

# 1 Evaluating metabolome-wide causal effects on risk for 2 psychiatric and neurodegenerative disorders

3 Lachlan Gilchrist<sup>1,2</sup>, Julian Mutz<sup>1</sup>, Pirro Hysi<sup>3,4,5</sup>, Cristina Legido-Quigley<sup>6,7</sup>, Sulev  
4 Koks<sup>8,2</sup>, Cathryn M. Lewis<sup>1,9,\*</sup> and Petroula Proitsi<sup>10,11,\*</sup>

5  
6 <sup>1</sup> Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology &  
7 Neuroscience, King's College London, London, United Kingdom

8 <sup>2</sup> Perron Institute for Neurological and Translational Science, Perth, Australia

9 <sup>3</sup> Department of Ophthalmology, King's College London, London, UK

10 <sup>4</sup> Department of Twins Research and Genetic Epidemiology, King's College London, London, UK

11 <sup>5</sup> Sørlandet Sykehus Arendal, Arendal, Norway

12 <sup>6</sup> Systems Medicine, Institute of Pharmaceutical Science, Life Science & Medicine, King's College  
13 London, London, UK

14 <sup>7</sup> Steno Diabetes Center Copenhagen, Copenhagen, Denmark

15 <sup>8</sup> Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, Australia

16 <sup>9</sup> Department of Medical and Molecular Genetics, King's College London, London, United Kingdom

17 <sup>10</sup> Centre for Preventive Neurology, Wolfson Institute of Population Health, Queen Mary's University of  
18 London

19 <sup>11</sup> Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and  
20 Neuroscience, King's College London, London, United Kingdom

21  
22 \* equally contributing authors/co-corresponding authors: Dr. Petroula Proitsi  
23 ([p.proitsi@qmul.ac.uk](mailto:p.proitsi@qmul.ac.uk)) and Prof. Cathryn Lewis ([cathryn.lewis@kcl.ac.uk](mailto:cathryn.lewis@kcl.ac.uk))

24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## Abstract

41

42 Evidence indicates phenotypic and biological overlap between psychiatric and  
43 neurodegenerative disorders. Further identification of underlying mutual and unique biological  
44 mechanisms may yield novel multi-disorder and disorder-specific therapeutic targets. The  
45 metabolome represents an important domain for target identification as metabolites play  
46 critical roles in modulating a diverse range of biological processes. Here, we used Mendelian  
47 randomisation (MR) to test the causal effects of ~1000 plasma metabolites and ~300  
48 metabolite ratios on anxiety, bipolar disorder, depression, schizophrenia, amyotrophic lateral  
49 sclerosis, Alzheimer's disease, Parkinson's disease and multiple sclerosis. In total, 85 causal  
50 effects involving 77 unique metabolites passed FDR correction and robust sensitivity analyses  
51 (IVW-MR OR range: 0.73-1.48;  $p_{FDR} < 0.05$ ). No evidence of reverse causality was identified.  
52 Multivariate analyses implicated sphingolipid metabolism in psychiatric disorder risk and  
53 carnitine derivatives in risk for amyotrophic lateral sclerosis and multiple sclerosis. However,  
54 polygenic risk scores for prioritised metabolites showed limited prediction in the UK Biobank.  
55 Downstream colocalisation in regions containing influential variants identified greater than  
56 suggestive evidence ( $PP.H_4 \geq 0.6$ ) for a shared causal variant for 29 metabolite/psychiatric  
57 disorder trait-pairs on chromosome 11 at the *FADS* gene cluster. Most of these metabolites  
58 were lipids containing linoleic or arachidonic acid. Additional colocalisation was identified  
59 between the ratio of histidine-to-glutamine, glutamine, Alzheimer's disease and *SPRYD4* gene  
60 expression on chromosome 12. Although no single metabolite had a causal effect on a  
61 psychiatric and a neurodegenerative disease, results suggest a broad effect of lipids across  
62 brain disorders. Metabolites identified here may help inform future targeted interventions.

## 63 1. Introduction

64 Psychiatric disorders show substantial co-morbidity<sup>1</sup> and shared genetic  
65 architecture<sup>2</sup>. Additionally, they are common both prior to and within  
66 neurodegenerative diseases<sup>3-7</sup> – themselves genetically related<sup>8,9</sup> – and are indicated  
67 as risk factors for their later onset<sup>10-12</sup>. Genomic analyses suggest that psychiatric and  
68 neurodegenerative risk are partially underpinned by shared risk loci and causal brain  
69 transcripts and proteins<sup>13-17</sup>. By further investigating shared biological risk  
70 mechanisms, novel interventions may be identified with effectiveness for specific or  
71 multiple disorders.

72 Metabolomics – the study of small molecules linked to metabolism – is a  
73 promising tool for such investigation as it represents the stage of the omics cascade  
74 closest to the phenotype<sup>18</sup>. Importantly, metabolites are strong drug target  
75 candidates as many are obtained dietarily and modulate diverse biological  
76 processes including gene expression, RNA metabolism and protein activity<sup>19-21</sup>.  
77 Obtaining samples is minimally invasive as thousands of metabolites can be detected  
78 in plasma, serum or urine<sup>22</sup>.

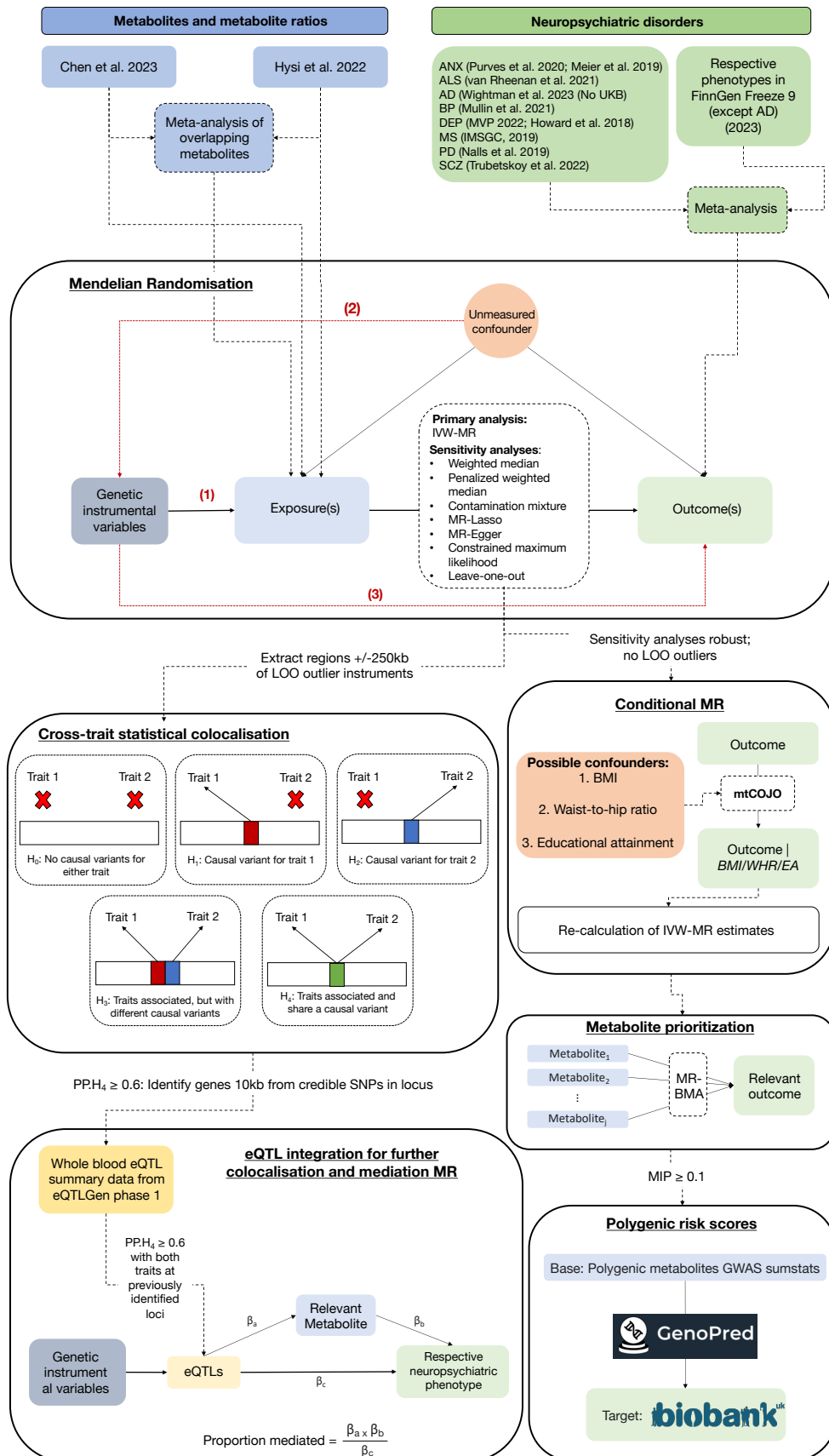
79 Similar metabolite groups, such as ceramides, are indicated to play a role in  
80 both psychiatric and neurodegenerative disorders<sup>23-27</sup>. However, observational  
81 studies may be confounded by reverse causation<sup>28</sup>. Importantly for psychiatric and  
82 neurodegenerative disorders, the metabolome can be perturbed by disease  
83 pathology and/or pharmaceutical interventions<sup>29-33</sup>. As such, following onset and  
84 treatment discerning the direction of effect and identifying causal metabolite targets  
85 is challenging. These limitations can be addressed using Mendelian randomisation  
86 (MR), a statistical framework that leverages genetic variants as instrumental variables  
87 (IVs) to proxy an exposure and infer its causal effect on an outcome of interest<sup>34</sup>. As  
88 risk alleles are randomly assorted and fixed from conception, MR is less vulnerable  
89 to reverse causality than classical epidemiological approaches<sup>28,35</sup>. Beneficially, this  
90 approach can be performed using summary statistics from large genome-wide  
91 association studies (GWAS) (two-sample MR (2SMR)), allowing researchers to  
92 perform increasingly powerful analyses<sup>36</sup>. This approach applied to the metabolome  
93 has already identified several causal risk metabolites for Alzheimer's disease, bipolar  
94 disorder, major depressive disorder, and multiple sclerosis<sup>37-41</sup>.

95 However, no study has yet been conducted to identify shared and unique  
96 metabolomic risk factors across the psychiatric and neurodegenerative disorder  
97 spectrum. In this study, we assess the causal effect of ~1000 unique plasma  
98 metabolites and ~300 metabolite ratios on four psychiatric disorders (anxiety, bipolar  
99 disorder, depression and schizophrenia) and four neurodegenerative diseases  
100 (amyotrophic lateral sclerosis (ALS), Alzheimer's disease (AD), Parkinson's disease

101 (PD) and multiple sclerosis (MS)). We further prioritise key metabolites using  
102 Bayesian model averaging, identify regions of shared causal variants with statistical  
103 colocalisation – integrating expression quantitative trait loci (eQTLs) – and assess the  
104 predictive performance of polygenic scores for risk metabolites on incident diseases  
105 in the UK Biobank (UKB).

106 **2. Methods**

107 An overview of the analyses is shown in Figure 1.



108

109 *Figure 1: Analysis flowchart. In the Mendelian randomisation section, the three key assumptions are*  
110 *illustrated and can be described as: (1) Instruments must be robustly associated with the exposure –*  
111 *usually defined as genome-wide statistical significance of the variants; (2) Instruments are not*  
112 *associated with a confounder and (3) Instruments are not associated with the outcome via pathways*  
113 *other than through the exposure, for example through horizontal pleiotropy. These assumptions are*  
114 *known as the (1) relevance, (2) independence and (3) exclusion restriction assumption, respectively.*  
115 *ANX = anxiety; ALS = amyotrophic lateral sclerosis; AD = Alzheimer's disease; BP = bipolar disorder;*  
116 *DEP = depression; MS = multiple sclerosis; PD = Parkinson's disease; SCZ = schizophrenia*

117

## 118 2.1 GWAS summary statistics

### 119 2.1.1 Neuropsychiatric disorders

120 An overview of the GWAS summary statistics is shown in **Table 1**. These consisted of  
121 individuals of European ancestry only. For each disorder, we performed inverse  
122 variance weighted (IVW) genome-wide meta-analysis across samples<sup>9,42–51</sup> using  
123 METAL<sup>52</sup> to maximise sample size and statistical power. Details on each sample can  
124 been seen in the original papers. For AD, we used summary statistics that did not  
125 contain family history-based proxy phenotyping from the UK Biobank given its noted  
126 influence on the effect direction of downstream analyses<sup>53</sup>. The *APOE* region  
127 (chr19:45,020,859–45,844,508 (GRCh37)) was excluded from the AD GWAS due to  
128 known pleiotropic effects on relevant non-AD diseases such as heart disease and  
129 hypercholesterolaemia<sup>54</sup>. It was not meta-analysed with FinnGen freeze 9 as with the  
130 others as the sample already contained samples from a previous FinnGen freeze.  
131 Genetic correlations between these disorders – calculated using Linkage  
132 Disequilibrium Score Regression (LDSC)<sup>55</sup> in GenomicSEM<sup>56</sup> v.0.0.5 – are described  
133 and plotted in **Supplementary Material 1** and **2**, respectively. Full output tables are  
134 in **Supplementary Tables 1** and **2**. Effective sample sizes were calculated as per  
135 Grotzinger et al.<sup>57</sup>, providing comparable sample size estimates representing  
136 equivalently powered GWAS with a 1:1 case/control ratio.

