

Genetic structure of major depression symptoms across clinical and community cohorts

Mark J Adams¹, Jackson G Thorp², Bradley S Jermy³, Alex S F Kwong^{1,4}, Kadri Kõiv⁵, Andrew D Grotzinger^{6,7}, Michel G Nivard⁸, Sally Marshall⁹, Yuri Milaneschi¹⁰, Bernhard T Baune^{11,12,13}, Bertram Müller-Myhsok^{14,15,16}, Brenda WJH Penninx¹⁰, Dorret I Boomsma¹⁷, Douglas F Levinson¹⁸, Gerome Breen^{19,20}, Giorgio Pistis²¹, Hans J Grabe²², Henning Tiemeier^{23,24}, Klaus Berger²⁵, Marcella Rietschel²⁶, Patrik K Magnusson²⁷, Rudolf Uher²⁸, Steven P Hamilton²⁹, Susanne Lucae³⁰, Kelli Lehto⁵, Qingqin S Li³¹, Enda M Byrne³², Ian B Hickie³³, Nicholas G Martin², Sarah E Medland², Naomi R Wray^{34,35}, Elliot M Tucker-Drob^{36,37}, Estonian Biobank Research Team³⁸, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Cathryn M Lewis^{19,39}, Andrew M McIntosh^{1,40}, Eske M Derks²

¹ Division of Psychiatry, University of Edinburgh, Edinburgh, UK

² Mental Health and Neuroscience, QIMR Berghofer Medical Research Institute, Brisbane, QLD, AU

³ Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, FI

⁴ MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

⁵ Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, EE

⁶ Department of Psychology and Neuroscience, University of Colorado at Boulder, Boulder, CO, US

⁷ Institute for Behavioral Genetics, University of Colorado at Boulder, Boulder, CO, US

⁸ Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, NL

⁹ Centre for Genomic & Experimental Medicine, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, UK

¹⁰ Department of Psychiatry, Amsterdam Public Health and Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, NL

¹¹ Department of Psychiatry, University of Melbourne, Melbourne, VIC, AU

¹² Florey Institute of Neuroscience and Mental Health, University of Melbourne, Melbourne, VIC, AU

¹³ Department of Psychiatry, University of Münster, Münster, NRW, DE

¹⁴ Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, BY, DE

¹⁵ Munich Cluster for Systems Neurology (SyNergy), Munich, BY, DE

¹⁶ Institute of Population Health, University of Liverpool, Liverpool, UK

¹⁷ Department of Biological Psychology & Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, NL

¹⁸ Department of Psychiatry & Behavioral Sciences, Stanford University, Stanford, CA, US

¹⁹ Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK

²⁰ NIHR Maudsley Biomedical Research Centre, King's College London, London, UK

²¹ Department of Psychiatry, Lausanne University Hospital and University of Lausanne, Prilly, VD, CH

²² Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald,

MV, DE

²³ Child and Adolescent Psychiatry, Erasmus University Medical Center Rotterdam, Rotterdam, NL

²⁴ Social and Behavioral Science, Harvard T.H. Chan School of Public Health, Boston, MA, US

²⁵ Institute of Epidemiology and Social Medicine, University of Münster, Münster, NRW, DE

²⁶ Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, BW, DE

²⁷ Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SE

²⁸ Psychiatry, Dalhousie University, Halifax, NS, CA

²⁹ Psychiatry, Kaiser Permanente Northern California, San Francisco, CA, US

³⁰ Max Planck Institute of Psychiatry, Munich, BY, DE

³¹ Neuroscience Therapeutic Area, Janssen Research and Development, LLC, Titusville, NJ, US

³² Child Health Research Centre, University of Queensland, Brisbane, QLD, AU

³³ Brain and Mind Centre, University of Sydney, Sydney, NSW, AU

³⁴ Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD, AU

³⁵ Queensland Brain Institute, University of Queensland, Brisbane, QLD, AU

³⁶ Department of Psychology, University of Texas at Austin, Austin, TX, US

³⁷ Population Research Center, University of Texas at Austin, Austin, TX, US

³⁸ Andres Metspalu, Lili Milani, Tõnu Esko, Reedik Mägi, Mari Nelis & Georgi Hudjashov

³⁹ Department of Medical & Molecular Genetics, King's College London, London, UK

⁴⁰ Institute for Genomics and Cancer, University of Edinburgh, Edinburgh, UK

Abstract

Diagnostic criteria for major depressive disorder allow for heterogeneous symptom profiles but genetic analysis of major depressive symptoms has the potential to identify clinical and aetiological subtypes. There are several challenges to integrating symptom data from genetically-informative cohorts, such as sample size differences between clinical and community cohorts and various patterns of missing data. We conducted genome-wide association studies of major depressive symptoms in three clinical cohorts that were enriched for affected participants (Psychiatric Genomics Consortium, Australian Genetics of Depression Study, Generation Scotland) and three community cohorts (Avon Longitudinal Study of Parents and Children, Estonian Biobank, and UK Biobank). We fit a series of confirmatory factor models with factors that accounted for how symptom data was sampled and then compared alternative models with different symptom factors. The best fitting model had a distinct factor for *Appetite/Weight* symptoms and an additional measurement factor that accounted for missing data patterns in the community cohorts (use of Depression and Anhedonia as gating symptoms). The results show the importance of assessing the directionality of symptoms (such as hypersomnia versus insomnia) and of accounting for study and measurement design when meta-analysing genetic association data.

Introduction

Major depressive disorder (MDD) is a mood disorder characterized by low mood, loss of interest or pleasure (anhedonia), irritable affect, biological symptoms (psychomotor agitation/slowing, altered sleep patterns, changes in appetite or weight), negative thought

content, and associated loss of function. To qualify for a diagnosis of major depression, the standard diagnostic classification systems (American Psychiatric Association, 2000, 2013; World Health Organization, 1992) require one of two cardinal symptoms plus at least four other symptoms to be present. Although conceptualized as a single disorder, the diagnostic criterion for MDD can be met with any combination of these other symptoms. For the DSM-5, this entails that there are 227 symptom profiles that would lead to a diagnosis of major depression (Zimmerman et al., 2015). When considering all potential symptom states (such as increasing versus decreasing appetite) the number of possible symptom profiles blooms into the thousands (Fried & Nesse, 2015a).

A single categorical phenotype---that might mask a multitude of separate disorder types---stymies the testing of correlates and treatments. Network analysis has shown that MDD symptoms are not all equally related to each other (Borsboom & Cramer, 2013) and latent class analysis has been used to identify several MDD subtypes with differing patterns of symptoms and differential association with demographic, psychological, and physical health factors (Lamers et al., 2010). However, the potential concealed heterogeneity within the MDD diagnosis does have an upper bound: only around one quarter of the potential symptom profiles are actually observed (Fried & Nesse, 2015a; Zimmerman et al., 2015). This suggests there is both regularity and variation in symptom presentation.

Analysing individual symptoms is one way to unwrap the heterogeneity of MDD (Cai et al., 2020; Fried & Nesse, 2015b). Phenotypic studies have derived and tested factor structures of MDD symptoms (Elhai et al., 2012; Krause et al., 2008, 2010) and twin models have been used to separate genetic from environmental sources of symptom covariance (Kendler et al., 2013). These models grouped symptoms together in two or three factors, which broadly contrast psychological versus somatic symptoms. The primary difference among the proposed two factor structures is whether psychological symptoms including anhedonia and concentration problems group with the cognitive/affective symptoms or with the somatic symptoms. Three factor models have instead posited splitting the psychological symptoms into affective and cognitive components. Clinical subtypes are also part of diagnostic criteria and these have been used to classify depression profiles that are differentially associated with specific clinical, behavioural and biological correlates (Milaneschi et al., 2020; Penninx et al., 2013).

More recently, genetic studies of depressive symptoms have updated the findings from twin models using data from genome-wide association studies (GWAS). A confirmatory factor analysis of genetic covariance estimates obtained from GWAS results on current depressive symptoms showed that a psychological and somatic factor had the best fit to the data (Thorp et al., 2020). The detection of genetic correlates specific to each symptom implies that symptoms may have differing genetic causes and consequences, even if the symptoms themselves are highly genetically correlated.

Understanding the genetic architecture of MDD symptoms is complicated by symptom ascertainment. In clinical samples, symptom data is often only available on affected participants, and is thus conditioned on having been diagnosed with depression. Conditioning data presence on a diagnosis can induce downward bias in correlations amongst the symptoms comprising that diagnosis. However, when symptom data is missing in controls, imputing the absence of the symptom in controls and including them in the analysis has the potential to recapitulate the signal from a case/control analysis rather than reveal genetic variance that is unique to each symptom. In community cohorts, participants are typically screened for the presence of cardinal symptoms (depressed mood and anhedonia) and only participants who report at least one cardinal symptom are assessed for other symptoms of depression, which also leads to high

levels of missing symptom data in these cohorts. Moreover, because community samples often contain symptom but not diagnostic information, many GWAS purporting to investigate MDD may actually be better characterized as investigating a broader dysphoria continuum rather than MDD specifically (Flint, 2023). Because community cohorts tend to have a larger sample size than clinical cohorts, meta-analysing all data together therefore has the potential to dilute information on case subtypes.

In this study we sought to uncover the genetic structure of depression symptoms while accounting for how samples were recruited and how symptoms were assessed. We did this by conducting GWAS of individual symptoms of depression, testing factor models to investigate genetic heterogeneity as a function of sample ascertainment (Clinical vs Community) and measurement (with or without screening based on cardinal/gating symptoms). Finally, we assessed the validity of the identified latent factors of depression by estimating genetic correlations with external traits.

Specifically, we conducted GWAS of symptom data in six cohorts and meta-analysed them in groups based on sample ascertainment. The first group (the “Clinical” cohorts) consisted of clinical cases from the Psychiatric Genomics Consortium MDD cohorts, participants from the Australian Genetics of Depression study who were recruited based on depression diagnosis, and participants from Generation Scotland who met DSM criteria for depression. The second group (the “Community” cohorts) consisted of the Avon Longitudinal Study of Parents and Children, Estonian Biobank, and UK Biobank, and thus contained data on participants who were not recruited with respect to depression status. Using the two sets of meta-analysed symptom GWASs, we first constructed and tested factor models that accounted for how the samples were recruited (Clinical versus Community) and how symptoms were assessed (such as gating symptoms in the Community cohorts). After understanding the measurements structure of the symptom GWASs, we then compared alternative factor models for the symptoms based on previous literature and diagnostic specifiers for depressive disorders. Using the best fitting overall model, we tested for shared and specific genetic correlates with other psychiatric, behavioral, and metabolic phenotypes that have known genetic links to MDD.

Methods

Samples and symptom measures

We analysed depression symptom data in six studies: the Psychiatric Genomics Consortium, the Australian Genetics of Depression Study, Generation Scotland, the Avon Longitudinal Study of Parents and Children, Estonian Biobank, and UK Biobank. Table 1 describes the number of participants with and without each symptom for each grouping of studies that were analysed. See Supplementary Material for information on genotyping and imputation.

Data from the Psychiatric Genomics Consortium (PGC) was drawn from 23 cohorts in the Wave 1 and Wave 2 datasets of the Major Depressive Disorder Working Group (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, 2013; Wray et al., 2018). Symptoms were assessed by trained interviewers using structured diagnostic instruments and DSM checklists. Because information on symptom presence was not available for control participants in most cohorts, participants with a diagnosis of depression were selected for analysis (N = 12,821).

The Australian Genetics of Depression Study (AGDS) (Byrne et al., 2020; Mitchell et al., 2022) is a study of depression and therapeutic response recruited using nationwide prescribing history and through publicity targeting adults who are or had ever been treated for clinical depression (N = 20,689). Symptoms experienced during the participant’s worst period of depression were assessed using the Composite International Diagnostic Interview (CIDI) Short Form (Hickie et al., 2001) and administered through an online questionnaire. Because the study was enriched for participants with a history of being diagnosed with or treated for depression, AGDS was grouped as a Clinical cohort.

Generation Scotland: Scottish Family Health Study (GS:SFHS) is a study of 7,000 families recruited from the general population of Scotland (Smith et al., 2012). Participants who screened reported seeking help for emotional or psychiatric problems were administered an in-person structured interview (Fernandez-Pujals et al., 2015; Smith et al., 2012); and a subset participated in an online follow-up that included a CIDI (Composite International Diagnostic Interview) questionnaire. Symptom data was analysed on participants who met DSM criteria for depression at either time point (N = 3,493).

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a UK-based population birth cohort (Boyd et al., 2013). Participants were from the children sample (N = 13,988) with symptoms present during the last two weeks assessed using the Clinical Interview Schedule Revised (CIS-R) (Lewis et al., 1992) collected during clinical visits at ages 18 and 24. Participants were considered to have had a symptom if they reported it at either measurement occasion.

Estonian Biobank (EstBB) is a population health cohort recruited from medical practitioners in Estonia (Leitsalu et al., 2015). Participants responded to a CIDI questionnaire of depression symptoms during the Mental Health online Survey (MHoS) recontact. Participants were first screened of the presence of low mood or anhedonia and then asked about symptoms during the worst period of depression (N = 84,079).

UK Biobank (UKB) is a population health cohort recruited from general practitioners in the United Kingdom (Sudlow et al., 2015). Lifetime depression symptoms were assessed during online recontact and taken from the CIDI portion of the Mental Health Questionnaire (Davis et al., 2020) (UKB-MHQ, N=157,366) and from assessments of low mood and anhedonia from the baseline touchscreen questionnaire (UKB Touchscreen, N=222,061). For the CIDI, low mood and anhedonia were used as gating symptoms, where participants had to endorse at least one to be asked about the other symptoms.

		Cohort N Symptom Present : Absent Sample Prevalence		
Symptom	Abbr.	<u>Clinical cohorts meta</u>	<u>Community cohorts meta</u>	<u>UKB Touchscreen</u>
1. Depressed mood	Dep	21681 : 1748 93%	107956 : 99480 52%	71964 : 57130 56%
2. Anhedonia	Anh	24732 : 2801 90%	181113 : 126167 39%	46952 : 80366 37%

3a. Weight loss / decrease in appetite	AppDec	9265 : 14594 39%	39453 : 36497 52%	0 : 0
3b. Weight gain / increase in appetite	AppInc	7902 : 13167 38%	22612 : 36489 38%	0 : 0
4a. Insomnia	SleDec	18917 : 6573 74%	73144 : 19851 79%	0 : 0
4b. Hypersomnia	SleInc	10586 : 11050 49%	20125 : 20055 50%	0 : 0
5a. Psychomotor agitation	MotoInc	10447 : 12372 46%	113 : 3181 3%	0 : 0
5b. Psychomotor slowing	MotoDec	12701 : 11214 53%	299 : 2995 9%	0 : 0
6. Fatigue	Fatig	23941 : 2497 91%	85304 : 16736 84%	0 : 0
7. Feelings of worthlessness / guilt	Guilt	21921 : 3888 85%	61757 : 43570 59%	0 : 0
8. Diminished concentration	Conc	23974 : 2386 91%	75190 : 23416 76%	0 : 0
9. Recurrent thoughts of death or suicide	Sui	18170 : 9609 65%	46984 : 58885 44%	0 : 0

Table 1. Sample size counts and sample prevalences of presence and absence of each symptom for participants used in the genetic analyses. Meta-analysis of Clinical (PGC, AGDS, GenScot), Community (ALSPAC, EstBB, UKB-MHQ) and UKB Touchscreen.

Genome-wide association meta-analysis

Genome-wide association study (GWAS) analyses were conducted on each symptom separately in the cohorts (PGC, AGDS, GS:SFHS, ALSPAC, EstBB, UKB-MHQ) on participants who had genetic similarity with each other and the 1000 Genomes European reference. Participants in UKB who clustered with other reference populations were not analysed because sample sizes did not meet the threshold for LD score estimation ($N > 5000$). See Supplementary Material for more information on the individual study GWASs. We meta-analysed the GWAS summary statistics based on the ascertainment design. PGC, AGDS, and GS:SFHS were meta-analysed together to form the “Clinical” symptom summary statistics; and ALSPAC, EstBB, and UKB-MHQ were meta-analysed together for the “Community” summary statistics. We performed the meta-analyses using Ricopili (Lam et al., 2020) and calculated SNP-based heritability using LD Score Regression (LDSC) (Bulik-Sullivan et al., 2015). For input into LDSC we set the sample size equal to the sum of effective sample sizes of each cohort in the meta-analysis and then specified sample prevalences of 50% (Grotzinger et al., 2022). Symptoms' population prevalences were estimated for the Clinical cohorts by multiplying the observed sample prevalence by the prevalence of MDD (15%) and for the Community cohorts by multiplying by the proportion of participants in the UKB MHQ sample who were positive on either one the gating symptoms. We

assessed significant associations in the meta-analysed summary statistics at $p < 5 \times 10^{-8} / 22$ (the number of meta-analyses conducted) or at $p < 5 \times 10^{-8}$ with prior association or biological evidence at the locus.

