Genetics of the mood spectrum disorders, symptoms, and measures

Mundy, Jessica

Awarding institution:
King's College London

Genetics of the mood spectrum:

disorders, symptoms, and measures

Jessica Mundy

Social, Genetic and Developmental Psychiatry Centre

King's College London

Submitted for the degree of

Doctor of Philosophy in Statistical Genetics

Supervised by Professor Gerome Breen,

Professor Sir Robin Murray, and Dr Evangelos Vassos

December 2022

Abstract

Pathological forms of mood have been documented for centuries. Mood disorder symptomatology varies enormously between individuals and can be viewed as a spectrum from depression to mania. The two most common disorders within the mood spectrum are major depressive disorder and bipolar disorder. The precise mechanisms that give rise to these disorders are largely unknown, but we now understand that both genetic and environmental factors increase individual risk. Improving knowledge of risk factors for mood disorders is crucial because they are leading drivers of disability and can substantially impact the quality of life of the affected individual.

Recent advances in technologies aimed at mapping people's inherited DNA have facilitated a deeper exploration of the genetic basis of mood disorders. Genome-wide association studies, which utilise genotype data, have been pivotal in confirming mood disorders' polygenic, heritable nature, as well as demonstrating that genetic risk factors are shared between psychiatric disorders. Genome-wide association studies of complex traits/diseases require sample sizes in the thousands to effectively capture the small effects of individual genetic variants. This is an even bigger priority for mood disorders due to their highly polygenic genetic architectures. Another factor is the level of detail included in the mood disorder phenotype, partly because this dictates sample size, but also because trait heterogeneity influences statistical power.

Research studies and biobanks that collect self-reported data on participants' psychiatric health, in addition to DNA samples, have facilitated the cumulation of samples sufficient for genetic studies of mood disorders. The UK Biobank, the Genetic Links to Anxiety and Depression study, and the COVID-19 Psychiatry and Neurological Genetics study are three UK-based studies that offer an opportunity to apply statistical genetics methods to self-reported data on disorders and symptoms within the mood spectrum. Effectively studying the genetics of any trait is rooted in the validity of the way it is measured, and there are many possible modes of measuring the mood spectrum with self-reported data. Given the sharp rise in genetic studies of mood disorders and the growing acceptance of their heritability, it is timely to evaluate approaches to measuring this disorder spectrum, to increase statistical power and to maximise the chance of replicable findings.

In this thesis, three empirical chapters are presented that explore three approaches and evaluate their utility for genetics research. The three approaches are: diagnostic subtypes, continuous

measures, and analyses at the symptom-level (including symptom subgroups and individual symptoms). The first empirical study (**chapter 2**) focuses on refined phenotyping approaches to explore the relationship between self-reported trauma and major depressive disorder. This chapter calculates the genetic overlap between various subtypes of this mood disorder and posttraumatic stress disorder to examine whether they share a genetic basis for trauma sensitivity. The second empirical study (**chapter 3**) investigates whether the Mood Disorder Questionnaire, a widely used screening tool for bipolar disorder, can be leveraged to construct a continuous measure of mania. This chapter examines whether this mania phenotype is valid for genome-wide association studies. The final empirical chapter (**chapter 4**) also applies a continuous measure to a specific mood symptom: anhedonia. Anhedonia has been posited as a risk factor for treatment resistance in individuals with major depressive disorder. This chapter examines whether the two mood phenotypes share genetic risk factors.

Genome-wide association studies hold great promise for improving the lives of individuals affected by mood disorders. However, the quality of their findings depends on the quality of the phenotypes examined. The final chapter (**chapter 5**) draws conclusions from the three studies together, and comments on the lessons learnt for phenotyping the mood spectrum for genetic studies. The thesis finds that the heterogeneity of mood disorders and their symptomatology can be accurately captured through various types of self-reported data and phenotyping strategies, but there are important caveats to this. An evaluation of the suitability of clinical tools when used for self-reported data collection should be prioritised. Also, data incorporated into continuous measures of mood psychopathology, such as composite symptom scores or staging models, requires careful consideration to reduce phenotypic noise and maximise statistical power. The hope is that the lessons presented in the thesis will be useful for other researchers who endeavour to study the genetic basis of mood disorders in the future.

Table of contents

Abstract **2**

Table of contents **4**

List of figures **7**

List of supplementary figures **9**

List of tables **13**

List of supplementary tables **14**

Acknowledgements **19**

Statement of authorship **21**

Publications arising from the chapters in this thesis **23**

Additional publications **24**

Chapter 1. Background **26**

History of the mood spectrum *26*

Mood disorders today *29*

Epidemiology *34*

Prevalence 34

Gender differences 35

Age of onset 35

Clinical course 36

Treatment and recovery 39

Burden of disease and mortality 40

Comorbidity *42*

Measuring the mood spectrum *45*

Diagnostic interviews 46

Algorithm-defined diagnoses 47

Screeners and symptom-based measures 47

Single-item self-reported clinical diagnoses 48

Self-reported disorder/treatment seeking for related mental health problems 49

The genetic basis of mood disorders *49*

A whistle-stop tour of the history of complex trait/disease genetics 50

An introduction to the genome-wide association study (GWAS) 54

Lessons learnt from genome-wide association studies 58
Mood disorders and rare variants 66
Measuring the mood spectrum: challenges in genetics research *67*
Aims and structure of the thesis *70*
References *72*

Chapter 2. Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder 89
Abstract *90*
Introduction *90*
Methods *91*
Results *94*
Discussion *95*
References *98*

Chapter 3. Genetic examination of the Mood Disorder Questionnaire and its relationship with bipolar disorder 100
Abstract *101*
Introduction *101*
Methods *104*
Results *111*
Discussion *118*
References *126*

Chapter 4. Exploring the genetics of anhedonia and its relationship with treatment response in major depressive disorder 130
Abstract *131*
Introduction *132*
Methods *134*
Results *142*
Discussion *150*
References *156*

Chapter 5. General discussion 161
Three lessons learnt from this thesis *161*
Lesson 1. Refined phenotypes offer insight into genetics as a driver of comorbidity 161
Lesson 2. Continuous measures can increase statistical power in a GWAS if used appropriately .. 166
Lesson 3. Clinical tools are not always suitable for data collection via self-reports 169
Limitations *173*

The study samples lack generalisability ... 173
The GWASs presented in this thesis only included individuals of European ancestries ... 175
Absence of replication in external data sets ... 178

Remaining challenges and future directions ... 178
Medical record linkage in the GLAD study ... 178
Diagnostic categories versus spectrum of mental illness ... 179
Is it possible to marry the two positions? ... 181
Integration of multimodal data to investigate the mood spectrum ... 184
Public education about risk factors for psychiatric disorders ... 185

Conclusion ... 189

References ... 190

Appendix 1. Supplementary material for chapter 1 ... 199

Supplementary material ... 199

Appendix 2. Supplementary material for chapter 2 ... 203

Supplementary material ... 203

Supplementary tables ... 216

Appendix 3. Supplementary material for chapter 3 ... 227

Supplementary material ... 227

Supplementary figures ... 237

Supplementary tables ... 267

Appendix 4. Supplementary material for chapter 4 ... 306

Supplementary material ... 306

Supplementary figures ... 316

Supplementary tables ... 334

Appendix 5. Perceptions of Psychiatric Risk (PerPsych) project ... 347

Open Science Framework pre-registration ... 347
PerPsych: Lived experience ... 348
PerPsych: Mental health professionals, students, and trainees ... 357
References ... 362

Script for animated video ... 364

Appendix 6. Data cleaning example ... 385

List of figures

Chapter 1:

Figure 1.1. Mood disorders act on a continuum of symptomatology from mania to depression, with euthymia (no mood disturbance) in the middle.
Figure 1.2. A comparison of the global burden of mental disorders.
Figure 1.3. Four phenotyping approaches depicted as a swimming pool of varying depth.
Figure 1.4. Individuals of European ancestries constitute a disproportionate amount of genome-wide association study (GWAS) samples.
Figure 1.5. Type I error, type II error, and statistical power.
Figure 1.6. The “mental health jar” analogy developed by Professor Jehannine Austin.
Figure 1.7. Diagram of UK-based studies included in the thesis.

Chapter 2:

Figure 2.1. Venn diagram showing participant overlap between the four major depressive disorder (MDD) categories in UK Biobank Mental Health Questionnaire (MHQ) respondents who met criteria for lifetime MDD (N=29,471). The MDD categories include MDD with reported trauma (N=13,391), MDD without reported trauma (N=9486), recurrent MDD (N=17,449) and single-episode MDD (N=12,022).

Chapter 3:

Figure 3.1. Distribution of Mood Disorder Questionnaire (MDQ) scores.
Figure 3.2. Factor analysis models of the Mood Disorder Questionnaire (MDQ).
Figure 3.3. Significant genetic correlations.

Chapter 4:

Figure 4.1. A-C: Plots of the three measures of anhedonic symptoms calculated for participants in the COVID-19 Psychiatry and Neurological Genetics (COPING) study (N=13,334).

Figure 4.2. Significant genetic correlations between the participants' highest anhedonic symptoms and external traits.

Chapter 5:

Figure 5.1. Mood disorders are heterogeneous.
Figure 5.2. A proposed spectrum of treatment response in major depressive disorder (MDD).
Figure 5.3. Hypothetical research study framework.
Figure 5.4. Flow-chart of the study design for the "lived experience" arm of the Perceptions of Psychiatric Risk (PerPsych) project.
Figure 5.5. Example question from the "lived experience" arm of the Perceptions of Psychiatric Risk (PerPsych) survey.
Figure 5.6. Positive feedback about the test animated video in the "lived experience" arm of the Perceptions of Psychiatric Risk (PerPsych) project.

List of supplementary figures

Appendix 3. Supplementary figures for chapter 3:

Figure S3.1. Correlations between concurrent Mood Disorder Questionnaire (MDQ) items in individuals affected by major depressive disorder (MDD) and/or an anxiety disorder.
Figure S3.2. Correlations between lifetime Mood Disorder Questionnaire (MDQ) items in individuals affected by major depressive disorder (MDD) and/or an anxiety disorder.
Figure S3.3. Correlations between lifetime Mood Disorder Questionnaire (MDQ) items in individuals unaffected by major depressive disorder (MDD) and/or an anxiety disorder.
Figure S3.4. Exploratory factor analysis (EFA): one factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.5. Exploratory factor analysis (EFA): two factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.6. Exploratory factor analysis (EFA): three factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.7. Exploratory factor analysis (EFA): four factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.8. Exploratory factor analysis (EFA): one factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.9. Exploratory factor analysis (EFA): two factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.10. Exploratory factor analysis (EFA): three factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.11. Exploratory factor analysis (EFA): four factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.
Figure S3.12. Exploratory factor analysis (EFA): one factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.
Figure S3.13. Exploratory factor analysis (EFA): two factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.
Figure S3.14. Exploratory factor analysis (EFA): three factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.

Figure S3.15. Exploratory factor analysis (EFA): four factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.

Figure S3.16. Raw factor scores from factor analysis of 12 concurrent manic symptoms measured by the Mood Disorder Questionnaire (MDQ) in affected participants.

Figure S3.17. Raw factor scores from factor analysis of 12 lifetime manic symptoms measured by the Mood Disorder Questionnaire (MDQ) in affected participants.

Figure S3.18. Principal component analysis (PCA) plots.