Category	Phenotype	GWAS sample	N Cases	N Controls	Total cases	Total controls	$N_{eff}$
Psychiatric	Anxiety	Purves et al.	25,453	58,113	70,228	354,864	226,035
		Meier et al. (IPSYCHE)	4,584	19,225			
		Finngen Freeze 9 (KRA_PSY_ANXIETY)	40,191	277,526			
	Bipolar disorder	Mullins et al.	41,917	371,549	48,923	651,818	129,398
		Finngen freeze 9 (F5_BIPO)	7,006	329,192			
	Depression	Million Veterans Project	83,810	166,405	298,038	824,848	825,105
		Psychiatric Genomics Consortium/UK Biobank (Howard et al.)	170,756	329,443			
		Finngen freeze 9 (F5_DEPRESSIO)	43,280	329,192			
	Schizophrenia	Trubetskoy et al.	53,386	77,258	66,447	354,784	176,178
Finngen freeze 9 (KRA_PSY_SCHIZODEL)		13,061	277,526				
Neurodegenerative	Alzheimer's disease				39,918	358,140	89,308
		Wightman et al. (excl. UKB)	39,918	358,140			
	Amyotrophic lateral sclerosis	van Rheenen et al.	27,205	110,881	27,686	281,548	115,084
		Finngen freeze 9 (G6_ALS)	483	170,667			
	Multiple Sclerosis	International Multiple Sclerosis Consortium	47,429	68,374	49,611	442,361	120,692
		Finngen freeze 9	2,182	373,987			
	Parkinson's disease	Nalls et al.	33,674	449,056	37,441	822,098	120,692
Finngen freeze 9		4,235	373,042				

138 *Table 1: Previously conducted genome-wide association study summary statistics for psychiatric and*  
139 *neurodegenerative disorders used in this study.*

140

### 141 2.1.2 Plasma metabolites

142 We obtained metabolite summary statistics from two studies conducted in non-  
143 overlapping samples: (1) 1,091 individual metabolites and 309 metabolite ratios  
144 measured in 8,299 unrelated individuals from the Canadian Longitudinal Study of  
145 Aging (CLSA)<sup>58</sup> and (2) 722 individual metabolites measured in 8,809 individuals in  
146 the NIHR UK Bioresource cohort<sup>59</sup>. In both studies, metabolites were detected and  
147 quantified in plasma by Metabolon Inc. using ultra-high performance liquid  
148 chromatography-tandem mass spectroscopy. Unidentified metabolites were  
149 excluded from analyses. We performed genome-wide meta-analysis as above for 431  
150 metabolites when the same metabolite was available in both samples for a meta-  
151 analysed sample size of 17,038.

152

### 153 2.1.3 GWAS standardisation

154 Prior to meta-analysis, summary statistics were standardised with MungeSumstats<sup>60</sup>  
155 version in R (v.4.2.1) using dbsnp 144 and the  
156 BSgenome.Hsapiens.1000genomes.hs37d5 reference genome available through  
157 Bioconductor (v.3.1.3). Missing rsids were obtained, duplicate and multi-allelic

158 variants excluded, and alleles aligned to the reference, flipping effect and frequency  
159 direction accordingly. Where necessary, MungeSumstats was used to lift over  
160 genome build coordinates from GRCh38 to GRCh37 via the UCSC Genome Browser  
161 chain file.

162

## 163 2.2 Mendelian randomisation

### 164 2.2.1. Metabolite instrument selection

165 Independent instrumental variables (IVs) were initially selected at genome-wide  
166 significance ( $p \leq 5 \times 10^{-8}$ ), clumping at an  $r^2$  threshold of 0.001 within 10,000 kb of the  
167 lead variants. Clumping was conducted using the ieugwasr package (v.0.1.4), PLINK  
168 (v.1.9)<sup>61</sup> binaries and the European sample of the 1000 Genomes phase 3 reference  
169 panel<sup>62</sup> ( $N = 503$ ), restricted to minor allele frequency (MAF)  $> 0.01$ . This is the same  
170 reference panel used by the MRC IEU OpenGWAS database<sup>63</sup>  
171 (<http://files.mrcieu.ac.uk/ld/1kg.v3.tgz>).

172 Where less than five IVs were available for any given exposure, a lower  $p$ -value  
173 threshold of  $5 \times 10^{-6}$  was used for IV selection as per previous analyses<sup>64</sup>. If less than  
174 5 IVs were available at the  $5 \times 10^{-6}$  threshold, this metabolite was excluded. Where an  
175 IV was not available for the outcome, proxy IVs (i.e., variants in linkage disequilibrium  
176 (LD) with the original instrument at  $r^2 > 0.8$ ) were identified using snappy (v.1.0)  
177 (<https://gitlab.com/richards-lab/vince.forgetta/snappy>) and same reference panel.  
178 Proxies were selected using the highest  $r^2$  value and closest genomic position. Prior  
179 to analysis, exposure and outcome were harmonised to the same effect allele, and  
180 strand ambiguous palindromic variants (MAF  $> 0.42$ ) dropped. Instrument strength  
181 was assessed via their F-statistic ( $\beta_{\text{exposure}}^2 / \text{SE}_{\text{exposure}}^2$ ), with weak instruments (F-  
182 statistic  $< 10$ ) excluded. If less than 5 IVs remained, the metabolite was again  
183 excluded. Instrument measurement error was assessed with the  $I^2_{G-X}$  statistic ( $< 0.9$   
184 suggestive of measurement error)<sup>65</sup>.

185 Following the above steps the effects of 428 meta-analysed metabolites were  
186 tested, plus 740 and 132 metabolites from the CLSA and NIHR UK Bioresource  
187 studies, respectively, for a total of 1300 unique metabolites/metabolite ratios.

188

### 189 2.2.2. MR analyses

190 We conducted the primary analyses using inverse variance weighted MR (IVW-  
191 MR) with multiplicative random effects as recommended<sup>66</sup>. If the IVW-MR estimate  
192 was statistically significant after Benjamini-Hochberg false discovery rate (FDR)  
193 correction<sup>67</sup> ( $p_{FDR} \leq 0.05$ , corrected for total number of tests *specific to each* outcome),  
194 six sensitivity analyses were conducted to assess its robustness to violations of the  
195 pleiotropy assumption and instrument validity. We implemented methods that assume



196 a majority valid instruments (weighted median<sup>68</sup>; penalised weighted median<sup>69</sup>),  
197 plurality valid instruments (constrained maximum likelihood (cML)<sup>70</sup>; contamination  
198 mixture model MR (MR-ContMix)<sup>71</sup>), and that exclude invalid outliers (MR-Lasso<sup>72</sup>).  
199 Additionally, we used MR-Egger<sup>73</sup>, which gives consistent causal estimates even if all  
200 instruments are not valid, provided pleiotropic effects are not correlated with variant-  
201 exposure associations. This is known as the Instrument Strength Independent of  
202 Direct Effect (InSIDE) assumption and is MR-Egger-specific. To pass sensitivity  
203 criteria, all sensitivity analysis point estimates needed to be directionally concordant  
204 with the IVW estimate and the majority ( $\geq 4$ ) statistically significant at  $p \leq 0.05$ . Where  
205 a metabolite passed sensitivity criteria, we performed reverse IVW-MR as above with  
206 the outcome as the exposure.

207 Pleiotropy was assessed using the MR-Egger intercept test (no evidence of  
208 pleiotropy:  $p_{Egger-Intercept} \geq 0.05$ ) and heterogeneity using Cochran's Q (no evidence of  
209 heterogeneity:  $p_{Q-Stat} \geq 0.05$ ). Metabolites passing sensitivity were further inspected  
210 using leave-one-out (LOO) analysis to assess whether causal estimates were driven  
211 by the inclusion of a single influential variant. Metabolites that passed sensitivity  
212 criteria and had no LOO outlier variant were considered polygenic metabolites and  
213 are referred to as such henceforth. Those passing sensitivity but with a LOO outlier  
214 variant were considered single instrument metabolites.

215

## 216 **2.2 Further analysis of polygenic metabolites**

### 217 **2.3.1 Conditional analysis**

218 Hysi et al.<sup>59</sup> noted that many metabolites are genetically correlated with phenotypes  
219 related to body composition and educational level. To assess whether any identified  
220 effects were driven in part by the genetic correlations between the metabolites and  
221 these phenotypes, we performed conditional GWAS of the neuropsychiatric disorders  
222 using multi-trait conditional and joint analysis (mtCOJO)<sup>74</sup> to remove the per-SNP  
223 effects of body mass index (BMI), BMI-adjusted waist-to-hip ratio (WHR), and  
224 educational attainment (EA).

225 We obtained GWAS summary statistics for BMI ( $N = 806,834$ ) and BMI-  
226 adjusted waist-to-hip ratio ( $N = 697,734$ ) from the GIANT consortium<sup>75,76</sup>. We  
227 additionally obtained summary statistics for educational attainment (EA) from the EA4  
228 GWAS meta-analysis, excluding 23andMe, from the Social Science Genetic  
229 Association Consortium (SSGAC)<sup>77</sup> ( $N = 765,283$ ).

230 IVW-MR estimates were recalculated for significant metabolite-  
231 neuropsychiatric disorder pairs post-mtCOJO and compared to the original  
232 estimates. Forest plots were visually inspected for evidence of attenuation between  
233 original and adjusted estimates.

234

### 235 **2.3.2 Metabolite prioritisation using MR-BMA**

236 Metabolites are highly genetically correlated<sup>58</sup>. Consequently, univariate MR  
237 estimates – which assume exposures are independent – may be biased by the  
238 pleiotropic effects of other risk metabolites. Such issues may be addressed using MR  
239 with Bayesian model averaging (MR-BMA)<sup>78</sup>. MR-BMA is a multivariable MR  
240 approach that ranks of causal effects among related exposures according to their  
241 independent causal signal. Unlike conventional multivariable MR methods, MR-BMA  
242 is well suited for the highly correlated nature of high-throughput exposures such as  
243 metabolites<sup>78</sup>. Instruments were selected by pooling those used in IVW-MR for all per-  
244 disorder statistically significant metabolites, excluding correlated instruments as  
245 before ( $r^2 > 0.001$ , window = 10,000 kb). We use a prior probability of 0.1 and a prior  
246 variance of 0.25 as previously<sup>78,79</sup>. Metabolites were ranked according to their  
247 marginal posterior probability (MIP), with  $MIP \geq 0.1$  interpreted as causal. To avoid  
248 multicollinearity between the exposures, we calculated a genetic correlation matrix  
249 using the linkage disequilibrium score regression (LDSC)<sup>80</sup> function of  
250 GenomicSEM<sup>56</sup>. As per the original MR-BMA study, where point estimates  $\geq 0.95$ , we  
251 excluded the metabolite correlated with the most other metabolites (if  $> one$ ) or  
252 retained only the metabolite with the most-significant IVW-MR estimate in the MR-  
253 BMA analysis (if only one).

254

### 255 **2.3.2 Cross-trait polygenic score (PGS) analysis**

256 To assess whether any of the polygenic metabolites prioritised in MR-BMA be used  
257 for disease prediction, we calculated polygenic scores (PGS) – the weighted sum of  
258 an individual's risk alleles<sup>81</sup> – within the reference standardised framework of the  
259 GenoPred pipeline<sup>82</sup> using individual level data from the UK Biobank (UKB)<sup>83</sup>. Within  
260 GenoPred, allele weights were calculated using MegaPRS<sup>84</sup>, a PGS method that uses  
261 the BLD-LDAK heritability model and improves prediction by incorporating  
262 information on allele frequency, LD and various functional annotations. MegaPRS has  
263 shown strong predictive performance compared to other PGS methods<sup>85,86</sup>.

264 UKB participants were excluded from analyses if they had unusual levels of  
265 heterozygosity, a call missing rate  $>2\%$  or discordant phenotypic and genetic sex ( $F_x$   
266  $< 0.9$  for males,  $F_x > 0.5$  for females). Individuals were filtered for relatedness up to  
267 3<sup>rd</sup> degree relatives based on KING kinship estimates ( $r < 0.044$ )<sup>87</sup> using  
268 GreedyRelate (v1.2.1) to remove one individual from each related pair, maximising  
269 sample size (<https://gitlab.com/choishingwan/GreedyRelated>). Analyses were  
270 restricted to individuals of European ancestry. The final sample contained 381,564  
271 individuals prior to any age-based exclusions.

