

1 Investigating the genetic relationship between 2 depression symptoms and Alzheimer's Disease in 3 clinically diagnosed and proxy cases

4
5 Lachlan Gilchrist^{a,b,c}, Thomas P. Spargo^{a,d,e}, Rebecca E. Green^f, Jonathan R.I. Coleman^{b,e},
6 David M. Howard^b, Jackson G. Thorp^g, Brett Adey^b, Jodie Lord^a, Helena L. Davies^b, Jessica
7 Mundy^b, Abigail ter Kuile^{b,e}, Molly R. Davies^{b,e}, Christopher Hübel^{b,h,i}, Shannon Bristow^{b,e},
8 Sang Hyuck Lee^{b,e}, Henry Rogers^{b,e}, Charles Curtis^{b,e}, Gursharan Kalsi^{b,e}, Ryan Arathimos^b,
9 Anne Corbettⁱ, Clive Ballard^j, Helen Brookerⁱ, Byron Creese^j, Dag Aarsland^k, Adam
10 Hampshire^l, Latha Velayudhan^k, Thalia C. Eley^{b,e}, Gerome Breen^{b,e}, Alfredo Iacoangeli^{d,a,e},
11 Sulev Koks^{c,m}, Cathryn M. Lewis^{b,n}, Petroula Proitsi^{a*} on behalf of the Alzheimer's Disease
12 Neuroimaging Initiative, AddNeuroMed, and GERAD1 Consortium

13
14 ^a Department of Basic and Clinical Neuroscience, King's College London, London, United Kingdom

15 ^b Social, Genetic and Developmental Psychiatry, King's College London, London, United Kingdom

16 ^c Perron Institute for Neurological and Translational Science, Perth, Australia

17 ^d Department of Biostatistics and Health Informatics, King's College London, London, United Kingdom

18 ^e NIHR Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust
19 and King's College London, King's College London, London, United Kingdom

20 ^f MRC Unit for Lifelong Health and Aging, University College London, London, United Kingdom

21 ^g Translational Neurogenomics Laboratory, QIMR Berghofer Medical Research Institute, Brisbane,
22 Australia

23 ^h National Centre for Register-based Research, Aarhus BSS Business and Social Sciences, Aarhus
24 University, Aarhus, Denmark

25 ⁱ Department of Pediatric Neurology, Charité – Universitätsmedizin Berlin, Berlin, Germany

26 ^j College of Medicine and Health, University of Exeter, Exeter, United Kingdom

27 ^k Old Age Psychiatry, King's College London, London, United Kingdom

28 ^l Faculty of Medicine, Department of Brain Sciences, Imperial College London, London, United Kingdom

29 ^m Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, Australia

30 ⁿ Department of Medical and Molecular Genetics, King's College London, London, United Kingdom

31

32

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

34

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

Abstract

36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67

Depression is a risk factor for 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 assessed global/local genetic correlations and statistical colocalisation between depression phenotypes 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 the predictiveness of depression phenotype 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 significant genetic correlation was identified with AD GWAS containing clinical-only cases. Colocalisation was not identified at loci contains local genetic correlation 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 statistically significant results.

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

68 1. INTRODUCTION

69 Epidemiological studies suggest that a diagnosis of depression is a risk factor for the
70 later development of dementia¹⁻⁶, of which Alzheimer's disease (AD) is the most
71 common form, accounting for approximately 80% of the over 40 million global cases⁷.
72 Establishing the underlying mechanisms by which depression confers increased risk
73 for AD offers a pathway by which new interventions might be implemented and the
74 global dementia burden reduced⁸.

75 As twin studies have demonstrated, both depression and AD are substantially
76 heritable – approximately 40% and 80% respectively^{9,10}. Further, large-scale genome-
77 wide association studies (GWAS) have demonstrated high polygenicity, identifying
78 over 70 genomic risk loci for AD and nearly 200 for depression¹¹⁻¹⁷. It is therefore
79 possible that their phenotypic association is in part due to shared genetic
80 architecture. However, results from previous investigations into the genetic overlap
81 between the two disorders have been mixed. For example, some findings indicate
82 non-significant genetic overlap^{18,19}, others a significant – if modest – genetic
83 correlation of ~16-17% and a risk increasing causal effect of depression on AD²⁰⁻²².

84 According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-
85 5), diagnosis of a major depressive episode requires the presence of at least five of
86 a possible nine symptoms for ≥ 2 weeks, including one of the two cardinal symptoms
87 – depressed mood or anhedonia²³. Potentially hundreds of symptom combinations
88 are possible to meet this diagnosis criteria²⁴⁻²⁶. As such, this heterogeneity poses
89 challenges to researchers seeking to better understand differences in the genetic
90 contribution to depression and its subtypes^{27,28}. However, the decomposition of
91 depression into individual symptoms has provided insight into unique patterns of
92 genome-wide significant loci and cross-trait genetic associations, as demonstrated
93 in a recent GWAS of depression symptoms on the Patient Health Questionnaire (PHQ-
94 9) by Thorp et al.²⁹.

95 A number of studies suggest that anhedonia may be a better predictor of
96 dementia than depressed mood³⁰⁻³². Further, several depression symptoms, such as
97 appetite changes, psychomotor dysfunction, and sleep disruption are commonly
98 observed in non-depressed dementia patients³³⁻³⁵. Taking this into account alongside
99 the mixed nature of previous findings examining the genetic overlap between
100 depression and AD, it is possible that leveraging depression symptom level genetic
101 information may offer greater insight into the disorders shared genetic architecture.

102 However, any association between depression and AD must also consider the
103 potential influence of differences in case/control ascertainment in AD GWAS. A review
104 by Escott-Price et al.³⁶ notes that recent large-scale AD GWAS contain a relatively
105 small proportion of clinically ascertained cases/controls, with a large percentage of

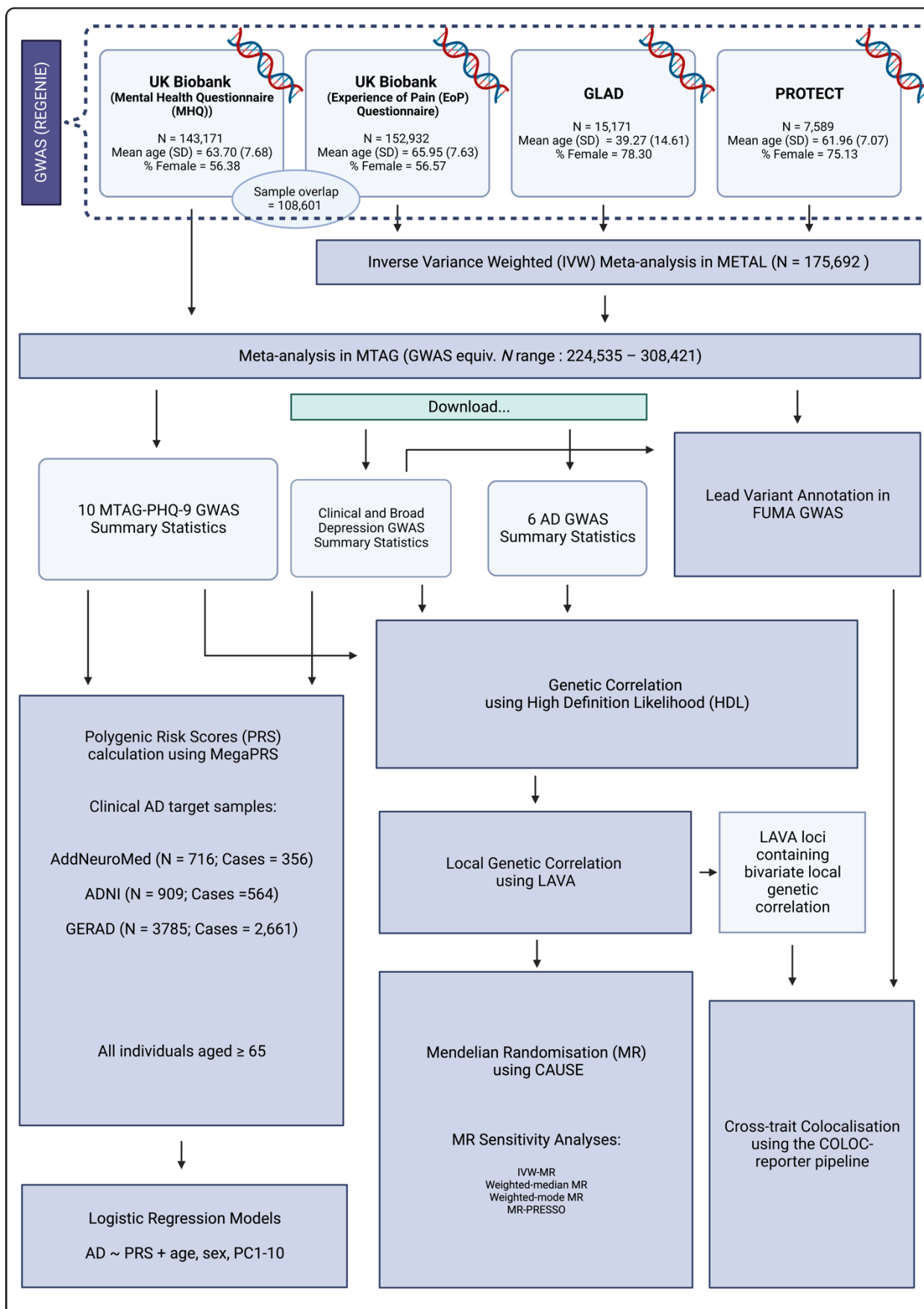
106 cases ascertained by proxy – that is, cases and controls are defined as individuals
107 with and without a self-reported parental history of AD/dementia, respectively. The
108 combination of clinical and proxy samples in AD GWAS meta-analyses has proved
109 an effective way of boosting sample size and variant discovery^{15–17,37}. However,
110 evidence suggests that this has come at the expense of specificity in regard to
111 genomic risk loci and an apparent stagnation in the percentage of variance explained
112 by common variants³⁶. Most importantly for cross-trait analysis, recent studies
113 indicate that the direction of Mendelian Randomisation (MR) causal estimates for AD
114 risk factors on AD can be in the opposite direction depending on whether the AD
115 outcome GWAS contains clinical and proxy cases/controls or is more strictly clinically
116 ascertained^{38–40}.

117 To address these points, we perform the first and largest genome-wide meta-
118 analyses of PHQ-9 depression symptom items using data from the Genetic Links to
119 Anxiety and Depression (GLAD) Study⁴¹, the PROTECT Study^{42,43} and two
120 questionnaires from UK Biobank (UKB)⁴⁴. We obtained summary statistics from
121 previous large-scale GWAS for clinical¹¹ and broad¹² depression, and six AD GWAS.
122 Specifically, three with clinical+proxy case/control ascertainment^{15–17}, one with proxy-
123 only³⁷, and two with clinical-only^{14,45}. We used these GWAS to assess the presence,
124 strength and differences in genetic overlap between depression, depression
125 symptoms and AD, with the additional aim of better understanding the influence of
126 different AD case ascertainment strategies on associations.

127
128
129
130
131
132
133
134
135
136
137
138
139
140
141

142 **2. METHODS**

143 For an analysis flowchart, see Figure 1.



144

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

153

154

155 **2.1 GWAS in the UK Biobank, GLAD and PROTECT**

156 **2.1.1 Patient Health Questionnaire-9 (PHQ-9) phenotypes**

157 The PHQ-9 is a well-validated clinical screening questionnaire used to assess
158 depression symptom severity on nine individual symptoms in the DSM-IV^{46,47}. The
159 severity of each symptom is measured by the self-reported persistence of that
160 symptom over the preceding two weeks on a scale of 0 to 3. Scores of 3 indicate an
161 individual experienced that symptom nearly every day, 2 indicates an individual
162 experienced that symptom more than half the days, 1 indicates an individual
163 experienced that symptom for several days and 0 indicates no experience of that
164 symptom at all. The sum of an individual's scores over all nine items (sum-score)
165 ranges from 0-27. For an overview of the PHQ-9 items and response distribution for
166 each sample, see **Table 1**. Sum-score distributions can be seen in **Supplementary**
167 **Table 1**.