Figure S3.19. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the concurrent manic symptom sum score measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568).

Figure S3.20. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of concurrent energy/activity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568).

Figure S3.21. Quantile-quantile (QQ) plot and Manhattan plot of the genome-wide association study (GWAS) results of concurrent cognitive factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568).

Figure S3.22. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the concurrent impulsivity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568).

Figure S3.23. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime manic symptom sum score measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859).

Figure S3.24. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime energy/activity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859).

Figure S3.25. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime cognitive factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859).

Figure S3.26. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime impulsivity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859).

Figure S3.27. Scatter plot of affected participants' lifetime manic symptoms measured by the Mood Disorder Questionnaire (MDQ) and current posttraumatic stress disorder (PTSD) symptoms

measured by the 6 item PTSD Checklist (PCL-6). Lifetime manic symptoms were scored 0-12 and current PTSD symptoms were scored 6-30.

Figure S3.28. Flow-chart detailing how Genetic Links to Anxiety and Depression (GLAD) Study and COVID-19 Psychiatry and Neurological Genetics Study participants from the NIHR Bioresource (COPING NBR) were categorised as either “affected” or “unaffected’ by major depressive disorder (MDD) and/or any anxiety disorder. MDQ=Mood Disorder Questionnaire.

Appendix 4: Supplementary figures for chapter 4:

Figure S4.1. Phenotypic correlation matrix between the participants’ highest, lowest, and mean anhedonic symptoms.

Figure S4.2. Phenotypic correlation matrix between the participants’ highest anhedonic symptoms, their depressive and anxiety symptoms at that time (i.e., concurrent symptoms), and their concurrent score on the single anhedonia item from the nine item Patient Health Questionnaire (PHQ9).

Figure S4.3. Principal component analysis (PCA) plots.

Figure S4.4. Quantile-quantile (QQ) plot and Manhattan plot of Single Nucleotide Polymorphism (SNP) associations with the COVID-19 Psychiatry and Neurological Genetics (COPING) participants' maximum anhedonic symptoms (sum score) (N=13,433).

Figure S4.5. Quantile-quantile (QQ) plot and Manhattan plot of gene associations with the COVID-19 Psychiatry and Neurological Genetics (COPING) participants' maximum anhedonic symptoms (sum score) (N=13,433).

Figure S4.6. Quantile-quantile (QQ) plot and Manhattan plot of Single Nucleotide Polymorphism (SNP) associations with the Genetic Links to Anxiety and Depression (GLAD) study participants’ treatment-resistant depression sum score (N=11,234).

Figure S4.7. Quantile-quantile (QQ) plot and Manhattan plot of gene associations with the Genetic Links to Anxiety and Depression (GLAD) study participants’ treatment-resistant depression sum score (N=8,165).

Figure S4.8. Quantile-quantile (QQ) plot and Manhattan plot of Single Nucleotide Polymorphism (SNP) associations with the meta-analysed treatment-resistant depression phenotype.

Figure S4.9. Quantile-quantile (QQ) plot and Manhattan plot of gene associations with the meta-analysed treatment-resistant depression phenotype.

Figure S4.10. Flow-chart of participant exclusions in the COVID-19 Psychiatry and Neurological Genetics (COPING) study sample.

Figure S4.11. Flow-chart of participant exclusions in the Genetic Links to Anxiety and Depression (GLAD) study sample.

List of tables

Chapter 2:

Table 2.1. Information about the four posttraumatic stress disorder (PTSD) and four major depressive disorder (MDD) genome-wide association study (GWAS) summary statistics, including the original publication, characteristics of the sample, number (N) of cases and controls in original GWAS, liability scale SNP-based heritability (h^2_{SNP}) and standard error (S.E.) from High Definition Likelihood.

Table 2.2. Descriptive statistics for age, sex and current depression severity of the study sample of individuals who met criteria for lifetime major depressive disorder (MDD) in UK Biobank Mental Health Questionnaire (MHQ) respondents.

Table 2.3. Difference in reporting rates of traumatic life events between individuals with recurrent and single-episode major depressive disorder (MDD) in UK Biobank Mental Health Questionnaire (MHQ) respondents (N=29,471).

Table 2.4. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (S.E.) and 95% confidence intervals (lower and upper CI) of (1) UK Biobank posttraumatic stress disorder (UKB-PTSD), (2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD) and (3) Million Veteran Program PTSD (MVP-PTSD) with the four major depressive disorder (MDD) categories.

Chapter 3:

Table 3.1. Hypomanic/manic symptoms assessed by the Mood Disorder Questionnaire (MDQ).

Table 3.2. Demographic information for participants included in the three sets of analyses of the Mood Disorder Questionnaire (MDQ).

Chapter 4:

Table 4.1. Demographic descriptive statistics for the two analytical study samples (anhedonic symptoms and staged treatment-resistant depression).

Table 4.2. Single Nucleotide Polymorphism (SNP)-based heritability estimates (h^2_{SNP}).

List of supplementary tables

Appendix 2. Supplementary tables for chapter 2:

Table S2.1. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (SE) and 95% confidence intervals (CI) of Psychiatric Genomics Consortium 2 PTSD (PGC2-PTSD) with the four major depressive disorder (MDD) categories.

Table S2.2. Difference in reporting rates of traumatic life events, assessed by the Mental Health Questionnaire (MHQ), between individuals with recurrent and single-episode major depressive disorder (MDD) in UK Biobank MHQ respondents who met criteria for lifetime MDD (*N*=29,471).

Table S2.3. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (SE), and 95% confidence intervals (CI) between the four posttraumatic stress disorder (PTSD) phenotypes: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD).

Table S2.4. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (SE), and 95% confidence intervals (CI) between the four major depressive disorder (MDD) categories: 1) MDD with reported trauma, 2) MDD without reported trauma, 3) recurrent MDD, 4) single-episode MDD.

Table S2.5. Genetic correlation results from the High Definition Likelihood (HDL) block-jackknife analysis of posttraumatic stress disorder (PTSD) and the four major depressive disorder (MDD) categories. The four PTSD phenotypes include: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD).

Table S2.6. Linkage Disequilibrium Score Regression (LDSC) genetic correlation estimates (r_g) and standard errors (SE) and 95% confidence intervals (CI) of 1) UK Biobank posttraumatic stress disorder (PTSD), 2) Psychiatric Genomics Consortium (PGC) 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD) with the four major depressive disorder (MDD) categories.

Table S2.7. Genetic correlation results from the Linkage Disequilibrium Score Regression (LDSC) block-jackknife analysis of posttraumatic stress disorder (PTSD) and the four major depressive disorder (MDD) categories. The four PTSD phenotypes include: 1) UK Biobank PTSD (UKB-

PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD).

Table S2.8. Results table from PRSice analysis. Polygenic Risk Scores (PRS) for posttraumatic stress disorder (PTSD) were calculated based on the Million Veteran Program (MVP-PTSD) summary statistics.

Appendix 3. Supplementary tables for chapter 3:

Table S3.1. Descriptive statistics of the manic symptom phenotypes (sum scores) derived from the Mood Disorder Questionnaire (MDQ) in participants with and without a self-reported diagnosis of bipolar disorder by a professional.

Table S3.2. Statistics from testing assumptions for factor analysis of the Mood Disorder Questionnaire (MDQ) items.

Table S3.3. Item loadings for all factor solutions for the exploratory factor analysis (EFA) of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.4a. Fit statistics for each factor solution in exploratory factor analysis (EFA) of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.4b. Correlations between each factor in the best-fitting model (three factor solution) identified by exploratory factor analysis (EFA) of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants. EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

Table S3.5. Confirmatory Factor Analysis (CFA) fit statistics of the three factor model of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.6. Item loadings for all factor solutions for the exploratory factor analysis (EFA) of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.7a. Fit statistics for each exploratory factor analysis (EFA) of the 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.7b. Correlations between each factor in the best-fitting model (three factor solution) identified by exploratory factor analysis (EFA) of 12 lifetime manic symptoms derived from the

Mood Disorder Questionnaire (MDQ) in affected participants. EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

Table S3.8. Confirmatory Factor Analysis (CFA) fit statistics of the three factor model of 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.9. Item loadings for all factor solutions for the exploratory factor analysis (EFA) of the 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in unaffected participants.

Table S3.10. Fit statistics for each exploratory factor analysis (EFA) of the 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in unaffected participants.

Table S3.11. Information about each psychiatric and behavioural trait included in genetic correlations with quantitative manic symptom phenotypes.

Table S3.12. Genetic correlations between the four concurrent manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.13. Genetic correlations between the four lifetime manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected participants.

Table S3.14. Genetic correlations between concurrent manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected participants and 36 psychiatric and behavioural phenotypes. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the extended 1000 Genomes Linkage Disequilibrium reference panel.

Table S3.15. Genetic correlations between lifetime manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected participants and 36 psychiatric and behavioural phenotypes. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the extended 1000 Genomes Linkage Disequilibrium reference panel.

Table S3.16. Single Nucleotide Polymorphism (SNP) based heritability estimates (h^2_{SNP}), standard errors (SE), lambda GC, and mean chi-square statistics of the manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) calculated by Linkage Disequilibrium Score Regression (LDSC).

Table S3.17. Block-jackknife results comparing genetic correlations between 15 traits and the lifetime manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ). The block-jackknife tested two genetic correlations for a significant difference from zero.

Table S3.18a. Sub-cohorts of the National Institute for Health and Care Research (NIHR) BioResource who were included in the "affected" study sample (N=5,380).

Table S3.18b. Sub-cohorts of the National Institute for Health and Care Research (NIHR) BioResource who were included in the "unaffected" study sample (N=6,308).

Appendix 4. Supplementary tables for chapter 4:

Table S4.1. Anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30) and questions included in the Maudsley Staging Method (MSM).

Table S4.2. Descriptive statistics of the three measures of anhedonic symptoms measured using the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30).

Table S4.3. Descriptive statistics of staged treatment-resistant depression measured using the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM).

Table S4.4. Genome-wide Complex Trait Analysis (GCTA) Single Nucleotide Polymorphism (SNP)-based heritability estimates (h^2_{SNP}), standard errors (SE), and z-scores of anhedonic symptoms and staged treatment-resistant depression.

Table S4.5. Linkage Disequilibrium Score Regression (LDSC) Single Nucleotide Polymorphism (SNP)-based heritability estimates (h^2_{SNP}), standard errors (SE), lambda GC, and mean chi-square statistic of anhedonic symptoms and staged treatment-resistant depression.

Table S4.6. Genetic correlation between participants' mean anhedonic symptoms and staged treatment-resistant depression. Genetic correlation was estimated using the Genome-wide Complex Trait Analysis (GCTA) software.

Table S4.7. Information about each psychiatric and behavioural trait included in genetic correlations with highest anhedonic symptoms.

Table S4.8. Genetic correlations between participants' highest anhedonic symptoms and 42 psychiatric and behavioural traits. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the 1000 Genomes Linkage Disequilibrium reference panel.

Table S4.9. Genetic correlations between the single PHQ9 anhedonia item in the UK Biobank (Ward et al. 2019) and five psychiatric and behavioural traits which were significantly genetically correlated with anhedonic symptoms in the the COVID-19 Psychiatry and Neurological Genetics (COPING) study, three measures of treatment-resistant depression, and two measures of antidepressant response.

