and loci were suggestive of colocalisation with Jansen et al.\(^{15} \) (\( \text{PP.H}_4 > 0.6 \)). In a follow-up analysis, we assessed statistical colocalisation +/-250kb \textit{TMEM106B} (chr7:12000920-12532993) between these four AD GWAS and all remaining depression phenotypes. Additional evidence of colocalisation was identified at \textit{TMEM106B} between fatigue and the Bellenguez et al.\(^{17} \), Wightman et al.\(^{16} \), and Wightman et al. (excluding the UKB)\(^{45} \) AD GWAS (Supplementary Table 15), and was suggestive for Jansen et al.\(^{15} \) (Supplementary Table 16). Evidence of colocalisation was also suggestive for psychomotor changes with Bellenguez et al.\(^{17} \), Wightman et al.\(^{16} \), and Wightman et al. (excluding the UKB)\(^{45} \) (\( \text{PP.H}_4 > 0.6 \)).
Figure 4: LocusZoom plots of the transmembrane protein 106B (TMEM106B) gene region containing evidence of colocalisation (PP.Hs ≤ 0.8). for (A.) broad depression, (B.) appetite changes, (C.) depressed mood, (D.), the PHQ-9 sum score, (E.) Bellenguez et al., (F.) Wightman et al. and (G.) Wightman et al. excluding the UK Biobank. The most significant variant for each phenotype is labelled in their respective plot.

3.5 Mendelian Randomisation

We conducted 144 MR tests between the depression phenotypes and AD – 72 in each direction. In CAUSE, no significant causal association was identified between any of the depression items and AD in either direction even at nominal significance (Supplementary Table 17).

F-statistics indicated that instrument strength was sufficient (\(F_{\text{Mean}}\) range: 22.43 – 63.36; \(F_{\text{Min}}\) range: 20.84 – 31.56; \(F_{\text{Max}}\) range: 26.37 – 402.86), as did \(I^2\) for instrument suitability for MR-Egger (\(I^2\) range: 0.91 – 0.98). A \(p_{\text{FDR}} \leq 0.05\) was applied in each of the other MR methods to correct for the 144 tests conducted, after which no statistically significant associations were observed in any method. Three nominally significant causal associations where a depression item was the exposure (\(\beta\) range = 0.21 – 1.59; \(p\)-value range = 0.01 – 2e-3) and two where an AD GWAS was the exposure (\(\beta\) range = 0.018 – 0.02; \(p\)-value range = 0.01 – 0.02) were detected by
IVW-MR, of which three were nominally significant in weighted-median MR. No significant pleiotropy (MR-Egger intercept test p-value ≤ 0.05) or heterogeneity (IVW Cochran’s Q p-value ≤ 0.05) was detected for these five tests. However, nominally significant estimates were only observed between exposures and outcomes where the AD GWAS contained proxy+clinical or proxy-only cases/controls and where the GWAS used for both exposure and outcome were either majority or entirety derived from the UK Biobank, suggesting sample overlap as a potential source of bias. MR-PRESSO detected and excluded significant outliers in 42 tests, returning outlier-corrected causal estimates. Results remained non-significant. Full MR results can be viewed in Supplementary Table 18.

3.6 Polygenic risk scores (PRS)

No statistically significant associations were detected between any of the depression phenotype PRS and AD case/control status in any of the three AD target samples ($\rho_{DR} \leq 0.05$, corrected within each sample), with only the suicidal thoughts PRS negatively nominally associated with AD case/control in the ADNI cohort (Nagelkerke’s pseudo-r$^2 = 0.01$, OR [95% CI] = 0.83 [0.72–0.94], $\beta = -0.19$, SE = 0.07, p-value = $7.90 \times 10^{-3}$). Exclusion of the APOE region had no effect on results (Supplementary Table 19).

4. Discussion

In this study, we performed the first and largest genome-wide meta-analysis of PHQ-9 depression symptom items to date (GWAS equivalent N range: 224,535 – 308,421), identifying 37 genomic risk loci. Follow-up analyses examining the genetic overlap between depression/depression symptoms and AD identified 20 significant global correlations and 15 significant local correlations at 14 loci across six AD GWAS with varying proportion of clinical case/control ascertainment. Significant global genetic correlation were primarily with AD GWAS containing proxy cases and controls. No colocalisation was identified at any of the regions of local genetic correlation. However, there was strong evidence of colocalisation between depression several phenotypes and AD in the region of $TMEM106B$. Overall, MR did not suggest that depression or its symptoms were causal for AD, nor vice versa, while polygenic risk scores for depression phenotypes were not predictive of AD case/control status in three clinical AD samples.