272 Case/control status was defined using the first occurrence data fields from the  
273 UKB (**Supplementary Table 3**), where cases were ascertained using primary care  
274 Read v2 or Read CTV3 codes, hospital inpatient ICD-9 and ICD-10 codes, death  
275 record ICD-10 codes and medical condition codes self-reported at UKB assessment  
276 centre visits. Individuals were excluded if the code had no event date or if the date  
277 of event was before, the same year as or on the participants date of birth, or a date  
278 in the future. Associations between metabolite-PGS and respective neuropsychiatric  
279 disease was assessed using logistic regression, controlling for age, sex, genotyping  
280 batch, assessment centre and the first ten genetic principal components from the  
281 UKB. Within each model, the comparison group was restricted to individuals with no  
282 occurrence of any of the neuropsychiatric disorders. To provide age-appropriate  
283 controls for AD and PD, samples were restricted for those analyses to individuals  
284 aged  $\geq 60$  and  $\geq 50$  years, respectively. For DEP and ANX cases, individuals with an  
285 occurrence of BP or SCZ were excluded. The variance explained by each PGS was  
286 assessed using Nagelkerke's pseudo- $r^2$ . Sample sizes ranged from 140,154 to  
287 356,763, depending on the disorder, with case numbers ranging from 600 (ALS) to  
288 46,900 (DEP) (mean age range = 56.65-64.13; %male range = 46.32-49.35).

289

## 290 **2.4 Further analysis of single instrument metabolites**

### 291 **2.4.1 Statistical colocalisation**

292 Using LOO analysis, we identified IVs that significantly attenuated the IVW-MR  
293 estimate when excluded. We extracted regions  $\pm 250$  kb of these IVs and performed  
294 statistical colocalisation for each metabolite and the respective neuropsychiatric  
295 disorder within the COLOC-reporter pipeline<sup>88</sup>. Regions of interest were extracted and  
296 harmonised for each phenotype to contain only shared variants on the reference  
297 panel. Colocalisation was performed using the *coloc.abf* function of the COLOC  
298 package<sup>89</sup>. A posterior probability of 0.8 for hypothesis four (PP.H<sub>4</sub>) – presence of a  
299 shared causal variant between the phenotypes – was taken as evidence of  
300 colocalisation (PP.H<sub>4</sub>  $\geq 0.6$  taken as suggestive).

301

### 302 **2.4.2 eQTL colocalisation**

303 For regions of colocalisation (including suggestive), we identified genes located  
304 within 10 kb of credible SNPs using Ensembl and biomaRt (v.2.54.0). For each gene,  
305 we extracted *cis*-eQTLs from the whole blood eQTL summary data of the eQTLGen  
306 Consortium<sup>90</sup> ( $N = 31,684$ ) and performed colocalisation between the eQTLs and both  
307 phenotypes in each metabolite-neuropsychiatric disorder pair as above.

308

309

## 310 2.4.2 eQTL MR

311 For genes colocalising with both the metabolite and neuropsychiatric disorder a trait-  
312 pair, we extracted independent instruments for gene expression from the eQTL  
313 summary statistics ( $p \leq 5 \times 10^{-8}$ ;  $r^2 \leq 0.001$ ; window = 10,000 kb) and used IVW-MR  
314 to assess the causal effect of gene expression levels on the outcome and exposure.  
315 We examined the consistency of the gene expression causal effects using Summary-  
316 based Mendelian Randomisation (SMR)<sup>91</sup>, extracting eQTL probes for the relevant  
317 gene from blood-based eQTL summary datasets from Westra et al.<sup>92</sup> ( $N = 3,511$ ) and  
318 Lloyd-Jones et al.<sup>93</sup> ( $N = 2,765$ ). Heterogeneity was assessed using the  
319 HEterogeneity In Dependent Instruments (HEIDI) test, with  $p_{\text{HEIDI}} < 0.05$  indicative of  
320 heterogeneity. This is essentially a test of colocalisation.

321 Previous research has indicated that the causal effect of gene expression on  
322 a complex trait can be mediated by the effect of metabolites within a transcript-  
323 metabolite-trait causal triplet<sup>94</sup>. Where gene expression levels showed a significant  
324 causal effect on both an outcome and a metabolite, we calculated the proportion  
325 mediated using the product of coefficients method as per previous analyses<sup>95</sup>.  
326 Standard errors were estimated using the delta method and  $p$ -values drawn from a  
327 z-score distribution.

328

## 329 3. Results

### 330 3.1 Mendelian randomisation

331 We conducted a total of 10,327 tests of the causal effect of the metabolites and ratios  
332 on the eight neuropsychiatric outcomes (instrument  $N$ -range: 5-27). Of these, 1695  
333 tests were conducted with instruments clumped at  $5 \times 10^{-8}$  and 8632 with instruments  
334 clumped at  $5 \times 10^{-6}$ . For single instruments, F-statistics ranged from 10.11 to 5625.88  
335 (mean per-test F-statistics range: 20.77-1163.90). Analyses were thus not impacted  
336 by weak instruments. Per-test  $I^2_{G-X}$  ranged from 0.45 to 1.000. Only 94 tests (0.009%),  
337 involving 13 unique metabolites, were conducted with  $I^2_{G-X} < 0.9$ , suggestive of  
338 measurement error. Only 27 tests were conducted using instruments with  $I^2_{G-X} \leq 0.8$ ,  
339 involving six metabolites: 2-hydroxyarachidate, 4-hydroxy-2-oxoglutaric acid, 4-  
340 vinylguaiacol sulphate, glycolithocholate, and O-cresol sulphate.

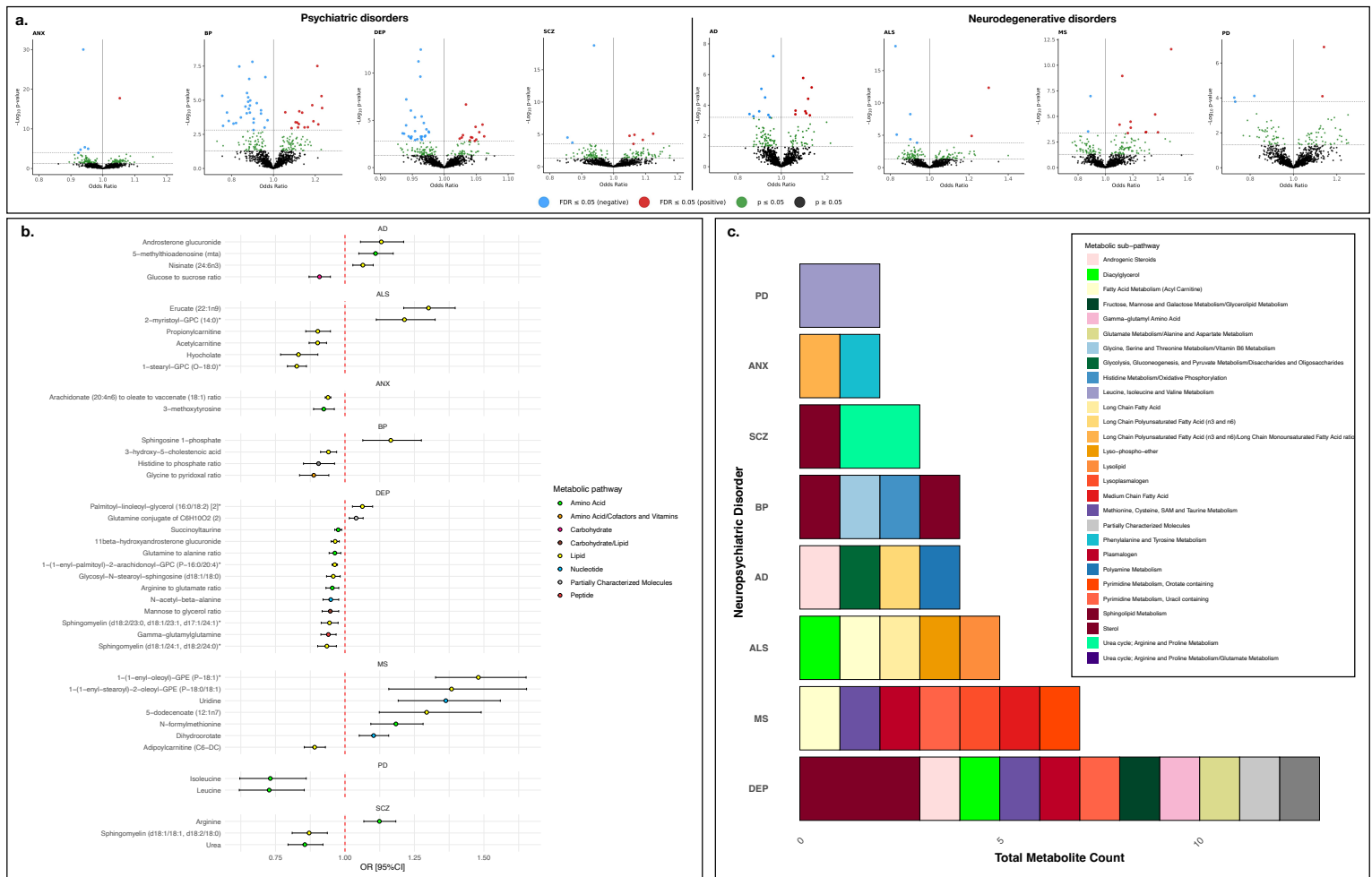
341 Primary IVW-MR analyses identified 138 causal effects involving 113 unique  
342 metabolites after outcome-specific FDR correction (**Supplementary Table 4; Figure**  
343 **2a**). Of these, 85 metabolite-outcome pairs passed our sensitivity criteria, involving  
344 77 unique metabolites (**Supplementary Table 5**). All neuropsychiatric disorders had  
345 two or more metabolites with statistically significant effects. No reverse effects were  
346 detected for these metabolite-outcome pairs after FDR-correction (**Supplementary**  
347 **Table 6**). IVW-MR odds ratios (OR) for metabolite-neuropsychiatric disorder trait-

348 pairs ranged from 0.73 to 1.48 ( $p$ -value range:  $1.52 \times 10^{-3}$  -  $9.67 \times 10^{-31}$ ;  $p_{FDR}$  range:  
349  $4.45 \times 10^{-2}$  -  $1.26 \times 10^{-27}$ ). However, MR-Egger intercept tests detected significant  
350 pleiotropy for eight of the 85 metabolite-outcome pairs ( $p_{Egger-Intercept} < 0.05$ )  
351 (**Supplementary Table 7**). Cochran's Q heterogeneity tests detected significant  
352 heterogeneity for 15 ( $p_{Q-Stat} < 0.05$ ) (**Supplementary Table 8**). All metabolite-outcome  
353 pairs with  $p_{Egger-Intercept} < 0.05$  and 14 out of 15 with  $p_{Q-Stat} < 0.05$  were identified as  
354 influenced by single influential variants in the LOO analysis (**Supplementary Table**  
355 **9**). Only the effect of the ratio of mannose to glycerol on depression ( $p_{Q-Stat} = 0.043$ )  
356 was not driven by a single influential variant.

357 Excluding metabolites with effect estimates primarily influenced by single  
358 variants left a total of 41 metabolite-neuropsychiatric disorder trait-pairs with  
359 polygenic metabolites – all unique. Of these, 13 were identified with depression,  
360 seven with MS, six with ALS, four each with AD and bipolar disorder, three with  
361 schizophrenia, and two each with anxiety and PD (**Figure 2b**). These metabolites  
362 were linked to nine broad metabolic classes (super-pathways), counting the  
363 combinations of pathways of metabolite ratios as unique pathways (**Figure 2b**;  
364 **Supplementary Table 10**). Over half ( $N = 22$ ) were lipids, and the rest amino acids  
365 ( $N = 10$ ), nucleotides ( $N = 3$ ), peptides ( $N = 1$ ), carbohydrates ( $N = 1$ ) or ratios  
366 involving amino acid/cofactors and vitamins ( $N = 1$ ), amino acid/energy ( $N = 1$ ) and  
367 carbohydrate/lipid ( $N = 1$ ) pathways. The remaining metabolite – glutamine conjugate  
368 of C6H10O2 (2) – is considered only partially characterised.

369 These metabolites were implicated in 28 unique sub-pathways representing  
370 the metabolic/biochemical subclasses related to the metabolites (**Figure 2c**;  
371 **Supplementary Table 10**). The largest group of metabolites was related to  
372 sphingolipid metabolism ( $N = 5$ ). All metabolites involved in sphingolipid metabolism  
373 had effects on psychiatric disorder risk. Sphingosine 1-phosphate has a risk  
374 increasing effect on bipolar disorder (OR [95%CI] = 1.17 [1.06-1.28],  $p$ -value =  $9.49$   
375  $\times 10^{-4}$ ,  $p_{FDR} = 0.03$ ), while glycosyl-N-stearoyl-sphingosine (d18:1/18:0),  
376 sphingomyelin (d18:1/24:1, d18:2/24:0) and sphingomyelin (d18:2/23:0, d18:1/23:1  
377 and d17:1/24:1) has risk decreasing effects on depression (OR range: 0.93-0.96,  $p$ -  
378 value range:  $2.38 \times 10^{-4}$  –  $9.91 \times 10^{-4}$ ;  $p_{FDR}$  range: 0.02-0.04). Additionally,  
379 sphingomyelin (d18:1/18:1, d18:2/18:0) has a risk decreasing effect on schizophrenia  
380 (OR [95%CI] = 0.87 [0.81-0.94],  $p$ -value =  $1.89 \times 10^{-4}$ ,  $p_{FDR} = 0.04$ ).