Confirmatory factor analysis of Genetic Covariance Structure

We fit confirmatory genetic factor analysis models to the meta-analysed ascertainment cohort (i.e., Clinical and Community-based) and UKB Touchscreen summary statistics for each symptom using Genomic SEM (Grotzinger et al., 2019). We first fit a common factor model, where all symptoms load on a single factor as a baseline, using symptoms with a non-negative LDSC heritability (Model A). To explore how sample ascertainment influenced the genetic correlations among the symptoms, we fit a series of models that captured various aspects of the sampling and measurement processes. We then used these results to inform the construction of models that grouped the symptoms based on previous findings and diagnostic criteria. We assessed relative model fit using Akaike Information Criterion (AIC) to pick the best model and absolute model fit with Standardized Root Mean Square Residual (SRMR) to determine how well the model was capturing the genetic correlations among symptoms. We also examined residual correlations to understand what aspects of symptom structure were not being captured. Factor structures are listed in Supplementary Table S4 and illustrated in Supplementary Figure S1.

Ascertainment/measurement models

The most pertinent measurement difference among the symptoms was which meta-analysed cohorts the symptom came from, so we created a two-factor model where all symptoms from the same cohorts (Clinical or Community) loaded on the same factor (Model B). The next model considered the effect of the cardinal symptoms as gating items in UK Biobank and posited a general MDD factor that all the symptoms loaded on alongside an uncorrelated Gating factor with loadings from just the Community and UKB Touchscreen *low mood* and *anhedonia* symptoms (Model C). The Gating factor would therefore isolate variation associated with differences across the full non-clinical (dysphoria) to clinical spectrum. Symptoms not loading on the gating factor (i.e., those for which data are conditional on the presence of the two gating symptoms) represent variation within the more severe region of the spectrum and are thus more directly comparable to analyses of data from cases only. We then combined the Clinical-Community and Gating models to create a three-factor model (Model D).

Symptom models

Based off the best measurement model, we then fit models that grouped symptoms into two or three factors based on previous findings from phenotypic, twin, and Genomic SEM models and from diagnostic criteria. The two factor models grouped symptoms into Psychological and Somatic (Model E); Psychological and Neurovegetative (Model F); or Affective and Neurovegetative (Model G) factors (Elhai et al., 2012; Krause et al., 2008, 2010; Thorp et al., 2020). The Affective factor contained symptoms *low mood*, *feelings of guilt*, and *suicidality*. The Psychological factors broadened the Affective factor to include the symptoms *anhedonia* and/or *loss of concentration*. The Somatic factor included the *appetite*, *sleep*, *fatigue*, and *psychomotor* symptoms. The Neurovegetative factors incorporated the somatic symptoms while also including *loss of concentration* and/or *anhedonia*. A three factor model (Model H) loaded symptoms onto cognitive (*feelings of guilt*, *loss of concentration*, *suicidality*), mood (*low mood*, *anhedonia*, *feelings of guilt*), and neurovegetative (*appetite*, *sleep*, *fatigue*, *psychomotor*) (Kendler et al., 2013).

We also fit factor models that disaggregated symptoms that involved an increasing or decreasing change (*appetite/weight, sleep, psychomotor*). One such model (Model I) was based on previous findings that identified factors for Appetite (*appetite/weight decrease and increase*), vegetative (*hypersomnia, psychomotor slowing, fatigue, concentration*) and Cognitive/Mood (*low mood, anhedonia, insomnia, psychomotor agitation, feelings of guilt, suicidality*) (van Loo et al., 2022). Finally, we considered a three-factor model (Model J) based on diagnostic criteria of melancholic depression (*anhedonia, insomnia, psychomotor agitation, appetite/weight decrease, feelings of guilt*) and atypical depression (*hypersomnia, appetite/weight increase, psychomotor slowing, fatigue*), with the remaining symptoms loading on an Affective/Cognitive factor (*low mood, suicidality, loss of concentration*).

Genomic factor meta-analysis

We conducted multivariate meta-analyses of symptoms in Genomic SEM (Grotzinger et al., 2019). Because of low power in some of the Clinical and Community symptoms summary statistics, we were unable to fully test SNP effects on symptom factors. We therefore fit a model with single common factor meta-analysis across well-powered symptoms ($h_{\text{SNP}}^2 > 0, g_{\text{cov}}$ intercept $> 1, \bar{\chi}^2 > 1.02$) and tested for SNP heterogeneity at the level of individual symptoms. A genomic factor meta-analysis estimates SNP associations in two structural models: a common pathway model where the SNP is associated with each symptom through its effect on the factor and an independent pathway model where the SNP is associated independently with each symptom, bypassing the common factor. The SNP coefficients from the common pathway model act as a meta-analysis of the symptom summary statistics while accounting for sample overlap. A comparison of fit between the common and independent pathway models yields a heterogeneity statistic for each SNP, Q_{SNP} , indicating whether the SNP's association varies between symptoms.

Regression Analysis

Using the best fitting models, we tested how the factors were related to correlates of depression. We selected phenotypes that are known to genetically correlate with depression, including psychiatric disorders (anxiety disorder, bipolar disorder, PTSD, schizophrenia); depression defined through clinical ascertainment (major depressive disorder) and through broader or more minimal definitions (major depression); and other health, behavioural, and social phenotypes (see Supplementary Materials for list of studies). We tested whether the other phenotypes had specific genetic correlations with each symptom factor. We did this by first fitting single regressions of a phenotype on each symptom factor. We then compared this to a multiple regression of the phenotype on all symptom factors simultaneously. We used Benjamini-Yekutieli FDR adjustment to correct for multiple testing (Benjamini & Yekutieli, 2001).

Results

Genome-wide association and meta-analyses

We conducted GWAS for each symptom separately in all cohorts and meta-analysed within sample ascertainment groups (*Clinical* or enriched cohorts: PGC, AGDS, GS:SFHS; *Community* cohorts: ALSPAC, EstBB, UKB-MHQ) (Supplementary Table S1). Two associations met the stringent multiple testing burden ($p < 5 \times 10^{-8} / 22$). One was an intron in *FTO* (ENSG00000140718, alpha-ketoglutarate dependent dioxygenase, a gene involved in food intake) associated with *Weight gain* in the Community cohorts. The other was associated with

Anhedonia in the Community cohorts and was an intron variant in an uncharacterised non-coding RNA gene (LOC105379109/ENSG00000251574) and in a region previously associated with neuroticism, depression, and subjective well-being.

At the genome-wide significance threshold ($p < 5 \times 10^{-8}$) there were three associations that were also supported by prior evidence. There were two associations with *Depressed mood* in the Community cohorts: an intron in COMP (ENSG00000105664, cartilage oligomeric matrix protein) also near *CRTC1* (ENSG00000105662, CREB regulated transcription coactivator 1, a gene that regulates metabolism); and an intron in an uncharacterised gene (LOC107986777) regionally associated with depression. An upstream variant for an uncharacterised long intergenic non-protein coding RNA (LINC01938) was associated with Community *Anhedonia* and in a region previously associated with neuroticism and major depressive disorder.

LDSC-estimated heritabilities were primarily in the 0.025–0.1 range (Figure 1, Supplementary Table S3). Many of the symptoms in the Clinical cohorts (*Depressed mood*, *Anhedonia*, *Fatigue*, *Concentration*) had negative heritabilities and the psychomotor symptoms from the Community cohorts did not meet the sample size inclusion criteria ($N_{\text{Eff}} > 5000$).

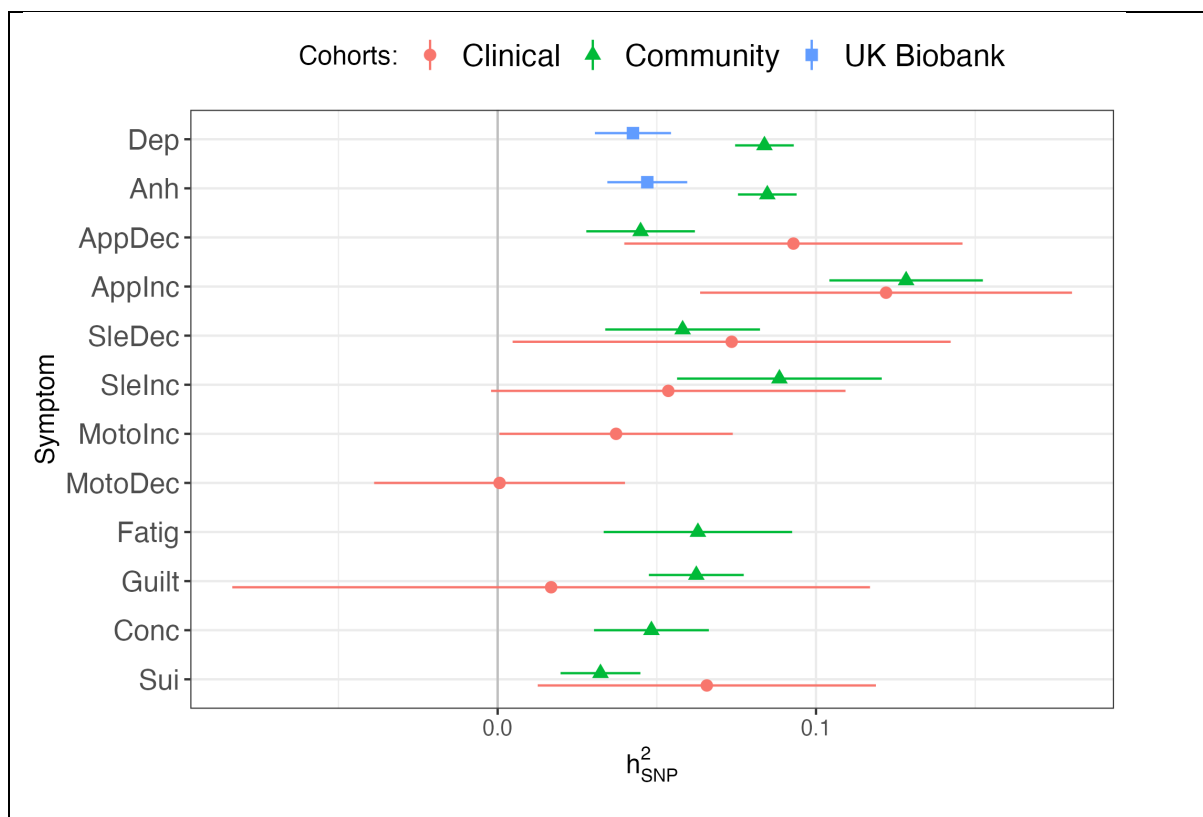


Figure 1. LDSC-estimated heritabilities (h^2_{SNP}) on the liability scale of depression symptoms (abbreviations are listed in Table 1) for summary statistics that met inclusion criteria ($N_{\text{Eff}} > 5000$, $h^2_{\text{SNP}} > 0$). Clinical = PGC + AGDS + GS:SFHS meta-analysis, Community = ALSPAC + EstBB + UKB-MHQ meta-analysis, UK Biobank = UKB-Touchscreen GWAS.

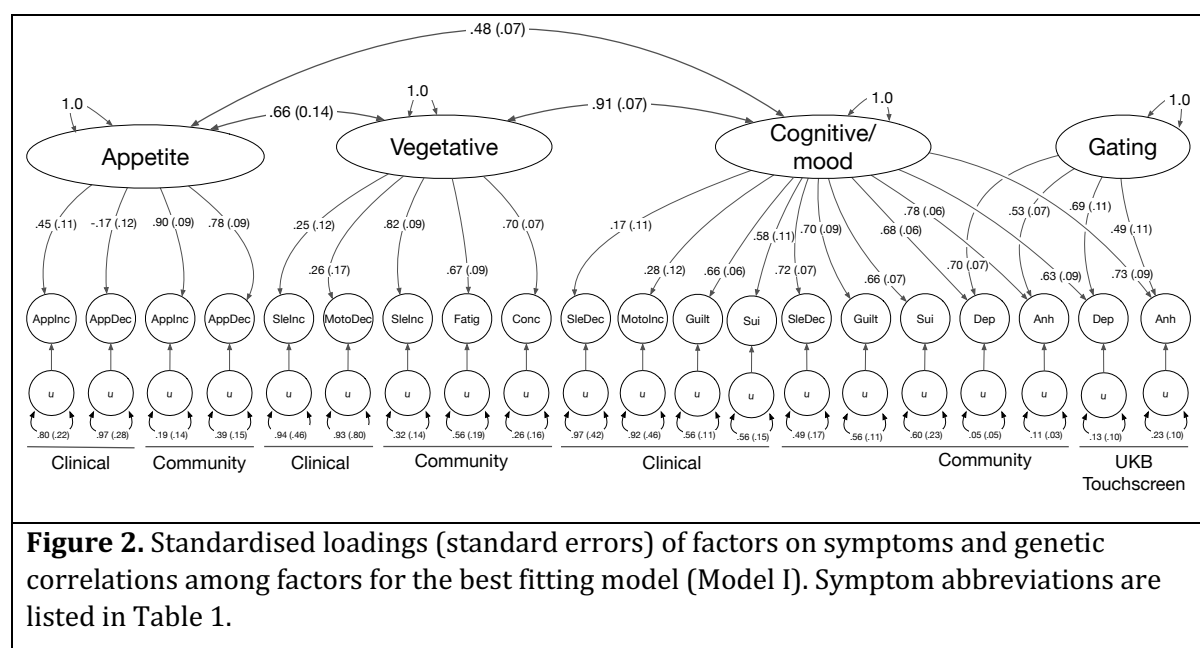
Confirmatory factor analysis

We brought forward symptoms from the Clinical and Community cohorts meta-analyses and the UKB Touchscreen assessment that had a h_{SNP}^2 greater than 0 and sample sizes > 5000 for confirmatory factor analysis (Supplementary Table S4, Supplementary Figure S1a-l).

A common factor model (A) of the symptoms showed poor fit (CI=0.932, SMR=0.169, AIC = 5355). A model (B) with separate factors for Clinical and Community cohort symptoms had slightly poorer fit (AIC = 5369) and yielded a genetic correlation between the two factors of $r_g = 0.63 \pm 0.14$, $p = 1.3 \times 10^{-5}$. An alternative model (C) that only split off the Community and UKB-Touchscreen *mood* and *anhedonia* symptoms into an orthogonal factor, capturing these symptoms use as gating items in EstBB and UKB-MHQ, showed substantially improved fit (AIC = 3229). A model (D) combining the sample factors with the orthogonal Gating factor also improved model fit (AIC = 3285) and led to a nominal increase in the genetic correlation between the Clinical and Community factors to $r_g = 0.75 \pm 0.17$, $p = 6.9 \times 10^{-6}$.

We then tested whether models that grouped symptoms together across cohorts fit better than the factor models based on sampling methodology. Because the addition of an orthogonal Gating factor improved model fit so substantially, the symptom-oriented models all included this factor. The best fitting of the symptom models was Model I which included factors capturing Appetite, Vegetative, and Cognitive/Mood symptoms (Figure 2). The Appetite factor had a genetic correlation of $r_g = 0.66 \pm 0.09$ ($p = 1.4 \times 10^{-12}$) with the Vegetative factor and $r_g = 0.48 \pm 0.07$ ($p = 3.4 \times 10^{-13}$) with the Cognitive/Mood factor, while the Vegetative and Cognitive/Mood factors were more highly correlated, $r_g = 0.91 \pm 0.07$ ($p = 7.8 \times 10^{-40}$).

None of the models fully captured the genetic correlations among the symptoms, as indicated by high SRMR (Model D = 0.149, Model I = 0.147). An inspection of the residual genetic correlations (Supplementary Figure S3d) indicated correlations between the same symptoms across the two cohorts (e.g., Clinical *appetite decrease* with Community *appetite decrease*) were not fully represented by the factor structure. We thus tested how adding residual correlations between symptoms that included from both cohorts (*appetite decrease*, *appetite increase*, *insomnia*, *hypersomnia*, and *suicidality*) improved absolute model fit. The addition of these residual correlations lowered SRMR to 0.140.