Table S4.10. Sub-cohorts of the National Institute for Health and Care Research (NIHR) BioResource who took part in the COVID-19 Psychiatry and Neurological Genetics (COPING) study and were included in the anhedonic symptoms study sample (N=7,843).

Acknowledgements

The research conducted in this PhD thesis would not have been possible without the participants of the UK Biobank, the GLAD study, and the COPING study. I enormously admire any person who willingly gives up their time and energy to share their personal experiences in the name of research. I would also like to thank the members of the NIHR BioResource team (past and present) who truly are the backbone of the GLAD and COPING studies. You have all worked so tirelessly to recruit participants and provide us with the data to carry out our studies. Thank you for being so patient with me, and for happily answering my many, many questions.

I would like to express an enormous amount of gratitude to my three fantastic PhD supervisors: Gerome, Robin, and Vangelis. The PhD has been a personal journey for me, as well as an academic one, and you have all been there through the happiest high points as well as my lowest moments.

To Gerome, you're always up for exploring new ideas and open to 'curve ball' suggestions, and I am sure this is what makes you such an incredible scientist. Not only did you offer me the chance to do a PhD with you, but you also supported me in setting up and launching the PerPsych study. Thank you for giving me the freedom to do that - PerPsych has been a highlight of my time as a PhD student. I often feel like I have learnt more in the last four years than in all the years that came before. That is very much down to you and the opportunities you gave to me to grow and develop as a scientist.

To Robin, thank you for being a constant and solid support. I always looked forward to our Zoom catch ups, mainly because you had a knack for making me feel calm. It really has been an honour to work with you, to hear about your varied experiences as a psychiatrist, and to receive your helpful input on my papers.

To Vangelis, you have been such an important emotional support to me over the last couple of years. I really don't know where I would be without you. Thank you for always being so understanding, incredibly generous with your time, and for making me feel like a valued member of the SGDP. You have helped to motivate me so many times and I am very thankful for the day that you agreed to jump on board as my third supervisor.

My experience at the SGDP has been shaped by so many dedicated, intelligent, kind, and wonderful people in both the TNG team and EDIT lab. I am so grateful for your input to my PhD, whether that be sharing code, receiving feedback on projects in team meetings, or reading manuscript drafts. I would like to say an extra special thanks to Thalia - you have been an emotional rock for me on multiple occasions and I honestly don't think I would have got through the PhD without your guidance. Thank you for being willing to lend an ear on countless occasions, for sharing your experiences as a woman in research, and for teaching me to reframe imposter syndrome as a strength and turn it into something positive.

Without a doubt, the best part of my PhD was getting to know six incredible women: Abi, Helena, Lissy, Katie, and Anna. I cherish our friendship; you all inspired me in a million wonderful ways and kept me going on so many occasions. I am endlessly proud of each of you and am so excited to see where our paths lead us in the future (hopefully always in each other's lives).

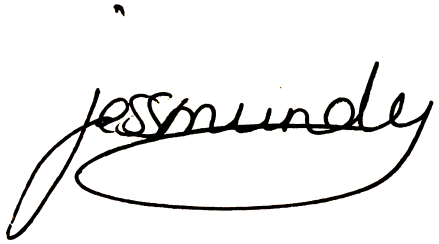
To Dan and Kelly, I lived with you both from the very beginning of the PhD all the way through to the end. You cheered me up when I was finding things tough, celebrated with me when things were going right, and reminded me to feel proud of myself (even for the small things). You were unwaveringly supportive from the get-go, and it has meant such a lot to have you both alongside me throughout this journey.

To my lovely family - Mum, Dad, and Charlie, I credit you with all the things I have been able to achieve. You supported me financially when I needed it during the PhD, have always been so interested in the research that I've been doing, and have given me so many words of encouragement. All those actions have contributed to where I am today.

Finally, to Finn, you have given me unquantifiable amounts of love, patience, and understanding over the last few years. You motivate me every day, thank you for everything.

Statement of authorship

All work presented in this thesis is my own except where acknowledged in the text. Data collection for study samples in **chapters 2-4** was completed by the respective research teams (the UK Biobank and the National Institute of Health and Care Research [NIHR] BioResource). I contributed to the phenotypic data cleaning pipeline for the data included in **chapter 3** and **chapter 4**: the Genetic Links to Anxiety and Depression (GLAD) study, the COVID-19 Psychiatry and Neurological Genetics (COPING) study, and the Eating Disorders Genetics Initiative (EDGI) (an example cleaning script is presented in **appendix 6**). Genetic data quality control (QC) for study samples in **chapters 2-4** was completed by the respective research teams. I conducted additional QC for the genotype and imputed data in **chapter 3** and **chapter 4**. Data collection for the Perceptions of Psychiatric Risk (PerPsych) project (mentioned in **chapter 5**) was completed by myself, another PhD student (HLD), and the NIHR BioResource research team. The studies presented in **chapters 2-4** were conceived and carried out by myself (JM), as first author, in collaboration with colleagues included in the author lists presented at the start of each chapter. Chapter-specific author contribution summaries are shown below, and acknowledgements are included in the respective chapters.



Jessica Mundy

Chapter 2:

JM, JRIC. and GB were responsible for study conception and design. JRIC, GB, MBS, The Million Veteran Program and the PTSD working group of the Psychiatric Genomics Consortium were responsible for the acquisition of the data. JM, JRIC and MS were responsible for data analysis. All authors were involved in the interpretation of the data. JM was responsible for the drafting of the paper, under the close supervision of JRIC and GB All authors read, edited and approved the final manuscript before submission. All authors agree to be accountable for all aspects of the

work, and in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated.

Chapter 3:

JM, GB, and EV, were responsible for the study conception and design. GB, TCE, GK, HBR, MRD, BNA, MH, and AM were responsible for the acquisition of the data. JM, HLD, CH, MRD, SHL, JRIC, were responsible for data cleaning. JM carried out data analysis with input from CH, EV, and GB. JM was responsible for the drafting and revising of the manuscript, under the close supervision of GB and EV. RBM, HLD, MRD, BNA, CH, MH, TCE, RMM, EV, and GB reviewed the manuscript. All authors agreed to be accountable for all aspects of the work, and in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated.

Chapter 4:

JM, GB, EV, and KY, were responsible for the study conception and design. GB, TCE, GK, HBR, MRD, BNA, MH, and AM were responsible for the acquisition of the data. JM, KLP, HLD, CH, MRD, SHL, JRIC, ZA, ATK, and JZK were responsible for data cleaning. JM carried out data analysis with input from CH, EV, and GB. JM was responsible for the drafting and revising of the manuscript, under the close supervision of GB and EV.

Publications arising from the chapters in this thesis

Chapter 2 is a published paper that has undergone peer-review:

Mundy, J., Hübel, C., Gelernter, J., Levey, D., Murray, R. M., Skelton, M., Stein, M. B., Vassos, E., Breen, G. and Coleman, J. R. I. (2021). Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder. *Psychological Medicine*. Cambridge University Press, pp. 1–10. doi: 10.1017/S0033291721000830.

Chapter 3 is currently accepted pending minor revisions at Neuropsychiatric Genetics Part B, of the American Journal of Medical Genetics:

Mundy, J., Hübel, C., Adey, B. A., Davies, H. L., Davies, M. R., Coleman, J. R. I., Hotopf, M., Kalsi, G., Lee, S.H., McIntosh, A. M., Rogers, H. C., Eley, T. C., Murray, R.M., Vassos, E., Breen, G. Latent factor structure of the Mood Disorder Questionnaire and analysis of its reliability and genetic validity. *Neuropsychiatric Genetics Part B, of the American Journal of Medical Genetics.*

Chapter 4 is in preparation:

Mundy, J., Purves, K. L., Young, K. S., Adey, B. A., Armour, C., Davies, H. L., Davies, M. R., Coleman, J. R. I., Hotopf, M., Hübel, C., Kalsi, G., Lee, S.H., McIntosh, A. M., ter Kuile, A. R., Rogers, H. C., Eley, T. C., Murray, R.M., Vassos, E., Breen, G. The genetics of anhedonia.

Additional publications

Co-authored published papers:

Mundy, J.*, Davies, H. L.*, Radu, M., Austin, J., Vassos. E., Eley, T. C., Breen, G., Moldovan, R. (2022). Research priorities in psychiatric genetic counselling: how to talk to children and adolescents about genetics and psychiatric disorders, *European Journal of Human Genetics.*
** Joint first author*

Davies, H. L., Hübel, C., Herle, M., Kakar, S., **Mundy, J.**, … Treasure, J., Breen, G. (2022). Risk and protective factors for new-onset binge eating, low weight, and self-harm symptoms in >35,000 individuals in the UK during the COVID-19 pandemic. *International Journal of Eating Disorders.*

Rayner, C., Coleman, J. R. I., Skelton, M., Armour, C., Bradley, J.R., Buckman, J. E. J., Davies, M., Hirsch, C., Hotopf, M., … **Mundy, J.,** … Breen, G., Eley, T. (2022). Patient characteristics associated with retrospectively self-reported treatment outcomes following psychological therapy for anxiety or depressive disorders - a cohort of GLAD Study participants. *BMC Psychiatry*.

Young, K. S., Purves, K. L., Hübel, C., Davies, M. R., Thompson, K. N., Bristow, S., Krebs, G., Danese, A., Hirsch, A. C., … **Mundy**, **J.,** … Eley, T. C., Breen, G. (2022). Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK. *Psychological Medicine.*

Davies, M. R., Buckman, J. E., Adey, B. N., Armour, C., Bradley, J. R., … **Mundy, J.**, … Breen, G., Eley, T. C. (2022). Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts, *Journal of Anxiety Disorders*.

Peel, A. J., Armour, C., Buckman, J. E. J., Coleman, J. R. I., … **Mundy, J**, … Danese, A., Eley, T. C. (2021). Comparison of depression and anxiety symptom networks in reporters and non-reporters of lifetime trauma in two samples of differing severity, *Journal of affective disorders reports.*

Di Gessa, G., Maddock, J., Green, M. J., Thompson, E. J., McElroy, E., Davies, H. L., **Mundy, J.**, …Porteous, D. J., Patalay, P. (2022). Pre-pandemic mental health and disruptions to healthcare, economic and housing outcomes during the COVID-19 pandemic: evidence from 12 UK longitudinal studies, *The British Journal of Psychiatry*.

Co-authored papers in press:

Davies, M. R., Glen, K., **Mundy, J.**, …Breen, G., Eley, T. C. (Accepted/In press). Factors associated with anxiety disorder comorbidity, *Journal of Affective Disorders.*

Co-authored papers in preparation:

Hübel, C., Johansson, T., **Mundy, J.**, Lin, Y., Petersen, L. V., Isomaa, R., Birgegård A., Comorbid posttraumatic stress disorder aggravates impulsive eating disorder symptoms and suicidality in a nationwide clinical eating disorder sample

Chapter 1. Background

Mood fluctuations, or changes in our background emotional state, are a normal and expected part of everyday life. Sometimes, changes in our mood may have no obvious trigger. Moods affect our thoughts and feelings and are often perceived as either “good” or “bad” (Thayer, 1990). In reality, mood is a continuum rather than dichotomous, ranging from low mood, known as **depression**, to euphoric or hyperactive mood, known as **mania**, with a plethora of symptoms in between these two poles. For some people, their mood may fall at the extreme ends of the spectrum for sustained periods of time. This may start to impact their quality of life, even when feeling “up” or euphoric. These individuals may be suffering from a “mood disorder”. The mood disorder category includes both depressive disorders (involving unipolar low mood) and bipolar disorders (involving low mood and periods of mania).