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183
184
185
186

Depression Symptom	PHQ-9 Question:	UK Biobank Data Field (MHQ/EoP)	Cohort	PHQ-9 Response			
	How often in the last 2 weeks have you been bothered by any of the following problems?			0 (Not at all)	1 (Several days)	2 (more than half the days)	3 (Nearly every day)
Anhedonia	Little interest or pleasure in doing things	20514/120104	UK Biobank (MHQ)	117045 (81.75%)	21088 (14.73%)	2965 (2.07%)	2073 (1.45%)
			UK Biobank (EoP)	116330 (76.07%)	29743 (19.45%)	4115 (2.69%)	2744 (1.79%)
			GLAD	4108 (27.08%)	6429 (42.31%)	2502 (16.49%)	2142 (14.12%)
			PROTECT	6194 (81.62%)	1130 (14.89%)	127 (1.67%)	138 (1.82%)
			Total	243677 (76.42%)	58380 (18.31%)	9709 (3.04%)	7097 (2.23%)
Appetite Changes	Poor appetite or overeating	20511/120108	UK Biobank (MHQ)	117505 (82.07%)	18521 (12.94%)	3860 (2.70%)	3285 (2.29%)
			UK Biobank (EoP)	121071 (79.17%)	22432 (14.67%)	5404 (3.53%)	4025 (2.63%)
			GLAD	4242 (27.96%)	4459 (29.39%)	3057 (20.15%)	3413 (22.50%)
			PROTECT	6364 (83.86%)	971 (12.79%)	145 (1.91%)	109 (1.44%)
			Total	249182 (78.15%)	46383 (14.55%)	12466 (3.90%)	10832 (3.40%)
Concentration Problems	Trouble concentrating on things	20508/120110	UK Biobank (MHQ)	117951 (82.07%)	20386 (14.24%)	2713 (1.89%)	2121 (1.48%)
			UK Biobank (EoP)	123920 (81.03%)	23441 (15.33%)	3258 (2.13%)	2313 (1.51%)
			GLAD	4665 (30.75%)	5200 (34.28%)	2891 (19.06%)	2415 (15.92%)
			PROTECT	6374 (83.99%)	1022 (13.47%)	116 (1.53%)	77 (1.01%)
			Total	252910 (79.32%)	50049 (15.70%)	8978 (2.82%)	6926 (2.17%)
Depressed Mood	Feeling down, depressed, or hopeless	20510/120105	UK Biobank (MHQ)	112187 (78.36%)	26516 (18.52%)	2690 (1.88%)	1778 (1.24%)
			UK Biobank (EoP)	119004 (77.81%)	28914 (18.91%)	2998 (1.96%)	2016 (1.32%)
			GLAD	3595 (23.70%)	6746 (44.47%)	2538 (16.73%)	2292 (15.11%)
			PROTECT	6108 (80.48%)	1329 (17.51%)	92 (1.21%)	60 (0.79%)
			Total	240894 (75.55%)	63505 (19.92%)	8318 (2.61%)	6146 (1.93%)
Fatigue	Feeling tired or having little energy	20519/120107	UK Biobank (MHQ)	71903 (50.39%)	56170 (39.23%)	7595 (5.30%)	7503 (5.24%)
			UK Biobank (EoP)	74733 (48.87%)	61571 (40.26%)	8820 (5.77%)	7808 (5.11%)
			GLAD	1335 (8.80%)	4971 (32.77%)	3530 (23.27%)	5335 (35.17%)
			PROTECT	3715 (48.95%)	3164 (41.69%)	359 (4.73%)	351 (4.63%)
			Total	151686 (47.57%)	125876 (39.48%)	20304 (6.37%)	20997 (6.58%)
Feelings of Inadequacy	Feeling bad about yourself – or that you are a failure or have let yourself or your family down	20507/120109	UK Biobank (MHQ)	116091 (81.09%)	21503 (15.02%)	2878 (2.01%)	2699 (1.89%)
			UK Biobank (EoP)	123702 (80.89%)	22756 (14.88%)	3399 (2.22%)	3075 (2.01%)
			GLAD	3986 (26.27%)	5131 (33.82%)	2800 (18.46%)	3254 (21.45%)
			PROTECT	6095 (80.31%)	1260 (16.60%)	128 (1.69%)	106 (1.40%)
			Total	249874 (78.36%)	50650 (15.88%)	9205 (2.89%)	9134 (2.86%)
Psychomotor Changes	Moving or speaking so slowly that other people could have noticed, or the opposite – being so fidgety or restless that you have been moving around a lot more than usual	20518/120111	UK Biobank (MHQ)	135592 (94.71%)	5805 (4.05%)	1025 (0.72%)	749 (0.52%)
			UK Biobank (EoP)	141662 (92.63%)	8285 (5.42%)	1676 (1.10%)	1309 (0.86%)
			GLAD	9932 (65.47%)	3254 (21.45%)	1215 (8.01%)	770 (5.08%)
			PROTECT	7363 (97.02%)	182 (2.40%)	22 (0.29%)	22 (0.29%)
			Total	294549 (92.37%)	17526 (5.5%)	3938 (1.24%)	2850 (0.89%)
Sleep Problems	Trouble falling asleep or staying asleep, or sleeping too much	20517/120106	UK Biobank (MHQ)	73219 (51.14%)	48936 (34.18%)	10007 (6.99%)	11009 (7.69%)
			UK Biobank (EoP)	70799 (46.29%)	56955 (37.24%)	12245 (8.01%)	12933 (8.46%)
			GLAD	2577 (16.99%)	5043 (33.24%)	3173 (20.91%)	4378 (28.86%)
			PROTECT	3708 (48.86%)	2749 (36.22%)	570 (7.51%)	562 (7.41%)
			Total	150303 (47.14%)	113683 (35.65%)	25995 (8.15%)	28882 (9.06%)
Suicidal Thoughts	Thoughts you would be better off dead or of hurting yourself in some way	20513/120112	UK Biobank (MHQ)	137381 (95.96%)	4801 (3.35%)	566 (0.40%)	423 (0.30%)
			UK Biobank (EoP)	146843 (96.02%)	4880 (3.19%)	686 (0.45%)	523 (0.34%)
			GLAD	9116 (60.09%)	3713 (24.47%)	1215 (8.01%)	1127 (7.43%)
			PROTECT	7392 (97.40%)	172 (2.27%)	13 (0.17%)	12 (0.16%)
			Total	300732 (94.31%)	13566 (4.25%)	2480 (0.78%)	2085 (0.65%)

187

188 **Table 1:** Patient Health Questionnaire (PHQ-9) derived depression symptom phenotypes, their
189 corresponding question on the PHQ-9, and the distribution and percentage of responses for each of
190 the four samples on which genome-wide association analysis was performed.

191

192

193 **2.1.2 Study population**

194 In each GWAS sample, individuals were only retained if they had reported European
195 ancestry and had provided a valid response to all PHQ-9 items. Individuals were
196 excluded if they had reported a previous professional diagnosis of schizophrenia,
197 psychosis, mania, hypomania, bipolar, or manic depression (UKB Field ID: 20544) or
198 a previous prescription of medication for a psychotic experience (UKB Field ID:
199 20466).

200 **2.1.3 GWAS software**

201 All GWAS were conducted using REGENIE v3.1.3⁴⁸. In step one of REGENIE, ridge
202 regression is applied to a subset of quality controlled (QC'd) variants to fit, combine
203 and decompose a set of leave-one-chromosome-out (LOCO) predictions. QC for step
204 one was undertaken using PLINK v1.9⁴⁹. In step two, imputed variants are tested for
205 association with the phenotype. LOCO predictions from step one are included as
206 covariates to control for proximal contamination. For all GWAS, genotyping batch,
207 sex, age, and age-squared were included as covariates, as were the maximum
208 available genetic principal components (PCs) for GLAD (10 PCs) and PROTECT (20
209 PCs) to control for population stratification. For the UK Biobank analyses, 16 PCs were
210 included as recommended by Privé et al.⁵⁰. Assessment centre was also included as
211 a covariate for UK Biobank analyses.

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

218 **2.1.4 GWAS: UK Biobank (UKB)**

219 UKB is a large-scale biomedical database and research resource consisting of
220 ~500,000 individuals with data across a broad range of phenotypes, including mental
221 health outcomes⁴⁴. Individuals in UKB have been genotyped on the custom UK
222 Biobank Axiom or UKBiLEVE arrays, with imputed data available for ~90 million
223 variants imputed with IMPUTE2 using the Haplotype Reference Consortium (HRC)⁵¹
224 and combined UK10K + 1000 Genomes Phase 3 reference panels⁵².

225 UKB participants have completed the PHQ-9 in two online surveys. In total,
226 157,345 individuals provided responses as part of the Mental Health Questionnaire
227 (UKB-MHQ) (Category: 136) between 2016 and 2017, and 167,199 individuals had
228 provided responses as part of the Experience of Pain Questionnaire (UKB-EoP)
229 (Category: 154) between 2019 and 2020.

230 After filtering for self-reported European ancestry, valid PHQ-9 responses, and
231 previous diagnosis/prescription exclusions, 144,630 (UKB-MHQ) and 155,027 (UKB-
232 EoP) individuals remained prior to genetic QC for REGENIE. In step one, single-
233 nucleotide polymorphisms (SNPs) with a call rate > 98%, minor allele frequency
234 (MAF) > 1%, and Hardy-Weinberg equilibrium test $p > 1 \times 10^{-8}$ were retained, as were
235 individuals with variant missingness < 2%, no unusual levels of heterozygosity, and
236 not mismatched on sex. Individuals were retained if they were determined to be of
237 European ancestry based on 4-means clustering on the first two principal
238 components.

239 For the final GWAS analyses, 143,171 (mean age [SD] = 63.70 [7.68]; %female
240 = 56.38%) and 152,932 (mean age [SD] = 65.95 [7.63]; %female = 56.57%)
241 individuals proceeded from the MHQ and EoP questionnaires respectively. Of these,
242 108,601 individuals had provided responses on both questionnaires. In step two, a
243 total of 9,746,698 imputed variants were retained with MAF ≥ 0.01 and imputation
244 INFO score ≥ 0.7 .

245 **2.1.5 GWAS: Genetic Links to Anxiety and Depression (GLAD) Study**

246 The GLAD study has the specific goal of recruiting a large cohort of re-contactable
247 individuals with anxiety or depression into the National Institute for Health and Care
248 Research (NIHR) Mental Health BioResource with genetic, environmental and
249 phenotypic data collected⁴¹. Genotyping for GLAD was conducted using the UK
250 Biobank v2 Axiom array and imputed using the TopMed imputation pipeline⁵³.

251 After filtering for self-reported European ancestry, valid PHQ-9 responses, and
252 previous diagnosis/prescription exclusions, 15,472 individuals remained prior to
253 genetic QC for REGENIE step one. Genotype data was provided by the study team
254 and had been filtered to retain SNPs with a genotype call rate > 95%, MAF > 1%,
255 Hardy-Weinberg equilibrium test $p > 1 \times 10^{-10}$, and individuals with genotype
256 missingness < 5%. Individuals were additionally excluded if they had unusual levels
257 of heterozygosity, mismatched on sex and of non-European ancestry based on 4-
258 means clustering. A total of 15,171 individuals (mean age [SD] = 39.27 [14.61];
259 %female = 78.30%) were retained for the final analysis. In step two, a total of
260 13,979,187 imputed variants with MAF ≥ 0.001 and INFO ≥ 0.7 were analysed.

261 **2.1.6 GWAS: PROTECT Study**

262 PROTECT is an online registry of ~25,000 UK-based individuals that aims to track
263 cognitive health in older adults. Individuals were only considered eligible for inclusion
264 in PROTECT if they were older than 50, had no previous dementia diagnosis and had
265 internet access. Genetic data are available alongside phenotypic data for ~10,000 of
266 the participants. These individuals were genotyped on the Illumina Infinium Global

267 Screening Array and imputed on the 1000 Genomes reference panel⁵⁴ using the
268 Michigan imputation server and genotype phasing using Eagle.

269 After filtering for self-reported European ancestry, valid PHQ-9 responses, and
270 previous diagnosis/prescription exclusions, 7,589 individuals remained for genetic
271 QC for step one of REGENIE. Genetic data in PROTECT had been previously QC'd
272 prior to imputation to only retain individuals and variants with a call rate > 98%, Hardy-
273 Weinberg equilibrium test $p > 0.00001$ and excluding unusual heterozygosity⁴².
274 Variants used in step one were down-sampled from the imputed data using a snplist
275 from the Illumina Infinium Global Screening Array provided by the PROTECT
276 investigators. Variants were retained if they had MAF > 1%. After mismatched sex
277 and 4-means clustering ancestry exclusions, a total of 7,589 individuals (mean age
278 [SD] = 61.96 [7.07]; %female = 75.13%) proceeded to step two. In step two,
279 9,388,534 imputed variants with MAF ≥ 0.001 and imputation INFO score ≥ 0.7 were
280 analysed.

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301 **2.2 GWAS summary statistics**

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

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

304

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

309

310 **2.2.1 Clinical and broad depression**

311 To examine potential differences in genetic overlap with AD between depression as
 312 a disorder compared to individual depression symptoms, summary statistics for two
 313 previously conducted GWAS of clinical and broad depression were obtained from
 314 the Psychiatric Genomics Consortium (PGC) (<https://pgc.unc.edu/researchers/download-results/>). For clinical depression, we used a subsample of the
 315 Major Depressive Disorder (MDD) GWAS by Wray et al.¹¹ that excluded samples from
 316 the UKB and 23andMe, and contained only individuals for whom case ascertainment
 317 was defined through structured diagnostic interview or electronic health records. For
 318 the broad definition depression GWAS, we used a subsample of the depression
 319 GWAS by Howard et al.¹², which also excluded samples from 23andMe. In addition
 320

321 to clinical cases and controls used by Wray et al.¹¹, this broad depression GWAS
322 included individuals in the UKB for whom case-control ascertainment was based on
323 self-reported responses to the questions “Have you ever seen a general practitioner
324 for nerves, anxiety tension or depression?” and “Have you ever seen a psychiatrist
325 for nerves anxiety, tension or depression?”.

326 **2.2.2 Alzheimer’s disease**

327 Summary statistics were obtained from six previously conducted AD GWAS: three
328 with proxy+clinical, one with proxy-only and two with clinical-only case
329 ascertainment. All three of the proxy+clinical AD GWAS – Bellenguez et al.¹⁷,
330 Wightman et al.¹⁶ and Jansen et al.¹⁵ – and the proxy-only AD GWAS – Marioni et al.³⁷
331 – used data from the UKB for proxy AD samples.

332 There are some key differences in the way these AD GWAS define proxy cases
333 and controls. Bellenguez et al.¹⁷ define proxy cases/controls as a binary phenotype,
334 whereby individuals reporting a parent with AD or dementia are considered cases
335 and those reporting no parental history are considered controls. Wightman et al.¹⁶
336 and Jansen et al.¹⁵ instead define proxy cases/control as a continuous phenotype,
337 summing the number of parents an individual has reported with dementia and down-
338 weighting unaffected parents by their age (or age of death).

339 For the proxy-only Marioni et al.³⁷ GWAS, summary statistics were obtained
340 from a meta-analysis of paternal and maternal AD. Here, proxy phenotyping was
341 based on the self-report of either maternal or paternal AD, including the parent’s age
342 at time of reporting/age of death as a covariate.

343 Summary statistics from a clinical-only subsample of the GWAS by Wightman
344 et al. that excluded proxy cases/controls from the UKB⁴⁵ were obtained from the
345 authors. Summary statistics for a final clinical-only AD GWAS were obtained from
346 Stage 1 of the GWAS by Kunkle et al.¹⁴.

347 **2.3 Summary statistic standardisation**

348 Summary statistics from all 40 depression symptom GWAS, the two depression
349 GWAS and the six AD GWAS, were standardised using the MungeSumstats⁵⁵ in R
350 version 4.2.1. Using dbSNP 141 and the BSgenome.Hsapiens.1000genomes.hs37d5
351 reference genome, missing rsIDs were corrected, duplicates and multi-allelic variants
352 removed, effect alleles and the direction of their effects aligned to the reference
353 genome, and variants filtered at INFO score ≥ 0.7 and MAF ≥ 0.01 . The GLAD Study
354 and Bellenguez et al.¹⁷ summary statistics were lifted over from GRCh38 to GRCh37.

355 **2.4 SNP heritability**

356 SNP heritability (h^2_{SNP}) estimates were calculated for all GWAS used in this
357 study with Linkage Disequilibrium Score Regression (LDSC)^{56,57}. Briefly, LDSC
358 calculates h^2_{SNP} by regressing the effect sizes from GWAS summary statistics on their

359 LD score as computed in a reference panel – in this case HapMap3 variants
360 contained within the European sample of 1000 Genomes Phase 3. Liability scale h^2_{SNP}
361 was calculated naively from the standardised depression GWAS and AD GWAS using
362 a 15% and 5% population prevalence respectively^{11,16}. Heritability Z -scores were
363 calculated for all phenotypes by dividing the h^2_{SNP} estimates by their standard error.