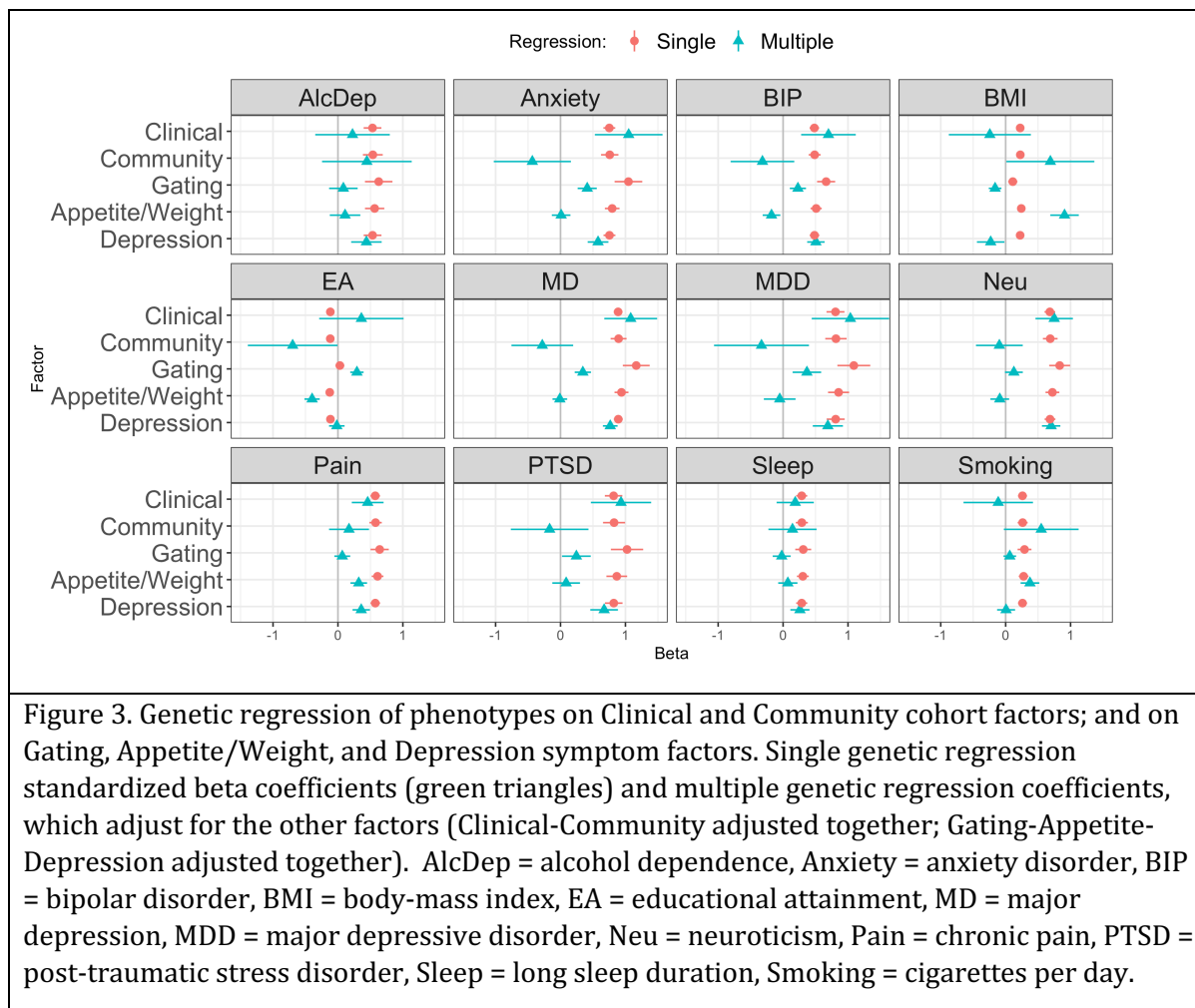
Multivariate meta-analysis of symptoms

Because many of the symptom summary statistics were low powered, we were unable to conduct a multivariate meta-analysis using the genetic factors. Alternatively, to test for SNP effects specific to each symptom, we conducted a multivariate meta-analysis of well-powered symptoms summary statistics. The common factor had three genome-wide significant associated variants: an intron variant in BRINP2 (ENSG00000198797, BMP/retinoic acid inducible neural specific 2, a regulator of neuronal differentiation); the upstream variant that was associated with Community *Anhedonia* symptoms; and an intron variant in LRRC37A3 (ENSG00000176809, leucine rich repeat containing 37 member A3) An intron in the FTO gene showed substantial heterogeneity in the common factor meta-analysis (Supplementary Table S7).

Genetic multiple regression

To determine whether each MDD symptom factor had specific genetic relationships with twelve phenotypes, we conducted genetic multiple regressions in GenomicSEM using the Clinical–Community sample and Appetite–Vegetative–Cognitive/Mood symptom factor models. Because the Vegetative and Cognitive/Mood symptom factors had a high genetic correlation, we combined these into one factor, which we labelled Depression. We first calculated single regressions of each phenotype on each of the factors separately, where the standardized coefficient indicated the overall shared genetics with each factor. We then then fit a genetic multiple regression with the two models, where the standardized coefficient represents the unique genetic relationship of the phenotype with each factor, after adjusting for shared overlap with the other factors (e.g., a phenotype's genetic relationship with the Clinical factor after adjusting for the Community factor, and vice versa; and its genetic relationship with the Appetite factor after adjusting for the Depression and Gating factors, and vice versa). In the single regression (unadjusted) analysis, the genetic relationship of each phenotype with all of the factors were in the same direction with the exception of educational attainment which had a negative relationship with most of the factors (at $p < 0.0005$) but a positive yet non-significant relationship with the Gating factor (Figure 3, Supplementary Table S7). After adjusting for shared overlap with the Clinical factor, the Community factor did not have any specific relationships with the phenotypes, while the Clinical factor had specific, positive genetic correlations with anxiety, bipolar disorder, major depression, major depressive disorder, neuroticism, PTSD, and chronic pain.

When adjusting for the Depression and Gating factors, the factor for the Appetite symptoms had specific genetic correlations with BMI and smoking. The factor for the remaining Depression symptoms was specifically correlated with alcohol dependence, anxiety, neuroticism, and long sleep duration. The Gating factor did not have any specific genetic correlates after adjusting for the other factors, but it did share with the Depression factor positive genetic correlations with bipolar disorder, major depression and major depressive disorder, and PTSD after adjustment for the Appetite factor (Figure 3, Supplementary Table S7).



Discussion

We used genome-wide association data to analyse the genetic relationships among symptoms of depression based on cohort sampling and symptom content and to estimate whether the genetic factors had specific correlates with other phenotypes. We analysed data from two sets of cohorts: Clinical cohorts that were ascertained to have depression through clinical or interview assessments or were recruited preferentially on a history of treatment for depression; and Community cohorts that were not recruited based on disease status (but for which symptom data was typically conditioned based on endorsement of cardinal gating symptoms). We conducted GWAS of major depression symptoms in each cohort then meta-analysed within the Clinical and Community groups.

We identified loci associated with individual major depression symptoms and with a common genetic factor of the symptoms. Several associations from the individual symptoms and common factor meta-analysis (rs7515828, rs30266, s6884321, rs150046352) have been identified previously in GWAS of or unipolar depression (EFO ID EFO_0003761) (Sollis et al., 2023) or in meta-analyses of major depressive disorder (Als et al., 2022; Howard et al., 2019; Levey et al., 2021; Wray et al., 2018) SNPs associated with *Appetite / weight increase* have primarily come up in GWAS of body mass index and related traits (Elsworth et al., 2020;

Hoffmann et al., 2018; Howe et al., 2022; Yengo et al., 2018) but another SNP in the FTO gene has also been associated with atypical subtypes (Milaneschi et al., 2014).

Many symptoms in the Clinical cohorts had heritability estimates that were zero or negative. This is not unexpected. Selecting individuals based on their phenotype (that is, that they are or have been affected by depression) would result in a sample that has a lower genetic variance for traits that contribute to the determination of the phenotype (Falconer & Mackay, 1996). Additionally, the symptoms with negative heritabilities (*depressed mood, anhedonia, fatigue, concentration problems* in the clinical cohorts) also had high endorsement rates (85–94%), and power to detect heritability is reduced the further the sample prevalence deviates from 50% (Lee et al., 2011). In contrast, the other symptoms had more equal endorsement rates (34–75%) among affected participants. Although the low heritabilities of symptoms from the Clinical cohorts limited the comprehensiveness of alternative factor models that could be tested, we did find congruence between the Clinical and Community cohort symptoms, with a high genetic correlation between their respective factors. We also showed that model fit was substantially improved by modeling the use of cardinal symptoms (*Low mood* and *Anhedonia*) as gating items for surveys of depression symptoms. Among the models that grouped symptoms together without consideration for symptom direction, we found broad support for a split between psychological and somatic symptoms identified in previous phenotypic (Elhai et al., 2012) and genetic (Thorp et al., 2020) analyses. When directional symptoms were portioned out based on diagnostic specifiers, we found that a three-factor model capturing Appetite, Vegetative, and Cognitive/Mood symptoms (van Loo et al., 2022) had the best fit among all models considered. The correlations among the factors indicated that the Vegetative and Cognitive/Mood symptoms should be grouped together, with only the Appetite symptoms making up a possibly different dimension of depression. However, the Clinical cohort symptoms had low loadings in both the sample-based and symptom-based models (except for the Clinical *Appetite/Weight* and *Suicidality* symptoms), and thus the model fit was driven primarily by capturing the structure among the Community cohort symptoms. This observation is consistent with the fact that the Clinical cohorts are more selected than the community cohorts, and that conditioning data presence on a diagnosis can induce downward bias in correlations amongst the symptoms that aggregate to form the diagnosis. Similar attenuation, albeit to a lesser degree, may be expected for items in community samples whose presence was conditioned on endorsement of cardinal symptoms.

Despite these limitations, the Clinical factor was genetically correlated with all the other phenotypes selected for comparison. A multiple genetic correlation analysis showed that the Clinical and Community factors had a shared genetic relationship with alcohol dependence, bipolar disorder, BMI, educational attainment, MDD, chronic pain, sleep, and smoking. The symptom-based factor model showed discriminative validity between the Appetite symptoms and the rest of the symptoms of depression through relationships with correlates of depression. A positive genetic correlation between increase in appetite/weight with BMI has previously been shown with PGC cohorts (Milaneschi et al., 2017) and in UKB (Badini et al., 2022), and our findings show that this result holds even when adjusting for genetic overlap with other symptoms.

Our results demonstrate the challenges and insights associated with considering symptoms of depression separately. In particular, substantial care must be taken to consider how samples are ascertained (clinical versus community recruitment), how symptoms are measured (the use of gating items in symptom inventories), and including assessments of item direction (e.g., insomnia versus hypersomnia) when modelling the genetic structure of depression symptoms. However, the evaluation of direction was limited to a small subset of symptoms and did not

include distinctions such, as low versus irritable mood, or included only partial assessments, such as weight but not appetite changes being assessed in UKB. The coverage of features of atypical and melancholic depression was likewise incomplete. For example, several diagnostic features of the atypical specifier were not included, such as mood reactivity, sensitivity to interpersonal rejection, and leaden paralysis. We also only examined subtypes defined by symptom profiles and not other sources of heterogeneity such as onset, life event exposure, or treatment outcomes (Harald & Gordon, 2012) which may also have a differential biological and genetic basis (Beijers et al., 2019; Milaneschi et al., 2020; Nguyen et al., 2022). The strongest genetic associations were between symptoms of weight/appetite change and genes linked to satiety and metabolism. This highlights the need to phenotype somatic symptoms (weight or sleep changes and fatigue) outside of the context of mental health assessments, so that their specific role in depression can be better isolated. Likewise, the use of gating symptoms makes it difficult to fully capture the range of genetic risk between everyday dysphoria and differences among affected individuals. While the results support the idea that depression is heterogeneous, the genetic liability for symptom profiles and comorbidities can be captured in relatively few dimensions.

Code and data availability

Primary code is available from the Psychiatric Genomics Consortium (PGC) GitHub Repository (<https://github.com/psychiatric-genomics-consortium/mdd-symptom-gwas/>) and meta-analysed summary statistics are available for download from the PGC website (<https://www.med.unc.edu/pgc/download-results/>). Individual-level PGC data is available by application to the PGC Data Access Committee (<https://www.med.unc.edu/pgc/shared-methods/>). Data from Estonian Biobank (<https://genomics.ut.ee/en/content/estonian-biobank>), UK Biobank (<https://www.ukbiobank.ac.uk>) and ALSPAC (<http://www.bristol.ac.uk/alspac/>) are available to bonafide researchers upon application. Data from AGDS is available for collaboration by contacting NGM (Nick.Martin@qimrberghofer.edu.au).

Acknowledgments

AGDS supported by National Health and Medical Research Council (NHMRC) (1086683, 1145645, 1078901, 1087889, 1173790); ALSPAC supported by the Medical Research Council and Wellcome Trust (217065/Z/19/Z). Estonian Biobank supported by the European Union through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012). NTR-NESDA supported by Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL; 184.021.007 and 184.033.111), National Institutes of Health (NIH, R01D0042157-01A, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Funding for NTR is acknowledged from NWO-MW 904-61-193; NWO 985-10-002; NWO 904-61-090; Royal Netherlands Academy of Science Professor Award (PAH/6635) to DIB; European Research Council (ERC-230374). Funding for the infrastructure of the NESDA study (www.nesda.nl) was obtained from the Netherlands Organization for Scientific Research (Geestkracht program grant 10-000-1002); the Center for Medical Systems Biology (CSMB, NWO Genomics), VU University Medical Center, GGZ inGeest, Leiden University Medical Center, Leiden University, GGZ Rivierduinen, University Medical Center Groningen, University of Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Rob Giel Onderzoekscentrum. MJA, ASFK, and AMMc are supported by the Wellcome Trust (104036/Z/14/Z, 220857/Z/20/Z). SEM is supported by NHMRC APP1172917, APP1138514 and MRF1200644. ADG, and MGN and EMTD were supported by National Institute of Mental

Health grant R01MH120219. KL and KK were supported by the Estonian Research Council grant PSG615. UKB analysis conducted under project 4844. This work made use of the NL Genetic Cluster Computer (<http://www.geneticcluster.org>) hosted by SURFsara and resources provided by the Edinburgh Compute and Data Facility (ECDF) (<http://www.ecdf.ed.ac.uk/>). The Psychiatric Genomics Consortium (PGC) has received major funding from the National Institute of Mental Health and the National Institute on Drug Abuse (U01 MH109528, U01 MH109532, U01 MH094421, U01 MH085520). This paper represents independent research part-funded by the NIHR Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London.

For the purposes of open access, the author has applied a Creative Commons Attribution 4.0 International Public License (CC BY 4.0) to any Accepted Author Manuscript version arising from this submission.

Declarations

CL sits on the SAB for Myriad Neuroscience, has received consultancy fees from UCB, and speaker fees from SYNLAB. HJG has received travel grants and speaker's honoraria from Fresenius Medical Care, Neuraxpharm, Servier and Janssen Cilag as well as research funding from Fresenius Medical Care.