History of the mood spectrum

Pathological forms of mood have always existed and the mechanisms that give rise to them have been debated for centuries. Much like most human health phenomena, the earliest theories were based on beliefs in religion or the supernatural (Clark *et al.*, 2017). The first natural-science theory of depression, then called “melancholia”, can be traced back to the Ancient Greeks (Mondimore, 2005). Hippocrates was also one of the first individuals to note that melancholia only became pathological when it was present for sustained periods of time, and that grief could lead to despondency and depression. In “The Nature of Man”, Hippocrates (460-370 BC) proposed the existence of four different types of bodily fluid which were naturally balanced in well individuals but, when in disequilibrium, could lead to mental illness. These were called the “four humours": yellow bile, black bile, blood, and phlegm (Jouanna and Allies, 2012). These were later co-opted by philosopher Galen (129-216 AD) who proposed four categories of temperament based on the humours: choleric, melancholic, sanguine, and phlegmatic respectively. Each temperament had an associated pathology based on having a disproportionate amount of the bodily fluid. Here is where we see some of the earliest notions of “melancholia” (Clark *et al.*, 2017). Ancient Greece was not the only civilisation that was host to early thinkers about the mood spectrum. Evidence from the Islamic Golden Age (8th - 14th centuries) showed that physicians, such as Ibn Sina (also

known as Avicenna), wrote about melancholia, regarding it as a disease of the brain, heart, and blood, rather than a purely mental disorder (Yousofpour *et al.*, 2015).

The origins of what we know today as "bipolar disorder", which involves swings from depressed to manic mood states, can also be traced back to the Ancient Greeks. Arataeus of Cappadocia (~2 AD) noted the distinction between melancholia, which presented as *"sorrow and despondency"*, and mania, which presented as *"anger and sometimes joy"* in his writing (Kotsopoulos, 1986). Despite these vast differences in emotional state, he conceptualised them as two parts of the same disease course:

> *"It appears to me that melancholy is the commencement and part of mania."*
> (Quoted in Kotsopoulos, 1986)

He also wrote about the complexities of mania, whilst maintaining that its various manifestations nonetheless represented a single disease entity:

> *"There are infinite forms of mania but the disease is one..... If mania is associated with joy, the patient may laugh, play, dance night and day, and go to the market crowned as if victor in some contest of skill. If it is associated with anger, the patient may tear his clothes, kill his keepers, and lay violent hands upon himself... Some, if intelligent and educated, believe they are experts in astronomy, philosophy or poetry...while some uneducated may have strange delusions… If the illness gets serious, the patient may become excitable, suspicious, and irritable… If aroused to anger, he may become wholly mad and run unrestrainedly, roar aloud, flee the haunts of men and go to the wilderness to live by himself."*
> (Quoted in Kotsopoulos, 1986)

However, it is believed that these descriptions of mental impairments were not specific to depression and mania, but also included other forms of insanity, such as the symptoms experienced in schizophrenia (Kotsopoulos, 1986; Paykel, 2008).

In the 17th century, we see evidence that "melancholia" was used to refer to emotional problems that are more aligned with our modern view of depression. For example, in 1621 an Oxfordian Monk named Richard Burton wrote a book called the "Anatomy of Melancholia". Here, he described a range of psychiatric symptoms which we now know as depression. Burton speculated about the possible causes, which spanned religion and the supernatural (e.g., devils and witches) to social antecedents such as poverty, parenting, and employment (Burton, 1989).

The birth of modern psychiatry occurred in the 19th century. French psychiatrist Louis Delasiauve used the term “depression” to indicate a state of sadness for the first time in 1856. Following this, the term started to be applied more broadly. Bipolar disorder was also gaining interest in France during this time. In 1854, Jules Baillargar (1809-1890) used the phrase *“folie à doble forme”* (meaning “madness of two forms”) with regard to psychiatric patients experiencing switches in mood states, just as Arataeus of Cappadocia had noted many centuries previously. In the same year, Jean Pierre Falret (1794-1870) described *“folie circulaire”* (meaning “circular madness”) and, also in a manner similar to Arataeus of Cappadocia, identified depression with mania as the same disease entity, rather than separate psychiatric conditions (Paykel, 2008).

During the 19th century, two German physicians were also advancing the study of mood disorders. Wilhelm Griesinger (1817-1868) proposed that melancholia was a somatic disorder with neurobiological origins (Jansson, 2011). He also maintained that melancholia was one of the most treatable psychiatric disorders, and was an advocate of treating psychiatric patients within the community rather than containing them in asylums (Rössler, 1992; Rössler, Riecher-Rössler and Meise, 1994). Karl Ludwig Kahlbaum (1829-1899) wrote about cyclical insanity and depression in their milder forms. Kahlbaum also proposed that mental disorders should be studied over the course of a long time period to identify the prodromal state, acute state, remission, and recovery (‘Karl Ludwig Kahlbaum, M.D. 1828–1899’, 1999). These two physicians were strong influences on German psychiatrist Emil Kraepelin (1856-1926).

Kraepelin studied the case notes of psychiatric patients in immense detail during the late 19th century. This led him to propose that specific combinations of symptoms could be used to devise diagnostic categories which distinguished different forms of mental illness from each other. Kraepelin coined the term “involutional melancholia” in his earlier work, which included depressed states along with a number of other psychiatric symptoms (Hoch and MacCurdy, 1922), and “manic-depressive illness” in his later work (Paykel, 2008). Kraepelin separated “dementia praecox” (now known as schizophrenia) from manic-depressive illness. He proposed that dementia praecox was degenerative, severe, and unremitting, and this difficult clinical course led to permanent impairment. On the other hand, manic-depressive illness was episodic rather than persistent, and had comparatively better outcomes (Ebert and Bär, 2010). Despite some resistance, Kraepelin’s ideas, namely the classification of disorders based clinical phenomenology, had a gargantuan influence over much of psychiatry as we know it today. In the 1960s, Jules Angst, Carlos Perris, and George Winokur distinguished unipolar depressive

disorders from bipolar disorders and suggested that they were separate disease entities (Angst and Marneros, 2001). This is how the two extremes of the mood spectrum are conceptualised today. Major depressive disorder (the most common clinically recognised depressive disorder) and bipolar disorder are the focuses of this thesis because they are the most common disorders within the mood spectrum.

Mood disorders today

In the modern day, the scientific method has allowed researchers to uncover the cause(s) of thousands of human diseases. Yet, despite concerted research efforts over hundreds of years, the aetiology of mood disorders remains largely a mystery. Nonetheless, we have a fairly solid appreciation of their respective risk factors, which include a combination of biological factors (such as hormones and inherited genetic variants) and psychosocial and environmental exposures (such as stressful life events, trauma exposure, or a lack of social support) (Johnson and Kizer, 2002; Shih, Belmonte and Zandi, 2004; Otte *et al.*, 2016; Dahl *et al.*, 2017). Such a formulation of risk factors is known as the "biopsychosocial model" of medical conditions, first proposed by George Engel in 1977 (Borrell-Carrió, Suchman and Epstein, 2004).

Knowledge of potential risk factors, in the absence of direct causal mechanisms, means that it is currently not possible to diagnose a mood disorder based on pathology or aetiology. For instance, there is no known biomarker that could inform a clinician about whether their patient currently has one type of mood disorder or another (Miller, Johnson and Eisner, 2009). Furthermore, despite having a heritable basis, there is currently no genetic test to confirm a diagnosis (Palk *et al.*, 2019). Instead, clinical phenomenology is the primary method for classifying psychiatric disorders. Observable behaviours (signs) and self-reported emotions (symptoms) are used to construct operational criteria for multiple discrete psychiatric syndromes (Clark *et al.*, 2017; De Aquino and Ross, 2017; Smoller *et al.*, 2019).

Reflecting the work of Angst, Perris, and Winokur in the 60s, major depressive disorder and bipolar disorder have their own nosological categories in the most widely used diagnostic systems: the fifth edition of the **Diagnostic Statistical Manual of Mental Disorders (DSM-5)** and the tenth revision of the **International Classification of Diseases (ICD-10)**. The DSM states that one or more depressive episodes are experienced in both major depressive disorder and bipolar disorder type II, and are often experienced in bipolar disorder type I. A depressive episode

can involve several symptoms including low/depressed mood, **anhedonia**, feelings of guilt or worthlessness, suicidal ideation, and other non-mood cognitive and **psychomotor symptoms**. The DSM-5 diagnostic criteria for major depressive disorder are presented in **appendix 1**. Note that major depressive disorder is sometimes referred to as "unipolar depression". "Major depressive disorder" will be used throughout this thesis apart from situations where unipolar depression is needed to distinguish from the depression experienced in bipolar disorder.

In addition to one or more depressive episodes, a diagnosis of bipolar disorder is made on the basis of a single lifetime episode of mania for type I or hypomania for type II (American Psychiatric Association, 2013). Mania can be considered on the opposite end of the mood spectrum to depression, involving feelings of elation, euphoria, being hyperactive and/or irritable, making impulsive or risky decisions, being unusually sociable, and taking on lots of new activities. The DSM-5 diagnostic criteria for bipolar disorder type I and II are presented in **appendix 1**. Criteria for hypomania is identical to that for mania but in a milder form and with a shortened minimum duration (four days versus seven days). Note also that a depressive episode is not a requirement for a diagnosis of bipolar disorder type I, while it is for type II (American Psychiatric Association, 2013).

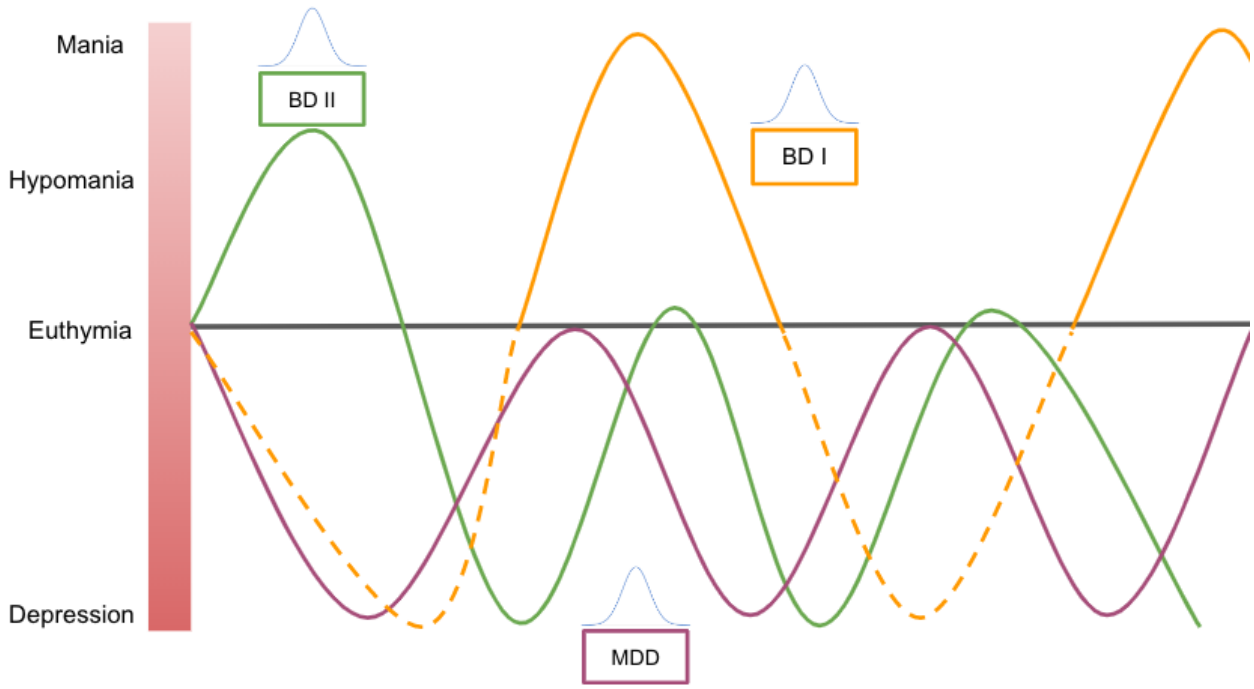
It is worth noting that the "mood disorder" category encompasses a plethora of psychiatric syndromes other than major depressive disorder and bipolar disorder. For instance, depression that does not meet criteria for major depressive disorder, premenstrual dysphoric disorder, persistent depressive disorder (also known as dysthymia), seasonal affective disorder, and disruptive mood regulation disorder, and cyclothymic disorder are other types of mood disorders (American Psychiatric Association, 2013).