381



383

384

385 *Figure 2: (a). Volcano plots for anxiety (ANX), bipolar disorder (BP), depression (DEP), schizophrenia*  
 386 *(SCZ), Alzheimer's disease (AD), amyotrophic lateral sclerosis (ALS), multiple sclerosis (MS) and*  
 387 *Parkinson's disease (PD) showing odds ratios and  $-\log_{10}$  p-values from inverse variance weighted*  
 388 *Mendelian randomisation (IVW-MR) for all tested metabolites. (b). Forest plot of FDR statistically*  
 389 *significant metabolites associated with neuropsychiatric outcomes in IVW-MR and passing sensitivity*  
 390 *tests, with no leave-one-out (LOO) outliers (i.e., polygenic metabolites). Point estimates indicate odds*  
 391 *ratios with bars representing 95% confidence intervals (95%CI). Point estimates for each metabolite*  
 392 *are coloured according to their metabolic super-pathway. (c). Stacked histograms showing total count*  
 393 *of FDR statistically significant polygenic metabolites associated with each neuropsychiatric disorder,*  
 394 *with fill colours indicating metabolite sub-pathways. AD = Alzheimer's disease; ALS = Amyotrophic*  
 395 *Lateral Sclerosis; ANX = anxiety; BP = Bipolar Disorder; DEP = Depression; MS = Multiple Sclerosis;*  
 396 *PD = Parkinson's disease; SCZ = Schizophrenia*

397

398 The strongest negative effect was observed for the amino acid leucine and PD  
 399 (OR [95%CI] = 0.73 [0.62-0.85],  $p = 9.64 \times 10^{-5}$ ;  $p_{FDR} = 3.13 \times 10^{-2}$ ). Isoleucine showed  
 400 similar effects (OR = 0.73 [0.62-0.86];  $p = 1.65 \times 10^{-4}$ ;  $p_{FDR} = 4.28 \times 10^{-2}$ ). The strongest  
 401 positive effect was observed for the lipid 1-(1-enyl-oleoyl)-GPE (P-18:1) and MS (OR

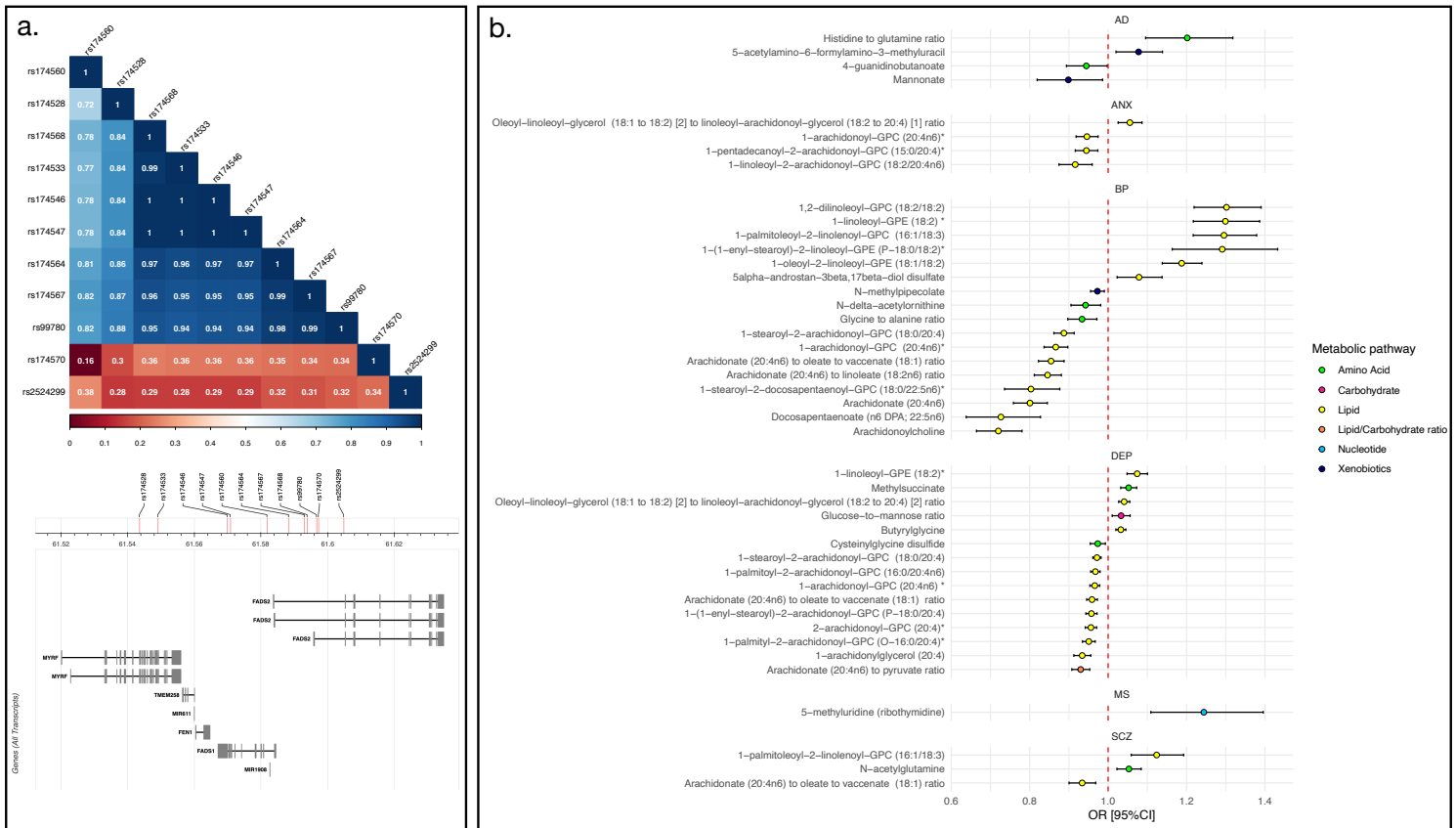
402 = 1.48 [1.33-1.65];  $p$ -value =  $2.71 \times 10^{-12}$ ;  $p_{FDR} = 3.42 \times 10^{-9}$ ). The most statistically  
403 significant effect was identified for the ratio of arachidonate (20:4n6) to oleate to  
404 vaccenate (18:1) and anxiety (OR = 0.93 [0.93-0.95],  $p$ -value =  $9.67 \times 10^{-31}$ ,  $p_{FDR} =$   
405  $1.26 \times 10^{-27}$ ).

406 The remaining 44 metabolite-neuropsychiatric disorder trait-pairs had  
407 attenuated effects in the LOO analysis, involving 36 unique metabolites and six of the  
408 neuropsychiatric outcomes (**Supplementary Table 11**). Most of the effects were  
409 observed in relation to psychiatric disorders ( $N = 40$ ), primarily with BP ( $N = 16$ ) and  
410 DEP ( $N = 16$ ). Another four were observed with ANX and three with SCZ. These  
411 metabolite-psychiatric trait-pairs involved 33 unique metabolites, 25 of which were  
412 lipids and one a ratio of lipid/carbohydrate. These lipids were involved in 12 unique  
413 sub-pathways, the largest group of which was related to phospholipid metabolism ( $N$   
414 = 6). The remainder were amino acids ( $N = 5$ ), a carbohydrate ( $N = 1$ ) and a  
415 xenobiotic ( $N = 1$ ), involved in seven unique sub-pathways. Of those remaining, four  
416 unique single instrument metabolites had a causal effect AD – two amino acids and  
417 two xenobiotics – and a nucleotide showed causal effect on MS.

418 A total of 31 of the LOO associations were driven by 11 influential variants on  
419 chromosome 11 within a ~500 kb window (61293499-61854782 bp). This region  
420 contains the fatty acid desaturase (*FADS*) gene cluster, which is linked to the  
421 regulation of fatty acid levels and their circulation<sup>96</sup>. All metabolites influenced by  
422 these variants were lipids and associated with psychiatric disorder risk ( $N_{ANX} = 4$ ,  $N_{BP}$   
423 = 13,  $N_{DEP} = 12$  and  $N_{SCZ} = 2$ ). Between-variant LD estimates for the influential  
424 instruments in this region were calculated using LDlinkR (v.1.3)<sup>97</sup> and ranged from  
425 0.16-1, with  $r^2 \geq 0.72$  for all variants except rs14570 and rs2524299 which showed a  
426 maximum  $r^2$  of 0.34 and 0.38, respectively (**Figure 3a; Supplementary Table 12**).

427 We used Wald ratio tests to re-estimate the effect of single instrument  
428 metabolites on the relevant outcome using the influential instrument only. F-statistics  
429 ranged from 105.37 to 5625.88. Wald estimates were directionally consistent with  
430 IVW-MR estimates and remained significant at  $p \leq 0.05$ , although the protective  
431 effects of 4-guanidinobutanoate and mannonate on AD were only nominally  
432 significant ( $p = 0.04$  and  $0.02$ , respectively) (**Figure 3b; Supplementary Table 13**).  
433 Regions containing influential instruments were further investigated using  
434 colocalisation.

435



436

437 *Figure 3: (a). The  $r^2$  estimates between influential single variants on chromosome 11 alongside their*  
 438 *genomic location. (b). Recalculated Wald ratio estimates for metabolites with influential single variants*  
 439 *using the influential variant only. Point estimates indicate odds ratios with bars representing 95%*  
 440 *confidence intervals (95%CI). Point estimates for each metabolite are coloured according to their*  
 441 *metabolic super-pathway. AD = Alzheimer's disease; ANX = anxiety; BP = Bipolar Disorder; DEP =*  
 442 *Depression; MS = Multiple Sclerosis; SCZ = Schizophrenia*

443

### 444 3.2 Further analysis of polygenic metabolites

#### 445 3.2.1 Conditional analysis

446 Conditioning neuropsychiatric outcomes on BMI, BMI-adjusted WHR or EA did not  
 447 attenuate the IVW-MR effect estimates for causally associated metabolites when re-  
 448 estimated (**Supplementary Table 15; Supplementary Material 3**). This suggests that  
 449 the metabolite associations were not influenced by any correlation between the  
 450 genetic instrument's effect and the effects of these traits as captured in the  
 451 neuropsychiatric GWAS.

452

#### 453 3.2.2 Metabolite prioritisation using MR-BMA

454 Prior to the per-outcome multivariate MR modelling in MR-BMA, we excluded the less  
 455 statistically significant of any metabolite pair correlated  $\geq 95\%$  (**Supplementary Table**  
 456 **15**). As such, we excluded 5-methylthioadenosine (mta) from the AD model due to its  
 457 correlation with the glucose to sucrose ratio, 5-dodecenoate (12:1n7) from the MS



458 model due to its correlation with N-formylmethionine and uridine, 2-myristoyl-GPC  
459 (14:0)\* from the ALS model due to its correlation with 1-stearoyl-GPC (O-18:0)\* and the  
460 glutamine to alanine ratio from the DEP model due to its correlation with gamma-  
461 glutamylglutamine.

462 For AD, anxiety, MS, PD and schizophrenia, all metabolites had MIPs  $\geq 0.1$ ,  
463 suggesting unique causal contributions to risk. For bipolar disorder, sphingosine 1-  
464 phosphate, the ratio of histidine to phosphate and the ratio of glycine to pyridoxal  
465 passed the MIP threshold, with 3-hydroxy-5-cholestenoic acid just under (MIP =  
466 0.098). However, of the 12 metabolites included for depression, only the two  
467 sphingomyelins (d18:1/24:1, d18:2/24:0 and d18:2/23:0, d18:1/23:1, d17:1/24:1) and  
468 1-(1-enyl-palmitoyl)-2-arachidonoyl-GPC (P-16:0/20:4) passed. For ALS, only  
469 acetylcarnitine passed the MIP threshold. Acetylcarnitine had the highest overall MIP  
470 of all metabolites tested (MIP = 0.893) (**Supplementary Table 16**).

471

### 472 3.2.3 Polygenic score (PGS) analysis

473 No MR-BMA prioritised metabolite PGS was associated with any of their respective  
474 outcomes after FDR correction for 23 tests ( $p_{FDR} \geq 0.05$ ). Nominal associations were  
475 observed between depression and the PGS for sphingomyelin (d18:2/23:0,  
476 d18:1/23:1, d17:1/24:1) and 1-(1-enyl-palmitoyl)-2-arachidonoyl-GPC (P-16:0/20:4)  
477 with identical effect sizes (OR[95%CI] = 0.989 [0.979-0.998],  $p$ -value range =  $2.131$   
478  $\times 10^{-2}$  -  $2.517 \times 10^{-2}$ ) and between the PGS for leucine and PD (OR [95%CI] = 0.962  
479 [0.93- 0.996,  $p$ -value =  $2.814 \times 10^{-2}$ ). These effects were directionally consistent with  
480 the MR estimates but had non-relevant predictive ability given small Nagelkerke's  
481 pseudo- $r^2$  estimates (Nagelkerke's pseudo- $r^2_{MAX} = 1.445 \times 10^{-4}$ ) (**Supplementary**  
482 **Table 17**).

483

## 484 3.3 Further analysis of single instrument metabolites

### 485 3.3.1 Colocalisation at loci of influential instruments

486 Of the 44 colocalisation tests conducted in regions +/- 250 kb of the LOO influential  
487 variants, 26 showed evidence of colocalisation (PP.H<sub>4</sub>  $\geq 0.8$ ) and a further five had  
488 suggestive evidence (PP.H<sub>4</sub>  $\geq 0.6$ ) (**Supplementary Table 18**). All except one of these  
489 regions were on chromosome 11 within a ~500\_kb window (61293499-61854782 bp).  
490 All disorders colocalising with metabolites in this region were psychiatric and every  
491 psychiatric disorder examined showed colocalisation with at least one metabolite.  
492 Colocalisation has been previously observed between the omega-3 fatty acid  
493 docosahexaenoic acid (DHA) and depression within the same region<sup>39</sup>.