References

- Als, T. D., Kurki, M., Grove, J., Voloudakis, G., Therrien, K., Tasanko, E., Nielsen, T. T., Naamanka, J., Veerapen, K., Levey, D., Bendl, J., Bybjerg-Grauholm, J., Zheng, B., Demontis, D., Rosengren, A., Athanasiadis, G., Bækved-Hansen, M., Qvist, P., Walters, B., ... Børghlum, A. (2022). *Identification of 64 new risk loci for major depression, refinement of the genetic architecture and risk prediction of recurrence and comorbidities* [Preprint]. *Genetic and Genomic Medicine*. <https://doi.org/10.1101/2022.08.24.22279149>
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (4th ed., text revision). American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Association.
- Badini, I., Coleman, J. R. I., Hagenaaars, S. P., Hotopf, M., Breen, G., Lewis, C. M., & Fabbri, C. (2022). Depression with atypical neurovegetative symptoms shares genetic predisposition with

- immuno-metabolic traits and alcohol consumption. *Psychological Medicine*, 52(4), 726–736. <https://doi.org/10.1017/S0033291720002342>
- Beijers, L., Wardenaar, K. J., van Loo, H. M., & Schoevers, R. A. (2019). Data-driven biological subtypes of depression: Systematic review of biological approaches to depression subtyping. *Molecular Psychiatry*, 24(6), 888–900. <https://doi.org/10.1038/s41380-019-0385-5>
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4), 1165–1188. <https://doi.org/10.1214/aos/1013699998>
- Borsboom, D., & Cramer, A. O. J. (2013). Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annual Review of Clinical Psychology*, 9(1), 91–121. <https://doi.org/10.1146/annurev-clinpsy-050212-185608>
- Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., & Davey Smith, G. (2013). Cohort Profile: The ‘Children of the 90s’—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 42(1), 111–127. <https://doi.org/10.1093/ije/dys064>
- Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., Patterson, N., Daly, M. J., Price, A. L., & Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47(3), 291–295. <https://doi.org/10.1038/ng.3211>
- Byrne, E. M., Kirk, K. M., Medland, S. E., McGrath, J. J., Colodro-Conde, L., Parker, R., Cross, S., Sullivan, L., Statham, D. J., Levinson, D. F., Licinio, J., Wray, N. R., Hickie, I. B., & Martin, N. G. (2020). Cohort profile: The Australian genetics of depression study. *BMJ Open*, 10(5), e032580. <https://doi.org/10.1136/bmjopen-2019-032580>

- Cai, N., Choi, K. W., & Fried, E. I. (2020). Reviewing the genetics of heterogeneity in depression: Operationalizations, manifestations and etiologies. *Human Molecular Genetics*, 29(R1), R10–R18. <https://doi.org/10.1093/hmg/ddaa115>
- Davis, K. A. S., Coleman, J. R. I., Adams, M., Allen, N., Breen, G., Cullen, B., Dickens, C., Fox, E., Graham, N., Holliday, J., Howard, L. M., John, A., Lee, W., McCabe, R., McIntosh, A., Pearsall, R., Smith, D. J., Sudlow, C., Ward, J., ... Hotopf, M. (2020). Mental health in UK Biobank – development, implementation and results from an online questionnaire completed by 157 366 participants: A reanalysis. *BJPsych Open*, 6(2), e18. <https://doi.org/10.1192/bjo.2019.100>
- Elhai, J. D., Contractor, A. A., Tamburrino, M., Fine, T. H., Prescott, M. R., Shirley, E., Chan, P. K., Slembariski, R., Liberzon, I., Galea, S., & Calabrese, J. R. (2012). The factor structure of major depression symptoms: A test of four competing models using the Patient Health Questionnaire-9. *Psychiatry Research*, 199(3), 169–173. <https://doi.org/10.1016/j.psychres.2012.05.018>
- Elsworth, B., Lyon, M., Alexander, T., Liu, Y., Matthews, P., Hallett, J., Bates, P., Palmer, T., Haberland, V., Smith, G. D., Zheng, J., Haycock, P., Gaunt, T. R., & Hemani, G. (2020). *The MRC IEU OpenGWAS data infrastructure* [Preprint]. *Genetics*. <https://doi.org/10.1101/2020.08.10.244293>
- Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Pearson.
- Fernandez-Pujals, A. M., Adams, M. J., Thomson, P., McKechnie, A. G., Blackwood, D. H. R., Smith, B. H., Dominiczak, A. F., Morris, A. D., Matthews, K., Campbell, A., Linksted, P., Haley, C. S., Deary, I. J., Porteous, D. J., MacIntyre, D. J., & McIntosh, A. M. (2015). Epidemiology and Heritability of Major Depressive Disorder, Stratified by Age of Onset, Sex, and Illness Course in Generation Scotland: Scottish Family Health Study (GS:SFHS). *PLoS ONE*, 10(11), e0142197. <https://doi.org/10.1371/journal.pone.0142197>
- Flint, J. (2023). The genetic basis of major depressive disorder. *Molecular Psychiatry*. <https://doi.org/10.1038/s41380-023-01957-9>

- Fried, E. I., & Nesse, R. M. (2015a). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, *172*, 96–102. <https://doi.org/10.1016/j.jad.2014.10.010>
- Fried, E. I., & Nesse, R. M. (2015b). Depression sum-scores don't add up: Why analyzing specific depression symptoms is essential. *BMC Medicine*, *13*(1), 72. <https://doi.org/10.1186/s12916-015-0325-4>
- Grotzinger, A. D., Fuente, J. de la, Privé, F., Nivard, M. G., & Tucker-Drob, E. M. (2022). Pervasive Downward Bias in Estimates of Liability-Scale Heritability in Genome-wide Association Study Meta-analysis: A Simple Solution. *Biological Psychiatry*, S0006322322013166. <https://doi.org/10.1016/j.biopsych.2022.05.029>
- Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., Ip, H. F., Marioni, R. E., McIntosh, A. M., Deary, I. J., Koellinger, P. D., Harden, K. P., Nivard, M. G., & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513–525. <https://doi.org/10.1038/s41562-019-0566-x>
- Harald, B., & Gordon, P. (2012). Meta-review of depressive subtyping models. *Journal of Affective Disorders*, *139*(2), 126–140. <https://doi.org/10.1016/j.jad.2011.07.015>
- Hickie, I. B., Davenport, T. A., Hadzi-Paviovic, D., Koschera, A., Naismith, S. L., Scott, E. M., & Wilhelm, K. A. (2001). Development of a simple screening tool for common mental disorders in general practice. *Medical Journal of Australia*, *175*(S1). <https://doi.org/10.5694/j.1326-5377.2001.tb143784.x>
- Hoffmann, T. J., Choquet, H., Yin, J., Banda, Y., Kvale, M. N., Glymour, M., Schaefer, C., Risch, N., & Jorgenson, E. (2018). A Large Multiethnic Genome-Wide Association Study of Adult Body Mass Index Identifies Novel Loci. *Genetics*, *210*(2), 499–515. <https://doi.org/10.1534/genetics.118.301479>
- Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shiralil, M., Coleman, J. R. I., Hagenars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y.,

- Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., ... Major Depressive Disorder Working Group of the Psychiatric Genomics, C. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*, 343–352. <https://doi.org/10.1038/s41593-018-0326-7>
- Howe, L. J., Nivard, M. G., Morris, T. T., Hansen, A. F., Rasheed, H., Cho, Y., Chittoor, G., Ahlskog, R., Lind, P. A., Palviainen, T., van der Zee, M. D., Cheesman, R., Mangino, M., Wang, Y., Li, S., Klaric, L., Ratliff, S. M., Bielak, L. F., Nygaard, M., ... Davies, N. M. (2022). Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics*, *54*(5), 581–592. <https://doi.org/10.1038/s41588-022-01062-7>
- Kendler, K. S., Aggen, S. H., & Neale, M. C. (2013). Evidence for Multiple Genetic Factors Underlying DSM-IV Criteria for Major Depression. *JAMA Psychiatry*, *70*(6), 599. <https://doi.org/10.1001/jamapsychiatry.2013.751>
- Krause, J. S., Bombardier, C., & Carter, R. E. (2008). Assessment of depressive symptoms during inpatient rehabilitation for spinal cord injury: Is there an underlying somatic factor when using the PHQ? *Rehabilitation Psychology*, *53*(4), 513–520. <https://doi.org/10.1037/a0013354>
- Krause, J. S., Reed, K. S., & McArdle, J. J. (2010). Factor Structure and Predictive Validity of Somatic and Nonsomatic Symptoms From the Patient Health Questionnaire-9: A Longitudinal Study After Spinal Cord Injury. *Archives of Physical Medicine and Rehabilitation*, *91*(8), 1218–1224. <https://doi.org/10.1016/j.apmr.2010.04.015>
- Lam, M., Awasthi, S., Watson, H. J., Goldstein, J., Panagiotaropoulou, G., Trubetskoy, V., Karlsson, R., Frei, O., Fan, C.-C., De Witte, W., Mota, N. R., Mullins, N., Brügger, K., Lee, S. H., Wray, N. R., Skarabis, N., Huang, H., Neale, B., Daly, M. J., ... Ripke, S. (2020). RICOPILI: Rapid Imputation for COnsortias PIpeLIne. *Bioinformatics (Oxford, England)*, *36*(3), 930–933. <https://doi.org/10.1093/bioinformatics/btz633>

- Lamers, F., de Jonge, P., Nolen, W. A., Smit, J. H., Zitman, F. G., Beekman, A. T. F., & Penninx, B. W. J. H. (2010). Identifying Depressive Subtypes in a Large Cohort Study: Results From the Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry*, *71*(12), 1582–1589. <https://doi.org/10.4088/JCP.09m05398blu>
- Lee, S. H., Wray, N. R., Goddard, M. E., & Visscher, P. M. (2011). Estimating Missing Heritability for Disease from Genome-wide Association Studies. *The American Journal of Human Genetics*, *88*(3), 294–305. <https://doi.org/10.1016/j.ajhg.2011.02.002>
- Leitsalu, L., Haller, T., Esko, T., Tammesoo, M.-L., Alavere, H., Snieder, H., Perola, M., Ng, P. C., Mägi, R., Milani, L., Fischer, K., & Metspalu, A. (2015). Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *International Journal of Epidemiology*, *44*(4), 1137–1147. <https://doi.org/10.1093/ije/dyt268>
- Levey, D. F., Stein, M. B., Wendt, F. R., Pathak, G. A., Zhou, H., Aslan, M., Quaden, R., Harrington, K. M., Nuñez, Y. Z., Overstreet, C., Radhakrishnan, K., Sanacora, G., McIntosh, A. M., Shi, J., Shringarpure, S. S., 23andMe Research Team, the Million Veteran Program, Concato, J., Polimanti, R., & Gelernter, J. (2021). Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nature Neuroscience*, *24*(7), 954–963. <https://doi.org/10.1038/s41593-021-00860-2>
- Lewis, G., Pelosi, A. J., Araya, R., & Dunn, G. (1992). Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine*, *22*(2), 465–486. <https://doi.org/10.1017/S0033291700030415>
- Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, *18*, 497–511. <https://doi.org/10.1038/mp.2012.21>
- Milaneschi, Y., Lamers, F., Berk, M., & Penninx, B. W. J. H. (2020). Depression Heterogeneity and Its Biological Underpinnings: Toward Immunometabolic Depression. *Biological Psychiatry*, *88*(5), 369–380. <https://doi.org/10.1016/j.biopsych.2020.01.014>

- Milaneschi, Y., Lamers, F., Mbarek, H., Hottenga, J.-J., Boomsma, D. I., & Penninx, B. W. J. H. (2014). The effect of FTO rs9939609 on major depression differs across MDD subtypes. *Molecular Psychiatry*, *19*(9), 960–962. <https://doi.org/10.1038/mp.2014.4>
- Milaneschi, Y., Lamers, F., Peyrot, W. J., Baune, B. T., Breen, G., Dehghan, A., Forstner, A. J., Grabe, H. J., Homuth, G., Kan, C., Lewis, C., Mullins, N., Nauck, M., Pistis, G., Preisig, M., Rivera, M., Rietschel, M., Streit, F., Strohmaier, J., ... the Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. (2017). Genetic association of major depression with atypical features and obesity-related immunometabolic dysregulations. *JAMA Psychiatry*, *74*, 1214–1225. <https://doi.org/10.1001/jamapsychiatry.2017.3016>
- Mitchell, B. L., Campos, A. I., Whiteman, D. C., Olsen, C. M., Gordon, S. D., Walker, A. J., Dean, O. M., Berk, M., Hickie, I. B., Medland, S. E., Wray, N. R., Martin, N. G., & Byrne, E. M. (2022). The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes. *Biological Psychiatry*, *92*(3), 227–235. <https://doi.org/10.1016/j.biopsych.2021.10.021>
- Nguyen, T.-D., Harder, A., Xiong, Y., Kowalec, K., Hägg, S., Cai, N., Kuja-Halkola, R., Dalman, C., Sullivan, P. F., & Lu, Y. (2022). Genetic heterogeneity and subtypes of major depression. *Molecular Psychiatry*, *27*(3), 1667–1675. <https://doi.org/10.1038/s41380-021-01413-6>
- Penninx, B. W. J. H., Milaneschi, Y., Lamers, F., & Vogelzangs, N. (2013). Understanding the somatic consequences of depression: Biological mechanisms and the role of depression symptom profile. *BMC Medicine*, *11*, 129. <https://doi.org/10.1186/1741-7015-11-129>
- Smith, B. H., Campbell, A., Linksted, P., Fitzpatrick, B., Jackson, C., Kerr, S. M., Deary, I. J., MacIntyre, D. J., Campbell, H., & McGilchrist, M. (2012). Cohort profile: Generation Scotland: Scottish Family Health Study (GS: SFHS). The study, its participants and their potential for genetic research on health and illness. *Int J Epidemiol*, *42*(2), 689–700.
- Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., Ibrahim, A., Ji, Y., John, S., Lewis, E., MacArthur, J. A. L., McMahon, A., Osumi-Sutherland, D., Panoutsopoulou, K., Pendlington, Z., ... Harris, L. W. (2023). The NHGRI-

- EBI GWAS Catalog: Knowledgebase and deposition resource. *Nucleic Acids Research*, 51(D1), D977–D985. <https://doi.org/10.1093/nar/gkac1010>
- Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., & Collins, R. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med*, 12(3), e1001779. <https://doi.org/10.1371/journal.pmed.1001779>
- Thorp, J. G., Marees, A. T., Ong, J.-S., An, J., MacGregor, S., & Derks, E. M. (2020). Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9. *Psychological Medicine*, 50(14), 2385–2396. <https://doi.org/10.1017/S0033291719002526>
- van Loo, H. M., Aggen, S. H., & Kendler, K. S. (2022). The structure of the symptoms of major depression: Factor analysis of a lifetime worst episode of depressive symptoms in a large general population sample. *Journal of Affective Disorders*, 307, 115–124. <https://doi.org/10.1016/j.jad.2022.03.064>
- World Health Organization (Ed.). (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. World Health Organization.
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S.-A., Bækvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschøn, H. N., Bybjerg-Grauholm, J., ... Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), 668–681. <https://doi.org/10.1038/s41588-018-0090-3>
- Yengo, L., Sidorenko, J., Kemper, K. E., Zheng, Z., Wood, A. R., Weedon, M. N., Frayling, T. M., Hirschhorn, J., Yang, J., Visscher, P. M., & the GIANT Consortium. (2018). Meta-analysis of genome-wide association studies for height and body mass index in ~700000

individuals of European ancestry. *Human Molecular Genetics*, 27(20), 3641–3649.

<https://doi.org/10.1093/hmg/ddy271>

Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many

different ways do patients meet the diagnostic criteria for major depressive disorder?

Comprehensive Psychiatry, 56, 29–34.

<https://doi.org/10.1016/j.comppsy.2014.09.007>

Supplementary text: Genetic structure of major depression symptoms in clinical and population cohorts

Table of Contents

Genotyping and QC	1
Ethics statements	3
Confirmatory factor analysis model schematics	3
Supplementary Figure S1. Schematic drawings of the CFA models	3
Symptom genetic correlations	6
Supplementary Figure S2. Genetic correlations between symptoms	6
Supplementary Figure S3. Model implied and residual proportions of genetic correlations	6
External phenotype summary statistics	8
Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium	9
References	25

Genotyping and QC

Psychiatric Genomics Consortium (PGC). The analysis used data from 24 cohorts from the PGC MDD datasets that had symptom data on cases. Data was drawn from the following cohorts:

- BiDirect (bidi1)
- BOMA (boma)
- CoFams (cof3)
- PsyCoLaus (col3)
- GenRED (gens, grnd)
- GenPod/Newmeds (gep3)
- GSK (gsk2)
- Janssen (janpy)
- MPIP/MARS (mmi2, mmo4)
- NESDA/NTR (nes1)
- QIMR (qi3c, qi6c, qio2)
- RADIANT (rad3, rage, rai2, rau2, rde4)
- Rotterdam (rot4)
- SHIP (shp0)
- STAR*D (stm2)
- TwinGene (twg2)

The genotypes were processed through Ricopili (Lam et al., 2020) with the following QC: SNP missingness < 0.05; sample missingness < 0.02; autosomal heterozygosity deviation ($|F_{het}| < 0.2$); and SNP Hardy-Weinberg equilibrium ($P > 10^{-6}$ in controls, $P > 10^{-10}$ in cases). QC'd genotypes were then imputed to the 1000 Genomes Reference Panel (The 1000 Genomes Project Consortium, 2015). Information on cohort genotyping and additional processing steps is available in (Wray et al., 2018).

Australian Genetics of Depression Study (AGDS). Genotyping was conducted using the Illumina Infinium Global Screening Array platform and QC'd for unknown or ambiguous map position and strand alignment, missingness >5%, HWE < 1×10^{-6} , MAF<1%. Genotypes were imputed to HRCr1.1. Individuals were excluded with missing rate > 3%, inconsistent sex, or if deemed ancestry outliers from the European population (6 standard deviations from the first two genetic principal components from 1000 Genomes). Imputed genotype dosages were used for the analyses. GWAS was carried out in SAIGE (Zhou et al., 2018) using a generalized linear mixed model with genotyping batch and 10 PCs as covariates. Variants with MAF<1% and imputation accuracy score <0.7 were excluded.