The increased power of our PHQ-9 GWAS allowed for the identification of 28 more genomic risk loci than the previous analysis by Thorp et al. Several of the loci identified in this study have shown previous associations with related phenotypes. For example, $SHISA4$ – here identified in association with fatigue symptoms – has been previously implicated as playing a role in disrupted sleep and daytime
napping\textsuperscript{104}. The top variant for sleep problems at genomic risk loci 6 (\textit{MEIS1}) – rs113851554 (chr2: 66750564) – was also the top variant in a previous GWAS of insomnia and restless leg syndrome\textsuperscript{105}. Additionally, the obesity gene \textit{FTO}\textsuperscript{106} was identified as a genomic risk loci for appetite changes. Although the role of \textit{FTO} in depression is inconclusive\textsuperscript{107}, it has been recently linked to anxiety and depression symptoms in individuals with anorexia nervosa (AN)\textsuperscript{108}. Consequently, its identification in association with appetite change symptoms – a phenotype relevant to eating behaviours – suggests that symptom-based genetic analysis can help identify the phenotype-relevant biology of individual depression symptoms. Taken alongside the fact that only three of the genomic risk loci were shared between more than one depression symptom phenotype, our GWAS meta-analysis backs findings by Thorp et al.\textsuperscript{29} of genetic heterogeneity between depression symptoms.

Our findings also highlight genetic similarities between depression symptoms. For example, \textit{TMEM106B} – a gene identified in previous depression GWAS\textsuperscript{11,12} – was the nearest gene to lead variants for three PHQ-9 items – appetite changes (rs13234970), depressed mood (rs3807866) and the PHQ-9 sum score (rs12699338). \textit{TMEM106B} was strongly suggested as a causal gene in a recent multi-ancestry depression GWAS\textsuperscript{109}. Further, dysregulation of \textit{TMEM106B} expression has been implicated in association with MDD\textsuperscript{110} and with anxious and weight gain MDD subtypes – both associated with treatment resistance\textsuperscript{111}. \textit{TMEM106B} has also been implicated in self-reported diagnosis of anxiety disorder\textsuperscript{112}, neuroticism\textsuperscript{113} and in a latent factor GWAS of depressive, manic and psychotic symptoms/disorders\textsuperscript{114}, suggesting it is linked to psychiatric risk more generally.

That colocalisation is observed in the \textit{TMEM106B} region between multiple depression phenotypes and both proxy+clinical and clinical-only AD is therefore of particular interest. \textit{TMEM106B} plays a role in lysosomal function – particularly in motor neurons\textsuperscript{115} – and is classically considered a risk gene for frontotemporal dementia\textsuperscript{116}. It is has also been identified in recent AD GWAS\textsuperscript{16,17}, and is associated with brain aging, cognitive decline, and neurodegeneration across other brain disorders including Amyotrophic Lateral Sclerosis, Multiple Sclerosis and Parkinson’s Disease\textsuperscript{117–122}. Further, \textit{TMEM106B} has been linked to higher CSF levels of neurofilament light (NfL)\textsuperscript{123} – itself predictive of cognitive decline, brain atrophy and cortical amyloid burden in individuals with AD and Mild Cognitive Impairment (MCI)\textsuperscript{124}. A recent study also identified higher levels of plasma NfL in individuals with depression, although the sample size of this study was relatively small\textsuperscript{125}. Thus, colocalisation between depression phenotypes and AD at \textit{TMEM106B} indicates that depression may itself be genetically linked to overall brain health and resulting general dementia risk. Two previous studies have identified \textit{TMEM106B} as playing a role in both depression and AD\textsuperscript{20,21}. However, our study suggests that this overlap
may be driven by the genetic architecture of specific depression symptoms, highlighting the benefits of symptom-level genetic analysis.

A mixed picture is painted across our other downstream analyses. Depression/depression symptom PRS were not predictive of AD case/control status in three clinical samples. Additionally, we do not find evidence of any causal associations using MR. These MR findings are in line with a number of previous studies\(^{19,39}\), but in contradiction to that conducted by Harerimana et al.\(^{22}\). In this study, we do not recapitulate the causal association identified by Harerimana et al.\(^{22}\) for broad depression on AD as measured by Jansen et al.\(^{15}\) – even at a nominal level. This is possibly due to the exclusion of rare variants (MAF < 0.01), and as such future investigation into the role of rare variants in shared risk for depression and AD is warranted. However, the lack of evidence in these analyses versus the observed relationship in epidemiological studies suggests the presence of unidentified confounding. The investigation of possible confounding is an important step for future research to help us better understand the association between depression and dementia.