The diagnostic categories in the mood spectrum are not distinct. A person with major depressive disorder may develop bipolar disorder at a later time point (Kessing *et al.*, 2017; Musliner and Østergaard, 2018; Baryshnikov *et al.*, 2020), while someone with bipolar disorder type II may subsequently experience a full manic episode and then be diagnosed with type I (Angst et al., 2005). A portion of individuals may experience subthreshold manic or hypomanic symptoms throughout their life, sometimes without ever reaching clinical threshold for a bipolar disorder diagnosis (Cassano *et al.*, 2004; Fiedorowicz *et al.*, 2011; Merikangas *et al.*, 2011). Unipolar organic mania (i.e., mania without depression that is not induced by a substance) is uncommon (Daly, 1997). Some individuals with bipolar disorder experience both depressive and manic

symptoms within one episode. This is known as a “mixed episode” (Muneer, 2017). For example, they may feel a sense of guilt, worthlessness, and sadness whilst simultaneously feeling energised, sociable, and sleeping much less than they usually do. Therefore, mood disorders sit within a continuum of symptomatology. At the same time, the individual diagnostic categories are not binary entities (despite often considered as such), so the individual disorders can be considered spectra in and of themselves. Whether someone is unaffected, subthreshold/subsyndromal, affected, or severely affected depends on where they fall on a continuum of severity and functional impairment (which are related, but not strictly the same), and duration (Clark *et al.*, 2017). This is represented graphically in **figure 1.1**. A challenge in classifying mental disorders, including those in the mood spectrum, is deciding where to apply clinical thresholds. A further challenge is addressing whether diagnostic categories accurately reflect natural disease biology (Smoller *et al.*, 2019). Epidemiological and genetics research suggest that the boundaries delimiting psychiatric disorder categories, and between individuals who are “affected” and “unaffected”, are not as obviously distinct as the prevailing multinomial taxonomic systems of the DSM and ICD suggest. Thus, the **validity** of the categories in the DSM and ICD, which are continually relied upon for both clinical work and research, have been called into question time and time again through psychiatric history.

Figure 1.1. Mood disorders act on a continuum of symptomatology from mania to depression, with euthymia (no mood disturbance) in the middle.

Diagnoses are made on the basis of symptom combination, severity, functional impairment, and duration. While sitting along a continuum, diagnostic categories also contain a spectrum of severity and functional impairment. This is denoted by the normal distribution curves above each diagnostic category. Therefore, despite diagnoses appearing as “either/or”, mood disorders are not binary entities. “BD” refers to bipolar disorder and “MDD” refers to major depressive disorder. The dashed line represents the fact that depressive episodes, while common in bipolar disorder type I, are not required for a diagnosis.



Mood disorders show a high degree of variation, also known as “heterogeneity”. This means that one individual’s experience of a mood disorder can differ substantially from another individual’s. This heterogeneity can come in the form of symptom combinations, symptom severity, clinical course, **clinical specifiers**, treatment-response, and other factors. There are 227 different symptom clusters that could result in a diagnosis of major depressive disorder in the DSM-5 (Ostergaard, Jensen and Bech, 2011). A commonly cited paper by Fried et al., (2015) showed that, purely on the basis of symptoms and impairment, there are upwards of 1,000 permutations of symptom combinations for major depressive disorder (Fried and Nesse, 2015). Bipolar disorder is a similarly heterogeneous disorder, complicated further by the presence of hypomania/mania in addition to depressive episodes.

The contribution of genetics to mood disorders is gradually being uncovered due to advances in genomic technologies, rapid progressions in computational power, and continually expanding sample sizes. Genetics research has the potential to improve the way we diagnose, treat, and prevent mood disorders, and enhance knowledge of their causes among those affected, their families, and the general public. Several questions regarding the optimal approaches taken to classify mood disorders in genetics studies are yet to be fully addressed, especially given their heterogeneity (Fried and Nesse, 2015; Coombes *et al.*, 2020). In the following sections, I will introduce major depressive disorder and bipolar disorder in greater detail. First, I will describe

their epidemiology and, second, their patterns of comorbidity with mental and physical health conditions. Third, I will introduce various approaches to classifying mood disorders that are employed in research. Fourth, I will discuss evidence for the **heritable** and **polygenic** basis of mood disorders. Last, I will summarise challenges posed by research into the genetics of mood disorders and explain how this thesis endeavours to address these. At the end of each section, there is a glossary of key terms that are highlighted in **maroon**.

Glossary of key terms 1:

Depression: A period of persistent feelings of sadness or loss of interest in activities.

Mania: A period of noticeably elevated or hyperactive mood, energy or activity levels.

Diagnostic Statistical Manual of Mental Disorders (DSM): A handbook used by clinicians and psychiatrists to diagnose individuals with mental disorders. The DSM contains descriptions, symptoms and other criteria for a number of mental disorders. It is published by the American Psychiatric Association (APA). There have been numerous editions of the DSM since its creation in 1952. The latest edition to be published was in 2013 (the DSM-5) with a text-revision (DSM-5-TR) published in 2022. The DSM is primarily used by clinicians in the United States (US) (Clark *et al.*, 2017).

International Classification of Disease (ICD): A guide for healthcare professionals to understand the extent, causes, and consequences of human diseases. Each disease has a clinical code which is used for standardised diagnosing of patients. The ICD includes physical health diseases as well as mental health diseases. The latest edition to be published was in 2022 (ICD 11). The ICD is authored by the World Health Organisation (WHO) and is used by clinicians internationally (Clark *et al.*, 2017).

Anhedonia: The inability to experience joy or pleasure in normal daily life. Anhedonia is a core symptom of depression (American Psychiatric Association, 2013).

Psychomotor symptoms: An aspect of depression symptomatology involving speech, posture, eye-movements, speed and degree of movement.

Clinical specifiers: Extensions to clinical diagnoses of psychiatric disorders which allow for more specific classification of symptoms.

Heritable: The proportion of phenotypic variance (Vp) that is explained by variance in genetics (Vg) within a population is known as "heritability" (H^2). Therefore, a trait that is heritable is partly influenced by inherited genetics. Note that traits have no "true" heritability because they are dependent on the time of measurement, type of measurement, and environmental context (Visscher, Hill and Wray, 2008).

Polygenic: Common diseases are caused by polymorphisms at many genetic loci in the genome, each of which contributes a small effect. The sum of all of these small effects, plus any environmental exposures, contributes to liability on an individual level. By contrast, monogenic disorders are those caused by mutations in a single gene (Visscher *et al.*, 2021).

Validity: The quality of being logically or factually sound. Within psychiatry, validity refers to the extent that diagnostic categories, as in the DSM and ICD, truly reflect discrete disease entities (Jablensky, 2016). Within research, "validity" refers to the extent that the chosen method of measuring a construct (e.g., a trait or disease) in a study sample accurately reflects the trait or disease in question in the population.

Epidemiology

Prevalence

Worldwide, mood disorders overall affect 10-20% of individuals during the lifetime (Weissman *et al.*, 1996; Kessler *et al.*, 2005; Steel *et al.*, 2014). There are differences between major depressive disorder, bipolar disorder type I, and type II. Major depressive disorder is one of the most common psychiatric disorders globally. The most accurate prevalence estimates come from the World Mental Health (WMH) survey initiative, which used DSM-IV criteria to assess major depressive disorder in 18 countries. They reported the average 12-month prevalence, which is the proportion of individuals who have had major depressive disorder in the last year, at around 6% (Bromet *et al.*, 2011). Lifetime estimates, which refers to the proportion of individuals who have experienced major depressive disorder in their lifetime, are usually higher at around 20% (Andrade *et al.*, 2003; Alonso *et al.*, 2004). Differences in the prevalence of major depressive disorder exist between countries. For instance, in the WMH survey, 12-month estimates ranged from 2% in Japan to 10% in Brazil. But, when the countries were categorised based on income level, prevalence estimates were broadly similar in ten high income countries (HICs) and eight low-and-middle income countries (LMICs) at around 6% (Bromet *et al.*, 2011).

Bipolar disorder affects fewer individuals than major depressive disorder. In a WMH survey initiative, this time of eleven countries in the Americas, Asia, and Europe, the average 12-month prevalence rates were 0.4% for bipolar disorder type I and 0.3% for bipolar disorder type II. The lifetime prevalences were marginally higher at 0.6% and 0.4% respectively. The WMH survey initiative also categorised individuals as having subthreshold bipolar disorder if they had at least one symptom on the screening questions for mania but did not meet the full diagnostic criteria for hypomania. The 12-month and lifetime prevalences of subthreshold bipolar disorder were higher at 0.8% and 1.4% respectively (Merikangas *et al.*, 2011). As with major depressive disorder, prevalence estimates for bipolar disorder differed between countries. For instance, the US had the highest 12-month and lifetime prevalence of bipolar spectrum disorders (which included subthreshold bipolar disorder) at 2.8% and 4.4%, while India had the lowest estimates (both 0.1%) (Merikangas *et al.*, 2011).

Gender differences

Mood disorders disproportionately affect women (Weissman *et al.*, 1996; Steel *et al.*, 2014). On average, women are at a 2x greater liability of developing major depressive disorder than men (Bromet *et al.*, 2011; Albert, 2015). This pattern has been consistently observed across different countries and cultural backgrounds (Kuehner, 2003; Seedat *et al.*, 2009; Bromet *et al.*, 2011; GBD 2019 Mental Disorders Collaborators, 2022). This female preponderance begins after puberty and persists into old age (Kuehner, 2017). This does not imply that women are innately predisposed to depressive symptoms to a greater extent than men. Rather, a number of biological factors (e.g., genetic and hormonal) as well as psychosocial factors likely play a role.