Avon Longitudinal Study of Parents and Children (ALSPAC). ALSPAC children were genotyped using the Illumina HumanHap550 quad chip genotyping platforms. Individuals were excluded on the basis of gender mismatches; minimal or excessive heterozygosity; disproportionate levels of individual missingness (>3%) and insufficient sample replication (IBD < 0.8). Population stratification was assessed by multidimensional scaling analysis, removing samples that clustered outside the CEU HapMap2 population. SNPs with a minor allele frequency of < 1%, a call rate of < 95% or evidence for violations of Hardy-Weinberg equilibrium ($P < 5E-7$) were removed. Cryptic relatedness was measured as proportion of identity by descent (IBD > 0.1). Related subjects that passed all other quality control thresholds were retained during subsequent phasing and imputation. 9,115 subjects and 500,527 SNPs passed these quality control filters. Imputation of the target data was performed using Impute V2.2.2 against 1000 genomes reference panel (Phase 1, Version 3) (all polymorphic SNPs excluding singletons), using all 2186 reference haplotypes (including non-Europeans). This resulted in 28,699,419 SNPs, with 8,282,911 SNPs with a MAF >0.01 and info score of >0.8. Analysis were conducted using SNPTEST v2.5.2, adjusting for sex and the first 10 principal components of ancestry.

Generation Scotland (GS:SFHS). GWAS data was obtained using the Illumina OmniExpress array, and imputed using the Haplotype Research Consortium (HRC) dataset. Further details of methods here <https://pubmed.ncbi.nlm.nih.gov/28270201/>. GWAS was conducted in regenie with 4 PCs removing SNPs with MAC < 100, genotype missingness > 10%, INFO < 0.1, and HWE $p > 1e-15$.

Estonian Biobank (EstBB). The samples from the Estonian Biobank have been genotyped at the Genotyping Core Facility of the Institute of Genomics, University of Tartu using the Global Screening Array (GSAv1.0, GSAv2.0, and GSAv2.0_EST) from Illumina. Altogether 155,772 samples have been genotyped and PLINK format files exported using GenomeStudio v2.0.4. Individuals were excluded from the analysis if their call-rate was <95% or if the sex defined based on heterozygosity of the X chromosome did not match the sex in the phenotype data. Variants were excluded if the call-rate was < 95% and HWE p -value < $1e-4$ (autosomal variants only). Variant positions were updated to genome build 37 and all alleles were switched to the TOP strand using tools and reference files provided at <https://www.well.ox.ac.uk/~wrayner/strand/>. After QC the dataset contained 154,201 samples for imputation. Before imputation variants with MAF<1% and indels were removed. Prephasing was done using the Eagle v2.3 software. The number of conditioning haplotypes Eagle2 uses when phasing each sample was set to: --Kpbwt=20000. Imputation was done using Beagle v.28Sep18.793 with effective population size $n_e=20,000$. An Estonian population specific imputation reference of 2,297 WGS samples was used. The analysis was performed using the SAIGE software, including related individuals and adjusting for the first 10 principal components (PCs) of the genotype matrix, as well as for birth year, birth year squared and sex.

UK Biobank (UKB). Imputed genotypes were analysed from the version 3 release (Bycroft et al., 2018). Imputed genotypes were QC'd to INFO ≥ 0.1 , MAC ≥ 100 , HWE $P > 1e-10$, max alleles = 2, and duplicate markers removed. Association analysis was performed as a logistic regression in Plink2 (Chang et al., 2015) with genotyping array and 20 PCs as covariates.

Ethics statements

Ethical approval was obtained from the ALSPAC Ethics and Law Committee and the local research ethics committees (project number B3118). Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

All participants in AGDS provided informed consent that they had read and understood the study information sheets and to confirm that they would be willing to provide a saliva sample for genotyping and downstream generic analyses. All study protocols were approved by the QIMR Berghofer Medical Research Institute Human Research Ethics Committee - approval numbers P2118, P1309 and P2304.

The activities of the EstBB are regulated by the Human Genes Research Act, which was adopted in 2000 specifically for the operations of the EstBB. Individual-level data analysis in the EstBB was carried out under ethical approvals [1.1-12/2860 & 1.1-12/624] from the Estonian Committee on Bioethics and Human Research (Estonian Ministry of Social Affairs), using data according to the release application [3-10/GI-28207] from the Estonian Biobank.

Ethical approval for the GS:SFHS data collection was obtained from the Tayside Committee on Medical Research Ethics A (ref 05/S1401/89). Generation Scotland is currently approved as a Research Tissue Bank by the East of Scotland Research Ethics Service (ref 20/ES/0021).

UK Biobank received ethical approval from the Research Ethics Committee (reference 11/NW/0382).

Confirmatory factor analysis model schematics

Supplementary Figure S1. Schematic drawings of the CFA models

Schematics to illustrate factor structures of the models that were tested. See main text Table 1 for symptom abbreviations and Supplementary Tables S4,6 for factor structures and coefficients.