While previous studies have shown that the direction of effect in MR can change depending on whether the outcome AD GWAS contains proxy or clinical cases/controls\(^{38,39}\), this study is – to the best of our knowledge – the first to demonstrate a similar effect with genetic correlations. Of the significant genetic correlations we identified, 95% were identified in proxy+clinical or proxy-only AD GWAS. Where two previous studies\(^{20,22}\) identified a genetic correlation between depression and AD, it is noticeable that they used the Jansen et al. proxy+clinical AD GWAS as their primary outcome. Further, in the study by Harerimana et al.\(^{22}\) sensitivity analysis did not observe a significant genetic correlation using the clinical-only GWAS by Kunkle et al.\(^{14}\).

Exactly why depression/depression symptoms show differences in genetic correlation between proxy and clinical AD is a matter of interest. It is noted that most individuals with dementia are cared for by a family member while they are still living in the community\(^{126}\). Caregivers incur considerable lifestyle changes, emotional stress, and social strain\(^{127,128}\) and depression is reported in up to 50% of caregivers for individuals with dementia\(^{129}\). As such, the genetic correlations between depression/depression symptoms and proxy-AD may result from increased depression risk due to caregiver responsibilities, exacerbated by a genetic risk. However, this would not explain why no genetic correlations were identified with the Bellenguez et al.\(^{17}\) GWAS despite this study also containing proxy+clinical phenotyping. As mentioned, Bellenguez et al.\(^{17}\) define proxy cases/controls as a binary phenotype, whereas Wightman et al.\(^{16}\) and Jansen et al.\(^{15}\) define proxy cases/controls as a continuous phenotype. These phenotyping differences likely
partially explain the differences in the genetic correlation results, given that these AD GWAS all use the same data from the UK Biobank. However, the proxy-only Marioni et al. GWAS – with which genetic correlations were also observed – is a meta-analysis of maternal and paternal AD GWAS where parental age-at-diagnosis/age-of-death is controlled for in the GWAS model, instead of being used for weighting the proxy phenotype prior analysis. Taking this into account, and considering that depression is itself associated with all-cause mortality\textsuperscript{130}, it is possible that including age-at-diagnosis/age-of-death in proxy-AD phenotyping induces a form of bias in later cross-trait analyses when the other trait is itself associated with longevity. Further investigation of this issue is required.

Nonetheless, conflicting results such as these pose a problem to researchers seeking to identify genetic relationships between AD and its risk factors. Large differences in the presence or direction of effects depending on which AD GWAS is used to assess associations increases the difficulty in discerning true associations for the purpose of designing interventions. As such, we suggest that future genetic studies examining genetic overlap between AD risk factors and AD conduct primary analyses using a clinically ascertained AD phenotype. A range of AD GWAS with different proxy/clinical ascertainment could then be used to examine the consistency of results in relation to clinical AD. While this approach would limit researchers to AD GWAS with slightly smaller sample sizes for primary analyses, it would also ensure that results are driven not by AD proxy phenotyping alone.

This study has several limitations. Despite being the largest meta-analysis of PHQ-9 items to date, the ability to detect genome-wide significant variants was likely limited by small sample sizes relative to other psychiatric conditions. This includes depression itself, where sample sizes have exceeded one million\textsuperscript{13}. Larger samples are required for future studies if we are to truly elucidate the unique genetic architecture of individual depression symptoms. Additionally, our analyses were restricted to individuals of European ancestry. GWAS results may therefore have poor transferability to other ancestry groups. Although efforts are underway to address the issue of lack of diversity in human genetic studies, urgency is required so that studies can be conducted across ancestry groups using equivalent sample sizes. Our study makes also heavy use of data from the UKB. The UKB is known to be affected by healthy volunteer bias, and as a consequence is not fully representative of the wider population\textsuperscript{131}.

It is worth noting that this study focused on depression as a risk factor for AD. However, there is evidence that some forms of later life depression are in fact a prodromal phase of dementia onset\textsuperscript{132}, and may be related to levels of dementia biomarkers such as plasma Aβ\textsubscript{42} and CSF p-tau\textsuperscript{133}. As such, dementia-related depression may be biologically distinct from depression as a mental health disorder.
While sufficient sample sizes are likely difficult to obtain, future large scale genomic 
studies of dementia-related depression – as has been undertaken with psychosis in 
AD\textsuperscript{134} – would prove illuminating on this matter.