494 All implicated metabolites except 1-stearoyl-2-docosapentaenoyl-GPC  
495 (18:0/22:5n6)\* were complex fatty acid chains containing the fatty acids arachidonic

496 acid (AA) ( $N = 18$ ), linoleic acid (LA) ( $N = 5$ ), alpha-linolenic acid (ALA) ( $N = 1$ ) or  
497 both AA and LA ( $N = 4$ ). Higher levels of AA-containing metabolites were consistently  
498 protective in Wald estimates while higher levels of LA- or ALA-containing metabolites  
499 had risk increasing effects. Where ratios contained both AA- and LA-containing  
500 metabolites, higher levels of the AA relative to the LA was protective and higher LA  
501 metabolites relative to the AA risk increasing (**Supplementary Table 19**). This is  
502 concordant with previous metabolome-wide MR analysis of bipolar disorder<sup>41</sup> and  
503 observational analysis of depression<sup>98</sup>. The other region of colocalisation was  
504 identified on chromosome 12 (56615338-57115338) between the ratio of histidine-to-  
505 glutamine and AD (PP.H<sub>4</sub> = 0.92).

506 Two tests showed evidence of two distinct causal variants at the tested locus  
507 – the ratio of glucose to mannose and depression (chr2: 27480940-27980940; PP.H<sub>3</sub>  
508 = 0.92) and docosapentaenoate (n6 DPA; 22:5n6) and bipolar disorder (chr11:  
509 61347212-61847212; PP.H<sub>3</sub> = 0.99). Suggestive evidence for two distinct causal  
510 variants was also identified for butyrylglycine and depression (chr12:120925524-  
511 121425524; PP.H<sub>3</sub> = 0.70), methylsuccinate and depression (chr12:120926083-  
512 121426083; PP.H<sub>3</sub> = 0.63), and 5-methyluridine (ribothymidine) and MS  
513 (Chr22:50687969-51187969; PP.H<sub>3</sub> = 0.63). PP.H<sub>3</sub> ≥ 0.6 in these regions suggests  
514 that MR estimated effects arose from linkage disequilibrium, violating MRs horizontal  
515 pleiotropy assumption. Remaining colocalisation tests were inconclusive.

516

### 517 3.3.2 eQTL colocalisation

518 In total, nine genes were identified within 10 kb of credible SNPs at colocalising loci  
519 (**Supplementary Table 20**). For the colocalisation analyses at chr11:61293499-  
520 61854782, eQTLs for *MIR611* or *MIR1908* were not available in the eQTLGen  
521 summary statistics. As such, follow-up eQTL colocalisation in this region was  
522 conducted only with *MYRF*, *TMEM258*, *FEN1*, *FADS2* and *FADS1*. For the region  
523 chr12:56615338-57115338, eQTLs were not available for *MIP*, but were available for  
524 *GLS2* and *SPRYD4*. In total, 130 colocalisation tests were conducted with eQTLs. Of  
525 these, only eQTLs for *SPRYD4* colocalised with both a risk metabolite and its  
526 respective neuropsychiatric outcome: the ratio of histidine-to-glutamine (PP.H<sub>4</sub> =  
527 0.91) and AD (PP.H<sub>4</sub> = 0.91). Additional colocalisation was identified for *GLS2* with  
528 the ratio of histidine-to-glutamine (PP.H<sub>4</sub> = 0.99), but not AD (PP.H<sub>4</sub> = 0.42). *TMEM258*  
529 colocalised with depression (PP.H<sub>4</sub> = 0.86). Suggestive colocalisation was identified  
530 between *TMEM258* and anxiety and between *MYRF* and depression (PP.H<sub>4</sub> ≥ 0.6). Of  
531 the remaining tests, 113 indicated two distinct causal variants in the region (PP.H<sub>3</sub> ≥  
532 0.8), four were suggestive of only a causal variant for gene expression, and the  
533 remainder inconclusive (**Supplementary Table 21**).

534

### 535 3.3.3 eQTL MR for the histidine to glutamine ratio and AD

536 We conducted MR analyses for the effect of *SPRYD4* expression levels on the ratio  
537 of histidine-to-glutamine and AD using five independent instruments. *SPRYD4*  
538 expression levels were identified as increasing risk for AD (OR = 1.08 [1.02-1.16],  $\beta$   
539 [95%CI] = 0.08 [0.02-0.15],  $p$ -value = 0.007) and the level of the ratio of histidine to  
540 glutamine ( $\beta$  [95%CI] = 0.30 [0.18-0.42],  $p$ -value =  $1.16 \times 10^{-6}$ ). SMR estimates for  
541 the effect of *SPRYD4* expression were directionally consistent and statistically  
542 significant at  $p \leq 0.05$  for AD and the ratio of histidine-to-glutamine across eQTL  
543 panels. HEIDI tests did not indicate heterogeneity ( $p_{HEIDI} \geq 0.05$ ) (**Supplementary**  
544 **Table 22**). There was no evidence that the ratio of histidine to glutamine substantially  
545 mediates the effect of *SPRD4* expression on AD, with only a nominally significant  
546 association and wide confidence intervals (proportion mediated [95%CI] = 63.20%  
547 [1.80-124.59%],  $p$ -value = 0.043).

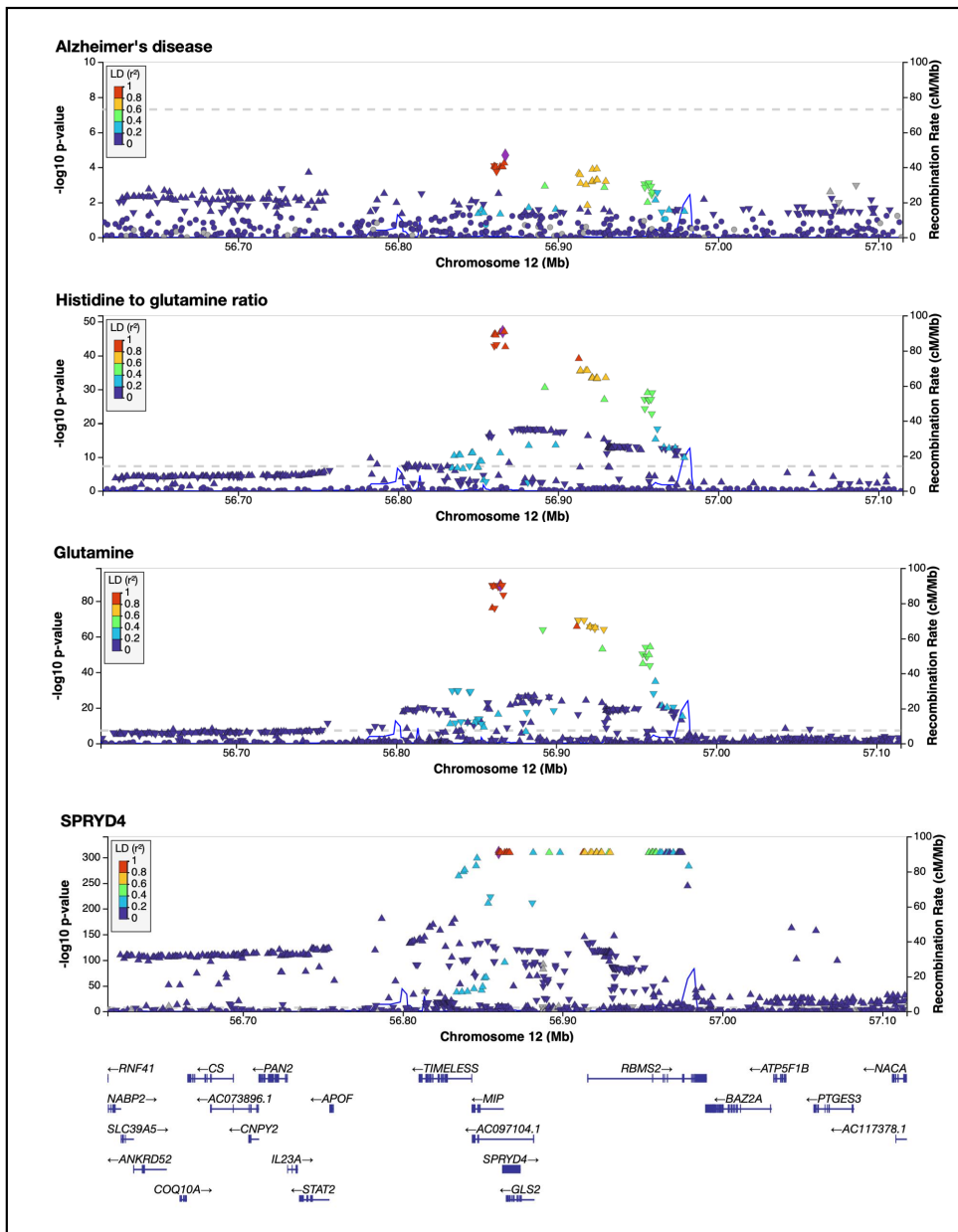
548

### 549 3.6 Further examination of histidine, glutamine, AD and *SPRYD4*

550 To further evaluate the association between the ratio of histidine to glutamine, we re-  
551 examined the IVW-MR estimates for the separate effects of histidine and glutamine  
552 on AD. No causal effect was observed for histidine on AD, not even nominally (OR =  
553 1.11 [0.99-1.27],  $\beta$  = 0.11 [-0.011-0.24],  $p$ -value = 0.07). However, there was a  
554 nominally significant negative causal effect for glutamine (OR = 0.87 [0.81-0.94],  $\beta$   
555 [95%CI] = -0.14 [-0.21 to -0.06],  $p$ -value = 0.0007,  $p_{FDR}$  = 0.052). Although glutamine  
556 passed post-hoc sensitivity criteria with no evidence of pleiotropy or heterogeneity  
557 (**Supplementary Table 23**), LOO identified the same influential variant on  
558 chromosome 12 as observed for the ratio of histidine-to-glutamine – the 3'-UTR  
559 variant rs2657879 (**Supplementary Table 24**). For glutamine, the effect of rs2657879  
560 was in the opposite direction to the histidine-to-glutamine ratio.

561 A Wald ratio test using only this instrument estimated a negative effect of  
562 glutamine on AD risk (OR = 0.84 [0.76-0.91],  $\beta$  = -0.19 [-0.28 to -0.09],  $p$ -value =  $9.65$   
563  $\times 10^{-5}$ ). Strong evidence of colocalisation was observed between glutamine and AD  
564 in the region (PP.H<sub>4</sub> = 0.92), and between glutamine and *SPRYD4* (PP.H<sub>4</sub> = 0.98)  
565 (**Supplementary Table 25**). Gene expression levels of *SPRYD4* were negatively  
566 associated with glutamine in the MR analysis ( $\beta$  = -0.35 [-0.52 to -0.18],  $p$ -value =  
567  $3.94 \times 10^{-5}$ ). Strong, directionally consistent effects of *SPRD4* expression on glutamine  
568 were observed in SMR (**Supplementary Table 26**). However, the HEIDI tests  
569 indicated heterogeneity ( $p_{HEIDI} \leq 0.05$ ). There was no evidence that glutamine  
570 mediated the effects of *SPRD4* expression on AD (proportion mediated [95%CI] =

571 75.81% [-0.36-151.99%],  $p$ -value = 0.051). LocusZoom plots for the ratio of histidine  
 572 to glutamine, glutamine, AD and *SPRYD4* gene expression are shown in Figure 4.  
 573



574 Figure 4: LocusZoom plots of the colocalising region ( $PPH_4 \geq 0.8$ ) chr12: 56615338-57115338 for  
 575 Alzheimer's disease, the ratio of histidine to glutamine, glutamine, and *SPRYD4* expression.

## 576 4. Discussion

577 We identified 77 unique plasma metabolites and a total of 85 sensitivity robust MR  
578 estimated causal effects on neuropsychiatric disorder risk. Over half of associations  
579 were driven by single influential variants. Of these, 30 metabolite-neuropsychiatric  
580 disorder trait-pairs had greater than suggestive colocalisation and 29 were lipids  
581 colocalising with a psychiatric disorder near the *FADS* gene cluster on chromosome  
582 11. No eQTL colocalisation was observed for any metabolite-outcome trait-pair in this  
583 region, suggesting the underlying mechanisms are not related to shared gene  
584 expression effects. Additional colocalisation was observed between AD glutamine-  
585 related metabolites and *SPRYD4* gene expression on chromosome 12. For the  
586 remaining 41 polygenic metabolites, multivariate analysis with MR-BMA prioritised 23  
587 metabolites and implicated sphingolipids and carnitine derivatives in psychiatric and  
588 neurodegenerative risk respectively. However, the predictive performance of  
589 metabolite polygenic scores was negligible in the UK Biobank.