364 **2.5 GWAS meta-analysis of depression symptoms**

365 To leverage the maximum genetic information available controlling for the
366 sample overlap between the UKB-MHQ and UKB-EoP samples, REGENIE output for
367 each PHQ-9 phenotype from the UKB-EoP, GLAD Study and PROTECT were first
368 subject to Inverse Variance Weighted (IVW) meta-analysis using METAL⁵⁸. All
369 available variants were included, for a total of 8,425,618 ($N = 175,692$). Multi-trait
370 Analysis of GWAS (MTAG)⁵⁹ v1.0.8 was then used to meta-analyse the METAL output
371 with the UKB-MHQ sample. While MTAG is commonly used for the joint genetic
372 analysis of multiple traits or multiple measurements of the same trait, by assuming the
373 heritability of included phenotypes are equal (*--equal-h2*) and their genetic correlation
374 is one (*--perfect-gencov*), MTAG performs an IVW meta-analysis of the same
375 measures of the same trait, accounting for sample overlap using the cross-trait
376 intercept from LDSC^{56,57}. Heritability estimates for all samples, plus the METAL meta-
377 analysis, are in **Supplementary Table 2**. Genetic correlations between the UKB-MHQ
378 and METAL GWAS are in **Supplementary Table 3**. For greater detail on the IVW
379 function of MTAG, see the online methods of the original MTAG paper⁵⁹. A total of
380 8,196,874 SNPs with MAF > 0.01 were present for MTAG analysis.

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

385

$$386 \quad N(\text{final}) = \frac{(N_1(\text{pre}) * N_1(\text{post})) + (N_2(\text{pre}) * N_2(\text{post}))}{N_1(\text{pre}) + N_2(\text{pre})}$$

387

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

391 **2.6 Genomic risk loci and gene annotation**

392 GWAS meta-analysis results were annotated using FUMA GWAS⁶⁰ v3.1.6a. Genome-
393 wide significance was set at $p \leq 5e-8$. Lead variants at genomic risk loci were defined
394 by clumping all variants correlated at $r^2 > 0.1$ 250kb either side. Clumping was
395 performed using the European sample of the 1000 Genomes Phase 3 reference

396 panel. Lead variants were mapped to genes within 10kb using positional mapping
397 and eQTLs from four brain (BrainSeq⁶¹, PsychENCODE⁶², CommonMind⁶³, and
398 BRAINEAC⁶⁴) and five blood (BloodeQTL⁶⁵, BIOS⁶⁶, eQTLGen cis and trans⁶⁷, Twins
399 UK⁶⁸ and xQTLServer⁶⁹) eQTL datasets, alongside all 54 tissue type eQTLs from GTEx
400 v8 (<https://gtexportal.org/home/tissueSummaryPage>).

401 **2.7 Genetic correlations**

402 Genetic correlation can be understood as the genome-wide correlation of genetic
403 effects between two phenotypes, and as such can be viewed as an estimate of
404 pleiotropy⁷⁰.

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

415 **2.8 Local genetic correlations**

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

420 LAVA⁷² was used to assess regions of local genetic correlation between each
421 depression phenotype and the six AD GWAS across 2495 semi-independent, pre-
422 defined LD-blocks of at least 2500 base pairs ([https://github.com/josefin-
423 werme/LAVA](https://github.com/josefin-werme/LAVA)). Loci-specific heritability estimates were calculated for each
424 phenotype for each block. If both a depression phenotype and an AD GWAS showed
425 significant local heritability at a specific locus (Bonferroni corrected p -value $\leq 2e-5$
426 (0.05/2495)), bivariate local genetic correlation was tested. Bivariate results were
427 considered significant at $p_{\text{FDR}} \leq 0.05$ correcting for the total number of bivariate tests.
428 Sample overlap was accounted for using an LDSC intercept matrix. Analysis was
429 restricted to the 5,531,969 non-strand-ambiguous variants shared across all GWAS
430 summary statistics and on the European ancestry 1000 Genomes Phase 3 reference
431 panel.

432
433

434 **2.9 Colocalisation**

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

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

440

441 H_0 : There is no causal variant in the region for either phenotype

442 H_1 : There is a causal variant in the region for the first phenotype

443 H_2 : There is a causal variant in the region for the second phenotype

444 H_3 : Distinct causal variants for each trait in the region

445 H_4 : Both traits share a causal variant.

446

447 Colocalisation analysis was conducted using the COLOC-reporter pipeline⁷³
448 (<https://github.com/ThomasPSpargo/COLOC-reporter>). COLOC-reporter extracts
449 variants in user-defined genomic regions and calculates the LD matrix for this region
450 from a user-defined reference panel. It then harmonises the summary statistics to
451 match the allele order of the reference panel, flipping effect directions accordingly.
452 Observed versus expected Z-scores are assessed using the diagnostic tools
453 provided in the susieR R package⁷⁴. Z-score outliers are omitted. Following this QC,
454 Sum of Single Effects (SuSiE) fine-mapping⁷⁴ is conducted to identify 95% credible
455 sets in these regions for each phenotype. Identification of credible sets in both
456 phenotype allows for the relaxation of the single causal variant assumption. All
457 possible credible sets are then assessed pairwise for shared signal between
458 phenotypes, improving resolution for colocalisation inference in regions containing
459 multiple signals⁷⁵. Should no 95% credible set be identified or only identified for one
460 phenotype, colocalisation under the single causal variant assumption is performed
461 using the *coloc.abf*⁷⁶. A posterior probability $\geq 80\%$ for H_4 was considered evidence
462 of colocalisation between two phenotypes ($PP.H_4 \geq 0.8$). The SuSiE model assumed
463 at most 10 causal variants ($L = 10$) per credible set. We used default priors.

464 Trait pairs with LAVA correlations significant at $p_{FDR} \leq 0.05$ were passed to
465 COLOC-reporter. Where the same depression phenotype showed nominally
466 significant local genetic correlation at the same locus but with different AD GWAS,
467 these phenotype pairs were included as a sensitivity analysis. Regions $\pm 250\text{kb}$ (r^2
468 ≥ 0.1) from lead variants at genome-wide significant loci from the MTAG-PHQ-9,
469 broad and clinical depression GWAS were also examined for evidence of
470 colocalisation with the 6 AD GWAS.

471

472 **2.10 Mendelian randomisation (MR)**

473 MR is a statistical method that uses genetic variants associated with an exposure as
474 instrumental variables to assess the causal effect of that exposure on an outcome of
475 interest⁷⁷. Two Sample MR framework estimates the causal relationships between an
476 exposure and outcome using GWAS summary statistics. Valid MR instruments are
477 defined by three key assumptions: (1) relevance – IVs are strongly associated with
478 the exposure of interest; (2) independence – there are no confounders in the
479 association between IVs and the outcome of interest; and (3) exclusion restriction –
480 instruments are not associated to the outcome other than via the exposure, for
481 example through horizontal pleiotropy⁷⁷.

482 Sample overlap is a known source of bias in Two Sample MR⁷⁸. Given the
483 likelihood of sample overlap between the depression phenotypes and the AD GWAS
484 containing proxy cases/controls due to participants from the UKB, MR analysis was
485 primarily conducted using Causal Analysis Using Summary Effect estimates
486 (CAUSE)⁷⁹ v1.2.0. CAUSE is a Bayesian MR method robust to sample overlap,
487 correlated and uncorrelated pleiotropy, also avoiding exclusion restriction
488 assumption violations⁸⁰.

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

496 Causal estimates were also calculated using the traditional IVW method. IVW
497 estimates are biased by the presence of horizontal pleiotropy⁸¹. Sensitivity analyses
498 were therefore conducted using MR-Egger, weighted-median and weighted-mode
499 MR. These methods allow for varying degrees of pleiotropy while providing unbiased
500 causal estimates⁸²⁻⁸⁴. MR-PRESSO⁸⁵ was also implemented. MR-PRESSO identifies
501 and excludes outlying instruments based on their contribution to heterogeneity and
502 provides a corrected causal estimate.

503 Pleiotropy was assessed using the MR-Egger intercept test⁸⁶ (significant
504 pleiotropy: $p \leq 0.05$). Heterogeneity tests were also conducted for IVW and MR-Egger
505 estimates using their respective Cochran's Q test⁸⁷ (significant heterogeneity: p -value
506 ≤ 0.05). Instrument strength was calculated via the minimum and mean F-statistic
507 (recommended F-statistic ≥ 10)⁸⁸ (β^2/SE^2). I^2 was calculated to ensure measurement
508 error was sufficiently low so as to ensure validity of results from MR-Egger
509 (recommended $I^2 \geq 0.90$)⁸⁹.

510 For IVW-MR, MR-Egger, weighted-median and weighted-mode MR,
511 instruments clumped at a $r^2 \geq 0.001$ and a p -value $\leq 5e-8$ within 10,000kb. Where no
512 instruments were available at $p \leq 5e-8$ or where < 5 instruments were available, a p
513 $\leq 5e-6$ threshold was used.

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

517 The *APOE* gene is known to be associated with non-AD phenotypes such as
518 cardiovascular disease⁹⁰ and type 2 diabetes (T2D)⁹¹. Both have been linked to
519 depression in previous MR analyses^{92,93}. As such, the inclusion of *APOE* would violate
520 MR's independence assumption. All MR analyses were therefore conducted
521 excluding variants in the APOE region (chr19:45,020,859–45,844,508 (GRCh37)) as
522 per Lord et al.⁹⁴.

523 **2.11 Polygenic risk scores (PRS)**

524 Polygenic risk scores (PRS) describe the sum of an individual's risk alleles, weighted
525 by their effect size⁹⁵. The ten PHQ-9-MTAG, clinical depression and broad definition
526 depression GWAS were processed for PRS using the BayesR-SS function contained
527 within MegaPRS and implemented in LDAK v5.2.1.⁹⁶ BayesR-SS assumes the BLD-
528 LDAK heritability model, which incorporates 65 genome annotations pertaining to
529 genomic features such as whether the variants are in coding regions or highly
530 conserved⁹⁶. Annotation files were obtained from the LDAK website
531 (<http://dougsspeed.com/blldlak/>). PRS calculation was restricted to the 1,217,311
532 HapMap3 SNPs, with strand ambiguous SNPs excluded.

533 PRS predictive utility for AD case/control status was assessed using logistic
534 regression in three clinically ascertained AD cohorts: AddNeuroMed and Dementia
535 Case Register Studies (ANM)⁹⁷ ($N_{cases} = 564$; $N_{controls} = 345$), the Alzheimer's Disease
536 Neuroimaging Initiative (ADNI) (<https://adni.loni.usc.edu>) ($N_{cases} = 356$; $N_{controls} = 360$),
537 and the Genetic and Environmental Risk in Alzheimer's Disease (GERAD1)
538 Consortium

539 (<https://portal.dementiasplatform.uk/CohortDirectory/Item?fingerPrintID=GERAD>)
540 ($N_{cases} = 2,661$; $N_{controls} = 1,124$) (Supplementary Table 4). All analyses controlled for
541 age, sex, and 10 PCs, and were restricted to individuals aged ≥ 65 to provide clean
542 controls. As a sensitivity analysis, PRS were also calculated in these cohorts
543 excluding the *APOE* region. Genetic QC and imputation steps for these cohorts can
544 be viewed in detail in the study by Lord et al.⁹⁸.

545
546
547

548 3. RESULTS

549 3.1 PHQ-9 genome-wide meta-analyses

550 The MTAG meta-analyses identified a total of 40 genomic risk loci between the ten
551 PHQ-9 phenotypes (GWAS equivalent N range: 224,535 – 308,421). Only one
552 depression symptom – suicidal thoughts – identified no genome-wide significant
553 variants. Three lead SNPs were shared with more than one PHQ-9 phenotype, leaving
554 a total of 37 unique genomic risk loci (**Table 3**). The significance of each of the lead
555 variants in each of the samples contributing to the meta-analysis can be seen in
556 **Supplementary Table 5**. eQTL mapping in FUMA mapped lead variants at genomic
557 risk loci to 76 genes (**Supplementary Table 6**). h^2_{SNP} for the MTAG-PHQ-9 GWAS
558 ranged from 1.12% for suicidal thoughts to 6.78% for the PHQ-9 sum-score. h^2_{SNP} Z -
559 scores were all > 4 (range: 6.59 – 18.50) (**Supplementary Table 7**), indicating
560 sufficient heritability to obtain reliable genetic correlation estimates in downstream
561 analyses⁵⁷. Genomic inflation factors (λ_{GC}) ranged from 1.0638 to 1.2156, with LDSC
562 intercepts ranging from 0.9997 to 1.0007, indicating inflation was due to polygenic
563 signal as opposed to confounding due to population stratification^{56,99}. Manhattan and
564 QQ plots can be viewed in **Figure 2**.