Figure S1a: Model A: Common factor

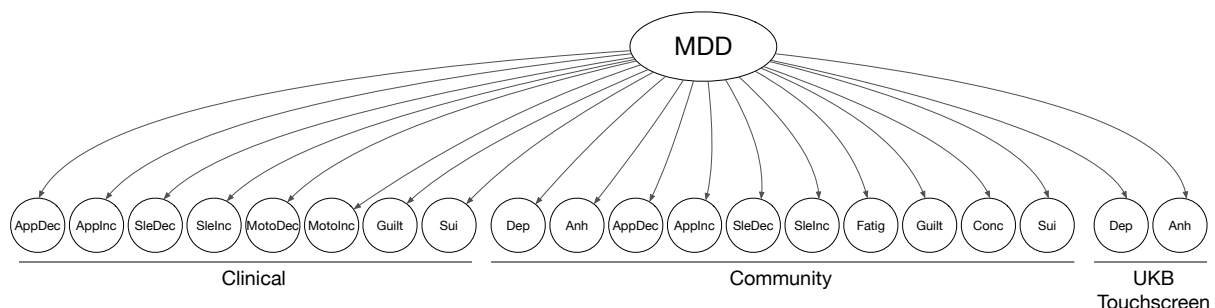


Figure S1b: Model B: Clinical and community factors

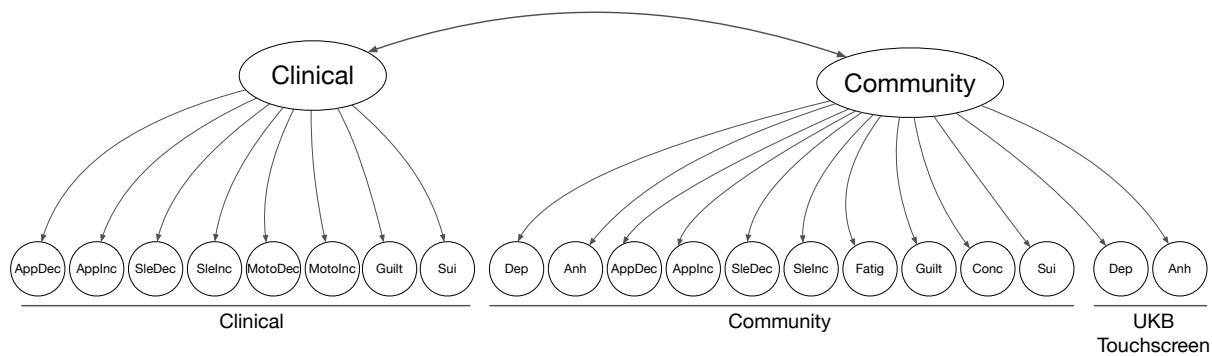


Figure S1c: Model C: Gating measurement factor

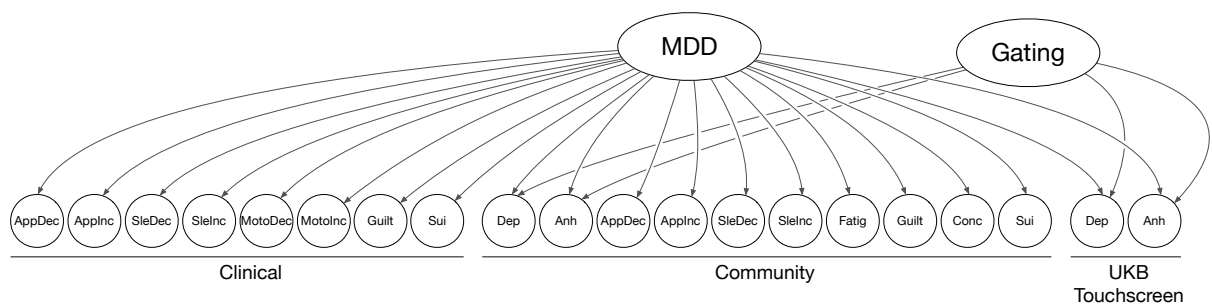


Figure S1d: Model D: Clinical-Community-Gating factors

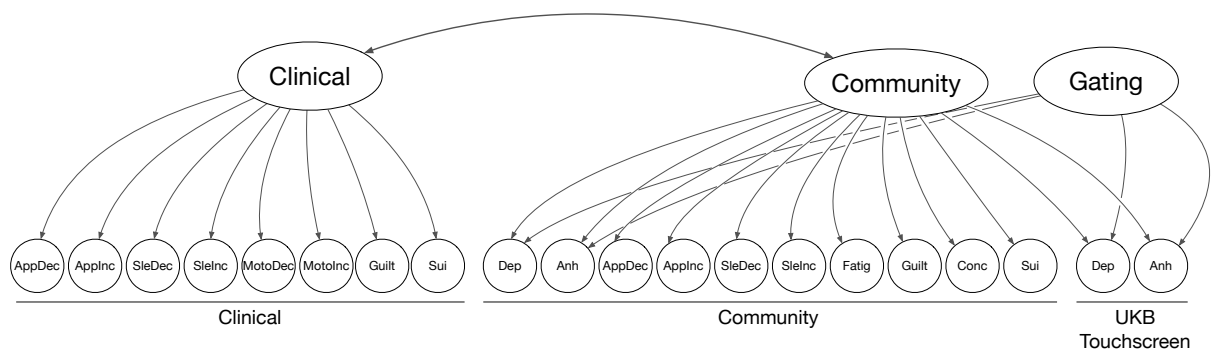


Figure S1e: Model E: Psychological-Somatic

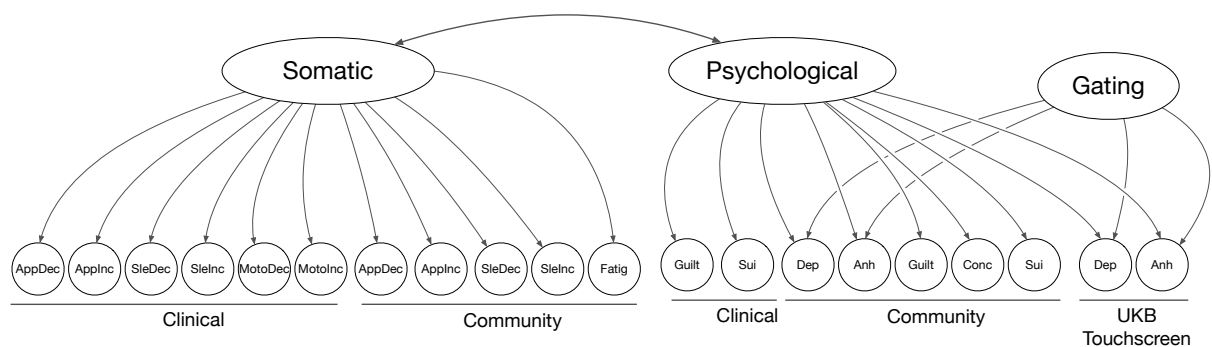


Figure S1f: Model F: Psychological- Neurovegetative

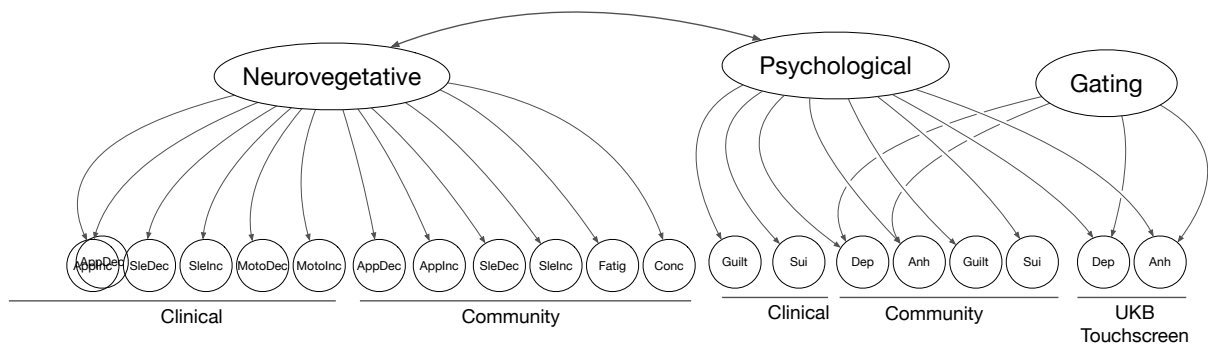


Figure S1g: Model G: Affective-Neurovegetative

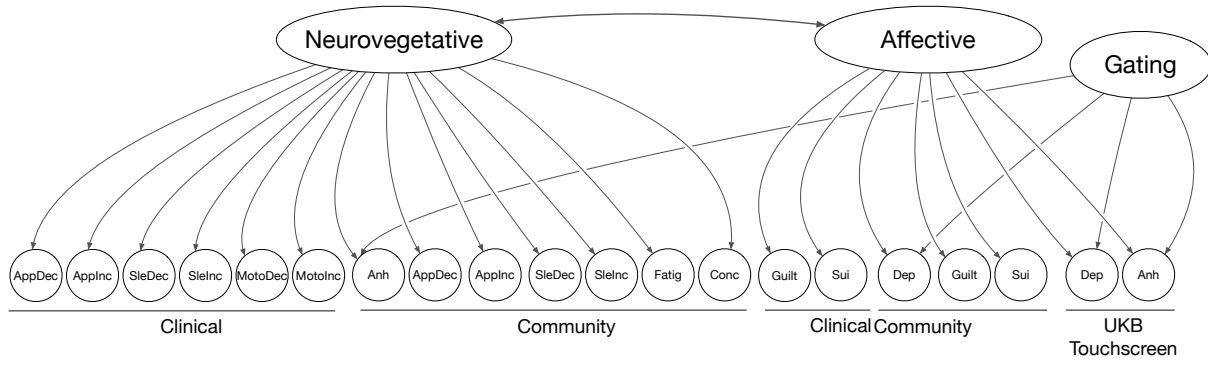


Figure S1h: Model H: Cognitive-Mood-Neurovegetative

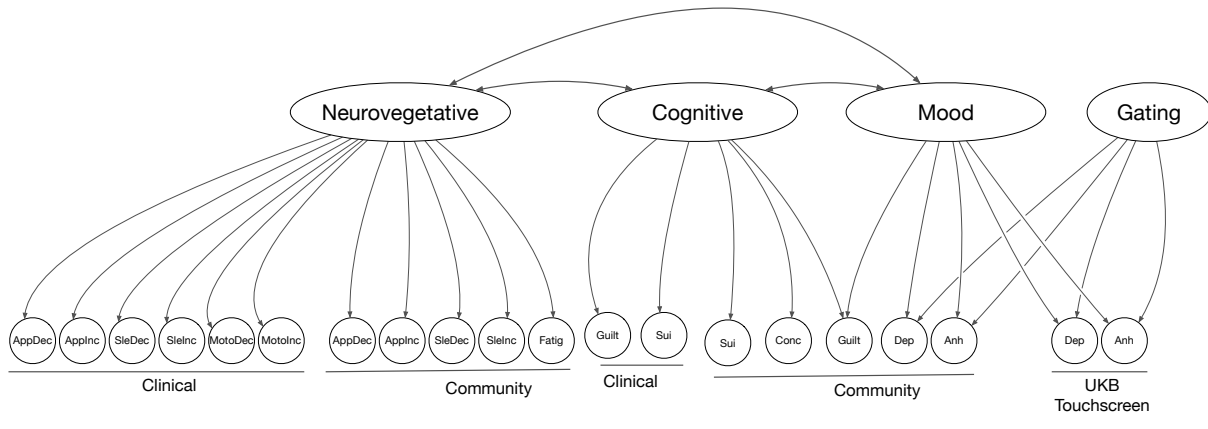


Figure S1i: Model I: Appetite-Vegetative-Cognitive/Mood

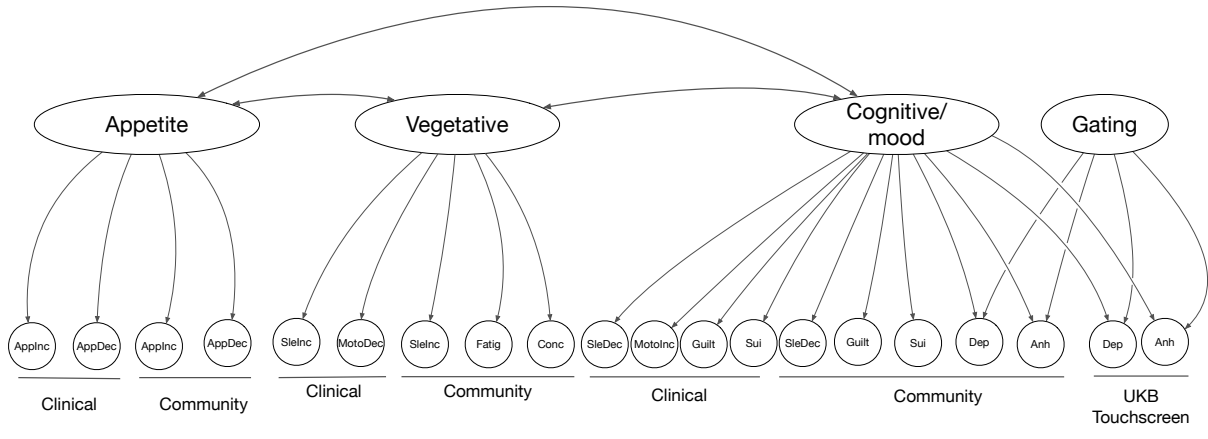
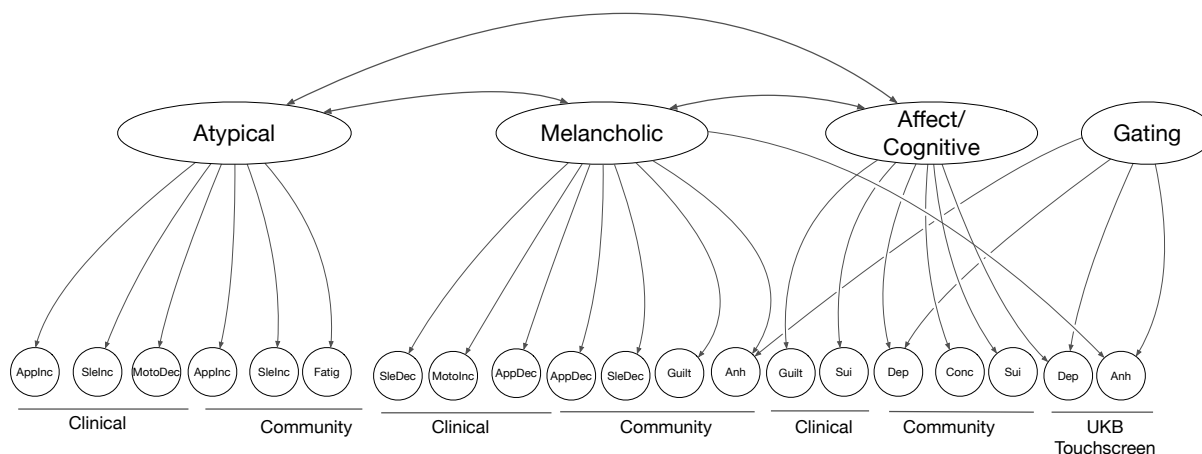
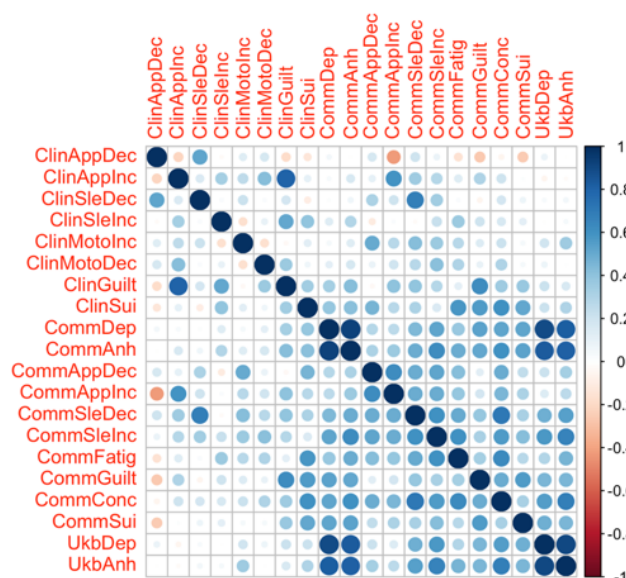


Figure S1j: Model J: Atypical-Melancholic-Affect/Cognitive



Symptom genetic correlations

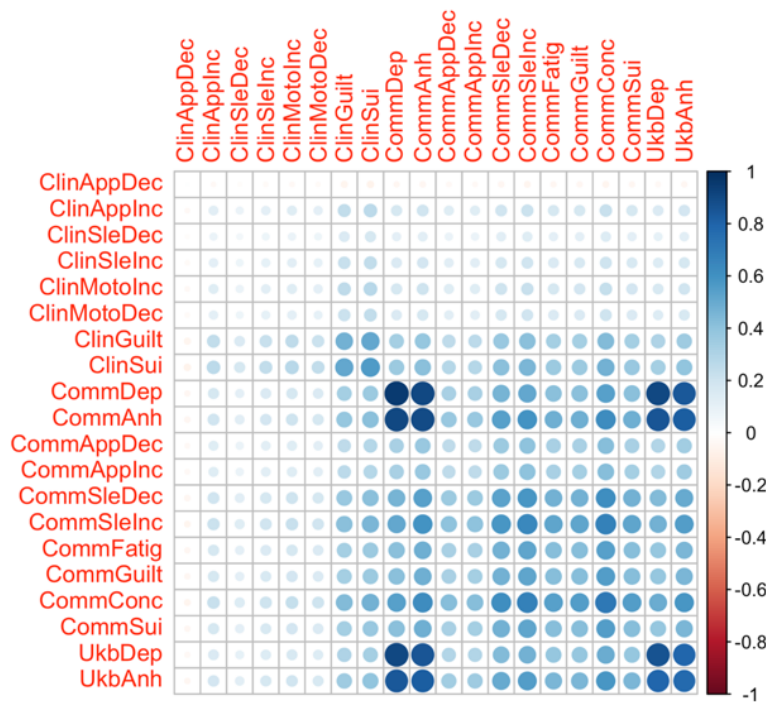
Supplementary Figure S2. Genetic correlations between symptoms



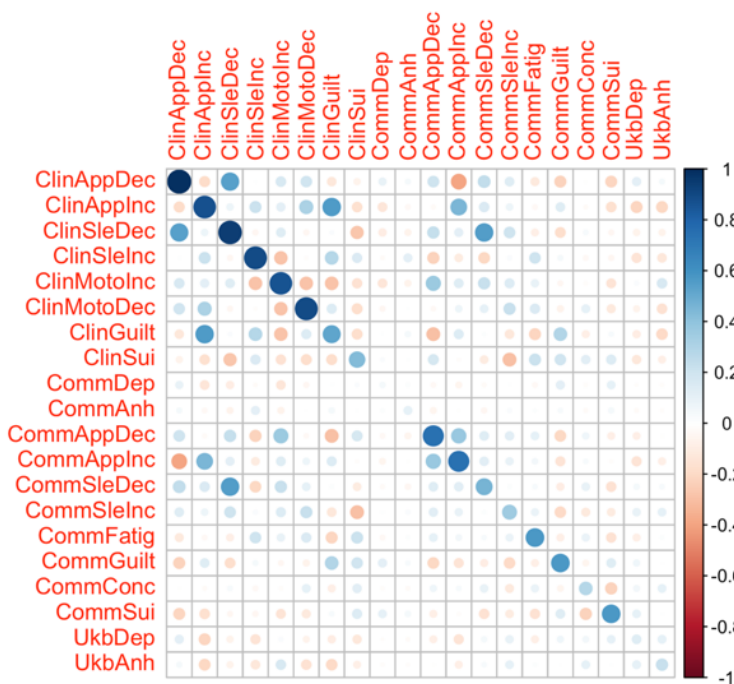
Supplementary Figure S3. Model implied and residual proportions of genetic correlations

Variance and covariances scaled by total genetic variance of each symptom.

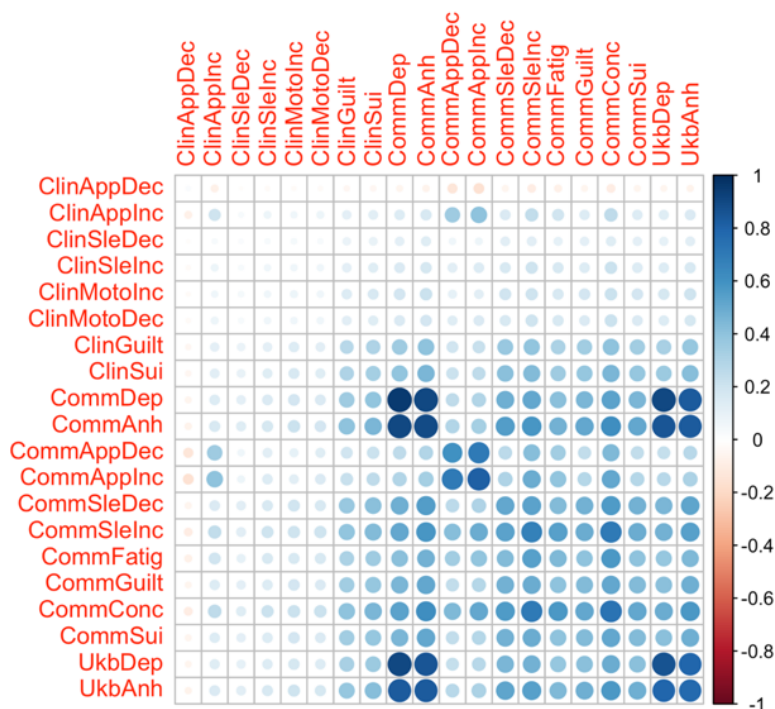
Supplementary Figure S3a. Model implied proportions of genetic correlations for Clinical-Community-Gating factors (Model D)



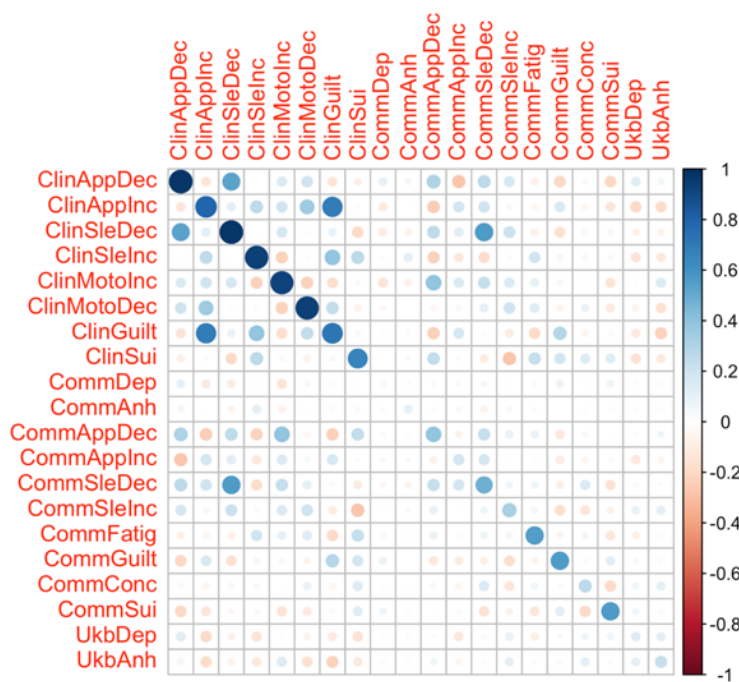
Supplementary Figure S3b. Model residual proportions of genetic correlations for Clinical-Community-Gating factors (Model D)



Supplementary Figure S3c. Model implied proportions of genetic correlations for Appetite-Vegetative-Cognitive/Mood factors (Model I)



Supplementary Figure S3d. Model residual proportions of genetic correlations for Appetite-Vegetative-Cognitive/Mood factors (Model I)



External phenotype summary statistics

For the genetic multiple regression analysis, we used the following summary statistics:

- Alcohol dependence (Walters et al., 2018)
- Anxiety (Grotzinger et al., 2022)
- Bipolar disorder (Mullins et al., 2021)
- Body mass index (Pulit et al., 2019)

- Educational attainment (Okbay et al., 2022)
- Major depression (Als et al., 2022)
- Major depressive disorder (Wray et al., 2018)
- Neuroticism (Nagel et al., 2018)
- Pain (multisite chronic pain) (Johnston et al., 2019)
- Post-traumatic stress disorder (Nievergelt et al., 2019)
- Sleep (long sleep duration) (Dashti et al., 2018)
- Smoking (cigarettes per day) (Liu et al., 2019)

Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium

Mark J Adams 1	Jian Zeng 16	Julia Boberg 36
Fabian Streit 2	Johan Zvrskovec 4, 8	Sigrid Børte 51, 52, 53
Swapnil Awasthi 3	Dag Aarsland 26	Emma Bränn 54
Brett N Adey 4	Ky'Era V Actkins 27	Alice Braun 55
Karmel W Choi 5, 6	Mazda Adli 3, 28	Thorsten Brodersen 56
V Kartik Chundru 7	Esben Agerbo 12, 29, 30	Søren Brunak 44
Jonathan RI Coleman 4, 8	Mareike Aichholzer 31	Mie T Bruun 57
Jerome C Foo 2	Tracy M Air 32	Pichit Buspavanich 58, 59
Olga Giannakopoulou 9	Allison Aiello 33	Jonas Bybjerg-Grauholm 60, 61
Alisha S M Hall 2, 10	Thomas D Als 30, 34, 35	Enda M Byrne 62
Jens Hjerling-Leffler 11	Evelyn Andersson 36	Archie Campbell 63, 64
David M Howard 4	Till F M Andlauer 37, 38	Megan L. Campbell 65
Christopher Hübel 4, 12, 13	Volker Arolt 39	Enrique Castela 66
Alex S F Kwong 1, 14	Helga Ask 40, 41	Jorge Cervilla 67, 68
Bochao Danae Lin 15	Sunita Badola 42	Boris Chaumette 69
Xiangrui Meng 9	Clive Ballard 43	Chia-Yen Chen 70
Guiyan Ni 16	Karina Banasik 44	Zhengming Chen 71, 72
Oliver Pain 17	Nicholas J Bass 9	Sven Cichon 73, 74, 75, 76
Gita A Pathak 18, 19	Aartjan T F Beekman 45	Lucía Colodro-Conde 24
Eva C Schulte 20, 21, 22, 23	Sintia Belangero 46	Anne Corbett 43
Jackson G Thorp 24	Elisabeth B Binder 38, 47	Elizabeth C Corfield 40, 77
Alicia Walker 16	Ottar Bjerkeset 48, 49	Baptiste Couvy-Duchesne 78
Shuyang Yao 25	Gyda Bjornsdottir 50	