590 Lipids – specifically polyunsaturated fatty acids (PUFAs) – have previously  
591 shown MR causal effects on depression, with docosahexaenoic acid (DHA)  
592 colocalising at the *FADS* gene cluster<sup>39</sup>. Our study highlights the broader significance  
593 of this region in linking lipid levels to risk across the psychiatric disorder spectrum.  
594 Here, we identify suggestive colocalisation between at least one lipid and every  
595 examined psychiatric disorder. Nearly all lipids colocalising with psychiatric  
596 disorders contained linoleic acid (LA) or arachidonic acid (AA) – both omega-6  
597 PUFAs. AA-containing lipids were consistently protective, while LA-containing lipids  
598 increased risk. These findings are consistent with other MR studies of bipolar  
599 disorder, depression and schizophrenia<sup>41,99,100</sup>. Our results suggest that these  
600 associations extend to ANX. Although these effects were primarily driven by shared  
601 causal variants on chromosome 11, several polygenic metabolites comprised of  
602 AA/LA-containing lipids also showed causal effect on psychiatric risk. The ratio of  
603 arachidonate (20:4n6)-to-oleate-to-vaccenate (18:1) and 1-(1-enyl-palmitoyl)-2-  
604 arachidonoyl-GPC (P-16:0/20:4) – both AA-containing – had risk decreasing effects  
605 on anxiety and depression respectively and were prioritised in multivariate analysis.  
606 In single instrument analyses, this ratio had risk decreasing effects on bipolar  
607 disorder, depression and schizophrenia, colocalising with all. These results suggest  
608 that effects of AA/LA-containing lipids in psychiatric risk may extend beyond effects  
609 on chromosome 11.

610 Circulating levels of lipids containing LA and AA may be promising intervention  
611 targets for psychiatric disorders. LA – the most abundant omega-6 PUFA in the  
612 western diet – is found in nearly all manufactured food through vegetable oils<sup>101</sup>, while  
613 AA is found in animal products such as meat, poultry, fish and eggs<sup>102</sup>. As free fatty

614 acids AA is a downstream product of LA, from which it is synthesised through  
615 desaturation and chain elongation<sup>102</sup> (**Supplementary Material 4**).

616 Free LA and AA are linked to inflammatory dysregulation<sup>103</sup> – itself implicated  
617 in psychiatric disorders<sup>104,105</sup>. AA is catalysed by cytochrome P450 enzymes to  
618 epoxyeicosatrienoic acids (EETs), which show anti-inflammatory effects and are  
619 suggested to be protective of neurological function<sup>106</sup>. The deleterious effects of LA-  
620 containing lipids compared to the protective effects of AA-containing lipids may  
621 implicate dysregulated conversion from LAs to AAs, contributing to elevated  
622 inflammation as less AA is available for conversion to anti-inflammatory EETs.  
623 However, LA-to-AA conversion is relatively low in humans and requires LA to be in  
624 free form following cleavage from lipid molecules via enzymes such as  
625 phospholipase A2 (PLA2). Although PLA2 action on fatty acids has been previously  
626 linked to inflammatory depression<sup>107</sup>, further investigation is required to elucidate  
627 mechanisms by which LA- and AA-containing lipids contribute to psychiatric risk.

628 The effects of the ratio of histidine-to-glutamine and glutamine on AD observed  
629 here are in line with previous MR evidence suggesting higher blood glutamine as  
630 protective<sup>37,108</sup>. Interestingly, glutamine is noted to exert a neuroprotective effect  
631 against  $\beta$ -amyloid aggregation<sup>109</sup> – a key drug target for AD<sup>110</sup>. This study provides  
632 further evidence for its protective effects on AD risk and suggests the association is  
633 driven by shared causal variants on chromosome 12. Further, we observe that the  
634 association between AD and glutamine may be partially driven by the mutual effect  
635 of *SPRYD4* gene expression. As *SPRYD4* has not been previously identified in AD  
636 GWAS it represents a novel locus of interest for future study.

637 Our study also provides further evidence for the role of sphingolipid  
638 metabolism in psychiatric disorders. Sphingolipids are fatty acids with roles in cell  
639 growth/death, inflammation, mitochondrial function and immune response<sup>111</sup>. We  
640 identified two sphingomyelins and glycosyl-N-stearoyl-sphingosine (d18:1/18:0) as  
641 protective for depression, with the sphingomyelins both prioritised in multivariate  
642 analysis. Further, sphingomyelin (d18:1/18:1, d18:2/18:0) was protective for  
643 schizophrenia and sphingosine-1-phosphate indicated to elevate risk of bipolar  
644 disorder. Sphingomyelins are central to ceramide metabolism, to which they are  
645 converted via sphingomyelinase and back via sphingomyelin synthase<sup>112</sup>  
646 (**Supplementary Material 5**). Ceramides are catabolised into sphingosine and  
647 eventually sphingosine-1-phosphate<sup>112</sup>. General dysfunction of sphingolipid  
648 metabolism, particularly ceramide aggregation, has been previously observed in  
649 psychiatric disorders<sup>23,24,111,112</sup> and reduced sphingomyelin levels may contribute to  
650 elevated neuroinflammation<sup>112</sup>. As such, the ceramide system is a proposed pathway  
651 for drug-based interventions for psychiatric disorders<sup>112</sup>. This study implicates  
652 sphingomyelins as possible protective targets.

653 Although fewer metabolomic patterns were observed in neurodegenerative  
654 disorders we did observe protective effects of several carnitine related metabolites –  
655 adipoylcarnitine for MS and acetylcarnitine and propionylcarnitine for ALS. All three  
656 are metabolised from L-carnitine, which plays a role in the transportation of fatty acids  
657 to mitochondria<sup>113</sup>, and when acetylated is suggested to slow disease progression for  
658 both disorders and AD<sup>114–116</sup>. Further, lower serum carnitine levels are observed in  
659 ALS<sup>117</sup> and are associated with increased fatigue in MS<sup>118</sup>. A randomised control trial  
660 (RCT) of acetylcarnitine supplementation in ALS suggests it may improve self-  
661 sufficiency in patients compared to placebo<sup>114</sup>. The importance of acetylcarnitine for  
662 ALS is further emphasised by our multivariate analysis, where it was the only  
663 metabolite prioritised for ALS and had the highest MIP of any metabolite examined.  
664 Interestingly, a larger phase two/three RCT of acetylcarnitine effects on biological and  
665 clinical outcomes in ALS is underway  
666 (<https://classic.clinicaltrials.gov/ct2/show/NCT06126315>). The remaining polygenic  
667 metabolites did not show clear functional relatedness, indicating disorder-specific  
668 metabolomic contributions – broadly confirmed in MR-BMA, where, for most  
669 disorders, all metabolites had a MIP  $\geq 0.1$ , indicating unique causal effects.

670 Although no specific metabolites were identified with an effect on both a  
671 psychiatric and neurodegenerative disorder, our results suggest the general  
672 importance circulating lipids in both categories. Dietary lipids are known to modulate  
673 neuroinflammation and play an important role in the brain by regulating neuronal and  
674 synaptic function<sup>119</sup>. Future studies exploring biological overlap in neuropsychiatric  
675 disorders may therefore wish to focus on lipidomic analysis for the purposes of further  
676 elucidating shared metabolic pathways.

677 Several disorder-specific metabolites warrant further discussion. For example,  
678 we replicated the risk-increasing effect of uridine on MS previously identified  
679 observationally and in MR<sup>40,120</sup>. Uridine is linked to pyrimidine metabolism with  
680 important biological functions including RNA synthesis and the regulation of glucose,  
681 lipid and amino acid levels<sup>121</sup>. Interesting, within pyrimidine biosynthesis, uridine is an  
682 eventual metabolic product of dihydroorotate<sup>121</sup>, also identified here as increasing  
683 risk for MS. Additionally, the branch chain amino acids (BCAAs) leucine and its  
684 isomer isoleucine were both protective for PD. Plasma levels are noted to be lower in  
685 cases than controls and correlated with increased functional disability<sup>122</sup>. In an RCT  
686 of PD/parkinsonism, supplementation with leucine and vitamin D enriched whey  
687 protein was significantly associated with increased lower body function and muscle  
688 mass retention<sup>123</sup>. Our findings provide further evidence of the protective effects of  
689 leucine/isoleucine in PD and suggest them as possible interventional targets.

690 In the UKB, PGS for prioritised polygenic metabolites showed limited  
691 prediction of their respective outcome disorder. Polygenic scores are noted to explain

692 only a small proportion of the variance even within the same phenotype<sup>124</sup>. As such,  
693 it is not entirely surprising that we are unable to detect significant associations cross-  
694 trait, particularly given the smaller effective target sample size available in the UKB  
695 compared to the large, meta-analytic samples sizes available from GWAS summary  
696 statistics for use in MR. To improve prediction, future studies might aggregate the  
697 genetic effects of multiple metabolites to better capture the total underlying metabolic  
698 risk for disease, for example through a multi-PGS framework<sup>125</sup>.

699 In conclusion, we performed metabolome-wide MR to test the causal effects  
700 of 1300 metabolites/metabolite ratios on eight neuropsychiatric disorders. Results  
701 implicate AA and LA-containing lipids across the psychiatric disorder spectrum, with  
702 shared causal variants identified near the *FADS* gene cluster. We also identified  
703 shared causal variants between glutamine metabolites and Alzheimer's disease, and  
704 identify a shared effect of *SPRD4* gene expression. Additionally, our results implicate  
705 sphingolipid metabolism in psychiatric disorders and carnitine related metabolites as  
706 protective for MS and ALS, alongside several promising target metabolites with  
707 disease-specific effects. More broadly, this study suggests that lipid metabolism  
708 plays an important role in both psychiatric and neurodegenerative risk. The  
709 metabolites identified in this study can help inform potential interventions and future,  
710 targeted studies of metabolic contributions to disease risk.

711

## 712 5. Limitations

713 This study used data from individuals of European ancestry only and as such may  
714 lack generalisability. Further, as previously noted<sup>79</sup>, although the exclusion of the  
715 *APOE* region from the AD analyses is necessary to avoid violating MR assumptions,  
716 this may result in false negatives due to the regions known role in lipid metabolism<sup>126</sup>.  
717 Conversely, we can have greater confidence that the metabolites identified here are  
718 not confounded by *APOE* effects. Given the importance of this region in AD and lipid  
719 metabolism, it deserves specific focus in future metabolomic work. Further, MR  
720 requires that several assumptions – such the availability of suitable instruments and  
721 absence of pleiotropic effects – be met to provide reliable results. Although we  
722 mitigate against these by including only strong instruments (F-statistic  $\geq 10$ ) and  
723 using a robust sensitivity criterion to delineate causal metabolites, results should be  
724 interpreted with caution prior to further triangulation.

725

## 726 6. Funding and acknowledgements

727 LG is funded by the King's College London DRIVE-Health Centre for Doctoral Training  
728 and the Perron Institute for Neurological and Translational. JM is supported by the  
729 King's Prize Fellowship. This study has been partly delivered through the National



730 Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre  
731 (BRC). PP is funded by an Alzheimer's Research UK Senior Research Fellowship.

732 This research has been conducted using the UK Biobank Resource under  
733 Application Number 82877. Ethical approval for the UK Biobank study has been  
734 granted by the National Information Governance Board for Health and Social Care  
735 and the NHS North West Multicentre Research Ethics Committee (11/NW/0382).  
736 Written informed consent was obtained from all participants by the UK Biobank. We  
737 thank the UK Biobank Team for collecting the data and making it available. We also  
738 thank the UK Biobank participants.

739 We also thank and acknowledge the contribution and use of the CREATE high-  
740 performance computing cluster at King's College London (King's College London.  
741 (2022). King's Computational Research, Engineering and Technology Environment  
742 (CREATE)).

743 We want to acknowledge the participants of the FinnGen study and the  
744 veterans of the Million Veterans Project, as well as the research teams for both  
745 studies.

746 For the purposes of open access, the author has applied a Creative Commons  
747 Attribution (CC BY) license to any Accepted Author Manuscript version arising from  
748 this submission.

749

## 750 **7. Data availability**

751 Data is available on reasonable request from the UK Biobank  
752 (<https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/contact-us>).

753 Summary statistics from FinnGen are available online  
754 ([https://www.finn-gen.fi/en/access\\_results](https://www.finn-gen.fi/en/access_results)) and from the Million Veterans Project via  
755 dbGaP request ([https://www.ncbi.nlm.nih.gov/projects/gap/cgi-  
756 bin/study.cgi?studyid=phs001672.v1.p1](https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?studyid=phs001672.v1.p1)).

757 GWAS summary statistics from the Psychiatric Genomics Consortium (PGC)  
758 for Alzheimer's disease, anxiety, bipolar disorder, depression and schizophrenia are  
759 available from their website (<https://pgc.unc.edu/for-researchers/download-results/>).

760 Summary statistics for anxiety from iPSYCH are also available online  
761 (<https://ipsych.dk/en/research/downloads>). For Parkinson's disease and ALS,  
762 summary statistics are available on GWAS Catalog at  
763 <https://www.ebi.ac.uk/gwas/studies/GCST009325> and  
764 <https://www.ebi.ac.uk/gwas/studies/GCST90027164> respectively. For multiple  
765 sclerosis, summary statistics are available through the IEU Open GWAS Project  
766 (<https://gwas.mrcieu.ac.uk/datasets/ieu-b-18/>). Summary statistics for metabolites  
767 measured from Chen et al.<sup>58</sup> are available on GWAS catalog

768 (<https://www.ebi.ac.uk/gwas/studies/>) under accession numbers GCST90199621-  
769 90201020. Access to summary statistics for the metabolites measured by Hysi et al.<sup>59</sup>  
770 were made available following request to the original authors.