Bipolar disorder was initially thought to affect men and women at equal rates (Dell'Osso, Cafaro and Ketter, 2021). Recent research has suggested that there are clear gender differences. In the WMH initiative, lifetime rates of bipolar disorder type I and subthreshold bipolar disorder were more common in men, while bipolar disorder type II was more common in women (Merikangas *et al.*, 2011). Comorbid psychiatric syndromes and the clinical course of bipolar disorder also differs between genders. For instance, pre-teen onset bipolar disorder, **rapid cycling**, depressive episodes, mixed episodes, and attempted or completed suicide are more common in women than men who have bipolar disorder. Comorbid substance use disorders and legal problems are more common in men (Arnold, 2003; Dell'Osso, Cafaro and Ketter, 2021).

Age of onset

Major depressive disorder involves depressive episodes which can appear for the first time at any age, although there are periods when the risk of experiencing a first episode is higher than in other periods. Mental health disorders, in general, usually begin in childhood or adolescence, but major depressive disorder tends to manifest later than this average. The prevalence of major depressive disorder in childhood is low (Maughan, Collishaw and Stringaris, 2013; Wilson *et al.*, 2015). Multiple studies across different countries and cultures show that the median age of onset is somewhere around 25 years, but the peak risk period ranges from anywhere in mid-late adolescence to the early forties (Kessler and Bromet, 2013; Park *et al.*, 2014; Solmi *et al.*, 2022). In HICs, the risk of developing major depressive disorder for the first time decreases slightly with age after early adulthood (Bromet *et al.*, 2011).

Individuals with bipolar disorder usually present with a depressive episode first (Musliner and Østergaard, 2018; Baryshnikov *et al.*, 2020). A study investigating average age of onset in bipolar disorder, which combined estimates from Europe and the US, reported a median age of onset between 23-30 years (Baldessarini *et al.*, 2010). In this study, bipolar disorder type I began, on average, 5.8 years earlier than bipolar disorder type II. Men had a younger age of onset than women for both subtypes (thus adding further evidence for gender differences). However, other large studies have reported younger ages of onset than these. One study reported that 59% developed symptoms in childhood or adolescence (Lish *et al.*, 1994). A more recent nationally representative study of nearly 10,000 people from the US reported that the average onset of bipolar disorder type I at 18.2 years, type II at 20.3 years, and subthreshold at 22.2 years (Merikangas *et al.*, 2007). The WMH survey initiative reported estimates more consistent with this 2007 study: bipolar disorder type I had an average age of onset at 18.4 years, type II at 20 years, and subthreshold bipolar disorder at 21.9 years (Merikangas *et al.*, 2011). In general, although a person can develop new symptoms at any age, the peak risk period for developing bipolar disorder spans late adolescence to early adulthood. A consistent pattern is that type I tends to develop earlier than type II, and both tend to develop earlier than subthreshold bipolar disorder. Following depression, individuals with bipolar disorder will experience mania or hypomania. There is variability in how mania manifests initially. For some, symptoms may begin suddenly within a few hours or days. For others, manic symptoms may gradually develop over weeks or months (Daly, 1997).

Disclaimer. Estimations of age of onset will differ depending on the definition of "onset" used by researchers.
Studies looking at the age at which depressive symptoms first presented in the individual will report earlier age of onset compared to studies which use age of first diagnosis as a marker for onset. For instance, in an international study which combined data from multiple countries, the median age of onset for depressive symptoms was 26 years, whereas the median age of first diagnosis was 31 years (Solmi *et al.*, 2022).

Clinical course

A major depressive episode can affect people for long or short periods over the course of their lifetime. Some individuals may experience only one episode and then recover (known as "single episode depression"). Some may experience more than one episode which is punctuated by

asymptomatic stretches (known as “recurrent depression”). For some individuals, their depressive symptoms may not get better or go away for very long periods (i.e., 24 months or more) and these individuals are known as having “chronic depression”. Within a single depressive episode, the likelihood of recovery reduces as the duration of the presenting episode increases (Keller *et al.*, 1992). In population-based studies, the mean duration of a major depressive episode falls somewhere between 12-30 weeks (Blazer *et al.*, 1994; Eaton *et al.*, 1997; Spijker *et al.*, 2002).

Recurrence is common. More than half of individuals with major depressive disorder who initially recover from their first episode will go on to experience a second episode. Around 80% of individuals who have two episodes will experience another (Burcusa and Iacono, 2007). Even after full recovery, remaining subthreshold symptoms are associated with a more chronic and challenging disease course (Judd *et al.*, 2000).

Similarly to major depressive disorder, episodes within bipolar disorder are often recurrent. As mentioned above, individuals with bipolar disorder usually present with a depressive episode first rather than a manic episode. While correct at the time of the presenting episode, a diagnosis of unipolar depression may lead to pharmacological intervention which is not suitable for “hidden” bipolar disorder patients (Tondo, Vázquez and Baldessarini, 2010). For instance, treatment with an antidepressant can induce a manic episode or rapid cycling. In a study of 4,000 participants diagnosed with bipolar disorder, 69% had been initially misdiagnosed, with the most common misdiagnosis being unipolar depression (Hirschfeld, Lewis and Vornik, 2003). Studies suggest that subthreshold manic symptoms predict later onset of bipolar disorder (Fiedorowicz *et al.*, 2011).

The average time between the onset of depression and a diagnosis of bipolar disorder can be anywhere between 7-10 years (Hirschfeld, Lewis and Vornik, 2003; Mantere *et al.*, 2004; Drancourt *et al.*, 2013), and potentially longer in cases of childhood onset (Post *et al.*, 2010). This is problematic because the negative consequences of bipolar disorder have the potential to be profoundly lessened if diagnosis and suitable treatment occurs early (Drancourt *et al.*, 2013; Ratheesh *et al.*, 2017), and the consequences of delayed diagnosis can include worsening symptoms and antidepressant-induced rapid cycling (Dunner, 2003). The under-recognition of bipolar disorder is important to bear in mind when thinking about the typical age of onset for bipolar disorder. Delayed diagnoses may cause estimations to be inaccurate depending on the way age of onset is defined (first symptoms versus first diagnosis).

A few studies have investigated rates of conversion to bipolar disorder from unipolar depression. Initially, it was believed that rates of conversion were consistent over time (i.e., a person is just as likely to develop mania one year after their first depressive episode as they are after ten years) (Angst et al., 2005). However, this theory has been debunked based on robust evidence from a number of large, nationally representative cohort and register-based studies. In a Finnish nationwide register study of all the people who had been hospitalised for depression between 1996-2022 (N=43,495), the conversion rate to bipolar disorder was 7.4% during a 15-year follow-up period (Baryshnikov *et al.*, 2020). Likewise, in a prospective cohort study of nearly 100,000 individuals who had received their first diagnosis of major depressive disorder in a psychiatric hospital in Denmark, conversion to bipolar disorder was 7-8% (Musliner and Østergaard, 2018). Both studies found that liability to conversion decreased with time since the first depressive episode, with liability being the highest in the first year. Studies outside of Scandinavia have reported estimates of the proportion who convert to bipolar disorder ranging 5-23% (James *et al.*, 2015; Ratheesh *et al.*, 2017) suggesting there may be variation between countries. A meta-analysis of 55 studies investigating this topic found that having a family history of bipolar disorder, earlier age of onset, and the presence of psychotic symptoms were consistent risk factors for conversion (Ratheesh *et al.*, 2017).

The clinical course of bipolar disorder can be profoundly difficult for the affected individual. Depressive, dysthymic, and mixed episodes tend to account for the majority of the disease burden (Tondo, Vázquez and Baldessarini, 2017). In the WMH survey, individuals with bipolar disorder reported greater symptom severity in depressive episodes than manic episodes. For instance, around half of participants reported “severe role impairment” in their manic periods, whereas nearly three-quarters reported this for their depressed periods (Merikangas *et al.*, 2011). The chance of recurrence is high. In the Systematic Treatment Enhancement Program for bipolar disorder (STEP-ipol) in the US, around 50% of the 1,469 patients who were studied experienced recurrences. Recurrent depressive episodes were almost twice as common as recurrent manic episodes (Perlis *et al.*, 2006). Thus, recurrent episodes exacerbate impairment as affected individuals tend to rate their depressive episodes as more disabling than their manic episodes.

Treatment and recovery

Widely used treatments for major depressive disorder include pharmacological intervention (e.g., antidepressant medication) and talking therapies (e.g., cognitive behavioural therapy). Data on time to remission vary depending on data collection type. In a Dutch prospective epidemiological study of 7,076 participants, 50% recovered within three months, 63% within half a year, 76% within one year, but 20% had not recovered within two years (meaning that 20% of the sample exhibited chronic depression) (Spijker *et al.*, 2002). This was congruent with a previous study which suggested that around 80% reach remission by the two year mark (Ormel *et al.*, 1993). A later Dutch study, this time with data from a clinical sample, also reported that nearly 80% of 267 patients with pure major depressive disorder (i.e., no comorbidities) had fully recovered within two years (Penninx *et al.*, 2011).

In clinical samples, the course of major depressive disorder appears to be less favourable because recurrences are higher, and likelihood of recovery is lower. For instance, in a systematic review of studies investigating the prevalence of recurrent major depressive disorder, they found that the likelihood of recurrence in participants from mental health care settings was 85% compared to 35% in participants recruited from the general population (Hardeveld *et al.*, 2010).

The treatment of bipolar disorder is complex due to the different phases of the disease. First-line treatments for an acute manic episode primarily aim for stabilisation of the patient's symptoms. Lithium, valproate and second-generation antipsychotics are commonly prescribed to achieve this. Lithium and valproate are then used as maintenance treatments (Gitlin, 2006). Antidepressants are also prescribed to individuals with bipolar disorder, although there is some concern that this can lead to a switch to mania (especially in type I) (Ghaemi, Lenox and Baldessarini, 2001; Bond *et al.*, 2008), even when the individual is also being treated with mood stabilising drugs (Undurraga *et al.*, 2012). However, a caveat is that there is difficulty disentangling switches due to mood-elevating drugs versus spontaneous switches as part of the natural course of the disease.

Despite remarkable advances in pharmacological treatments for bipolar disorder, it remains a life-long illness for many of those affected. Syndromal recovery (no longer meeting diagnostic criteria) is seen in up to 98% of treated patients, but full symptomatic recovery (near absence of symptoms on standardised rating scales) is achieved in only 26-43%. Functional recovery, which involves

the patient regaining near pre-illness levels of psychosocial, occupational, and residential functioning, lags far behind both syndromal and symptomatic recovery (potentially in as few as 24% of treated individuals) (Keck *et al.*, 1998; Tohen *et al.*, 2003; Huxley and Baldessarini, 2007; Wingo *et al.*, 2010). However, it should be noted that bipolar disorder patients often demonstrate a relatively high base level of subsyndromal symptoms, which may mean that complete recovery post-treatment is an unrealistic goal (Sachs and Rush, 2003). Compared to major depressive disorder, research into treatment-response in bipolar disorder is lacking (Baldessarini, Vázquez and Tondo, 2020).