Nick Craddock 79, 80	Robert C Free 108	Marcus Ining 123
Udo Dannlowski 39	He Gao 109	Erkki Isometsä 124
Gail Davies 81	Michael Gill 110	Rick Jansen 45
EJC de Geus 82	Maria Gilles 87	Yunxuan Jiang 125
Ian J Deary 81	Fernando S Goes 86	Ian Jones 80
Franziska Degenhardt 76, 83	Scott Douglas Gordon 24	Lisa A Jones 126
Abbas Dehghan 84, 85	Jakob Grove 30, 34, 35, 111	Lina Jonsson 127
J Raymond DePaulo 86	Daniel F Gudbjartsson 50, 112	Robert Karlsson 25
Michael Deuschle 87	Blanca Gutierrez 67, 68	Siegfried Kasper 128
Maria Didriksen 88	Tim Hahn 39	Kenneth S Kendler 129
Khoa Manh Dinh 89	Lynsey S Hall 80	Ronald C Kessler 130
Nese Direk 90	Thomas F Hansen 44, 60, 113	Stefan Kloiber 101, 123, 131, 132
Srdjan Djurovic 91, 92	Magnus Haraldsson 114	James A Knowles 133
Anna R Docherty 93, 94, 95	Catherina A Hartman 115	Nastassja Koen 65
Katharina Domschke 96	Alexandra Havdahl 40	Julia Kraft 55
Joseph Dowsett 88	Caroline Hayward 116	Henry R Kranzler 134, 135
Ole Kristian Drange 49, 97, 98, 99	Stefanie Heilmann-Heimbach 76	Kristi Krebs 136
Erin C Dunn 6, 100	Stefan Herms 74, 76	Theodora Kunovac Kallak 137
Gudmundur Einarsson 50	Ian B Hickie 117	Zoltán Kutalik 138, 139, 140
Thalia C Eley 4	Henrik Hjalgrim 118	Elisa Lahtela 141
Samar S M Elsheikh 101	Per Hoffmann 74, 76	Margit Hørup Larsen 88
Jan Engelmann 102	Georg Homuth 119	Eric J Lenze 142
Michael E Benros 60, 103, 104	Carsten Horn 120	Daniel F Levey 143, 144
Christian Erikstrup 89	Jouke-Jan Hottenga 82	Melissa Lewins 1
Valentina Escott-Price 80	David M Hougaard 60, 61	Glyn Lewis 9
Chiara Fabbri 4, 105	Iiris Hovatta 121	Liming Li 145, 146
Yu Fang 106	Qin Qin Huang 7	Kuang Lin 71
Sarah Finer 107	Floris Huider 82	Penelope A Lind 24
Josef Frank 2	Karen A Hunt 122	Donald J MacIntyre 1, 147, 148

Dean F MacKinnon 86	Sisse Rye Ostrowski 88, 166	Daniel J Smith 185
Wolfgang Maier 149	Vanessa K Ota 46	Erik Sørensen 88
Victoria S Marshe 101, 150	Michael J Owen 80	Sundararajan Srinivasan 186
Hamdi Mbarek 82	Richard Packer 167	Hreinn Stefansson 50
Peter McGuffin 4	Teemu Palviainen 141	Kari Stefansson 50, 187
Sarah E Medland 24	Pedro M Pan 168	Dan J. Stein 188
Susanne Meinert 39, 151	Carlos N Pato 169	Frederike Stein 189
Susan Mikkelsen 89	Michele T Pato 169	André Tadic 102, 190
Christina Mikkelsen 88, 152	Nancy L Pedersen 25	Henning Teismann 191
Yuri Milaneschi 45	Ole Birger Pedersen 170	Alexander Teumer 192
Iona Y Millwood 71, 72	Roseann E Peterson 129, 171	Anita Thapar 80, 193
Brittany L Mitchell 24	Wouter J Peyrot 45	Pippa A Thomson 64
Esther Molina 67, 153	James B Potash 86	Lise Wegner Thørner 88
Francis M Mondimore 86	Martin Preisig 66	Apostolia Topaloudi 194
Preben Bo Mortensen 12, 29, 30	Jorge A Quiroz 172	Ioanna Tzoulaki 84, 85, 195
Benoit H Mulsant 101, 131	Charles F Reynolds III 173	Monica Uddin 196
Joonas Naamanka 121	John P Rice 142	André G Uitterlinden 197
Jake M Najman 154	Giovanni A Salum 174	Henrik Ullum 88, 198, 199
Matthias Nauck 155, 156	Robert A Schoevers 175, 176	Daniel Umbricht 200
Igor Nenadi? 157	Andrew Schork 30, 177, 178	Robert J Ursano 201
Kasper R Nielsen 158	Thomas G Schulze 2, 21, 86, 179, 180	Sandra Van der Auwera 202
Ilja M Nolte 159	Tabea S Send 87	David A van Heel 122
Merete Nordentoft 60, 103, 104	Jianxin Shi 181	Albert M van Hemert 203
Markus M Nöthen 76	Engilbert Sigurdsson 114	Abirami Veluchamy 186
Mette Nyegaard 30, 160, 161, 162	Kritika Singh 27	Alexander Viktorin 25
Michael C O'Donovan 80	Grant C B Sinnamon 182	Henry Völzke 192
Asmundur Oddsson 50	Lea Sirignano 2	Agaz Wani 196
Catherine M Olsen 163, 164	Olav B Smeland 183, 184	G Bragi Walters 50
Hogni Oskarsson 165		Robin G Walters 71, 72

Sylvia Wassertheil-Smoller 204	Hilary Coon 94	Lili Milani 136
Myrna M Weissman 205, 206	William E Copeland 218	Ole Mors 30, 233
Jürgen Wellmann 191	Byron Creese 43	Daniel J Müller 101, 131, 132, 234
David C Whiteman 163	Lea K Davis 27	Bertram Müller-Myhsok 38, 235, 236
Derek Wildman 196	Eske M Derks 24	Albertine J Oldehinkel 115
Gonneke Willemsen 82	Enrico Domenici 219	Sara A Paciga 237
Alexander T Williams 167	Paul Elliott 84, 85, 195, 220	Colin NA Palmer 186
Bendik S Winsvold 51, 52, 207	Andreas J Forstner 73, 76	Peristera Paschou 194
Stephanie H Witt 2	Micha Gawlik 221	Brenda WJH Penninx 45
Ying Xiong 25	Joel Gelernter 19, 143, 222	Roy H Perlis 5, 6, 238
Lea Zillich 2	Hans J Grabe 202	Giorgio Pistis 66
John-Anker Zwart 51, 52, 53	Steven P Hamilton 223	Renato Polimanti 18, 19
23andMe Research Team 125	Kristian Hveem 224, 225, 226	David J Porteous 64
Estonian Biobank Research Team 136	Catherine John 167, 227	Danielle Posthuma 239, 240
HUNT All-In Psychiatry 208	Jaakko Kaprio 141	Ted Reichborn-Kjennerud 40
China Kadoorie Biobank Collaborative Group 209	Tilo Kircher 157	Andreas Reif 31
Genes & Health Research Team 210	Marie-Odile Krebs 228	Frances Rice 80, 241
Ole A Andreassen 183, 184, 211	Karoline Kuchenbaecker 9, 71	Roland Ricken 3
Bernhard T Baune 212, 213, 214	Mikael Landén 25, 127	Marcella Rietschel 2
Klaus Berger 191	Kelli Lehto 136	Margarita Rivera 67, 242
Dorret I Boomsma 82	Douglas F Levinson 229	Christian Rück 243
Anders D Børglum 30, 34, 35	Qingqin S Li 230	Catherine Schaefer 244
Gerome Breen 4, 8	Klaus Lieb 102	Srijan Sen 106, 245
Na Cai 215, 216, 217	Yi Lu 25	Alessandro Serretti 105
	Susanne Lucae 123	Alkistis Skalkidou 137
	Jurjen J Luykx 15, 231	Jordan W Smoller 5, 246, 247
	Patrik K Magnusson 25	Frederike Stein 189
	Nicholas G Martin 24	Murray B Stein 248, 249, 250
	Hilary C Martin 7	Patrick F Sullivan 25, 251
	Andrew McQuillin 9	Martin Tesli 40
	Christel M Middeldorp 62, 232	

Thorgeir E Thorgeirsson 50

Henning Tiemeier 252, 253

Nicholas J Timpson 14

Rudolf Uher 254

Jens R Wendland 42

Thomas Werge 60, 177, 199,
255, 256

Naomi R Wray 16, 257

Stephan Ripke 3, 246

Cathryn M Lewis 4, 258

Andrew M McIntosh 1, 25

- 1, Division of Psychiatry, University of Edinburgh, Edinburgh, UK
- 2, Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, BW, DE
- 3, Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Berlin, BE, DE
- 4, Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK
- 5, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, US
- 6, Department of Psychiatry, Harvard Medical School, Boston, MA, US
- 7, Human Genetics, Wellcome Sanger Institute, Hinxton, UK
- 8, NIHR Maudsley Biomedical Research Centre, King's College London, London, UK
- 9, Division of Psychiatry, University College London, London, UK
- 10, Department of Clinical Medicine, Aarhus University, Aarhus, DK
- 11, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, SE
- 12, National Centre for Register-based Research, Aarhus University, Aarhus, DK
- 13, Department of Pediatric Neurology, Charité – Universitätsmedizin Berlin, Berlin, BE, DE
- 14, MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
- 15, Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Centre, Maastricht, NL
- 16, Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD, AU
- 17, Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, King's College London, London, UK
- 18, Veterans Affairs Connecticut Healthcare System, West Haven, CT, US
- 19, Department of Psychiatry, Yale University School of Medicine, New Haven, CT, US
- 20, Department of Psychiatry, University of Munich, Munich, BY, DE
- 21, Institute of Psychiatric Phenomics and Genomics, University of Munich, Munich, BY, DE
- 22, Department of Psychiatry and Psychotherapy, University Hospital Bonn, Medical Faculty, University of Bonn, Bonn, DE
- 23, Institute of Human Genetics, University Hospital Bonn, Medical Faculty, University of Bonn, Bonn, DE
- 24, Mental Health and Neuroscience, QIMR Berghofer Medical Research Institute, Brisbane, QLD, AU
- 25, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SE
- 26, Old Age Psychiatry, King's College London, London, UK
- 27, Department of Medicine, Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, US

- 28, Department of Psychiatry and Psychotherapy, Fliedner Klinik Berlin, Berlin, BE, DE
- 29, Centre for Integrated Register-based Research, Aarhus University, Aarhus, DK
- 30, iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Aarhus, DK
- 31, Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Goethe University Frankfurt - University Hospital, Frankfurt am Main, DE
- 32, Discipline of Psychiatry, University of Adelaide, Adelaide, SA, AU
- 33, Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, US
- 34, Department of Biomedicine and Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, DK
- 35, Center for Genomics and Personalized Medicine, Aarhus University, Aarhus, DK
- 36, Department of Clinical Neuroscience, Karolinska Institutet,, SE
- 37, Department of Neurology, Klinikum rechts der Isar, Technical University of Munich, Munich, BY, DE
- 38, Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, BY, DE
- 39, Institute for Translational Psychiatry, University of Münster, Münster, NRW, DE
- 40, Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, NO
- 41, PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, NO
- 42, Research and Development, Takeda Pharmaceutical Company Limited, Cambridge, MA, US
- 43, Faculty of Health and Life Sciences, University of Exeter, Exeter, UK
- 44, Novo Nordisk Center for Protein Research, Department of Health Sciences, University of Copenhagen, Copenhagen, DK
- 45, Department of Psychiatry, Amsterdam Public Health and Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, NL
- 46, Morphology and Genetics, Universidade Federal de Sao Paulo, Sao Paulo, SP, BR
- 47, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, US
- 48, Faculty of Nursing and Health Sciences, NORD University, Levanger, NO
- 49, Department of Mental Health, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, TRD, NO
- 50, deCODE Genetics / Amgen, Reykjavik, IS
- 51, K. G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, TRD, NO

- 52, Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, NO
- 53, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, NO
- 54, Institute of Environmental Medicine, Unit of Integrative Epidemiology, Karolinska Institutet, Stockholm, SE
- 55, Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Berlin, DE
- 56, Department of Clinical Immunology, Roskilde University/Næstved Hospital, Roskilde, DK
- 57, Department of Clinical Immunology, Odense University Hospital, Odense, DK
- 58, Department of Psychiatry, Psychotherapy and Psychosomatics, Brandenburg Medical School Theodor Fontane, Neuruppin, BB, DE
- 59, Department of Psychiatry and Psychotherapy, Gender Research in Medicine, Institute of Sexology and Sexual Medicine, Charité – Universitätsmedizin Berlin, Berlin, BE, DE
- 60, iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Copenhagen, DK
- 61, Center for Neonatal Screening, Department for Congenital Disorders, Statens Serum Institut, Copenhagen, DK
- 62, Child Health Research Centre, University of Queensland, Brisbane, QLD, AU
- 63, Centre for Medical Informatics, Usher Institute, University of Edinburgh, Edinburgh, UK
- 64, Centre for Genomic & Experimental Medicine, Institute for Genetics and Cancer, University of Edinburgh, Edinburgh, UK
- 65, Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, SA
- 66, Department of Psychiatry, Lausanne University Hospital and University of Lausanne, Prilly, VD, CH
- 67, Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, ES
- 68, Department of Psychiatry, Faculty of Medicine and Institute of Neurosciences, Biomedical Research Centre (CIBM), University of Granada, Granada, ES
- 69, Université de Paris Cité, INSERM U1266, Institute of Psychiatry and Neuroscience of Paris, GHU Paris Psychiatry and Neuroscience, Paris, FR
- 70, Translational Biology, Biogen, Cambridge, MA, US
- 71, Nuffield Department of Population Health, University of Oxford, Oxford, UK
- 72, MRC Population Health Research Unit, University of Oxford, Oxford, UK
- 73, Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, DE
- 74, Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, CH
- 75, Institute of Medical Genetics and Pathology, University Hospital Basel, University of Basel, Basel, CH

- 76, Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, DE
- 77, Nic Waals Institute, Lovisenberg Diakonale Hospital, Oslo, NO
- 78, Centre for Advanced Imaging, University of Queensland, Saint Lucia, QLD, AU
- 79, Psychological Medicine, Cardiff University, Cardiff, WLS, UK
- 80, Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, WLS, UK
- 81, The Lothian Birth Cohorts, University of Edinburgh, Edinburgh, UK
- 82, Department of Biological Psychology & Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, NL
- 83, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Essen, University of Duisburg-Essen, Duisburg, DE
- 84, MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, UK
- 85, Imperial College Dementia Research Institute, Imperial College London, London, UK
- 86, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, US
- 87, Department of Psychiatry and Psychotherapy, Research Group Stress Related Disorders, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, BW, DE
- 88, Department of Clinical Immunology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, CPH, DK
- 89, Department of Clinical Immunology, Aarhus University Hospital, Aarhus, DK
- 90, Department of Psychiatry, Istanbul University, Istanbul, TR
- 91, Department of Medical Genetics, Oslo University Hospital, Oslo, OSL, NO
- 92, NORMENT, Department of Clinical Science, University of Bergen, Bergen, NO
- 93, Virginia Institute for Psychiatric & Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, US
- 94, Psychiatry Department / Huntsman Mental Health Institute, University of Utah School of Medicine, Salt Lake City, UT, US
- 95, Center for Genomic Research, University of Utah School of Medicine, Salt Lake City, UT, US
- 96, Department of Psychiatry and Psychotherapy, Medical Center, University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, DE
- 97, Division of Mental Health Care, St. Olavs Hospital, Trondheim University Hospital, Trondheim, TRD, NO
- 98, Department of Psychiatry, Sørlandet Hospital, Kristiansand, AG, NO
- 99, University of Oslo, NORMENT Centre, Institute of Clinical Medicine, Oslo, OSL, NO

- 100, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, US
- 101, Centre for Addiction and Mental Health, Toronto, ON, CA
- 102, Department of Psychiatry and Psychotherapy, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, DE
- 103, Mental Health Center Copenhagen, Mental Health Services Capital Region of Denmark, Copenhagen, DK
- 104, Faculty of Health Science, Department of Clinical Medicine, University of Copenhagen, Copenhagen, DK
- 105, Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, IT
- 106, Michigan Neuroscience Institute, University of Michigan, Ann Arbor, MI, US
- 107, Wolfson Institute of Population Health, Queen Mary University of London, London, UK
- 108, School of Computing and Mathematical Sciences, University of Leicester, Leicester, UK
- 109, Department of Epidemiology and Biostatistics, Imperial College London, London, UK
- 110, Discipline of Psychiatry, School of Medicine, Trinity College Dublin, Dublin, IE
- 111, Bioinformatics Research Centre, Aarhus University, Aarhus, DK
- 112, School of Engineering, University of Iceland, Reykjavik, IS
- 113, Danish Headache Centre, Department of Neurology, Rigshospitalet, Glostrup, DK
- 114, Faculty of Medicine, Department of Psychiatry, University of Iceland, Reykjavik, IS
- 115, Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, NL
- 116, MRC Human Genetics Unit, Institute for Genetics and Cancer, University of Edinburgh, Edinburgh, UK
- 117, Brain and Mind Centre, University of Sydney, Sydney, NSW, AU
- 118, Department of Epidemiology Research, Statens Serum Institut, Copenhagen, DK
- 119, Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University Medicine Greifswald, Greifswald, MV, DE
- 120, Roche Pharmaceutical Research and Early Development, Pharmaceutical Sciences, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Basel, CH
- 121, SleepWell Research Program and Department of Psychology and Logopedics, University of Helsinki, Helsinki, FI
- 122, Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
- 123, Max Planck Institute of Psychiatry, Munich, BY, DE
- 124, Department of Psychiatry, University of Helsinki, Helsinki, FI