771

## 772 **8. Conflicts of Interest**

773 CML is on the Scientific Advisory Board of Myriad Neuroscience and has received  
774 honoraria for consultancy from UCB. C.L.-Q. has received consultancy fees from  
775 Pfizer. C.L.-Q. has received honoraria, travel or speakers' fees from Biogen and  
776 research funds from Pfizer and Novo Nordisk. C.L.-Q. is the director of the company  
777 BrainLogia. All other authors declare no competing interests.

## 778 References

779

780 1. Plana-Ripoll, O. *et al.* Exploring comorbidity within mental disorders among a danish  
781 national population. *JAMA Psychiatry* **76**, 259–270 (2019).

782 2. Grotzinger, A. D. *et al.* Genetic architecture of 11 major psychiatric disorders at  
783 biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat. Genet.*  
784 **54**, 548–559 (2022).

785 3. Aarsland, D. *et al.* Neuropsychiatric symptoms in patients with Parkinson’s disease and  
786 dementia: frequency, profile and associated care giver stress. *J. Neurol. Neurosurg.*  
787 *Psychiatry* **78**, 36–42 (2007).

788 4. Turner, M. R., Goldacre, R., Talbot, K. & Goldacre, M. J. Psychiatric disorders prior to  
789 amyotrophic lateral sclerosis. *Ann. Neurol.* **80**, 935–938 (2016).

790 5. Ismail, Z. *et al.* Psychosis in Alzheimer disease - mechanisms, genetics and therapeutic  
791 opportunities. *Nat. Rev. Neurol.* **18**, 131–144 (2022).

792 6. Chertcoff, A. S. *et al.* Psychiatric comorbidity during the prodromal period in patients  
793 with multiple sclerosis. *Neurology* **101**, e2026–e2034 (2023).

794 7. Burchill, E. *et al.* The impact of psychiatric comorbidity on Parkinson’s disease  
795 outcomes: a systematic review and meta-analysis. *Lancet Reg. Health Eur.* **39**, 100870  
796 (2024).

797 8. Wainberg, M., Andrews, S. J. & Tripathy, S. J. Shared genetic risk loci between  
798 Alzheimer’s disease and related dementias, Parkinson’s disease, and amyotrophic  
799 lateral sclerosis. *Alzheimers Res. Ther.* **15**, 113 (2023).

800 9. Wightman, D. P. *et al.* The genetic overlap between Alzheimer’s disease, amyotrophic  
801 lateral sclerosis, Lewy body dementia, and Parkinson’s disease. *Neurobiol. Aging* **127**,  
802 99–112 (2023).

803 10. Gimson, A., Schlosser, M., Huntley, J. D. & Marchant, N. L. Support for midlife  
804 anxiety diagnosis as an independent risk factor for dementia: a systematic review. *BMJ*  
805 *Open* **8**, e019399 (2018).

806 11. Kuusimäki, T. *et al.* Increased risk of parkinson’s disease in patients with schizophrenia  
807 spectrum disorders. *Mov. Disord.* **36**, 1353–1361 (2021).

808 12. Richmond-Rakerd, L. S., D’Souza, S., Milne, B. J., Caspi, A. & Moffitt, T. E.  
809 Longitudinal Associations of Mental Disorders With Dementia: 30-Year Analysis of  
810 1.7 Million New Zealand Citizens. *JAMA Psychiatry* **79**, 333–340 (2022).

811 13. Andreassen, O. A. *et al.* Genetic pleiotropy between multiple sclerosis and  
812 schizophrenia but not bipolar disorder: differential involvement of immune-related gene  
813 loci. *Mol. Psychiatry* **20**, 207–214 (2015).

814 14. McLaughlin, R. L. *et al.* Genetic correlation between amyotrophic lateral sclerosis and  
815 schizophrenia. *Nat. Commun.* **8**, 14774 (2017).

816 15. Reynolds, R. H. *et al.* Local genetic correlations exist among neurodegenerative and  
817 neuropsychiatric diseases. *npj Parkinsons Disease* **9**, 70 (2023).

818 16. Wingo, T. S. *et al.* Shared mechanisms across the major psychiatric and  
819 neurodegenerative diseases. *Nat. Commun.* **13**, 4314 (2022).

820 17. Harerimana, N. V. *et al.* Genetic evidence supporting a causal role of depression in  
821 alzheimer’s disease. *Biol. Psychiatry* **92**, 25–33 (2022).

822 18. Patti, G. J., Yanes, O. & Siuzdak, G. Innovation: Metabolomics: the apogee of the  
823 omics trilogy. *Nat. Rev. Mol. Cell Biol.* **13**, 263–269 (2012).

824 19. Guijas, C., Montenegro-Burke, J. R., Warth, B., Spilker, M. E. & Siuzdak, G.  
825 Metabolomics activity screening for identifying metabolites that modulate phenotype.  
826 *Nat. Biotechnol.* **36**, 316–320 (2018).

- 827 20. Garana, B. B. & Graham, N. A. Metabolomics paves the way for improved drug target  
828 identification. *Mol. Syst. Biol.* **18**, e10914 (2022).
- 829 21. Rinschen, M. M., Ivanisevic, J., Giera, M. & Siuzdak, G. Identification of bioactive  
830 metabolites using activity metabolomics. *Nat. Rev. Mol. Cell Biol.* **20**, 353–367 (2019).
- 831 22. Gowda, G. A. N. *et al.* Metabolomics-based methods for early disease diagnostics.  
832 *Expert Rev Mol Diagn* **8**, 617–633 (2008).
- 833 23. Dinoff, A., Herrmann, N. & Lanctôt, K. L. Ceramides and depression: A systematic  
834 review. *J. Affect. Disord.* **213**, 35–43 (2017).
- 835 24. Brunkhorst-Kanaan, N. *et al.* Targeted lipidomics reveal derangement of ceramides in  
836 major depression and bipolar disorder. *Metab. Clin. Exp.* **95**, 65–76 (2019).
- 837 25. Filippatou, A. G. *et al.* Serum ceramide levels are altered in multiple sclerosis. *Mult.*  
838 *Scler.* **27**, 1506–1519 (2021).
- 839 26. Chowdhury, M. R., Jin, H. K. & Bae, J.-S. Diverse roles of ceramide in the progression  
840 and pathogenesis of alzheimer’s disease. *Biomedicines* **10**, (2022).
- 841 27. Vos, M., Klein, C. & Hicks, A. A. Role of ceramides and sphingolipids in parkinson’s  
842 disease. *J. Mol. Biol.* **435**, 168000 (2023).
- 843 28. Davey Smith, G. & Ebrahim, S. Epidemiology--is it time to call it a day? *Int. J.*  
844 *Epidemiol.* **30**, 1–11 (2001).
- 845 29. Webhofer, C. *et al.* Metabolite profiling of antidepressant drug action reveals novel  
846 drug targets beyond monoamine elevation. *Transl. Psychiatry* **1**, e58 (2011).
- 847 30. Pillinger, T. *et al.* Comparative effects of 18 antipsychotics on metabolic function in  
848 patients with schizophrenia, predictors of metabolic dysregulation, and association with  
849 psychopathology: a systematic review and network meta-analysis. *Lancet Psychiatry* **7**,  
850 64–77 (2020).
- 851 31. Shao, Y. *et al.* Comprehensive metabolic profiling of Parkinson’s disease by liquid  
852 chromatography-mass spectrometry. *Mol. Neurodegener.* **16**, 4 (2021).
- 853 32. Pu, J. *et al.* Effects of pharmacological treatment on metabolomic alterations in animal  
854 models of depression. *Transl. Psychiatry* **12**, 175 (2022).
- 855 33. Murgia, F. *et al.* Metabolomic Changes in Patients Affected by Multiple Sclerosis and  
856 Treated with Fingolimod. *Metabolites* **13**, (2023).
- 857 34. Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N. & Davey Smith, G.  
858 Mendelian randomization: using genes as instruments for making causal inferences in  
859 epidemiology. *Stat. Med.* **27**, 1133–1163 (2008).
- 860 35. Smith, G. D. & Ebrahim, S. “Mendelian randomization”: can genetic epidemiology  
861 contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.*  
862 **32**, 1–22 (2003).
- 863 36. Abdellaoui, A., Yengo, L., Verweij, K. J. H. & Visscher, P. M. 15 years of GWAS  
864 discovery: Realizing the promise. *Am. J. Hum. Genet.* **110**, 179–194 (2023).
- 865 37. Lord, J. *et al.* Disentangling independent and mediated causal relationships between  
866 blood metabolites, cognitive factors, and alzheimer’s disease. *Biological Psychiatry*  
867 *Global Open Science* **2**, 167–179 (2022).
- 868 38. Zacharias, H. U. *et al.* A metabolome-wide association study in the general population  
869 reveals decreased levels of serum laurycarnitine in people with depression. *Mol.*  
870 *Psychiatry* **26**, 7372–7383 (2021).
- 871 39. Davyson, E. *et al.* Metabolomic investigation of major depressive disorder identifies a  
872 potentially causal association with polyunsaturated fatty acids. *Biol. Psychiatry* **94**,  
873 630–639 (2023).
- 874 40. Ge, A., Sun, Y., Kiker, T., Zhou, Y. & Ye, K. A metabolome-wide Mendelian  
875 randomization study prioritizes potential causal circulating metabolites for multiple  
876 sclerosis. *J. Neuroimmunol.* **379**, 578105 (2023).

- 877 41. Stacey, D., Benyamin, B., Lee, S. H. & Hyppönen, E. A Metabolome-Wide Mendelian  
878 Randomization Study Identifies Dysregulated Arachidonic Acid Synthesis as a  
879 Potential Causal Risk Factor for Bipolar Disorder. *Biol. Psychiatry* (2024).  
880 doi:10.1016/j.biopsych.2024.02.1005
- 881 42. Purves, K. L. *et al.* A major role for common genetic variation in anxiety disorders.  
882 *Mol. Psychiatry* **25**, 3292–3303 (2020).
- 883 43. Meier, S. M. *et al.* Genetic Variants Associated With Anxiety and Stress-Related  
884 Disorders: A Genome-Wide Association Study and Mouse-Model Study. *JAMA*  
885 *Psychiatry* (2019). doi:10.1001/jamapsychiatry.2019.1119
- 886 44. Kurki, M. I. *et al.* FinnGen provides genetic insights from a well-phenotyped isolated  
887 population. *Nature* **613**, 508–518 (2023).
- 888 45. Mullins, N. *et al.* Genome-wide association study of more than 40,000 bipolar disorder  
889 cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829  
890 (2021).
- 891 46. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102  
892 independent variants and highlights the importance of the prefrontal brain regions. *Nat.*  
893 *Neurosci.* **22**, 343–352 (2019).
- 894 47. Levey, D. F. *et al.* Bi-ancestral depression GWAS in the Million Veteran Program and  
895 meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat.*  
896 *Neurosci.* **24**, 954–963 (2021).
- 897 48. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in  
898 schizophrenia. *Nature* **604**, 502–508 (2022).
- 899 49. van Rheenen, W. *et al.* Common and rare variant association analyses in amyotrophic  
900 lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-  
901 specific biology. *Nat. Genet.* **53**, 1636–1648 (2021).
- 902 50. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map  
903 implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, (2019).
- 904 51. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for  
905 Parkinson’s disease: a meta-analysis of genome-wide association studies. *Lancet*  
906 *Neurol.* **18**, 1091–1102 (2019).
- 907 52. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of  
908 genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
- 909 53. Liu, H. *et al.* Mendelian randomization highlights significant difference and genetic  
910 heterogeneity in clinically diagnosed Alzheimer’s disease GWAS and self-report proxy  
911 phenotype GWAX. *Alzheimers Res. Ther.* **14**, 17 (2022).
- 912 54. Lumsden, A. L., Mulugeta, A., Zhou, A. & Hyppönen, E. Apolipoprotein E (APOE)  
913 genotype-associated disease risks: a phenome-wide, registry-based, case-control study  
914 utilising the UK Biobank. *EBioMedicine* **59**, 102954 (2020).
- 915 55. Bulik-Sullivan, B. *et al.* LD Score regression distinguishes confounding from  
916 polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
- 917 56. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into  
918 the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525  
919 (2019).
- 920 57. Grotzinger, A. D., Fuente, J. de la, Privé, F., Nivard, M. G. & Tucker-Drob, E. M.  
921 Pervasive Downward Bias in Estimates of Liability-Scale Heritability in Genome-wide  
922 Association Study Meta-analysis: A Simple Solution. *Biol. Psychiatry* **93**, 29–36  
923 (2023).
- 924 58. Chen, Y. *et al.* Genomic atlas of the plasma metabolome prioritizes metabolites  
925 implicated in human diseases. *Nat. Genet.* **55**, 44–53 (2023).
- 926 59. Hysi, P. G. *et al.* Metabolome Genome-Wide Association Study Identifies 74 Novel