In conclusion, this study describes the largest genome-wide meta-analysis of 
PHQ-9 depression symptom items to date (GWAS equivalent \textit{N} range: 224,535 – 308,421), identifying 37 unique genomic risk loci. Genetic correlations between 
depression/depression symptoms and AD were primarily observed when the AD 
GWAS contained clinical+proxy or proxy-only AD case/control ascertainment. 
Further, this study does not support depression or its symptoms as being causal for 
AD. However, colocalisation in the \textit{TMEM106B} region between four depression 
phenotypes and AD across both proxy and clinical AD GWAS suggests future 
research into the shared biological mechanisms underlying the role of this locus in 
depression and AD are warranted.

5. Funding and Acknowledgements

LG is funded by the King’s College London DRIVE-Health Centre for Doctoral Training 
and the Perron Institute for Neurological and Translational. PP is funded by 
Alzheimer’s Research UK. SK is funded by MSWA and the Perron Institute. HLD 
acknowledges funding from the Economic and Social Research Council (ESRC). 
DMH is supported by a Sir Henry Wellcome Postdoctoral Fellowship (Reference 
213674/Z/18/Z).

This paper represents independent research part-funded by the NIHR Maudsley 
Biomedical Research Centre at South London and Maudsley NHS Foundation Trust 
and King’s College London. The views expressed are those of the author(s) and not 
necessarily those of the NIHR or the Department of Health and Social Care.

This research has been conducted using the UK Biobank Resource under 
Application Number 18177. We thank the UK Biobank Team for collecting the data 
and making it available. We also thank the UK Biobank participants.

We thank the GLAD Study volunteers for their participation, and gratefully 
acknowledge the NIHR BioResource centres, NHS Trusts and staff for their 
contribution. We thank the National Institute for Health Research, NHS Blood and 
Transplant, and Health Data Research UK as part of the Digital Innovation Hub 
Programme. This study presents independent research funded by the NIHR 
Biomedical Research Centre at South London and Maudsley NHS Foundation Trust 
and King’s College London. Further information can be found at
http://brc.slam.nhs.uk/about/core-facilities/bioresource. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the HSC R&D Division, King’s College London, or the Department of Health and Social Care.

The PROTECT study was funded/supported by the National Institute of Health and Care Research Exeter Biomedical Research Centre. PROTECT genetic data was funded in part by the University of Exeter through the MRC Proximity to Discovery: Industry Engagement Fund (External Collaboration, Innovation and Entrepreneurism: Translational Medicine in Exeter 2 (EXCITEME2) ref. MC_PC_17189). Genotyping was performed at deCODE Genetics. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Dept of Health and Social Care.

As data used in this study was obtained from the ADNI database, the ADNI investigators contributed to the conception of the sample and acquisition of this data, although were not participants in other parts of this study, such as conceptualisation, data analysis or writing. A full acknowledgment list of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Similarly, data was obtained from the GERAD1 Consortium and GERAD1 investigators contributed to the conception of the cohort and acquisition of this data, but did not participate in the conceptualisation, analysis or writing of this study. A full list of GERAD1 collaborators can be found in Supplementary Material 1.

We also thank and acknowledge the contribution and use of the CREATE high-performance computing cluster at King’s College London (King’s College London. (2022). King’s Computational Research, Engineering and Technology Environment (CREATE). Retrieved May 23, 2023, from https://doi.org/10.18742/rnvf-m076).

The analysis flowchart in Figure 1. was created and licenced with BioRender (www.biorender.com)

6. Conflicts of Interest

CML is on the Scientific Advisory Board of Myriad Neuroscience and has received honoraria for consultancy from UCB.

7. Data availability

All GWAS summary statistics generated in the process of conducting this study will be deposited in the NHGRI-EBI Catalog of human genome-wide association studies.
(GWAS Catalog) following peer-review and publication. This study was pre-registered on the Open Science Framework (https://osf.io/94q35/?view_only=e77f72d4100d47eea7f3ef07dfa9c059).