565
566
567
568
569
570
571
572

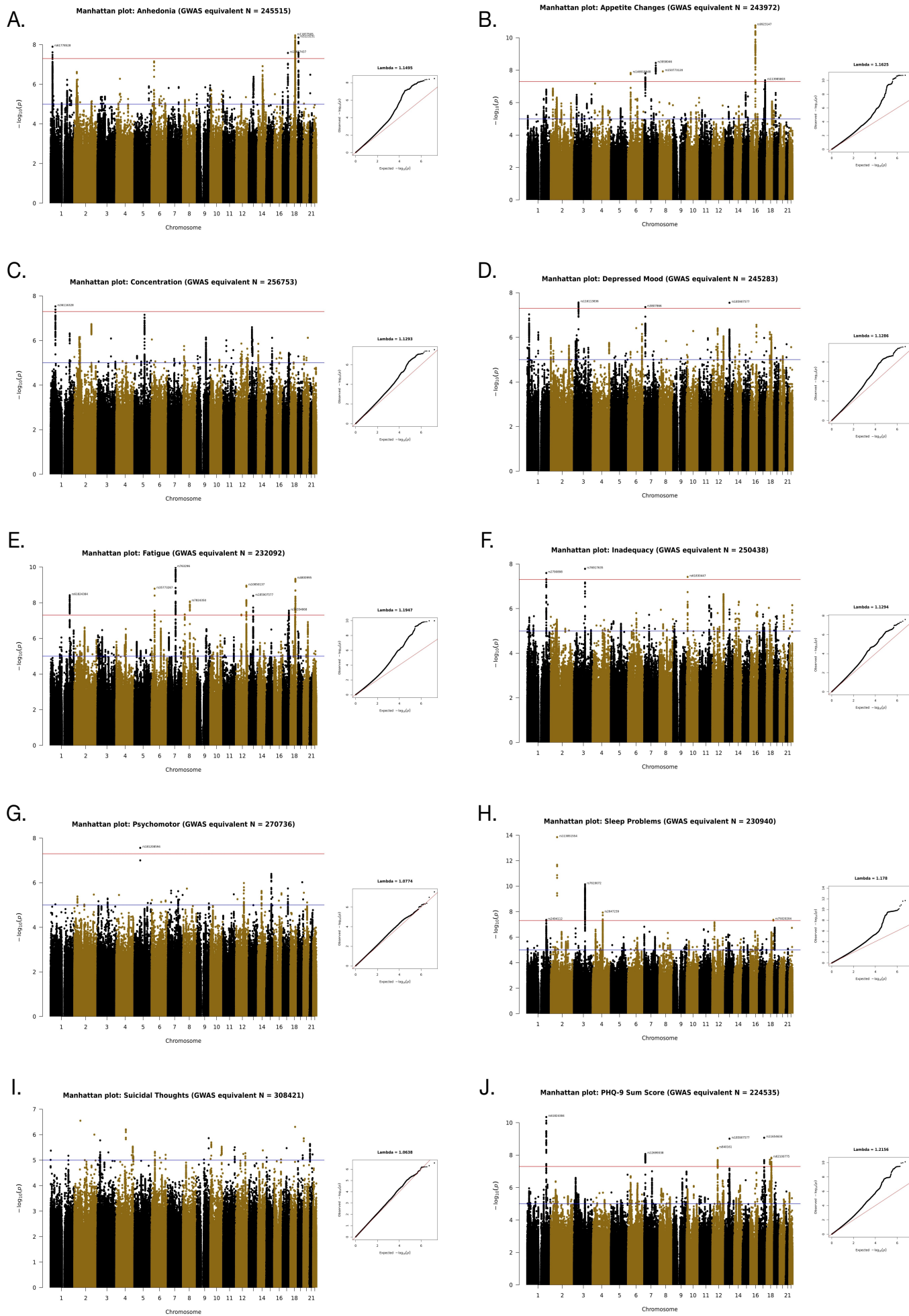
Genomic Locus	CHR	Position	Lead Variant	Phenotype	p-value	BETA	SE	Effect Allele	MAF	Nearest Gene	eQTL Genes
1	1	21059837	rs61778528	Anhedonia	1.27E-08	-0.011	0.002	T	0.384	RP5-930J4.2	KIF17, SH2D5, ECE1, HP1BP3, EIF4G3
2	1	51661974	rs36116328	Concentration Problems	2.94E-08	0.011	0.002	C	0.052	RP11-296A18.3	EPS15, TTC39A, RNF11, CDKN2C,
3	1	201787833	rs2494112	Sleep Problems	4.46E-08	0.011	0.002	A	0.486	NAV1:IPO9-AS1	NAV1, LMOD1, SHISA4, RNPEP,
4	1	201859213	rs61824384	Fatigue	3.81E-09	-0.012	0.002	G	0.385	SHISA4	NAV1, LMOD1, SHISA4, IPO9, TIMM17A, RNPEP
				PHQ-9 Sum Score	4.48E-11	-0.014	0.002	G	0.385	SHISA4	
5	1	201919674	rs2790898	Feelings of Inadequacy	2.54E-08	0.011	0.002	T	0.425	LMOD1	LMOD1, RNPEP, IPO9, SHISA4, TIMM17A
6	2	66750564	rs113851554	Sleep Problems	1.42E-14	0.016	0.002	T	0.053	MEIS1	MEIS1
7	3	48812594	rs116113836	Depressed Mood	2.74E-08	0.011	0.002	A	0.033	PRKAR2A	AMT, ZNF589, ATRIP, NME6, WDR6, P4HTM, QRICH1, NICN1, RNF123, GMPPB, LAMB2, USP4, MST1, TREX1, TMA7, KLHDC8B
8	3	117535095	rs76917635	Feelings of Inadequacy	1.65E-08	-0.012	0.002	G	0.194	SAMP:RP11-384F7	-
9	3	117822025	rs7619072	Sleep Problems	7.14E-11	0.013	0.002	C	0.229	RP11-384F7.1	LSAMP
10	4	106217349	rs2647239	Sleep Problems	1.17E-08	-0.012	0.002	G	0.420	RN7SL89P	PPA2, TET2
11	5	62392481	rs181208594	Psychomotor Changes	2.72E-08	0.011	0.002	G	0.017	RP11-586E1.1	-
12	6	28548674	rs146918648	Appetite Changes	1.49E-08	0.012	0.002	A	0.027	SCAND3	ZSCAN9, PGBD1, ZKSCAN3, ZNF165, TRIM27, ZSCAN23, SCAND3, ZSCAN12, C6orf100"
13	6	32555944	rs35773267	Fatigue	1.63E-09	0.013	0.002	T	0.372	HLA-DRB1	-
14	7	12250378	rs3807866	Depressed Mood	4.38E-08	0.011	0.002	A	0.398	TMEM106B	TMEM106B
15	7	12277442	rs13234970	Appetite Changes	1.56E-08	0.012	0.002	C	0.400	TMEM106B	-
16	7	12285140	rs12699338	PHQ-9 Sum Score	8.47E-09	0.012	0.002	T	0.405	TMEM106B	TMEM106B
17	7	82394712	rs763296	Fatigue	1.03E-10	0.013	0.002	G	0.495	PCLO	-
18	7	121955328	rs3958046	Appetite Changes	3.58E-09	0.012	0.002	T	0.389	CADPS2	-
19	8	19551329	rs13253689	Fatigue	4.55E-08	-0.011	0.002	C	0.243	CSGALNACT1	INTS10, CSGALNACT1, LPL
20	8	37211598	rs150773120	Appetite Changes	1.19E-08	0.012	0.002	A	0.012	RP11-527N22.1	-
21	8	73452488	rs7824350	Fatigue	8.58E-09	-0.012	0.002	G	0.410	KCNB2	-
22	10	11149133	rs61830447	Inadequacy	3.80E-08	0.011	0.002	A	0.243	CELF2	CELF2
23	12	57323523	rs840161	PHQ-9 Sum Score	3.71E-09	0.012	0.002	G	0.339	SDR9C7	METTL21B, TMEM194A, STAT6, NAB2, RBMS2, LRP1
24	12	57331741	rs725957	Fatigue	4.68E-08	0.011	0.002	A	0.394	SDR9C7	METTL21B, TMEM194A, STAT6, NAB2, RBMS2, LRP1, RDH16, DTX3, SDR9C7
25	12	109874794	rs10850137	Fatigue	1.11E-09	0.013	0.002	G	0.336	MYO1H	UBE3B, MMAB, ACACB, KCTD10, MVK, MYO1H, FOXN4, UNG
26	13	66718798	rs185907577	Depressed Mood	2.81E-08	0.011	0.002	G	0.103	MIR4704	-
				Fatigue	4.11E-09	0.012	0.002	G	0.103	MIR4704	
				PHQ-9 Sum Score	9.73E-10	0.013	0.002	G	0.103	MIR4704	
27	16	53801549	rs9923147	Appetite Changes	1.74E-11	0.014	0.002	T	0.443	FTO	FTO
28	17	56044110	rs11654634	PHQ-9 Sum Score	8.48E-10	0.013	0.002	C	0.067	VEZF1	CUEDC1, TEX14
29	17	56124979	rs11657437	Anhedonia	2.62E-08	0.011	0.002	A	0.105	RP11-159D12.10	VEZF1
30	17	65989971	rs113985803	Appetite Changes	4.18E-08	-0.011	0.002	A	0.226	RP11-855A2.5	BPTF, KPNA2, PSMD12
31	17	66045831	rs74354868	Fatigue	2.78E-08	0.011	0.002	A	0.229	KPNA2	BPTF, KPNA2, C17orf58, ARSG, PSMD12
32	18	35127427	rs1557341	PHQ-9 Sum Score	1.94E-08	-0.012	0.002	C	0.341	CELF4	-
33	18	50754093	rs62100775	Fatigue	1.92E-08	0.012	0.002	G	0.419	DCC	DCC
				PHQ-9 Sum Score	1.44E-08	0.012	0.002	G	0.419	DCC	
34	18	50829023	rs11877505	Anhedonia	3.40E-09	0.012	0.002	G	0.388	DCC	DCC
35	18	53415377	rs4800995	Fatigue	4.60E-10	0.013	0.002	A	0.214	RP11-397A16.1	-
36	18	71620555	rs75020204	Sleep Problems	4.34E-08	-0.011	0.002	C	0.071	RP11-25L3.3	-
37	19	7871720	rs3223231	Anhedonia	4.40E-09	-0.012	0.002	A	0.462	EXOSC3P2	-

573

574

575 **Table 3:** Genomic risk loci from the 10 MTAG-PHQ-9 genome-wide association meta-analyses. Nearest gene is
 576 based on positional mapping and ANNOVAR annotation. Expression quantitative trait loci (eQTL) mapping was
 577 limited to $p_{FDR} \leq 0.05$ in the selected eQTL datasets (see: Methods). Specific tissues with significant p_{FDR} can be
 578 seen in Supplementary Table 2.

57



580 **Figure 2:** *Manhattan and QQ-plots for each of the 10 MTAG-PHQ-9 GWAS meta-analyses. The red*
581 *lines on the Manhattan plots indicate genome-wide significance ($p \leq 5 \times 10^{-8}$), while the blue lines*
582 *indicates suggestive significance ($p \leq 1e-5$). Top variants at each genomic risk loci are annotated.*

583

584

585 **3.2 Genetic correlations**

586 Of the 72 bivariate genetic correlations calculated between the 12 depression
587 phenotypes and the six AD GWAS, 24 were nominally significant and 20 remained
588 significant at $p_{\text{FDR}} \leq 0.05$ (r_g range: -0.25 – 0.35; p -value range: 1.25×10^{-2} – 4.01×10^{-5} ;
589 p_{FDR} range: 4.5×10^{-2} – 1.9×10^{-3}). Of these, 19 were identified when the AD GWAS
590 in the pair contained either clinical+proxy cases and controls, or proxy-only cases
591 and controls (**Figure 3; Supplementary Table 8**). Only one p_{FDR} significant association
592 was found when using a clinical AD GWAS – between suicidal thoughts and
593 Wightman et al. ($r_g = -0.25$, $p = 6.78 \times 10^{-3}$, p -value = 6.78×10^{-3} ; $p_{\text{FDR}} = 3.48 \times 10^{-2}$).
594 All depression phenotypes were significantly genetically correlated with each other
595 (r_g range: 0.57–0.98; p -values $\leq 3.71 \times 10^{-23}$) (**Supplementary Table 9, Supplementary**
596 **Material 1**). Only one PHQ-9 symptom pair – concentration problems and
597 psychomotor changes – showed a genetic correlation that was not statistically
598 different from one (95% CI included one), indicating genetic heterogeneity across
599 depression symptoms.

600

601

602

603

604

605

606

607



608

609 **Figure 3:** Heatmap of High Definition Likelihood (HDL) calculated genetic correlations between
 610 depression items and the six Alzheimer's disease genome-wide association studies of varying
 611 proxy/clinical case/control ascertainment. The colour of each square indicates the strength of the
 612 correlation on a scale of -1 to 1. Fill size of the square indicates uncorrected p-value significance.
 613 Genetic correlations significant at $p_{FDR} \leq 0.05$ are circled.

614

615

616 3.3 Local genetic correlations

617 After univariate testing, a total of 4271 bivariate local genetic correlation tests were
 618 conducted across 324 genomic loci. Of these tests, 716 were nominally significant
 619 and 15 remained significant at $p_{FDR} \leq 0.05$ across 14 unique genomic loci (local r_g
 620 range: -0.81 – 0.82; p -value range: 1.48×10^{-4} – 4.2×10^{-6} ; p_{FDR} range = 4.22×10^{-2} –
 621 1.38×10^{-2}) (**Supplementary Table 10**). Of the 15 statistically significant tests, ten were
 622 identified when using clinical+proxy/proxy-only AD GWAS. No depression phenotype
 623 showed a statistically significant association at the same genomic locus with more
 624 than one AD GWAS. However, for ten of the 15 statistically significant tests, nominally
 625 significant local genetic correlation was observed between the depression
 626 phenotype and at least one additional AD GWAS at the same loci (**Supplementary**
 627 **Table 11**). Only locus 1790 (chr12:51769420– 53039987) showed p_{FDR} significant
 628 local genetic correlation with more than one depression phenotype – concentration
 629 problems and sleep problems, both with the clinical-only Wightman et al. GWAS. The
 630 number of positively and negatively correlated loci identified between each
 631 phenotype pair can be viewed in **Supplementary Table 12**.

632

633

3.4 Colocalisation

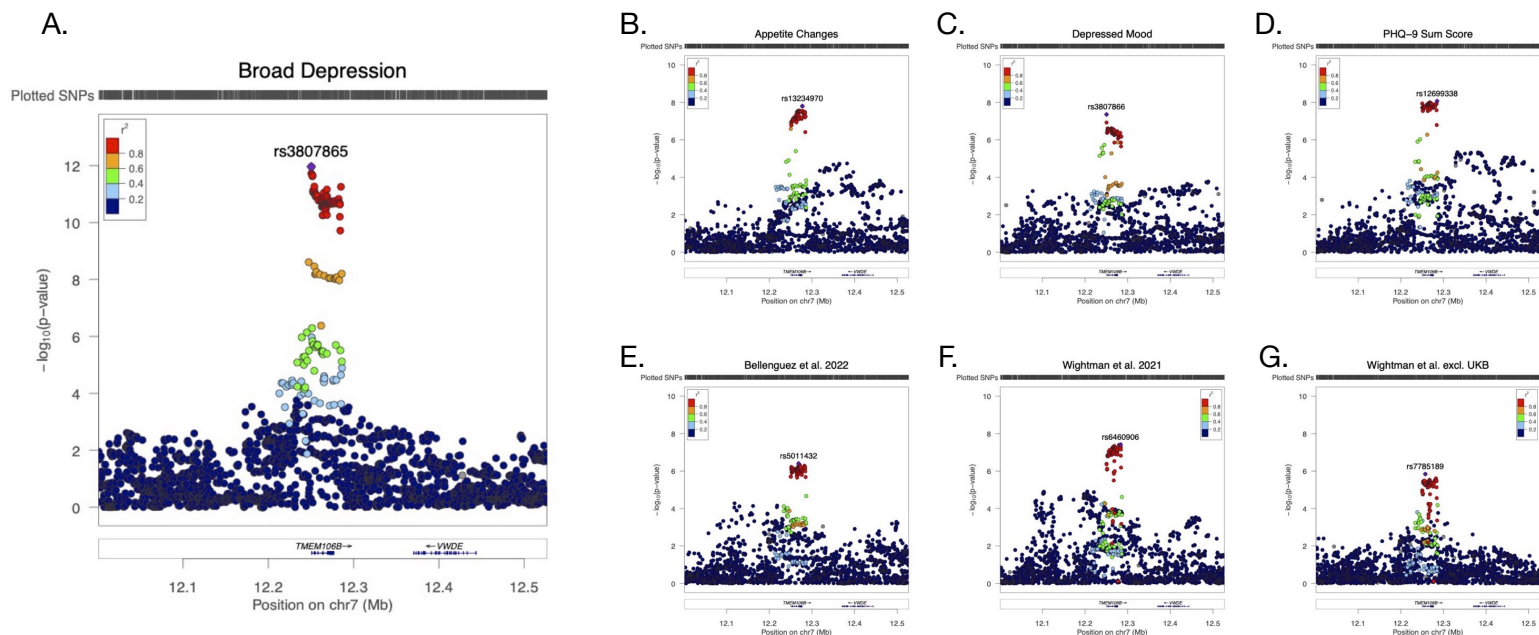
634 Following LAVA, 14 p_{FDR} significant regions of local genetic correlation were passed
635 to the COLOC-reporter pipeline across 15 depression-AD phenotype pairs. A further
636 14 colocalisation tests were conducted where nominally significant local genetic
637 correlation was observed at a p_{FDR} significant loci between the same depression
638 phenotype and a different AD GWAS. As such, a total of 29 statistical colocalisation
639 tests were conducted to follow up LAVA results. No 95% credible sets were identified
640 by SuSiE for any phenotype pairs in these regions. As such, all analyses were
641 conducted under the single causal variant assumption of *coloc.abf*. There was no
642 evidence of colocalisation in any of these loci (mean PP.H₄ = 0.59%) (**Supplementary**
643 **Table 11**). All but two of these tests indicated no causal variant present in either
644 phenotype (PP.H₀ > 0.8). The two tests in locus 319 (chr2:126754028-127895644),
645 indicated strong probability of a causal variant for the Kunkle et al.¹⁴ and Wightman
646 et al.⁴⁵ clinical AD GWAS (PP.H₂ > 0.9). This locus contains *BIN1*, a known risk gene
647 for AD involved in tau regulation^{100,101}.