- 927 Genomic Regions Influencing Plasma Metabolites Levels. *Metabolites* **12**, (2022).
- 928 60. Murphy, A. E., Schilder, B. M. & Skene, N. G. MungeSumstats: A Bioconductor  
929 package for the standardisation and quality control of many GWAS summary statistics.  
930 *Bioinformatics* (2021). doi:10.1093/bioinformatics/btab665
- 931 61. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and  
932 richer datasets. *Gigascience* **4**, s13742–015–0047–8 (2015).
- 933 62. 1000 Genomes Project Consortium *et al.* A global reference for human genetic  
934 variation. *Nature* **526**, 68–74 (2015).
- 935 63. Elsworth, B. *et al.* The MRC IEU OpenGWAS data infrastructure. *BioRxiv* (2020).  
936 doi:10.1101/2020.08.10.244293
- 937 64. Palmos, A. B. *et al.* Proteome-wide Mendelian randomization identifies causal links  
938 between blood proteins and severe COVID-19. *PLoS Genet.* **18**, e1010042 (2022).
- 939 65. Bowden, J. *et al.* Assessing the suitability of summary data for two-sample Mendelian  
940 randomization analyses using MR-Egger regression: the role of the I2 statistic. *Int. J.*  
941 *Epidemiol.* **45**, 1961–1974 (2016).
- 942 66. Burgess, S. *et al.* Guidelines for performing Mendelian randomization investigations:  
943 update for summer 2023. *Wellcome Open Res.* **4**, 186 (2019).
- 944 67. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and  
945 powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*  
946 *(Methodological)* **57**, 289–300 (1995).
- 947 68. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in  
948 Mendelian Randomization with Some Invalid Instruments Using a Weighted Median  
949 Estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
- 950 69. Burgess, S., Bowden, J., Dudbridge, F. & Thompson, S. G. Robust instrumental  
951 variable methods using multiple candidate instruments with application to Mendelian  
952 randomization. *arXiv* (2016). doi:10.48550/arxiv.1606.03729
- 953 70. Xue, H., Shen, X. & Pan, W. Constrained maximum likelihood-based Mendelian  
954 randomization robust to both correlated and uncorrelated pleiotropic effects. *Am. J.*  
955 *Hum. Genet.* **108**, 1251–1269 (2021).
- 956 71. Burgess, S., Foley, C. N., Allara, E., Staley, J. R. & Howson, J. M. M. A robust and  
957 efficient method for Mendelian randomization with hundreds of genetic variants. *Nat.*  
958 *Commun.* **11**, 376 (2020).
- 959 72. Rees, J. M. B., Wood, A. M., Dudbridge, F. & Burgess, S. Robust methods in  
960 Mendelian randomization via penalization of heterogeneous causal estimates. *PLoS*  
961 *One* **14**, e0222362 (2019).
- 962 73. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid  
963 instruments: effect estimation and bias detection through Egger regression. *Int. J.*  
964 *Epidemiol.* **44**, 512–525 (2015).
- 965 74. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred  
966 from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
- 967 75. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body  
968 mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–  
969 3649 (2018).
- 970 76. Pulit, S. L. *et al.* Meta-analysis of genome-wide association studies for body fat  
971 distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–  
972 174 (2019).
- 973 77. Okbay, A. *et al.* Polygenic prediction of educational attainment within and between  
974 families from genome-wide association analyses in 3 million individuals. *Nat. Genet.*  
975 **54**, 437–449 (2022).
- 976 78. Zuber, V., Colijn, J. M., Klaver, C. & Burgess, S. Selecting likely causal risk factors

- 977 from high-throughput experiments using multivariable Mendelian randomization. *Nat.*  
978 *Commun.* **11**, 29 (2020).
- 979 79. Lord, J. *et al.* Mendelian randomization identifies blood metabolites previously linked  
980 to midlife cognition as causal candidates in Alzheimer’s disease. *Proc. Natl. Acad. Sci.*  
981 *USA* **118**, (2021).
- 982 80. Vilhjálmsson, B. J. *et al.* Modeling linkage disequilibrium increases accuracy of  
983 polygenic risk scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
- 984 81. Choi, S. W., Mak, T. S.-H. & O’Reilly, P. F. Tutorial: a guide to performing polygenic  
985 risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).
- 986 82. Pain, O., Al-Chalabi, A. & Lewis, C. M. The genopred pipeline: A comprehensive and  
987 scalable pipeline for polygenic scoring. *medRxiv* (2024).  
988 doi:10.1101/2024.06.12.24308843
- 989 83. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data.  
990 *Nature* **562**, 203–209 (2018).
- 991 84. Zhang, Q., Privé, F., Vilhjálmsson, B. & Speed, D. Improved genetic prediction of  
992 complex traits from individual-level data or summary statistics. *Nat. Commun.* **12**, 4192  
993 (2021).
- 994 85. Ni, G. *et al.* A comparison of ten polygenic score methods for psychiatric disorders  
995 applied across multiple cohorts. *Biol. Psychiatry* **90**, 611–620 (2021).
- 996 86. Monti, R. *et al.* Evaluation of polygenic scoring methods in five biobanks shows larger  
997 variation between biobanks than methods and finds benefits of ensemble learning. *Am.*  
998 *J. Hum. Genet.* **111**, 1431–1447 (2024).
- 999 87. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association  
1000 studies. *Bioinformatics* **26**, 2867–2873 (2010).
- 1001 88. Spargo, T. P. *et al.* Statistical examination of shared loci in neuropsychiatric diseases  
1002 using genome-wide association study summary statistics. (2023).  
1003 doi:10.7554/eLife.88768
- 1004 89. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic  
1005 association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
- 1006 90. Vösa, U. *et al.* Large-scale cis- and trans-eQTL analyses identify thousands of genetic  
1007 loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–  
1008 1310 (2021).
- 1009 91. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts  
1010 complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
- 1011 92. Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of  
1012 known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).
- 1013 93. Lloyd-Jones, L. R. *et al.* The genetic architecture of gene expression in peripheral  
1014 blood. *Am. J. Hum. Genet.* **100**, 228–237 (2017).
- 1015 94. Auwerx, C. *et al.* Exploiting the mediating role of the metabolome to unravel transcript-  
1016 to-phenotype associations. *Elife* **12**, (2023).
- 1017 95. Yoshiji, S. *et al.* Proteome-wide Mendelian randomization implicates nephronectin as  
1018 an actionable mediator of the effect of obesity on COVID-19 severity. *Nat. Metab.* **5**,  
1019 248–264 (2023).
- 1020 96. Koletzko, B. *et al.* FADS1 and FADS2 polymorphisms modulate fatty acid metabolism  
1021 and dietary impact on health. *Annu. Rev. Nutr.* **39**, 21–44 (2019).
- 1022 97. Myers, T. A., Chanock, S. J. & Machiela, M. J. Ldlinkr: an R package for rapidly  
1023 calculating linkage disequilibrium statistics in diverse populations. *Front. Genet.* **11**,  
1024 157 (2020).
- 1025 98. Jansen, R. *et al.* The Metabolome-Wide Signature of Major Depressive Disorder. *Res.*  
1026 *Sq.* (2023). doi:10.21203/rs.3.rs-3127544/v1

- 1027 99. Zeng, L. *et al.* Causal effects of fatty acids on depression: Mendelian randomization  
1028 study. *Front. Nutr.* **9**, 1010476 (2022).
- 1029 100. Gao, Y. *et al.* Association between Arachidonic Acid and the Risk of Schizophrenia: A  
1030 Cross-National Study and Mendelian Randomization Analysis. *Nutrients* **15**, (2023).
- 1031 101. Whelan, J. & Fritsche, K. Linoleic acid. *Adv. Nutr.* **4**, 311–312 (2013).
- 1032 102. Tallima, H. & El Ridi, R. Arachidonic acid: Physiological roles and potential health  
1033 benefits - A review. *J. Advanc. Res.* **11**, 33–41 (2018).
- 1034 103. Innes, J. K. & Calder, P. C. Omega-6 fatty acids and inflammation. *Prostaglandins*  
1035 *Leukot Essent Fatty Acids* **132**, 41–48 (2018).
- 1036 104. Yuan, N., Chen, Y., Xia, Y., Dai, J. & Liu, C. Inflammation-related biomarkers in  
1037 major psychiatric disorders: a cross-disorder assessment of reproducibility and  
1038 specificity in 43 meta-analyses. *Transl. Psychiatry* **9**, 233 (2019).
- 1039 105. Perry, B. I. *et al.* Associations of immunological proteins/traits with schizophrenia,  
1040 major depression and bipolar disorder: A bi-directional two-sample mendelian  
1041 randomization study. *Brain Behav. Immun.* **97**, 176–185 (2021).
- 1042 106. Wang, L., Luo, G., Zhang, L.-F. & Geng, H.-X. Neuroprotective effects of  
1043 epoxyeicosatrienoic acids. *Prostaglandins Other Lipid Mediat* **138**, 9–14 (2018).
- 1044 107. Su, K.-P. *et al.* Phospholipase A2 and cyclooxygenase 2 genes influence the risk of  
1045 interferon-alpha-induced depression by regulating polyunsaturated fatty acids levels.  
1046 *Biol. Psychiatry* **67**, 550–557 (2010).
- 1047 108. Adams, C. D. Circulating glutamine and alzheimer’s disease: A mendelian  
1048 randomization study. *Clin Interv Aging* **15**, 185–193 (2020).
- 1049 109. Chen, J. & Herrup, K. Glutamine acts as a neuroprotectant against DNA damage, beta-  
1050 amyloid and H2O2-induced stress. *PLoS One* **7**, e33177 (2012).
- 1051 110. Sims, J. R. *et al.* Donanemab in Early Symptomatic Alzheimer Disease: The  
1052 TRAILBLAZER-ALZ 2 Randomized Clinical Trial. *JAMA* **330**, 512–527 (2023).
- 1053 111. Bernal-Vega, S., García-Juárez, M. & Camacho-Morales, A. Contribution of ceramides  
1054 metabolism in psychiatric disorders. *J. Neurochem.* **164**, 708–724 (2023).
- 1055 112. Zhuo, C. *et al.* Acid sphingomyelinase/ceramide system in schizophrenia: implications  
1056 for therapeutic intervention as a potential novel target. *Transl. Psychiatry* **12**, 260  
1057 (2022).
- 1058 113. Pekala, J. *et al.* L-carnitine--metabolic functions and meaning in humans life. *Curr*  
1059 *Drug Metab* **12**, 667–678 (2011).
- 1060 114. Beghi, E. *et al.* Randomized double-blind placebo-controlled trial of acetyl-L-carnitine  
1061 for ALS. *Amyotroph. Lateral Scler. Frontotemporal Degener.* **14**, 397–405 (2013).
- 1062 115. Pennisi, M. *et al.* Acetyl-L-Carnitine in Dementia and Other Cognitive Disorders: A  
1063 Critical Update. *Nutrients* **12**, (2020).
- 1064 116. Gharighnia, S., Omidi, A., Ragerdi Kashani, I., Sepand, M. R. & Pour Beiranvand, S.  
1065 Ameliorative effects of acetyl-L-carnitine on corpus callosum and functional recovery  
1066 in demyelinated mouse model. *Int. J. Neurosci.* 1–11 (2022).  
1067 doi:10.1080/00207454.2022.2107515
- 1068 117. Sarraf, P. *et al.* The correlation of the serum level of L-carnitine with disease severity in  
1069 patients with Amyotrophic lateral sclerosis. *J. Clin. Neurosci.* **89**, 232–236 (2021).
- 1070 118. Harirchian, M. H., Babaie, S., Keshtkaran, N. & Bitarafan, S. The association of serum  
1071 carnitine levels with severity of fatigue in patients with multiple sclerosis: A pilot  
1072 study. *CJN* **22**, 30–34 (2023).
- 1073 119. Custers, Emma, E. M., Kiliaan & Amanda, J. Dietary lipids from body to brain. *Prog.*  
1074 *Lipid Res.* **85**, 101144 (2022).
- 1075 120. Zahoor, I., Rui, B., Khan, J., Datta, I. & Giri, S. An emerging potential of metabolomics  
1076 in multiple sclerosis: a comprehensive overview. *Cell Mol. Life Sci.* **78**, 3181–3203



- 1077 (2021).  
1078 121. Zhang, Y., Guo, S., Xie, C. & Fang, J. Uridine metabolism and its role in glucose, lipid,  
1079 and amino acid homeostasis. *Biomed Res. Int.* **2020**, 7091718 (2020).  
1080 122. Zhang, Y. *et al.* Plasma branched-chain and aromatic amino acids correlate with the gut  
1081 microbiota and severity of Parkinson's disease. *npj Parkinsons Disease* **8**, 48 (2022).  
1082 123. Barichella, M. *et al.* Muscle-targeted nutritional support for rehabilitation in patients  
1083 with parkinsonian syndrome. *Neurology* **93**, e485–e496 (2019).  
1084 124. Lewis, C. M. & Vassos, E. Polygenic risk scores: from research tools to clinical  
1085 instruments. *Genome Med.* **12**, 44 (2020).  
1086 125. Albiñana, C. *et al.* Multi-PGS enhances polygenic prediction by combining 937  
1087 polygenic scores. *Nat. Commun.* **14**, 4702 (2023).  
1088 126. Yang, L. G., March, Z. M., Stephenson, R. A. & Narayan, P. S. Apolipoprotein E in  
1089 lipid metabolism and neurodegenerative disease. *Trends Endocrinol. Metab.* **34**, 430–  
1090 445 (2023).  
1091