Disclaimer. Treatment response and resistance is hugely complex and challenging to formally define.
Response to treatment is conventionally thought of as a clinically meaningful reduction in symptoms following some form of treatment (usually a reduction in symptoms of 50% or more although this threshold may differ depending on the patient). Remission refers to the absence, or near absence of symptoms, and recovery refers to the absence of symptoms for a sustained period (Frank *et al.*, 1991; Rush *et al.*, 2006). Remission is the goal of acute treatment, while recovery is the goal of long-term treatment. However, remission and recovery may not be achievable by all patients, and some may exhibit "resistance" to many forms of treatment. There is no universally accepted definition of treatment resistance for both major depressive disorder and bipolar disorder (Gitlin, 2006; Fekadu *et al.*, 2009; Fountoulakis, 2012; Sforzini *et al.*, 2021). Thus, establishing accurate estimations of the proportion of individuals who do not recover after a treatment course is challenging. Nonetheless, it is estimated that around one-third of individuals with bipolar disorder exhibit some form of treatment resistance (Sportiche *et al.*, 2017). Similarly, in major depressive disorder, it is estimated that 20-40% of individuals do not respond to their first antidepressant medication (Fava, 2003; Trivedi *et al.*, 2006) and 7-22% do not respond to at least two antidepressants (Rizvi *et al.*, 2014; Fekadu, Donocik and Cleare, 2018; Wigmore *et al.*, 2020).

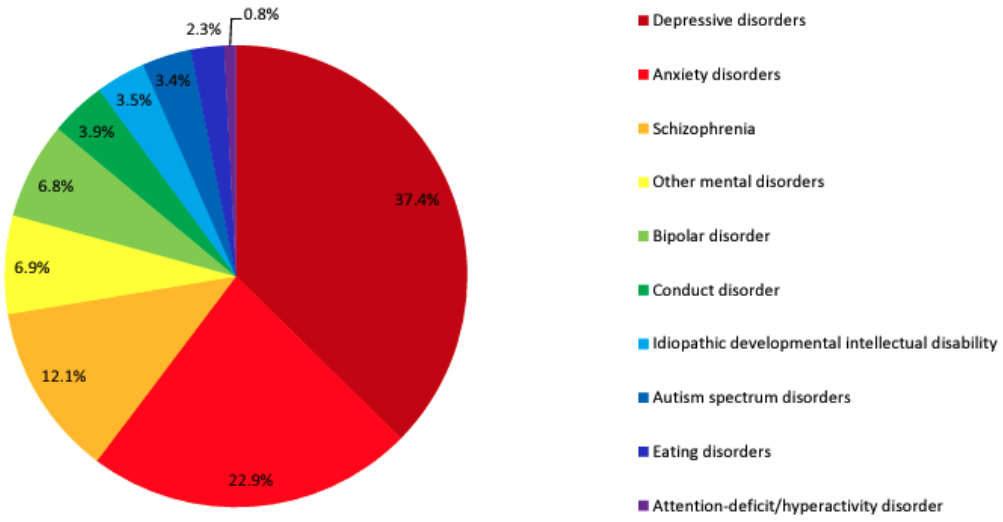
Burden of disease and mortality

Research shows that psychiatric disorders are leading causes of disability across the globe. When considered jointly, they are the second leading cause of disability behind ischemic heart disease (GBD 2019 Mental Disorders Collaborators, 2022). Disability is often quantified through **disability adjusted life years (DALYs)**. In 2019, mental health disorders were responsible for 125.3 million DALYs, which was a sizable increase since 1990 where they accounted for 80.8 million DALYs (GBD 2019 Mental Disorders Collaborators, 2022).

When comparing categories of psychiatric disorder, mood disorders are a serious public health concern regarding the DALYs that they account for. Depressive disorders consistently account

for the largest number (Reddy, 2010; Whiteford *et al.*, 2013; GBD 2019 Mental Disorders Collaborators, 2022) which is unsurprising given its status as the most common mental illness across the globe. In 2019, when the most recent Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) was conducted, depressive disorders accounted for the most DALYs due to mental disorders (37.5%), with anxiety disorders in second place (22.9%), and schizophrenia in third place (12.2%). Bipolar disorders were in fifth place (6.8%) (see **figure 1.2**). This pattern was broadly similar to the pattern observed in the 2010 GBD, where depressive disorders were in first place (accounting for 40.5% of DALYs) and bipolar disorders were in sixth place (accounting for 7.0% of DALYs) (Whiteford *et al.*, 2013). Major depressive disorder is expected to be the top cause of burden of disease by 2030 (Malhi and Mann, 2018)

Figure 1.2. A comparison of the global burden of mental disorders.
The pie chart shows the proportion of global mental disorder disability adjusted life years (DALYs) attributable to each disorder for both sexes and all ages in 2019. This pie chart was taken from the supplementary material of GBD 2019 Mental Disorders Collaborators (2022).



However, the 2019 GBD only included years of life lost to premature mortality (YLLs) in their estimation of DALYs for eating disorders. At a population level, individuals with mental health disorders have a reduced life expectancy of 10-20 years in comparison to the general population (Nordentoft *et al.*, 2013; Walker, McGee and Druss, 2015). Specifically, mood disorders are associated with exceptionally high susceptibility of early death (Osby *et al.*, 2001; Angst *et al.*,

2002). Thus, the true impact of mood disorders on overall DALYs in this 2019 GBD may be underestimated (GBD 2019 Mental Disorders Collaborators, 2022).

In a meta-review of 407 reviews, researchers found that individuals affected by major depressive disorder are 20x more likely to die by suicide than the general population, and those with bipolar disorder are 17-20x more likely (Chesney, Goodwin and Fazel, 2014; Miller and Black, 2020). In a Swedish national cohort study of nearly 7 million adults, women and men with bipolar disorder died, on average, 9.0 and 8.5 years earlier than the general population respectively (Crump *et al.*, 2013). Suicide is not the only cause of high levels of premature mortality in mood disorders. Medical comorbidities are also key drivers of this unfortunate statistic (Rowland and Marwaha, 2018). In the same Swedish cohort study mentioned previously, those affected by bipolar disorder (N=6,618) demonstrated increased mortality compared to unaffected individuals due to a range of physical health conditions such as CVD, diabetes mellitus, and respiratory diseases. Women with bipolar disorder also demonstrated increased mortality due to cancers (Crump *et al.*, 2013). Furthermore, in a study of all people who had been diagnosed with major depressive disorder or bipolar disorder between 1973-1995 in a hospital setting, also in Sweden, the most common cause of death was, in fact, a physical health disorder (CVD). Suicide was the second leading cause (Osby *et al.*, 2001).

Glossary of key terms 2:
Rapid cycling: In rapid cycling, an individual experiences four or more manic or depressive episodes within twelve months (Carvalho *et al.*, 2014).
Disability adjusted life years (DALYs): DALYs are calculated as the sum of years lost to disability (YLDs) and years of life lost to premature mortality (YLLs). The idea behind DALYs is to provide a metric which conveys the gap between the current health of the population and a standard life expectancy spent in full health (GBD 2019 Mental Disorders Collaborators, 2022).

Comorbidity

Mood disorders are disabling in and of themselves due to their symptoms and associated interference with daily functioning. But, as mentioned above, comorbid physical health disorders such as cancer, diabetes, and cardiovascular disease (CVD) hugely increase their burden (Gold *et al.*, 2020). A World Health Organisation (WHO) study of around 250,000 individuals from 60 countries in all continents of the world found that 9-23% of those with either angina, arthritis, asthma, or diabetes also had major depressive disorder compared to 3% without one of these

physical health disorders (Moussavi *et al.*, 2007). Likewise, multiple studies have documented a high prevalence of bipolar disorder concurrent with various medical conditions, such as CVD, obesity, thyroid problems, and type II diabetes (Osby *et al.*, 2001; Angst *et al.*, 2002; Soreca *et al.*, 2008; Kemp *et al.*, 2010).

Of course, comorbidity does not indicate directionality or causality. The experience of coping with a physical health disorder may lead to the development of a mood disorder (i.e., through stress or reduced quality of life). Equally, lifestyle correlates of mood disorders, such as an unhealthy or low quality diet, physical inactivity, smoking, low living standards or social deprivation, and reduced access to healthcare may make a person more likely to develop a comorbid physical health condition (Roshanaei-Moghaddam and Katon, 2009). For example, there are high levels of unemployment among individuals with bipolar disorder despite these individuals displaying relatively high academic achievement (Kupfer *et al.*, 2002; Kogan *et al.*, 2004). Therefore, an appreciation of the fact that individuals with mood disorders are more vulnerable to physical health conditions than those without is hugely important for improving overall health and mortality rates of this population. An example of this in action is the American Heart Association (AHA) who, in 2015, made a statement that children and adolescents with depression and bipolar disorder are at an increased risk of developing accelerated atherosclerosis and early CVD. Due to the high prevalence of these mood disorders, the AHA advised doctors to monitor such patients to prevent onset of these medical conditions (Goldstein *et al.*, 2015).

Comorbidity is the rule, rather than the exception, in psychiatry: individuals with mental illness often meet diagnostic criteria for another (Plana-Ripoll *et al.*, 2019). Research shows that nearly all mental disorders are twice as likely to occur in individuals with major depressive disorder compared to those without (Steffen *et al.*, 2020). Anxiety disorders are the most common comorbidity in major depressive disorder. This pattern has been observed in numerous studies of both in-patients and out-patients with estimates ranging 31-81% (Sartorius *et al.*, 1996; Olfson *et al.*, 1997; Hirschfeld, 2001; Zimmerman, Chelminski and McDermut, 2002; Lamers *et al.*, 2011; Plana-Ripoll *et al.*, 2019). Likewise, international studies demonstrate that major depressive disorder is more often comorbid with anxiety disorders than other types of psychiatric disorder. The WMH surveyed 74,000 adults in 24 countries and reported that 46% of those with major depressive disorder had a comorbid lifetime anxiety disorder (Kessler *et al.*, 2015). Also, a nationally representative study of nearly 10,000 adults in the US reported that 59% of individuals with major depressive disorder had comorbid anxiety (Kessler *et al.*, 2003).

Aside from anxiety disorders, comorbid substance use disorders (de Graaf *et al.*, 2002; Scott *et al.*, 2006; Kessler, Merikangas and Wang, 2007; Teesson, Slade and Mills, 2009; Lai *et al.*, 2015; Steffen *et al.*, 2020), personality disorders (Friborg *et al.*, 2014; Steffen *et al.*, 2020), and behaviour disorders (Steffen *et al.*, 2020) are common among individuals with major depressive disorder. In research from the Danish registers, comorbid mood and substance use disorders were associated with 3x mortality compared to experiencing a mood disorder alone (Plana-Ripoll *et al.*, 2020). Depressive symptoms are also the most common clinical correlates of schizophrenia (Siris, 2001), being apparent in 40% of those affected (Conley *et al.*, 2007).

Major depressive disorder with posttraumatic stress disorder (PTSD) is the most common comorbidity to develop following exposure to a traumatic event (Schindel-Allon *et al.*, 2010). Approximately 50% of individuals with PTSD have a comorbid diagnosis of major depressive disorder (Kessler *et al.*, 1995; Breslau *et al.*, 1997; Blanchard *et al.*, 1998; Rytwinski *et al.*, 2013). The high prevalence of this comorbidity has led some to suggest that it could represent a specific trauma-related disorder, or potentially even a subtype of PTSD (Flory and Yehuda, 2015), since comorbid individuals show distinct clinical profiles (e.g., more severe depressive symptoms and higher levels of distress, impairment, and suicidal ideation) (Blanchard *et al.*, 1998; Campbell *et al.*, 2007; Ramsawh *et al.*, 2014). The diathesis-stress model is helpful when thinking about the high prevalence of major depressive disorder-PTSD comorbidity (Monroe and Simons, 1991). This model was initially developed to explain how risk factors combine to influence whether someone develops schizophrenia (Rosenthal, 1963). It posits that environmental stressors may activate an innate vulnerability (diathesis), which turns the underlying potential for mental illness into a reality. In the context of major depressive disorder-PTSD comorbidity, the “stress” is exposure to a traumatic event: to receive a diagnosis of PTSD, the individual must have been exposed to trauma (American Psychiatric Association, 2013) and, while trauma is not a requirement for a major depressive disorder diagnosis, it is one of its strongest environmental predictors (Kessler, Davis and Kendler, 1997; Green *et al.*, 2010; Nanni, Uher and Danese, 2012). Variability in predisposition (diathesis) may be due to individual genetic differences. Support for this comes from recent research showing that PTSD has, in part, a genetic basis, despite many viewing the disorder as purely environmentally-triggered (Nievergelt *et al.*, 2019), and the fact that not everyone who is exposed to trauma develops PTSD, MDD, or both.