648 An additional 762 colocalisation tests were conducted with the six AD GWAS
649 using regions +/-250kb ($r^2 > 0.1$) lead variants from the MTAG-PHQ-9, broad and
650 clinical depression GWAS. SuSiE identified evidence of colocalisation in regions +/-
651 250kb of lead variants at genomic risk loci 14 (depressed mood), 15 (appetite
652 change), 16 (PHQ-9 sum score) and for broad depression at chr7:12000402-
653 12500402 (PP.H₄ range: 0.79 – 0.85), all with the same three AD GWAS – Bellenguez
654 et al.¹⁷, Wightman et al.¹⁶, and Wightman et al. (excluding the UKB)⁴⁵ (**Supplementary**
655 **Table 13**). These colocalisations were all in the region of the transmembrane protein
656 106B gene (*TMEM106B*) – visualised using LocusZoom¹⁰² in **Figure 4**. Colocalisation
657 was also identified for the same phenotype pairs at the same loci under the single
658 causal variant assumption of *coloc.abf* (**Supplementary Table 14**). The same
659 depression phenotypes and loci were suggestive of colocalisation with Jansen et al.¹⁵
660 (PP.H₄ > 0.6). In a follow-up analysis, we assessed statistical colocalisation +/-250kb
661 *TMEM106B* (chr7:12000920-12532993) between these four AD GWAS and all
662 remaining depression phenotypes. Additional evidence of colocalisation was
663 identified at *TMEM106B* between fatigue and the Bellenguez et al.¹⁷, Wightman et
664 al.¹⁶, and Wightman et al. (excluding the UKB)⁴⁵ AD GWAS (**Supplementary Table 15**),
665 and was suggestive for Jansen et al.¹⁵ (**Supplementary Table 16**). Evidence of
666 colocalisation was also suggestive for psychomotor changes with Bellenguez et al.¹⁷,
667 Wightman et al.¹⁶, and Wightman et al. (excluding the UKB)⁴⁵ (PP.H₄ > 0.6).

668

669

670



671 **Figure 4:** LocusZoom plots of the transmembrane protein 106B (*TMEM106B*) gene region containing
 672 evidence of colocalisation ($PP.H_4 \leq 0.8$) for (A.) broad depression, (B.) appetite changes, (C.)
 673 depressed mood, (D.) the PHQ-9 sum score, (E.) Bellenguez et al., (F.) Wightman et al. and (G.)
 674 Wightman et al. excluding the UK Biobank. The most significant variant for each phenotype is labelled
 675 in their respective plot.

676

677

678 3.5 Mendelian Randomisation

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

683 F-statistics indicated that instrument strength was sufficient (F_{Mean} range: 22.43
 684 – 63.36; F_{Min} range: 20.84 – 31.56; F_{Max} range: 26.37 – 402.86), as did I^2 for instrument
 685 suitability for MR-Egger (I^2 range: 0.91 – 0.98). A $p_{\text{FDR}} \leq 0.05$ was applied in each of
 686 the other MR methods to correct for the 144 tests conducted, after which no
 687 statistically significant associations were observed in any method. Three nominally
 688 significant causal associations where a depression item was the exposure (β range
 689 = 0.21 – 1.59; p -value range = 0.01 – $2e-3$) and two where an AD GWAS was the
 690 exposure (β range = 0.018 – 0.02; p -value range = 0.01 – 0.02) were detected by

691 IVW-MR, of which three were nominally significant in weighted-median MR. No
692 significant pleiotropy (MR-Egger intercept test p -value ≤ 0.05) or heterogeneity (IVW
693 Cochran's Q p -value ≤ 0.05) was detected for these five tests. However, nominally
694 significant estimates were only observed between exposures and outcomes where
695 the AD GWAS contained proxy+clinical or proxy-only cases/controls and where the
696 GWAS used for both exposure and outcome were either majority or entirely derived
697 from the UK Biobank, suggesting sample overlap as a potential source of bias. MR-
698 PRESSO detected and excluded significant outliers in 42 tests, returning outlier-
699 corrected causal estimates. Results remained non-significant. Full MR results can be
700 viewed in **Supplementary Table 18**.

701 **3.6 Polygenic risk scores (PRS)**

702 No statistically significant associations were detected between any of the depression
703 phenotype PRS and AD case/control status in any of the three AD target samples
704 ($p_{\text{FDR}} \leq 0.05$, corrected within each sample), with only the suicidal thoughts PRS
705 negatively nominally associated with AD case/control in the ADNI cohort
706 (Nagelkerke's pseudo- $r^2 = 0.01$, OR [95% CI] = 0.83 [0.72–0.94], $\beta = -0.19$, SE =
707 0.07, p -value = 7.90×10^{-3}). Exclusion of the APOE region had no effect on results
708 (**Supplementary Table 19, Supplementary Material 2**).

709

710 **4. Discussion**

711 In this study, we performed the first and largest genome-wide meta-analysis of PHQ-
712 9 depression symptom items to date (GWAS equivalent N range: 224,535 – 308,421),
713 identifying 37 genomic risk loci. Follow-up analyses examining the genetic overlap
714 between depression/depression symptoms and AD identified 20 significant global
715 correlations and 15 significant local correlations at 14 loci across six AD GWAS with
716 varying proportion of clinical case/control ascertainment. Significant global genetic
717 correlations were primarily with AD GWAS containing proxy cases and controls. No
718 colocalisation was identified at any of the regions of local genetic correlation.
719 However, there was strong evidence of colocalisation between depression several
720 phenotypes and AD in the region of *TMEM106B*. Overall, MR did not suggest that
721 depression or its symptoms were causal for AD, nor vice versa, while polygenic risk
722 scores for depression phenotypes were not predictive of AD case/control status in
723 three clinical AD samples.

724 The increased power of our PHQ-9 GWAS allowed for the identification of 28
725 more genomic risk loci than the previous analysis by Thorp et al.²⁹. Several of the loci
726 identified in this study have shown previous associations with related phenotypes.
727 For example, *SHISA4* – here identified in association with fatigue symptoms – has
728 been previously implicated as playing a role in disrupted sleep¹⁰³ and daytime

729 napping¹⁰⁴. The top variant for sleep problems at genomic risk loci 6 (*MEIS1*) –
730 rs113851554 (chr2: 66750564) – was also the top variant in a previous GWAS of
731 insomnia and restless leg syndrome¹⁰⁵. Additionally, the obesity gene *FTO*¹⁰⁶ was
732 identified as a genomic risk loci for appetite changes. Although the role of *FTO* in
733 depression is inconclusive¹⁰⁷, it has been recently linked to anxiety and depression
734 symptoms in individuals with anorexia nervosa (AN)¹⁰⁸. Consequently, its
735 identification in association with appetite change symptoms – a phenotype relevant
736 to eating behaviours – suggests that symptom-based genetic analysis can help
737 identify the phenotype-relevant biology of individual depression symptoms. Taken
738 alongside the fact that only three of the genomic risk loci were shared between more
739 than one depression symptom phenotype, our GWAS meta-analysis backs findings
740 by Thorp et al.²⁹ of genetic heterogeneity between depression symptoms.

741 Our findings also highlight genetic similarities between depression symptoms.
742 For example, *TMEM106B* – a gene identified in previous depression GWAS^{11,12} – was
743 the nearest gene to lead variants for three PHQ-9 items – appetite changes
744 (rs13234970), depressed mood (rs3807866) and the PHQ-9 sum score
745 (rs12699338). *TMEM106B* was strongly suggested as a causal gene in a recent multi-
746 ancestry depression GWAS¹⁰⁹. Further, dysregulation of *TMEM106B* expression has
747 been implicated in association with MDD¹¹⁰ and with anxious and weight gain MDD
748 subtypes – both associated with treatment resistance¹¹¹. *TMEM106B* has also been
749 implicated in self-reported diagnosis of anxiety disorder¹¹², neuroticism¹¹³ and in a
750 latent factor GWAS of depressive, manic and psychotic symptoms/disorders¹¹⁴,
751 suggesting it is linked to psychiatric risk more generally.

752 That colocalisation is observed in the *TMEM106B* region between multiple
753 depression phenotypes and both proxy+clinical and clinical-only AD is therefore of
754 particular interest. *TMEM106B* plays a role in lysosomal function – particularly in
755 motor neurons¹¹⁵ – and is classically considered a risk gene for frontotemporal
756 dementia¹¹⁶. It is has also been identified in recent AD GWAS^{16,17}, and is associated
757 with brain aging, cognitive decline, and neurodegeneration across other brain
758 disorders including Amyotrophic Lateral Sclerosis, Multiple Sclerosis and Parkinson's
759 Disease¹¹⁷⁻¹²². Further, *TMEM106B* has been linked to higher CSF levels of
760 neurofilament light (NfL)¹²³ – itself predictive of cognitive decline, brain atrophy and
761 cortical amyloid burden in individuals with AD and Mild Cognitive Impairment
762 (MCI)¹²⁴. A recent study also identified higher levels of plasma NfL in individuals with
763 depression, although the sample size of this study was relatively small¹²⁵. Thus,
764 colocalisation between depression phenotypes and AD at *TMEM106B* indicates that
765 depression may itself be genetically linked to overall brain health and resulting
766 general dementia risk. Two previous studies have identified *TMEM106B* as playing a
767 role in both depression and AD^{20,21}. However, our study suggests that this overlap

768 may be driven by the genetic architecture of specific depression symptoms,
769 highlighting the benefits of symptom-level genetic analysis.

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

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

794 Exactly why depression/depression symptoms show differences in genetic
795 correlation between proxy and clinical AD is a matter of interest. It is noted that most
796 individuals with dementia are cared for by a family member while they are still living
797 in the community¹²⁶. Caregivers incur considerable lifestyle changes, emotional
798 stress, and social strain^{127,128} and depression is reported in up to 50% of caregivers
799 for individuals with dementia¹²⁹. As such, the genetic correlations between
800 depression/depression symptoms and proxy-AD may result from increased
801 depression risk due to caregiver responsibilities, exacerbated by a genetic risk.
802 However, this would not explain why no genetic correlations were identified with the
803 Bellenguez et al.¹⁷ GWAS despite this study also containing proxy+clinical
804 phenotyping. As mentioned, Bellenguez et al.¹⁷ define proxy cases/controls as a
805 binary phenotype, whereas Wightman et al.¹⁶ and Jansen et al.¹⁵ define proxy
806 cases/controls as a continuous phenotype. These phenotyping differences likely

807 partially explain the differences in the genetic correlation results, given that these AD
808 GWAS all use the same data from the UK Biobank. However, the proxy-only Marioni
809 et al.³⁷ GWAS – with which genetic correlations were also observed – is a meta-
810 analysis of maternal and paternal AD GWAS where parental age-at-diagnosis/age-
811 of-death is controlled for in the GWAS model, instead of being used for weighting the
812 proxy phenotype prior analysis. Taking this into account, and considering that
813 depression is itself associated with all-cause mortality¹³⁰, it is possible that including
814 age-at-diagnosis/age-of-death in proxy-AD phenotyping induces a form of bias in
815 later cross-trait analyses when the other trait is itself associated with longevity. Further
816 investigation of this issue is required.

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

828 This study has several limitations. Despite being the largest meta-analysis of
829 PHQ-9 items to date, the ability to detect genome-wide significant variants was likely
830 limited by small sample sizes relative to other psychiatric conditions. This includes
831 depression itself, where sample sizes have exceeded one million¹³. Larger samples
832 are required for future studies if we are to truly elucidate the unique genetic
833 architecture of individual depression symptoms. Additionally, our analyses were
834 restricted to individuals of European ancestry. GWAS results may therefore have poor
835 transferability to other ancestry groups. Although efforts are underway to address the
836 issue of lack of diversity in human genetic studies, urgency is required so that studies
837 can be conducted across ancestry groups using equivalent sample sizes. Our study
838 makes also heavy use of data from the UKB. The UKB is known to be affected by
839 healthy volunteer bias, and as a consequence is not fully representative of the wider
840 population¹³¹.

841 It is worth noting that this study focused on depression as a risk factor for AD.
842 However, there is evidence that some forms of later life depression are in fact a
843 prodromal phase of dementia onset^{8,132}, and may be related to levels of dementia
844 biomarkers such as plasma A β 42 and CSF p-tau¹³³. As such, dementia-related
845 depression may be biologically distinct from depression as a mental health disorder.

846 While sufficient sample sizes are likely difficult to obtain, future large scale genomic
847 studies of dementia-related depression – as has been undertaken with psychosis in
848 AD¹³⁴ – would prove illuminating on this matter.

849 In conclusion, this study describes the largest genome-wide meta-analysis of
850 PHQ-9 depression symptom items to date (GWAS equivalent N range: 224,535 –
851 308,421), identifying 37 unique genomic risk loci. Genetic correlations between
852 depression/depression symptoms and AD were primarily observed when the AD
853 GWAS contained clinical+proxy or proxy-only AD case/control ascertainment.
854 Further, this study does not support depression or its symptoms as being causal for
855 AD. However, colocalisation in the *TMEM106B* region between four depression
856 phenotypes and AD across both proxy and clinical AD GWAS suggests future
857 research into the shared biological mechanisms underlying the role of this locus in
858 depression and AD are warranted.

859

860

861 5. Funding and Acknowledgements

862

863 LG is funded by the King's College London DRIVE-Health Centre for Doctoral Training
864 and the Perron Institute for Neurological and Translational. PP is funded by
865 Alzheimer's Research UK. SK is funded by MSWA and the Perron Institute. HLD
866 acknowledges funding from the Economic and Social Research Council (ESRC).
867 DMH is supported by a Sir Henry Wellcome Postdoctoral Fellowship (Reference
868 213674/Z/18/Z).