This study made use of publicly available analysis software:

- MungeSumstats (https://neurogenomics.github.io/MungeSumstats/articles/MungeSumstats.html)
- REGENIE (https://rgcgithub.github.io/regenie/)
- MTAG (https://github.com/JonJala/mtag)
- METAL (https://genome.sph.umich.edu/wiki/METAL_Documentation)
- FUMA GWAS (https://fuma.ctglab.nl)
- LDSC (https://github.com/bulik/ldsc)
- HDL (https://github.com/zhenin/HDL)
- LAVA (https://github.com/josefin-werme/LAVA)
- coloc (https://chr1swallace.github.io/coloc/)
- COLOC-reporter (https://github.com/ThomasPSpargo/COLOC-reporter)
- TwoSampleMR (https://mrcieu.github.io/TwoSampleMR/)
- MegaPRS (https://dougspeed.com/megaprs/)

This study was pre-registered on the Open Science Framework (https://osf.io/94q35/?view_only=e77f72d4100d47eea7f3ef07dfa9c059).
References


with the Risk of Dementia in the Cognitively Normal Elderly Individuals: A Prospective Cohort Study. Psychiatry Investig 16, 575–580.


1160 63. Fromer, M., Roussos, P., Sieberts, S.K., Johnson, J.S., Kavanagh, D.H., Perumal, T.M.,
1163 64. Ramasamy, A., Trabzuni, D., Guelfi, S., Varghese, V., Smith, C., Walker, R., De, T., UK
1165 (2014). Genetic variability in the regulation of gene expression in ten regions of the human
1168 Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic
1170 45, 1238–1243.
1171 66. Zhernakova, D.V., Deelen, P., Vermaat, M., van Iterson, M., van Galen, M., Arindarto,
1176 complex traits using blood eQTL meta-analysis. BioRxiv.
1180 47, 88–91.
1181 69. Ng, B., White, C.C., Klein, H.-U., Sieberts, S.K., McCabe, C., Patrick, E., Xu, J., Yu, L.,
1182 Gaiteri, C., Bennett, D.A., et al. (2017). An xQTL map integrates the genetic architecture of
1186 581.
1191 73. Spargo, T.P., Gilchrist, L., Hunt, G.P., Dobson, R.J., Proitsi, P.P., Al-Chalabi, A., Pain,
1192 O., and Iacoangeli, A. (2023). Statistical examination of shared loci in neuropsychiatric
1195 summary data with the “Sum of Single Effects” model. PLoS Genet. 18, e1010299.
1198 76. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace,
1199 C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic
1204 78. Burgess, S., Davies, N.M., and Thompson, S.G. (2016). Bias due to participant overlap in
1207 randomization accounting for correlated and uncorrelated pleiotropic effects using genome-
1260 98. Lord, J., Green, R., Choi, S.W., Hüb, C., Aarsland, D., Velayudhan, L., Sham, P.,
1261 Legido-Quigley, C., Richards, M., Dobson, R., et al. (2022). Disentangling independent and
1262 mediated causal relationships between blood metabolites, cognitive factors, and alzheimer’s
1267 100. Voskobiynyk, Y., Roth, J.R., Cochran, J.N., Rush, T., Carullo, N.V., Mesina, J.S.,
1270 101. Lambert, E., Saha, O., Soares Landeira, B., Melo de Farias, A.R., Hermant, X., Carrier,
1272 susceptibility gene BIN1 induces isoform-dependent neurotoxicity through early endosome
1274 102. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P.,
1275 Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of
1276 genome-wide association scan results. Bioinformatics 26, 2336–2337.
1277 103. Dashti, H.S., Daghlas, I., Lane, J.M., Huang, Y., Udler, M.S., Wang, H., Ollila, H.M.,
1280 104. Jansen, P.R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A.R., de
1282 wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional
1284 105. Hammerschlag, A.R., Stringer, S., de Leeuw, C.A., Sniekers, S., Taskesen, E.,
1286 Genome-wide association analysis of insomnia complaints identifies risk genes and genetic
1288 106. Loos, R.J.F., and Yeo, G.S.H. (2022). The genetics of obesity: from discovery to
1292 127, 630–637.
1295 psychopathological features of patients with Anorexia Nervosa. Gene 783, 145572.
1296 109. Li, Y., Dang, X., Chen, R., Wang, J., Li, S., Mitchell, B.L., Yao, Y.-G., Li, M., Li, T.,
1297 Zhang, Z., et al. (2023). Cross-ancestry genome-wide association study and systems-level
1298 integrative analyses implicate new risk genes and therapeutic targets for depression.
1299 MedRxiv.
1303 Transcriptome-wide association study of treatment-resistant depression and depression
1305 112. Koskinen, M.-K., and Hovatta, I. (2023). Genetic insights into the neurobiology of
1307 113. Nagel, M., Jansen, P.R., Stringer, S., Watanabe, K., de Leeuw, C.A., Bryois, J., Savage,