- 125, 23andMe Research Team, 23andMe, Inc., Sunnyvale, CA, US
- 126, Department of Psychological Medicine, University of Worcester, Worcester, UK
- 127, Institution of Neuroscience and Physiology, University of Gothenburg, Gothenburg, SE
- 128, Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, AT
- 129, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, US
- 130, Health Care Policy, Harvard Medical School, Boston, MA, US
- 131, Department of Psychiatry, University of Toronto, Toronto, ON, CA
- 132, Department of Pharmacology & Toxicology, University of Toronto, Toronto, ON, CA
- 133, Department of Genetics, Rutgers University, Piscataway, NJ, US
- 134, Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, US
- 135, Mental Illness Research, Education and Clinical Center, Crescenz VA Medical Center, Philadelphia, PA, US
- 136, Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, EE
- 137, Department of Women's and Children's Health, Uppsala University, Uppsala, SE
- 138, Department of Epidemiology and Health Systems, Center for Primary Care and Public Health, Lausanne, VD, CH
- 139, Swiss Institute of Bioinformatics, Lausanne, VD, CH
- 140, Department of Computational Biology, University of Lausanne, Lausanne, VD, CH
- 141, Institute for Molecular Medicine Finland - FIMM, University of Helsinki, Helsinki, FI
- 142, Department of Psychiatry, Washington University School of Medicine in St. Louis, St. Louis, MO, US
- 143, Psychiatry, Veterans Affairs Connecticut Healthcare System, West Haven, CT, US
- 144, Department of Psychiatry, Yale University, New Haven, CT, US
- 145, Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, CN
- 146, Peking University Center for Public Health and Epidemic Preparedness & Response, Peking University, Beijing, CN
- 147, Mental Health, NHS 24, Glasgow, UK
- 148, Royal Edinburgh Hospital, NHS Lothian, Edinburgh, UK
- 149, Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, DE
- 150, Center for Translational and Computational Neuroimmunology, Columbia University Medical Center, New York, NY, US

- 151, Institute for Translational Neuroscience, University of Münster, Münster, NRW, DE
- 152, Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health Science, Copenhagen University, Copenhagen, DK
- 153, Department of Nursing, Faculty of Health Sciences and Institute of Neurosciences, Biomedical Research Centre (CIBM), University of Granada, Granada, ES
- 154, School of Public Health, University of Queensland, Brisbane, QLD, AU
- 155, DZHK (German Centre for Cardiovascular Research), Partner Site Greifswald, Greifswald, MV, DE
- 156, Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, MV, DE
- 157, Department of Psychiatry, University of Marburg, Marburg, DE
- 158, Department of Clinical Immunology, Aalborg University Hospital, Aalborg, DK
- 159, Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, NL
- 160, Department of Health, Science and Technology, Aalborg University, Aalborg, DK
- 161, Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, DK
- 162, Department of Biomedicine-Human Genetics, Aarhus University, Aarhus, DK
- 163, Population Health, QIMR Berghofer Medical Research Institute, Brisbane, QLD, AU
- 164, The Fraser Institute, Faculty of Medicine, University of Queensland, Brisbane, QLD, AU
- 165, Humus, Reykjavik, IS
- 166, Department of Clinical Medicine, University of Copenhagen, Copenhagen, CPH, DK
- 167, Department of Population Health Sciences, University of Leicester, Leicester, UK
- 168, Department of Psychiatry, Universidade Federal de Sao Paulo, Sao Paulo, SP, BR
- 169, Department of Psychiatry, Rutgers University, Piscataway, NJ, US
- 170, Department of Clinical Immunology, Zealand University Hospital, Køge, DK
- 171, Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University, Brooklyn, NY, US
- 172, NMD Pharma, Lexington, MA, US
- 173, Psychiatry, University of Pittsburgh Medical Centre, Pittsburgh, PA, US
- 174, Psychiatry, Universidade Federal do Rio Grande do Sul, Porto Alegre, BR
- 175, Department of Psychiatry, University Medical Center Groningen, Groningen, NL
- 176, Research School of Behavioural and Cognitive Neurosciences (BCN), University of Groningen, Groningen, NL

- 177, Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Mental Health Services Capital Region of Denmark, Copenhagen, DK
- 178, Neurogenomics Division, The Translational Genomics Research Institute (TGEN), Phoenix, AZ, US
- 179, Human Genetics Branch, NIMH Division of Intramural Research Programs, Bethesda, MD, US
- 180, Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Goettingen, NI, DE
- 181, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, US
- 182, School of Medicine and Dentistry, James Cook University, Townsville, QLD, AU
- 183, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, OSL, NO
- 184, NORMENT, Institute of Clinical Medicine, University of Oslo, Oslo, OSL, NO
- 185, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK
- 186, Division of Population Health and Genomics, Ninewells Hospital and School of Medicine, University of Dundee, Dundee, UK
- 187, Faculty of Medicine, University of Iceland, Reykjavik, IS
- 188, SAMRC Unit on Risk & Resilience in Mental Disorders, Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, SA
- 189, Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, HE, DE
- 190, Department of Psychiatry, Psychotherapy and Psychosomatics, Dr. Fontheim Mentale Gesundheit, Liebenburg, DE
- 191, Institute of Epidemiology and Social Medicine, University of Münster, Münster, NRW, DE
- 192, Institute for Community Medicine, University Medicine Greifswald, Greifswald, MV, DE
- 193, Wolfson Centre for Young People's Mental Health, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, WLS, UK
- 194, Department of Biological Sciences, Purdue University, West Lafayette, IN, US
- 195, Imperial College BHF Centre for Research Excellence, Imperial College London, London, UK
- 196, Genomics Program, University of South Florida College of Public Health, Tampa, FL, US
- 197, Department of Internal Medicine, Erasmus University Medical Center Rotterdam, Rotterdam, NL
- 198, Management Section, Statens Serum Institut, Copenhagen, DK
- 199, Department of Clinical Medicine, University of Copenhagen, Copenhagen, DK
- 200, Xperimed LLC, Basel, CH
- 201, Psychiatry, USUHS, Bethesda, US

- 202, Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, MV, DE
- 203, Department of Psychiatry, Leiden University Medical Center, Leiden, NL
- 204, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, US
- 205, Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, US
- 206, Division of Epidemiology, New York State Psychiatric Institute, New York, NY, US
- 207, Department of Neurology, Oslo University Hospital, Oslo, NO
- 208, HUNT All-In Psychiatry
- 209, China Kadoorie Biobank Collaborative Group
- 210, Genes & Health Research Team
- 211, KG Jebsen Centre for Neurodevelopmental Research, University of Oslo, Oslo, OS, NO
- 212, Department of Psychiatry, University of Melbourne, Melbourne, VIC, AU
- 213, Florey Institute of Neuroscience and Mental Health, University of Melbourne, Melbourne, VIC, AU
- 214, Department of Psychiatry, University of Münster, Münster, NRW, DE
- 215, Computational Health Centre, Helmholtz Zentrum München, Neuherberg, DE
- 216, School of Medicine, Technical University of Munich, Munich, BY, DE
- 217, Helmholtz Pioneer Campus, Helmholtz Zentrum München, Neuherberg, DE
- 218, Department of Psychiatry, University of Vermont, Burlington, VT, US
- 219, Department of Cellular, Computational and Integrative Biology, Università degli Studi di Trento, Trento, IT
- 220, Imperial College Biomedical Research Centre, Imperial College London, London, UK
- 221, Department of Psychiatry, Psychosomatics and Psychotherapy, Julius-Maximilians-Universität Würzburg, Würzburg, DE
- 222, Department of Genetics, Department of Neuroscience, Yale University School of Medicine, New Haven, CT, US
- 223, Psychiatry, Kaiser Permanente Northern California, San Francisco, CA, US
- 224, K. G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, NO
- 225, HUNT Research Center, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, NO

- 226, Department of Research, Innovation and Education, St. Olavs Hospital, Trondheim University Hospital, Trondheim, NO
- 227, NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK
- 228, Pathophysiology of Psychiatric Diseases, INSERM, Univ Paris Cité, GHU Paris, Paris, FR
- 229, Department of Psychiatry & Behavioral Sciences, Stanford University, Stanford, CA, US
- 230, Neuroscience Therapeutic Area, Janssen Research and Development, LLC, Titusville, NJ, US
- 231, Second Opinion Outpatient Clinic, GGNet Mental Health, Warnsveld, NL
- 232, Child and Youth Mental Health Service, Children's Health Queensland Hospital and Health Service, Brisbane, QLD, AU
- 233, Psychosis Research Unit, Aarhus University Hospital-Psychiatry, Aarhus, DK
- 234, Department of Psychiatry, Psychosomatics and Psychotherapy, University Hospital of Würzburg, Würzburg, DE
- 235, Munich Cluster for Systems Neurology (SyNergy), Munich, BY, DE
- 236, University of Liverpool, Liverpool, UK
- 237, Human Genetics and Computational Biomedicine, Pfizer Global Research and Development, Groton, CT, US
- 238, Centre for Quantitative Health, Massachusetts General Hospital, Boston, MA, US
- 239, Child and Adolescent Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, NL
- 240, Complex Trait Genetics, Vrije Universiteit Amsterdam, Amsterdam, NL
- 241, Wolfson Centre for Young People's Mental Health, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK
- 242, Department of Biochemistry and Molecular Biology II, Faculty of Pharmacy and Institute of Neurosciences, Biomedical Research Centre (CIBM), University of Granada, Granada, ES
- 243, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, SE
- 244, Division of Research, Kaiser Permanente Northern California, Oakland, CA, US
- 245, Department of Psychiatry, University of Michigan, Ann Arbor, MI, US
- 246, Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, US
- 247, Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, US
- 248, Psychiatry, UCSD School of Medicine, La Jolla, CA, US
- 249, Public Health, UCSD School of Public Health, La Jolla, CA, US
- 250, Psychiatry, Veterans Affairs San Diego Healthcare System, San Diego, CA, US

251, Departments of Genetics and Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, US

252, Child and Adolescent Psychiatry, Erasmus University Medical Center Rotterdam, Rotterdam, NL

253, Social and Behavioral Science, Harvard T.H. Chan School of Public Health, Boston, MA, US

254, Psychiatry, Dalhousie University, Halifax, NS, CA

255, Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Copenhagen University Hospital, Mental Health Services, Copenhagen, DK

256, GLOBE Institute, Lundbeck Foundation Centre for Geogenetics, University of Copenhagen, Copenhagen, DK

257, Queensland Brain Institute, University of Queensland, Brisbane, QLD, AU

258, Department of Medical & Molecular Genetics, King's College London, London, UK

259, Institute for Genomics and Cancer, University of Edinburgh, Edinburgh, UK

References

- Als, T. D., Kurki, M., Grove, J., Voloudakis, G., Therrien, K., Tasanko, E., Nielsen, T. T., Naamanka, J., Veerapen, K., Levey, D., Bendl, J., Bybjerg-Grauholm, J., Zheng, B., Demontis, D., Rosengren, A., Athanasiadis, G., Bækved-Hansen, M., Qvist, P., Walters, B., ... Børghlum, A. D. (2022). *Identification of 64 new risk loci for major depression, refinement of the genetic architecture and risk prediction of recurrence and comorbidities* [Preprint]. *Genetic and Genomic Medicine*. <https://doi.org/10.1101/2022.08.24.22279149>
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209. <https://doi.org/10.1038/s41586-018-0579-z>
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Dashti, H. S., Redline, S., & Saxena, R. (2018). Polygenic risk score identifies associations between sleep duration and diseases determined from an electronic medical record biobank. *Sleep*, *42*, 1–10. <https://doi.org/10.1093/sleep/zsy247>
- Grotzinger, A. D., Mallard, T. T., Akingbuwa, W. A., Ip, H. F., Adams, M. J., Lewis, C. M., McIntosh, A. M., Grove, J., Dalsgaard, S., Lesch, K.-P., Strom, N., Meier, S. M., Mattheisen, M., Børghlum, A. D., Mors, O., Breen, G., iPSYCH, Mattheisen, M., Mors, O., ... Nivard, M. G. (2022). Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nature Genetics*, *54*(5), 548–559. <https://doi.org/10.1038/s41588-022-01057-4>
- Johnston, K. J. A., Adams, M. J., Nicholl, B. I., Ward, J., Strawbridge, R. J., Ferguson, A., McIntosh, A. M., Bailey, M. E. S., & Smith, D. J. (2019). Genome-wide association study of multisite

chronic pain in UK Biobank. *PLOS Genetics*, 15(6), e1008164.

<https://doi.org/10.1371/journal.pgen.1008164>

Lam, M., Awasthi, S., Watson, H. J., Goldstein, J., Panagiotaropoulou, G., Trubetskoy, V., Karlsson, R., Frei, O., Fan, C.-C., De Witte, W., Mota, N. R., Mullins, N., Brügger, K., Lee, S. H., Wray, N. R., Skarabis, N., Huang, H., Neale, B., Daly, M. J., ... Ripke, S. (2020). RICOPIII: Rapid Imputation for COnsortias PIpeLine. *Bioinformatics (Oxford, England)*, 36(3), 930–933.

<https://doi.org/10.1093/bioinformatics/btz633>

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C., Zhan, X., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Furlotte, N. A., ... HUNT All-In Psychiatry. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use.

Nature Genetics, 51(2), 237–244. <https://doi.org/10.1038/s41588-018-0307-5>

Mullins, N., Forstner, A. J., O’Connell, K. S., Coombes, B., Coleman, J. R. I., Qiao, Z., Als, T. D., Bigdeli, T. B., Børte, S., Bryois, J., Charney, A. W., Drange, O. K., Gandal, M. J., Hagenaars, S. P., Ikeda, M., Kamitaki, N., Kim, M., Krebs, K., Panagiotaropoulou, G., ... HUNT All-In Psychiatry. (2021). Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nature Genetics*, 53(6), 817–829.

<https://doi.org/10.1038/s41588-021-00857-4>

Nagel, M., Jansen, P. R., Stringer, S., Watanabe, K., de Leeuw, C. A., Bryois, J., Savage, J. E., Hammerschlag, A. R., Skene, N. G., Muñoz-Manchado, A. B., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Furlotte, N. A., Hinds, D. A., ... andMe Research, T. (2018). Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nature Genetics*, 50(7), 920–

927. <https://doi.org/10.1038/s41588-018-0151-7>

Nievergelt, C. M., Maihofer, A. X., Klengel, T., Atkinson, E. G., Chen, C.-Y., Choi, K. W., Coleman, J. R. I., Dalvie, S., Duncan, L. E., Gelernter, J., Levey, D. F., Logue, M. W., Polimanti, R., Provost, A.

- C., Ratanatharathorn, A., Stein, M. B., Torres, K., Aiello, A. E., Almlí, L. M., ... Koenen, K. C. (2019). International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nature Communications*, *10*(1), 4558. <https://doi.org/10.1038/s41467-019-12576-w>
- Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlsgog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., ... Young, A. I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437–449. <https://doi.org/10.1038/s41588-022-01016-z>
- Pulit, S. L., Stoneman, C., Morris, A. P., Wood, A. R., Glastonbury, C. A., Tyrrell, J., Yengo, L., Ferreira, T., Marouli, E., Ji, Y., Yang, J., Jones, S., Beaumont, R., Croteau-Chonka, D. C., Winkler, T. W., GIANT Consortium, Hattersley, A. T., Loos, R. J. F., Hirschhorn, J. N., ... Lindgren, C. M. (2019). Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Human Molecular Genetics*, *28*(1), 166–174. <https://doi.org/10.1093/hmg/ddy327>
- The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, *526*, 68–74. <https://doi.org/10.1038/nature15393>
- Walters, R. K., Adams, M. J., Adkins, A. E., Aliev, F., Bacanu, S.-A., Batzler, A., Bertelsen, S., Biernacka, J., Bigdeli, T. B., Chen, L.-S., Clarke, T.-K., Chou, Y.-L., Degenhardt, F., Docherty, A. R., Fontanillas, P., Foo, J., Fox, L., Frank, J., Giegling, I., ... Agrawal, A. (2018). Trans-ancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *BioRxiv*, 257311. <https://doi.org/10.1101/257311>
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S.-A., Bækvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschøn, H. N., Bybjerg-

Grauholm, J., ... Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), 668–681. <https://doi.org/10.1038/s41588-018-0090-3>

Zhou, W., Nielsen, J. B., Fritsche, L. G., Dey, R., Gabrielsen, M. E., Wolford, B. N., LeFaive, J., VandeHaar, P., Gagliano, S. A., Gifford, A., Bastarache, L. A., Wei, W.-Q., Denny, J. C., Lin, M., Hveem, K., Kang, H. M., Abecasis, G. R., Willer, C. J., & Lee, S. (2018). Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature Genetics*, 50(9), 1335–1341. <https://doi.org/10.1038/s41588-018-0184-y>