869

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

874

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

878

879 We thank the GLAD Study volunteers for their participation, and gratefully
880 acknowledge the NIHR BioResource centres, NHS Trusts and staff for their
881 contribution. We thank the National Institute for Health Research, NHS Blood and
882 Transplant, and Health Data Research UK as part of the Digital Innovation Hub
883 Programme. This study presents independent research funded by the NIHR
884 Biomedical Research Centre at South London and Maudsley NHS Foundation Trust
885 and King's College London. Further information can be found at

886 <http://brc.slam.nhs.uk/about/core-facilities/bioresource>. The views expressed are
887 those of the authors and not necessarily those of the NHS, the NIHR, the HSC R&D
888 Division, King's College London, or the Department of Health and Social Care.

889
890 The PROTECT study was funded/supported by the National Institute of Health and
891 Care Research Exeter Biomedical Research Centre. PROTECT genetic data was
892 funded in part by the University of Exeter through the MRC Proximity to Discovery:
893 Industry Engagement Fund (External Collaboration, Innovation and Entrepreneurism:
894 Translational Medicine in Exeter 2 (EXCITEME2) ref. MC_PC_17189). Genotyping was
895 performed at deCODE Genetics. The views expressed are those of the author(s) and
896 not necessarily those of the NIHR or the Dept of Health and Social Care.

897
898 As data used in this study was obtained from the ADNI database, the ADNI
899 investigators contributed to the conception of the sample and acquisition of this data,
900 although were not participants in other parts of this study, such as conceptualisation,
901 data analysis or writing. A full acknowledgment list of ADNI investigators can be found
902 at: [https://adni.loni.usc.edu/wp-](https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)
903 [content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf](https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

904
905 Similarly, data was obtained from the GERAD1 Consortium and GERAD1
906 investigators contributed to the conception of the cohort and acquisition of this data,
907 but did not participate in the conceptualisation, analysis or writing of this study. A full
908 list of GERAD1 collaborators can be found in **Supplementary Material 3**.

909
910 We also thank and acknowledge the contribution and use of the CREATE high-
911 performance computing cluster at King's College London (King's College London.
912 (2022). King's Computational Research, Engineering and Technology Environment
913 (CREATE). Retrieved May 23, 2023, from <https://doi.org/10.18742/rnvf-m076>).

914
915 The analysis flowchart in **Figure 1**. was created and licenced with BioRender
916 (www.biorender.com)

917 918 **6. Conflicts of Interest**

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

921 922 **7. Data availability**

923 All GWAS summary statistics generated in the process of conducting this study will
924 be deposited in the NHGRI-EBI Catalog of human genome-wide association studies

925 (GWAS Catalog) following peer-review and publication. This study was pre-registered
926 on the Open Science Framework
927 (https://osf.io/94q35/?view_only=e77f72d4100d47eea7f3ef07dfa9c059).

928

929 This study made use of publicly available analysis software:

930

931 MungeSumstats

932 (<https://neurogenomics.github.io/MungeSumstats/articles/MungeSumstats.html>)

933

934 REGENIE (<https://rgcgithub.github.io/regenie/>)

935

936 MTAG (<https://github.com/JonJala/mtag>)

937

938 METAL (https://genome.sph.umich.edu/wiki/METAL_Documentation)

939

940 FUMA GWAS (<https://fuma.ctglab.nl>)

941

942 LDSC (<https://github.com/bulik/ldsc>)

943

944 HDL (<https://github.com/zhenin/HDL>)

945

946 LAVA (<https://github.com/josefin-werme/LAVA>)

947

948 susieR (<https://stephenslab.github.io/susieR/index.html>)

949

950 coloc (<https://chr1swallace.github.io/coloc/>)

951

952 COLOC-reporter (<https://github.com/ThomasPSpargo/COLOC-reporter>)

953

954 CAUSE (<https://jean997.github.io/cause/index.html>)

955

956 TwoSampleMR (<https://mrcieu.github.io/TwoSampleMR/>)

957

958 MegaPRS (<https://dougsped.com/megaprs/>)

959

960 References

- 961
- 962 1. Byers, A.L., and Yaffe, K. (2011). Depression and risk of developing dementia. *Nat. Rev.*
963 *Neurol.* 7, 323–331.
- 964 2. Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne,
965 C., Burns, A., Cohen-Mansfield, J., Cooper, C., et al. (2020). Dementia prevention,
966 intervention, and care: 2020 report of the Lancet Commission. *Lancet* 396, 413–446.
- 967 3. Yu, O.-C., Jung, B., Go, H., Park, M., and Ha, I.-H. (2020). Association between dementia
968 and depression: a retrospective study using the Korean National Health Insurance Service-
969 National Sample Cohort database. *BMJ Open* 10, e034924.
- 970 4. Yang, W., Li, X., Pan, K.-Y., Yang, R., Song, R., Qi, X., Pedersen, N.L., and Xu, W.
971 (2021). Association of life-course depression with the risk of dementia in late life: A
972 nationwide twin study. *Alzheimers Dement.* 17, 1383–1390.
- 973 5. Richmond-Rakerd, L.S., D’Souza, S., Milne, B.J., Caspi, A., and Moffitt, T.E. (2022).
974 Longitudinal Associations of Mental Disorders With Dementia: 30-Year Analysis of 1.7
975 Million New Zealand Citizens. *JAMA Psychiatry* 79, 333–340.
- 976 6. Yang, L., Deng, Y.-T., Leng, Y., Ou, Y.-N., Li, Y.-Z., Chen, S.-D., He, X.-Y., Wu, B.-S.,
977 Huang, S.-Y., Zhang, Y.-R., et al. (2022). Depression, depression treatments, and risk of
978 incident dementia: A prospective cohort study of 354,313 participants. *Biol. Psychiatry*.
- 979 7. GBD 2016 Dementia Collaborators (2019). Global, regional, and national burden of
980 Alzheimer’s disease and other dementias, 1990-2016: a systematic analysis for the Global
981 Burden of Disease Study 2016. *Lancet Neurol.* 18, 88–106.
- 982 8. Dafsari, F.S., and Jessen, F. (2020). Depression-an underrecognized target for prevention
983 of dementia in Alzheimer’s disease. *Transl. Psychiatry* 10, 160.
- 984 9. Corfield, E.C., Yang, Y., Martin, N.G., and Nyholt, D.R. (2017). A continuum of genetic
985 liability for minor and major depression. *Transl. Psychiatry* 7, e1131.
- 986 10. Gatz, M., Reynolds, C.A., Fratiglioni, L., Johansson, B., Mortimer, J.A., Berg, S., Fiske,
987 A., and Pedersen, N.L. (2006). Role of genes and environments for explaining Alzheimer
988 disease. *Arch. Gen. Psychiatry* 63, 168–174.
- 989 11. Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A.,
990 Adams, M.J., Agerbo, E., Air, T.M., Andlauer, T.M.F., et al. (2018). Genome-wide
991 association analyses identify 44 risk variants and refine the genetic architecture of major
992 depression. *Nat. Genet.* 50, 668–681.
- 993 12. Howard, D.M., Adams, M.J., Clarke, T.-K., Hafferty, J.D., Gibson, J., Shiralil, M.,
994 Coleman, J.R.I., Hagenaars, S.P., Ward, J., Wigmore, E.M., et al. (2019). Genome-wide
995 meta-analysis of depression identifies 102 independent variants and highlights the importance
996 of the prefrontal brain regions. *Nat. Neurosci.* 22, 343–352.
- 997 13. Levey, D.F., Stein, M.B., Wendt, F.R., Pathak, G.A., Zhou, H., Aslan, M., Quaden, R.,
998 Harrington, K.M., Nuñez, Y.Z., Overstreet, C., et al. (2021). Bi-ancestral depression GWAS
999 in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new
1000 therapeutic directions. *Nat. Neurosci.* 24, 954–963.
- 1001 14. Kunkle, B.W., Grenier-Boley, B., Sims, R., Bis, J.C., Damotte, V., Naj, A.C., Boland, A.,
1002 Vronskaya, M., van der Lee, S.J., Amlie-Wolf, A., et al. (2019). Genetic meta-analysis of
1003 diagnosed Alzheimer’s disease identifies new risk loci and implicates A β , tau, immunity and
1004 lipid processing. *Nat. Genet.* 51, 414–430.
- 1005 15. Jansen, I.E., Savage, J.E., Watanabe, K., Bryois, J., Williams, D.M., Steinberg, S.,
1006 Sealock, J., Karlsson, I.K., Hägg, S., Athanasiu, L., et al. (2019). Genome-wide meta-analysis
1007 identifies new loci and functional pathways influencing Alzheimer’s disease risk. *Nat. Genet.*
1008 51, 404–413.
- 1009 16. Wightman, D.P., Jansen, I.E., Savage, J.E., Shadrin, A.A., Bahrami, S., Holland, D.,

- 1010 Rongve, A., Børte, S., Winsvold, B.S., Drange, O.K., et al. (2021). A genome-wide
1011 association study with 1,126,563 individuals identifies new risk loci for Alzheimer’s disease.
1012 *Nat. Genet.* *53*, 1276–1282.
- 1013 17. Bellenguez, C., Küçükali, F., Jansen, I.E., Kleindan, L., Moreno-Grau, S., Amin, N.,
1014 Naj, A.C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., et al. (2022). New insights
1015 into the genetic etiology of Alzheimer’s disease and related dementias. *Nat. Genet.* *54*, 412–
1016 436.
- 1017 18. Gibson, J., Russ, T.C., Adams, M.J., Clarke, T.K., Howard, D.M., Hall, L.S., Fernandez-
1018 Pujals, A.M., Wigmore, E.M., Hayward, C., Davies, G., et al. (2017). Assessing the presence
1019 of shared genetic architecture between Alzheimer’s disease and major depressive disorder
1020 using genome-wide association data. *Transl. Psychiatry* *7*, e1094.
- 1021 19. Andrews, S.J., Fulton-Howard, B., O’Reilly, P., Marcora, E., Goate, A.M., and
1022 collaborators of the Alzheimer’s Disease Genetics Consortium (2021). Causal associations
1023 between modifiable risk factors and the alzheimer’s phenome. *Ann. Neurol.* *89*, 54–65.
- 1024 20. Santos, F.C.D., Mendes-Silva, A.P., Nikolova, Y.C., Sibille, E., and Diniz, B.S. (2021).
1025 Genetic overlap between major depression, bipolar disorder and Alzheimer’s Disease.
1026 *MedRxiv*.
- 1027 21. Monereo-Sánchez, J., Schram, M.T., Frei, O., O’Connell, K., Shadrin, A.A., Smeland,
1028 O.B., Westlye, L.T., Andreassen, O.A., Kaufmann, T., Linden, D.E.J., et al. (2021). Genetic
1029 overlap between alzheimer’s disease and depression mapped onto the brain. *Front. Neurosci.*
1030 *15*, 653130.
- 1031 22. Harerimana, N.V., Liu, Y., Gerasimov, E.S., Duong, D., Beach, T.G., Reiman, E.M.,
1032 Schneider, J.A., Boyle, P., Lori, A., Bennett, D.A., et al. (2022). Genetic evidence supporting
1033 a causal role of depression in alzheimer’s disease. *Biol. Psychiatry* *92*, 25–33.
- 1034 23. American Psychiatric Association (2013). Depressive Disorders. In *Diagnostic and*
1035 *Statistical Manual of Mental Disorders*, (American Psychiatric Association), p.
- 1036 24. Fried, E.I., and Nesse, R.M. (2015). Depression is not a consistent syndrome: An
1037 investigation of unique symptom patterns in the STAR*D study. *J. Affect. Disord.* *172*, 96–
1038 102.
- 1039 25. Zimmerman, M., Ellison, W., Young, D., Chelminski, I., and Dalrymple, K. (2015). How
1040 many different ways do patients meet the diagnostic criteria for major depressive disorder?
1041 *Compr. Psychiatry* *56*, 29–34.
- 1042 26. Park, S.-C., Kim, J.-M., Jun, T.-Y., Lee, M.-S., Kim, J.-B., Yim, H.-W., and Park, Y.C.
1043 (2017). How many different symptom combinations fulfil the diagnostic criteria for major
1044 depressive disorder? Results from the CRESCEND study. *Nord J Psychiatry* *71*, 217–222.
- 1045 27. Cai, N., Choi, K.W., and Fried, E.I. (2020). Reviewing the genetics of heterogeneity in
1046 depression: operationalizations, manifestations and etiologies. *Hum. Mol. Genet.* *29*, R10–
1047 R18.
- 1048 28. Nguyen, T.-D., Harder, A., Xiong, Y., Kowalec, K., Hägg, S., Cai, N., Kuja-Halkola, R.,
1049 Dalman, C., Sullivan, P.F., and Lu, Y. (2022). Genetic heterogeneity and subtypes of major
1050 depression. *Mol. Psychiatry* *27*, 1667–1675.
- 1051 29. Thorp, J.G., Marees, A.T., Ong, J.-S., An, J., MacGregor, S., and Derks, E.M. (2020).
1052 Genetic heterogeneity in self-reported depressive symptoms identified through genetic
1053 analyses of the PHQ-9. *Psychol. Med.* *50*, 2385–2396.
- 1054 30. So, Y., Kim, K.W., Park, J.H., Lee, J.J., Lee, S.B., and TaeHui Kim (2012). P3-396:
1055 Anhedonia is associated with the risk of Alzheimer’s disease in elders with mild cognitive
1056 impairment: Results from the Korean Longitudinal Study on Health and Aging (KLOSHA).
1057 *Alzheimers Dement.* *8*, P594–P594.
- 1058 31. Lee, J.R., Suh, S.W., Han, J.W., Byun, S., Kwon, S.J., Lee, K.H., Kwak, K.P., Kim, B.J.,
1059 Kim, S.G., Kim, J.L., et al. (2019). Anhedonia and Dysphoria Are Differentially Associated

- 1060 with the Risk of Dementia in the Cognitively Normal Elderly Individuals: A Prospective
1061 Cohort Study. *Psychiatry Investig* 16, 575–580.
- 1062 32. Vaquero-Puyuelo, D., De-la-Cámara, C., Olaya, B., Gracia-García, P., Lobo, A., López-
1063 Antón, R., and Santabárbara, J. (2021). Anhedonia as a Potential Risk Factor of Alzheimer’s
1064 Disease in a Community-Dwelling Elderly Sample: Results from the ZARADEMP Project.
1065 *Int. J. Environ. Res. Public Health* 18,.
- 1066 33. Ikeda, M., Brown, J., Holland, A.J., Fukuhara, R., and Hodges, J.R. (2002). Changes in
1067 appetite, food preference, and eating habits in frontotemporal dementia and Alzheimer’s
1068 disease. *J. Neurol. Neurosurg. Psychiatry* 73, 371–376.
- 1069 34. Bailon, O., Roussel, M., Boucart, M., Krystkowiak, P., and Godefroy, O. (2010).
1070 Psychomotor slowing in mild cognitive impairment, Alzheimer’s disease and lewy body
1071 dementia: mechanisms and diagnostic value. *Dement. Geriatr. Cogn. Disord.* 29, 388–396.
- 1072 35. Koren, T., Fisher, E., Webster, L., Livingston, G., and Rapaport, P. (2022). Prevalence of
1073 sleep disturbances in people with dementia living in the community: A systematic review and
1074 meta-analysis. *Ageing Res Rev* 101782.
- 1075 36. Escott-Price, V., and Hardy, J. (2022). Genome-wide association studies for Alzheimer’s
1076 disease: bigger is not always better. *Brain Commun.* 4, faac125.
- 1077 37. Marioni, R.E., Harris, S.E., Zhang, Q., McRae, A.F., Hagenaars, S.P., Hill, W.D., Davies,
1078 G., Ritchie, C.W., Gale, C.R., Starr, J.M., et al. (2018). GWAS on family history of
1079 Alzheimer’s disease. *Transl. Psychiatry* 8, 99.
- 1080 38. Liu, H., Hu, Y., Zhang, Y., Zhang, H., Gao, S., Wang, L., Wang, T., Han, Z., Sun, B.-L.,
1081 and Liu, G. (2022). Mendelian randomization highlights significant difference and genetic
1082 heterogeneity in clinically diagnosed Alzheimer’s disease GWAS and self-report proxy
1083 phenotype GWAS. *Alzheimers Res. Ther.* 14, 17.
- 1084 39. Desai, R., John, A., Saunders, R., Marchant, N.L., Buckman, J.E.J., Charlesworth, G.,
1085 Zuber, V., and Stott, J. (2023). Examining the Lancet Commission risk factors for dementia
1086 using Mendelian randomisation. *BMJ Ment Health* 26,.
- 1087 40. European Alzheimer’s & Dementia Biobank Mendelian Randomization (EADB-MR)
1088 Collaboration, Luo, J., Thomassen, J.Q., Bellenguez, C., Grenier-Boley, B., de Rojas, I.,
1089 Castillo, A., Parveen, K., Küçükali, F., Nicolas, A., et al. (2023). Genetic associations
1090 between modifiable risk factors and alzheimer disease. *JAMA Netw. Open* 6, e2313734.
- 1091 41. Davies, M.R., Kalsi, G., Armour, C., Jones, I.R., McIntosh, A.M., Smith, D.J., Walters,
1092 J.T.R., Bradley, J.R., Kingston, N., Ashford, S., et al. (2019). The Genetic Links to Anxiety
1093 and Depression (GLAD) Study: Online recruitment into the largest recontactable study of
1094 depression and anxiety. *Behav. Res. Ther.* 123, 103503.
- 1095 42. Creese, B., Arathimos, R., Brooker, H., Aarsland, D., Corbett, A., Lewis, C., Ballard, C.,
1096 and Ismail, Z. (2021). Genetic risk for Alzheimer’s disease, cognition, and mild behavioral
1097 impairment in healthy older adults. *Alzheimers Dement (Amst)* 13, e12164.
- 1098 43. Arathimos, R., Fabbri, C., Vassos, E., Davis, K.A.S., Pain, O., Gillett, A., Coleman,
1099 J.R.I., Hanscombe, K., Hagenaars, S., Jermy, B., et al. (2022). Latent subtypes of manic
1100 and/or irritable episode symptoms in two population-based cohorts. *Br. J. Psychiatry* 221,
1101 722–731.
- 1102 44. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A.,
1103 Vukcevic, D., Delaneau, O., O’Connell, J., et al. (2018). The UK Biobank resource with deep
1104 phenotyping and genomic data. *Nature* 562, 203–209.
- 1105 45. Wightman, D.P., Savage, J.E., Tissink, E., Romero, C., Jansen, I.E., and Posthuma, D.
1106 (2023). The genetic overlap between Alzheimer’s disease, amyotrophic lateral sclerosis,
1107 Lewy body dementia, and Parkinson’s disease. *Neurobiol. Aging*.
- 1108 46. Kroenke, K., Spitzer, R.L., and Williams, J.B. (2001). The PHQ-9: validity of a brief
1109 depression severity measure. *J. Gen. Intern. Med.* 16, 606–613.

- 1110 47. Sun, Y., Fu, Z., Bo, Q., Mao, Z., Ma, X., and Wang, C. (2020). The reliability and
1111 validity of PHQ-9 in patients with major depressive disorder in psychiatric hospital. *BMC*
1112 *Psychiatry* 20, 474.
- 1113 48. Mbatchou, J., Barnard, L., Backman, J., Marcketta, A., Kosmicki, J.A., Ziyatdinov, A.,
1114 Benner, C., O'Dushlaine, C., Barber, M., Boutkov, B., et al. (2021). Computationally
1115 efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* 53, 1097–
1116 1103.
- 1117 49. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015).
1118 Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*
1119 4, 7.
- 1120 50. Privé, F., Luu, K., Blum, M.G.B., McGrath, J.J., and Vilhjálmsson, B.J. (2020). Efficient
1121 toolkit implementing best practices for principal component analysis of population genetic
1122 data. *Bioinformatics* 36, 4449–4457.
- 1123 51. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang,
1124 H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976
1125 haplotypes for genotype imputation. *Nat. Genet.* 48, 1279–1283.
- 1126 52. Chou, W.-C., Zheng, H.-F., Cheng, C.-H., Yan, H., Wang, L., Han, F., Richards, J.B.,
1127 Karasik, D., Kiel, D.P., and Hsu, Y.-H. (2016). A combined reference panel from the
1128 1000 Genomes and UK10K projects improved rare variant imputation in European and
1129 Chinese samples. *Sci. Rep.* 6, 39313.
- 1130 53. Taliun, D., Harris, D.N., Kessler, M.D., Carlson, J., Szpiech, Z.A., Torres, R., Taliun,
1131 S.A.G., Corvelo, A., Gogarten, S.M., Kang, H.M., et al. (2021). Sequencing of 53,831
1132 diverse genomes from the NHLBI TOPMed Program. *Nature* 590, 290–299.
- 1133 54. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison,
1134 E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., et al. (2015). A
1135 global reference for human genetic variation. *Nature* 526, 68–74.
- 1136 55. Murphy, A.E., Schilder, B.M., and Skene, N.G. (2021). MungeSumstats: A Bioconductor
1137 package for the standardisation and quality control of many GWAS summary statistics.
1138 *Bioinformatics*.
- 1139 56. Bulik-Sullivan, B., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia
1140 Working Group of the Psychiatric Genomics Consortium, Patterson, N., Daly, M.J., Price,
1141 A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from
1142 polygenicity in genome-wide association studies. *Nat. Genet.* 47, 291–295.
- 1143 57. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R.,
1144 ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia
1145 Nervosa of the Wellcome Trust Case Control Consortium 3, Duncan, L., et al. (2015). An
1146 atlas of genetic correlations across human diseases and traits. *Nat. Genet.* 47, 1236–1241.
- 1147 58. Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis
1148 of genomewide association scans. *Bioinformatics* 26, 2190–2191.
- 1149 59. Turley, P., Walters, R.K., Maghzian, O., Okbay, A., Lee, J.J., Fontana, M.A., Nguyen-
1150 Viet, T.A., Wedow, R., Zacher, M., Furlotte, N.A., et al. (2018). Multi-trait analysis of
1151 genome-wide association summary statistics using MTAG. *Nat. Genet.* 50, 229–237.
- 1152 60. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional
1153 mapping and annotation of genetic associations with FUMA. *Nat. Commun.* 8, 1826.
- 1154 61. Jaffe, A.E., Straub, R.E., Shin, J.H., Tao, R., Gao, Y., Collado-Torres, L., Kam-Thong,
1155 T., Xi, H.S., Quan, J., Chen, Q., et al. (2018). Developmental and genetic regulation of the
1156 human cortex transcriptome illuminate schizophrenia pathogenesis. *Nat. Neurosci.* 21, 1117–
1157 1125.
- 1158 62. Wang, D., Liu, S., Warrell, J., Won, H., Shi, X., Navarro, F.C.P., Clarke, D., Gu, M.,
1159 Emani, P., Yang, Y.T., et al. (2018). Comprehensive functional genomic resource and

- 1160 integrative model for the human brain. *Science* 362,
1161 63. Fromer, M., Roussos, P., Sieberts, S.K., Johnson, J.S., Kavanagh, D.H., Perumal, T.M.,
1162 Ruderfer, D.M., Oh, E.C., Topol, A., Shah, H.R., et al. (2016). Gene expression elucidates
1163 functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* 19, 1442–1453.
1164 64. Ramasamy, A., Trabzuni, D., Guelfi, S., Varghese, V., Smith, C., Walker, R., De, T., UK
1165 Brain Expression Consortium, North American Brain Expression Consortium, Coin, L., et al.
1166 (2014). Genetic variability in the regulation of gene expression in ten regions of the human
1167 brain. *Nat. Neurosci.* 17, 1418–1428.
1168 65. Westra, H.-J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J.,
1169 Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., et al. (2013). Systematic
1170 identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.*
1171 45, 1238–1243.
1172 66. Zhernakova, D.V., Deelen, P., Vermaat, M., van Iterson, M., van Galen, M., Arindrarto,
1173 W., van 't Hof, P., Mei, H., van Dijk, F., Westra, H.-J., et al. (2017). Identification of context-
1174 dependent expression quantitative trait loci in whole blood. *Nat. Genet.* 49, 139–145.
1175 67. Vösa, U., Claringbould, A., Westra, H.-J., Bonder, M.J., Deelen, P., Zeng, B., Kirsten, H.,
1176 Saha, A., Kreuzhuber, R., Kasela, S., et al. (2018). Unraveling the polygenic architecture of
1177 complex traits using blood eQTL meta-analysis. *BioRxiv*.
1178 68. Buil, A., Brown, A.A., Lappalainen, T., Viñuela, A., Davies, M.N., Zheng, H.-F.,
1179 Richards, J.B., Glass, D., Small, K.S., Durbin, R., et al. (2015). Gene-gene and gene-
1180 environment interactions detected by transcriptome sequence analysis in twins. *Nat. Genet.*
1181 47, 88–91.
1182 69. Ng, B., White, C.C., Klein, H.-U., Sieberts, S.K., McCabe, C., Patrick, E., Xu, J., Yu, L.,
1183 Gaiteri, C., Bennett, D.A., et al. (2017). An xQTL map integrates the genetic architecture of
1184 the human brain's transcriptome and epigenome. *Nat. Neurosci.* 20, 1418–1426.
1185 70. van Rheenen, W., Peyrot, W.J., Schork, A.J., Lee, S.H., and Wray, N.R. (2019). Genetic
1186 correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* 20, 567–
1187 581.
1188 71. Ning, Z., Pawitan, Y., and Shen, X. (2020). High-definition likelihood inference of
1189 genetic correlations across human complex traits. *Nat. Genet.* 52, 859–864.
1190 72. Werme, J., van der Sluis, S., Posthuma, D., and de Leeuw, C.A. (2022). An integrated
1191 framework for local genetic correlation analysis. *Nat. Genet.* 54, 274–282.
1192 73. Spargo, T.P., Gilchrist, L., Hunt, G.P., Dobson, R.J., Proitsi, P.P., Al-Chalabi, A., Pain,
1193 O., and Iacoangeli, A. (2023). Statistical examination of shared loci in neuropsychiatric
1194 diseases using genome-wide association study summary statistics. *MedRxiv*.
1195 74. Zou, Y., Carbonetto, P., Wang, G., and Stephens, M. (2022). Fine-mapping from
1196 summary data with the “Sum of Single Effects” model. *PLoS Genet.* 18, e1010299.
1197 75. Wallace, C. (2021). A more accurate method for colocalisation analysis allowing for
1198 multiple causal variants. *PLoS Genet.* 17, e1009440.
1199 76. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace,
1200 C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic
1201 association studies using summary statistics. *PLoS Genet.* 10, e1004383.
1202 77. Sanderson, E., Glymour, M.M., Holmes, M.V., Kang, H., Morrison, J., Munafò, M.R.,
1203 Palmer, T., Schooling, C.M., Wallace, C., Zhao, Q., et al. (2022). Mendelian randomization.
1204 *Nat. Rev. Methods Primers* 2, 6.
1205 78. Burgess, S., Davies, N.M., and Thompson, S.G. (2016). Bias due to participant overlap in
1206 two-sample Mendelian randomization. *Genet. Epidemiol.* 40, 597–608.
1207 79. Morrison, J., Knoblauch, N., Marcus, J.H., Stephens, M., and He, X. (2020). Mendelian
1208 randomization accounting for correlated and uncorrelated pleiotropic effects using genome-
1209 wide summary statistics. *Nat. Genet.* 52, 740–747.

- 1210 80. Hemani, G., Bowden, J., and Davey Smith, G. (2018). Evaluating the potential role of
1211 pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* *27*, R195–R208.
- 1212 81. Slob, E.A.W., and Burgess, S. (2020). A comparison of robust Mendelian randomization
1213 methods using summary data. *Genet. Epidemiol.* *44*, 313–329.
- 1214 82. Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with
1215 invalid instruments: effect estimation and bias detection through Egger regression. *Int. J.*
1216 *Epidemiol.* *44*, 512–525.
- 1217 83. Bowden, J., Davey Smith, G., Haycock, P.C., and Burgess, S. (2016). Consistent
1218 Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted
1219 Median Estimator. *Genet. Epidemiol.* *40*, 304–314.
- 1220 84. Hartwig, F.P., Davey Smith, G., and Bowden, J. (2017). Robust inference in summary
1221 data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.*
1222 *46*, 1985–1998.
- 1223 85. Verbanck, M., Chen, C.-Y., Neale, B., and Do, R. (2018). Detection of widespread
1224 horizontal pleiotropy in causal relationships inferred from Mendelian randomization between
1225 complex traits and diseases. *Nat. Genet.* *50*, 693–698.
- 1226 86. Burgess, S., and Thompson, S.G. (2017). Interpreting findings from Mendelian
1227 randomization using the MR-Egger method. *Eur J Epidemiol* *32*, 377–389.
- 1228 87. Burgess, S., Bowden, J., Fall, T., Ingelsson, E., and Thompson, S.G. (2017). Sensitivity
1229 Analyses for Robust Causal Inference from Mendelian Randomization Analyses with
1230 Multiple Genetic Variants. *Epidemiology* *28*, 30–42.
- 1231 88. Lawlor, D.A., Harbord, R.M., Sterne, J.A.C., Timpson, N., and Davey Smith, G. (2008).
1232 Mendelian randomization: using genes as instruments for making causal inferences in
1233 epidemiology. *Stat. Med.* *27*, 1133–1163.
- 1234 89. Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N.A., and
1235 Thompson, J.R. (2016). Assessing the suitability of summary data for two-sample Mendelian
1236 randomization analyses using MR-Egger regression: the role of the I² statistic. *Int. J.*
1237 *Epidemiol.* *45*, 1961–1974.
- 1238 90. Mahley, R.W. (2016). Apolipoprotein E: from cardiovascular disease to
1239 neurodegenerative disorders. *J. Mol. Med.* *94*, 739–746.
- 1240 91. Liu, S., Liu, J., Weng, R., Gu, X., and Zhong, Z. (2019). Apolipoprotein E gene
1241 polymorphism and the risk of cardiovascular disease and type 2 diabetes. *BMC Cardiovasc.*
1242 *Disord.* *19*, 213.
- 1243 92. Xuan, L., Zhao, Z., Jia, X., Hou, Y., Wang, T., Li, M., Lu, J., Xu, Y., Chen, Y., Qi, L., et
1244 al. (2018). Type 2 diabetes is causally associated with depression: a Mendelian randomization
1245 analysis. *Front Med* *12*, 678–687.
- 1246 93. Lu, Y., Wang, Z., Georgakis, M.K., Lin, H., and Zheng, L. (2021). Genetic liability to
1247 depression and risk of coronary artery disease, myocardial infarction, and other
1248 cardiovascular outcomes. *J. Am. Heart Assoc.* *10*, e017986.
- 1249 94. Lord, J., Jermy, B., Green, R., Wong, A., Xu, J., Legido-Quigley, C., Dobson, R.,
1250 Richards, M., and Proitsi, P. (2021). Mendelian randomization identifies blood metabolites
1251 previously linked to midlife cognition as causal candidates in Alzheimer’s disease. *Proc.*
1252 *Natl. Acad. Sci. USA* *118*,.
- 1253 95. Choi, S.W., Mak, T.S.-H., and O’Reilly, P.F. (2020). Tutorial: a guide to performing
1254 polygenic risk score analyses. *Nat. Protoc.* *15*, 2759–2772.
- 1255 96. Zhang, Q., Privé, F., Vilhjálmsson, B., and Speed, D. (2021). Improved genetic prediction
1256 of complex traits from individual-level data or summary statistics. *Nat. Commun.* *12*, 4192.
- 1257 97. Lovestone, S., Francis, P., Kloszewska, I., Mecocci, P., Simmons, A., Soyninen, H.,
1258 Spenger, C., Tsolaki, M., Vellas, B., Wahlund, L.-O., et al. (2009). AddNeuroMed--the
1259 European collaboration for the discovery of novel biomarkers for Alzheimer’s disease. *Ann.*

- 1260 N. Y. Acad. Sci. *1180*, 36–46.
- 1261 98. Lord, J., Green, R., Choi, S.W., Hübel, C., Aarsland, D., Velayudhan, L., Sham, P.,
1262 Legido-Quigley, C., Richards, M., Dobson, R., et al. (2022). Disentangling independent and
1263 mediated causal relationships between blood metabolites, cognitive factors, and alzheimer’s
1264 disease. *Biological Psychiatry Global Open Science* *2*, 167–179.
- 1265 99. Yang, J., Weedon, M.N., Purcell, S., Lettre, G., Estrada, K., Willer, C.J., Smith, A.V.,
1266 Ingelsson, E., O’Connell, J.R., Mangino, M., et al. (2011). Genomic inflation factors under
1267 polygenic inheritance. *Eur. J. Hum. Genet.* *19*, 807–812.
- 1268 100. Voskobiynyk, Y., Roth, J.R., Cochran, J.N., Rush, T., Carullo, N.V., Mesina, J.S.,
1269 Waqas, M., Vollmer, R.M., Day, J.J., McMahon, L.L., et al. (2020). Alzheimer’s disease risk
1270 gene BIN1 induces Tau-dependent network hyperexcitability. *Elife* *9*.
- 1271 101. Lambert, E., Saha, O., Soares Landeira, B., Melo de Farias, A.R., Hermant, X., Carrier,
1272 A., Pelletier, A., Gadaut, J., Davoine, L., Dupont, C., et al. (2022). The Alzheimer
1273 susceptibility gene BIN1 induces isoform-dependent neurotoxicity through early endosome
1274 defects. *Acta Neuropathol. Commun.* *10*, 4.
- 1275 102. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P.,
1276 Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of
1277 genome-wide association scan results. *Bioinformatics* *26*, 2336–2337.
- 1278 103. Dashti, H.S., Daghlas, I., Lane, J.M., Huang, Y., Udler, M.S., Wang, H., Ollila, H.M.,
1279 Jones, S.E., Kim, J., Wood, A.R., et al. (2021). Genetic determinants of daytime napping and
1280 effects on cardiometabolic health. *Nat. Commun.* *12*, 900.
- 1281 104. Jansen, P.R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A.R., de
1282 Leeuw, C.A., Benjamins, J.S., Muñoz-Manchado, A.B., Nagel, M., et al. (2019). Genome-
1283 wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional
1284 pathways. *Nat. Genet.* *51*, 394–403.
- 1285 105. Hammerschlag, A.R., Stringer, S., de Leeuw, C.A., Sniekers, S., Taskesen, E.,
1286 Watanabe, K., Blanken, T.F., Dekker, K., Te Lindert, B.H.W., Wassing, R., et al. (2017).
1287 Genome-wide association analysis of insomnia complaints identifies risk genes and genetic
1288 overlap with psychiatric and metabolic traits. *Nat. Genet.* *49*, 1584–1592.
- 1289 106. Loos, R.J.F., and Yeo, G.S.H. (2022). The genetics of obesity: from discovery to
1290 biology. *Nat. Rev. Genet.* *23*, 120–133.
- 1291 107. Zarza-Rebollo, J.A., Molina, E., and Rivera, M. (2021). The role of the FTO gene in the
1292 relationship between depression and obesity. A systematic review. *Neurosci. Biobehav. Rev.*
1293 *127*, 630–637.
- 1294 108. González, L.M., García-Herráiz, A., Mota-Zamorano, S., Flores, I., Albuquerque, D.,
1295 and Gervasini, G. (2021). Variants in the Obesity-Linked FTO gene locus modulates
1296 psychopathological features of patients with Anorexia Nervosa. *Gene* *783*, 145572.
- 1297 109. Li, Y., Dang, X., Chen, R., Wang, J., Li, S., Mitchell, B.L., Yao, Y.-G., Li, M., Li, T.,
1298 Zhang, Z., et al. (2023). Cross-ancestry genome-wide association study and systems-level
1299 integrative analyses implicate new risk genes and therapeutic targets for depression.
1300 MedRxiv.
- 1301 110. Dall’Aglia, L., Lewis, C.M., and Pain, O. (2021). Delineating the genetic component of
1302 gene expression in major depression. *Biol. Psychiatry* *89*, 627–636.
- 1303 111. Fabbri, C., Pain, O., Hagenars, S.P., Lewis, C.M., and Serretti, A. (2021).
1304 Transcriptome-wide association study of treatment-resistant depression and depression
1305 subtypes for drug repurposing. *Neuropsychopharmacology* *46*, 1821–1829.
- 1306 112. Koskinen, M.-K., and Hovatta, I. (2023). Genetic insights into the neurobiology of
1307 anxiety. *Trends Neurosci.* *46*, 318–331.
- 1308 113. Nagel, M., Jansen, P.R., Stringer, S., Watanabe, K., de Leeuw, C.A., Bryois, J., Savage,
1309 J.E., Hammerschlag, A.R., Skene, N.G., Muñoz-Manchado, A.B., et al. (2018). Meta-analysis

- 1310 of genome-wide association studies for neuroticism in 449,484 individuals identifies novel
1311 genetic loci and pathways. *Nat. Genet.* *50*, 920–927.
- 1312 114. Mallard, T.T., Linnér, R.K., Grotzinger, A.D., Sanchez-Roige, S., Seidlitz, J., Okbay,
1313 A., de Vlaming, R., Meddens, S.F.W., Bipolar Disorder Working Group of the Psychiatric
1314 Genomics Consortium, Palmer, A.A., et al. (2022). Multivariate GWAS of psychiatric
1315 disorders and their cardinal symptoms reveal two dimensions of cross-cutting genetic
1316 liabilities. *Cell Genomics* *2*,
- 1317 115. Lüningschrör, P., Werner, G., Stroobants, S., Kakuta, S., Dombert, B., Sinske, D.,
1318 Wanner, R., Lüllmann-Rauch, R., Wefers, B., Wurst, W., et al. (2020). The FTL risk factor
1319 TMEM106B regulates the transport of lysosomes at the axon initial segment of motoneurons.
1320 *Cell Rep.* *30*, 3506–3519.e6.
- 1321 116. Van Deerlin, V.M., Sleiman, P.M.A., Martinez-Lage, M., Chen-Plotkin, A., Wang, L.-
1322 S., Graff-Radford, N.R., Dickson, D.W., Rademakers, R., Boeve, B.F., Grossman, M., et al.
1323 (2010). Common variants at 7p21 are associated with frontotemporal lobar degeneration with
1324 TDP-43 inclusions. *Nat. Genet.* *42*, 234–239.
- 1325 117. Nicholson, A.M., and Rademakers, R. (2016). What we know about TMEM106B in
1326 neurodegeneration. *Acta Neuropathol.* *132*, 639–651.
- 1327 118. Feng, T., Lacrampe, A., and Hu, F. (2021). Physiological and pathological functions of
1328 TMEM106B: a gene associated with brain aging and multiple brain disorders. *Acta*
1329 *Neuropathol.* *141*, 327–339.
- 1330 119. Perneel, J., and Rademakers, R. (2022). Identification of TMEM106B amyloid fibrils
1331 provides an updated view of TMEM106B biology in health and disease. *Acta Neuropathol.*
1332 *144*, 807–819.
- 1333 120. Mao, F., Robinson, J.L., Unger, T., Posavi, M., Amado, D.A., Elman, L., Grossman, M.,
1334 Wolk, D.A., Lee, E.B., Van Deerlin, V.M., et al. (2021). TMEM106B modifies TDP-43
1335 pathology in human ALS brain and cell-based models of TDP-43 proteinopathy. *Acta*
1336 *Neuropathol.* *142*, 629–642.
- 1337 121. Vass, R., Ashbridge, E., Geser, F., Hu, W.T., Grossman, M., Clay-Falcone, D., Elman,
1338 L., McCluskey, L., Lee, V.M.Y., Van Deerlin, V.M., et al. (2011). Risk genotypes at
1339 TMEM106B are associated with cognitive impairment in amyotrophic lateral sclerosis. *Acta*
1340 *Neuropathol.* *121*, 373–380.
- 1341 122. Shafit-Zagardo, B., Sidoli, S., Goldman, J.E., DuBois, J.C., Corboy, J.R., Strittmatter,
1342 S.M., Guzik, H., Graff, S., and Nagra, R.M. (2022). TMEM106B is increased in Multiple
1343 Sclerosis plaques, and deletion causes accumulation of lipid after demyelination. *BioRxiv*.
- 1344 123. Hong, S., Dobricic, V., Ohlei, O., Bos, I., Vos, S.J.B., Prokopenko, D., Tijms, B.M.,
1345 Andreasson, U., Blennow, K., Vandenberghe, R., et al. (2021). TMEM106B and CPOX are
1346 genetic determinants of cerebrospinal fluid Alzheimer’s disease biomarker levels. *Alzheimers*
1347 *Dement.* *17*, 1628–1640.
- 1348 124. Dhiman, K., Gupta, V.B., Villemagne, V.L., Eratne, D., Graham, P.L., Fowler, C.,
1349 Bourgeat, P., Li, Q.-X., Collins, S., Bush, A.I., et al. (2020). Cerebrospinal fluid
1350 neurofilament light concentration predicts brain atrophy and cognition in Alzheimer’s
1351 disease. *Alzheimers Dement (Amst)* *12*, e12005.
- 1352 125. Chen, M.-H., Liu, Y.-L., Kuo, H.-W., Tsai, S.-J., Hsu, J.-W., Huang, K.-L., Tu, P.-C.,
1353 and Bai, Y.-M. (2022). Neurofilament light chain is a novel biomarker for major depression
1354 and related executive dysfunction. *Int. J. Neuropsychopharmacol.* *25*, 99–105.
- 1355 126. Brodaty, H., and Donkin, M. (2009). Family caregivers of people with dementia.
1356 *Dialogues Clin Neurosci* *11*, 217–228.
- 1357 127. Lindeza, P., Rodrigues, M., Costa, J., Guerreiro, M., and Rosa, M.M. (2020). Impact of
1358 dementia on informal care: a systematic review of family caregivers’ perceptions. *BMJ*
1359 *Support. Palliat. Care.*

- 1360 128. Victor, C.R., Rippon, I., Quinn, C., Nelis, S.M., Martyr, A., Hart, N., Lamont, R., and
1361 Clare, L. (2021). The prevalence and predictors of loneliness in caregivers of people with
1362 dementia: findings from the IDEAL programme. *Aging Ment. Health* 25, 1232–1238.
- 1363 129. Huang, S.-S. (2022). Depression among caregivers of patients with dementia:
1364 Associative factors and management approaches. *World J. Psychiatry* 12, 59–76.
- 1365 130. Chesney, E., Goodwin, G.M., and Fazel, S. (2014). Risks of all-cause and suicide
1366 mortality in mental disorders: a meta-review. *World Psychiatry* 13, 153–160.
- 1367 131. Fry, A., Littlejohns, T.J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins,
1368 R., and Allen, N.E. (2017). Comparison of Sociodemographic and Health-Related
1369 Characteristics of UK Biobank Participants With Those of the General Population. *Am. J.*
1370 *Epidemiol.* 186, 1026–1034.
- 1371 132. Brommelhoff, J.A., Gatz, M., Johansson, B., McArdle, J.J., Fratiglioni, L., and
1372 Pedersen, N.L. (2009). Depression as a risk factor or prodromal feature for dementia?
1373 Findings in a population-based sample of Swedish twins. *Psychol. Aging* 24, 373–384.
- 1374 133. Creese, B., and Ismail, Z. (2022). Mild behavioral impairment: measurement and
1375 clinical correlates of a novel marker of preclinical Alzheimer’s disease. *Alzheimers Res.*
1376 *Ther.* 14, 2.
- 1377 134. DeMichele-Sweet, M.A.A., Klei, L., Creese, B., Harwood, J.C., Weamer, E.A.,
1378 McClain, L., Sims, R., Hernandez, I., Moreno-Grau, S., Tárraga, L., et al. (2021). Genome-
1379 wide association identifies the first risk loci for psychosis in Alzheimer disease. *Mol.*
1380 *Psychiatry* 26, 5797–5811.
- 1381