There is also a high frequency of psychiatric comorbidity in individuals with bipolar disorder, particularly substance abuse disorders, anxiety, and personality disorders (Baldessarini, Vázquez and Tondo, 2020). In the 2007 National Comorbidity Survey replication in the US, 97% of individuals with bipolar spectrum disorders met criteria for a comorbid psychiatric illness and comorbidity with more than one illness was the norm (Merikangas *et al.*, 2007). Lower estimates were found in the 2011 WMH survey (which drew upon data from all around the world rather than just the US). There, three-quarters of participants with bipolar disorder also met diagnostic criteria for another psychiatric disorder. Consistent with the National Comorbidity Replication survey, comorbidity with multiple disorders was common. Half of comorbid participants met criteria for three or more disorders. As observed in individuals with major depressive disorder, anxiety was the most common (particularly panic attacks) (69%), followed by behaviour disorders (45%), and substance use disorders (37%). It was noted that patterns of comorbid anxiety disorders were broadly similar across different regions of the world, while patterns of comorbid behaviour disorders differed (higher prevalence in the US and New Zealand compared to other nations) (Merikangas *et al.*, 2011). In agreement with these findings, a systematic review and meta-analysis reported that individuals are 3x more likely to develop an anxiety disorder if they have bipolar disorder compared to unaffected individuals. The lifetime prevalence of anxiety disorders in individuals with bipolar disorder was reported at 45% while the lifetime prevalence in the general population is much lower (varying between 4-29%) (Kessler *et al.*, 2005; Bandelow and Michaelis, 2015; Pavlova *et al.*, 2015).

Measuring the mood spectrum

As discussed at the very beginning of this chapter, a diagnosis of a mood disorder is made on the basis of a plethora of observable behaviours (signs) and self-reported emotional experiences (symptoms) (Clark *et al.*, 2017). Placing individuals within a psychiatric category based on their unique combination of signs and symptoms is not a simple task. As well as a correct assessment of the individual's symptoms, a diagnosis by a mental health professional requires temporal information and knowledge of contextual factors. The process of identifying individuals who are affected and unaffected by psychiatric disorders is essential for research purposes as well as in the clinic. To achieve this in research, there are different approaches, known as "phenotyping methods", which can be adopted. These can be viewed on a spectrum from "deep" to "shallow"

(see **figure 1.3**). Each approach comes with its own merits and limitations, largely regarding a trade-off between cost effectiveness/efficiency and detail.

Diagnostic interviews

The gold-standard approach for identifying individuals affected by mood disorders is the diagnostic interview. This interview can be fully- or semi-structured, and can take place over the phone or in-person. A fully-structured interview involves no open-ended questions, while a semi-structured interview allows for more flexibility in the questions asked by the interviewer. Following a clinical interview, the participant's answers inform whether they are affected or unaffected by various psychiatric disorders according to pre-specified diagnostic criteria.

One of the most well known fully-structured interviews is the Composite International Diagnostic Interview (CIDI). Starting in 1979, the CIDI was the result of a collaboration between the WHO Division of Mental Health and the US Alcohol, Drug Abuse, and Mental Health administration task force. The aim was to develop diagnostic interviews that were based upon diagnostic criteria for 40 different psychiatric disorders from the ICD and DSM to be used in epidemiological studies (e.g., to chart their prevalence, comorbidity rates, risk factors, and societal consequences). It is fully-structured so that a non-clinically trained person can administer the interview (which is essential for collecting data at-scale). The average duration of the interview is one hour (Robins *et al.*, 1988).

An example of a semi-structured clinical interview is the Structured Clinical Interview for DSM-5 diagnosis (SCID-5), which is based on diagnostic criteria from the DSM-5. The interview must be administered by a clinician or trained mental health professional who is familiar with the criteria. However, individuals without a clinical background but who have diagnostic experience with a particular study population can be trained to use it (First *et al.*, 2015). For bipolar disorder, another semi-structured interview is the Schedule for Affective Disorders and Schizophrenia (SADS) (Endicott and Spitzer, 1978).

Clinical interviews yield in-depth data on symptom combinations, duration, and possibly relevant contextual information about the patient which leads to an accurate depiction of their mental health status based upon internationally accepted diagnostic criteria. As such, they provide reproducible assessments of which diagnostic category/categories a patient fits into (Haro *et al.*,

2006). However, they are costly and time-consuming to administer to large numbers of individuals which, accordingly, restricts the size of sample that can be obtained through this approach to data collection (Davis, Cullen, *et al.*, 2019).

Algorithm-defined diagnoses

Clinical interviews are usually conducted between two individuals (i.e., the participant and the interviewer). However, since the CIDI is fully-structured, the questions can also be administered via a self-answered survey. “Algorithm-defined diagnoses” can then be derived from the participants’ answers. These are also known as “strictly-defined” or “detailed diagnoses” (Cai, Kendler and Flint, 2018; Davis, Cullen, *et al.*, 2019; Davies *et al.*, 2022). This method of ascertaining cases and controls is hugely advantageous when collecting large quantities of data in a research setting because, while benefiting from some of the depth of information collected in clinical interviews, they are nowhere near as burdensome or expensive to obtain.

Screeners and symptom-based measures

A screener is a quick tool for assessing which symptoms an individual is currently experiencing, how severely impacted they are by them, and, in some cases, how likely they are to meet clinical criteria for a diagnosis, usually utilised in a clinical context (Rush *et al.*, 2006). Screening tools can be self-scored by the individual or rated by a mental health professional. Careful assessment of symptoms can aid clinicians in determining whether the presenting individual requires support and, if they do, which type of support is most suitable. There is an abundance of screening tools for the symptoms within the mood spectrum for children, adults, and older aged persons. They are usually separated into tools for depressive symptoms and manic symptoms individually, except for the General Behaviour Inventory (GBI) which measures both (Depue *et al.*, 1981). Screeners for depressive symptoms that can be self-scored include the nine item Patient Health Questionnaire (PHQ9) (Kroenke, Spitzer and Williams, 2001) and the Beck Depression Inventory (BDI) (Beck *et al.*, 1961). Screening tools to identify mania/hypomania among those presenting with mental health problems are particularly important due to significant delays in diagnosing bipolar disorder and associated negative outcomes (Zimmerman *et al.*, 2009). An example of a self-scored screener for manic symptoms is the Mood Disorder Questionnaire (MDQ) (Hirschfeld *et al.*, 2000).

Repeated assessment of symptom severity can be used for monitoring mental illness over time and evaluating progress (e.g., since receiving treatment). Tracking symptoms is a crucial process for adapting medication(s) where necessary. For instance, treatment-response in MDD is based upon a reduction of at least 50% in MDD severity on a standardised rating scale (Fava, 2003). Clinician-rated scales for MDD include the Montgomery-Åsberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979) and the Hamilton Depression Rating Scale (Ham-D) (Hamilton, 1960). For mania/hypomania, the most widely used clinician-rated measures to track symptoms are the Young Mania Rating Scale (YMRS) (Young *et al.*, 1978), the Bech-Rafaelsen Mania Rating Scale (MAS), and the mania subscale of the SADS. Self-reported measures include the Altman Self-Rating Mania (ASRM) Scale (Altman *et al.*, 1997) and the Self-Rating Mania Inventory (SRMI) (Shugar *et al.*, 1992).

It should be emphasised that screeners and symptom-based measures are not intended to provide diagnostic accuracy but alert clinicians to individuals who might warrant further investigation and support. The **sensitivity** and **specificity** of screening tools massively vary. For instance, when using a cut-off threshold of ≥10, the PHQ9 has both sensitivity and specificity of 88% for major depressive disorder (Kroenke, Spitzer and Williams, 2001). The MDQ has variable psychometric properties depending on the context (e.g., clinical or community) (Miller, Johnson and Eisner, 2009). Screeners and symptom-based measures can be applied in research settings to capture a picture of participants' current or past symptoms, especially if they are suited to being self-rated.

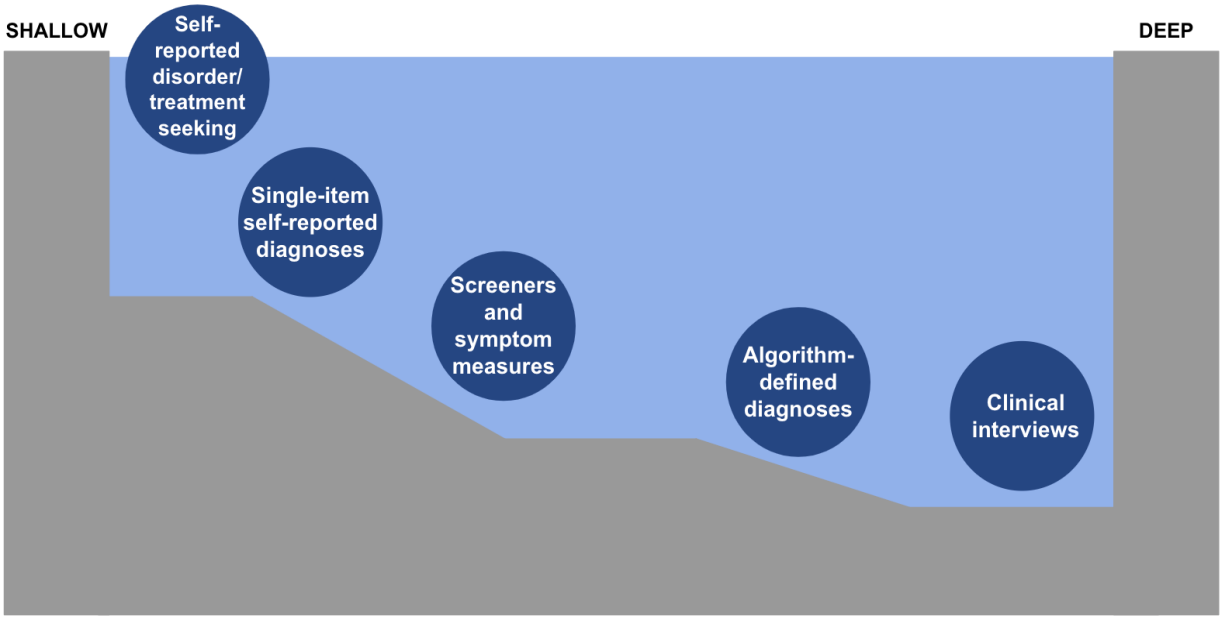
Single-item self-reported clinical diagnoses

At the shallow end of the phenotyping spectrum are self-reported clinical diagnoses. This approach relies on a single question, asked of participants in a questionnaire, about whether they have received a diagnosis of a specific disorder from a healthcare/medical professional at some point in their lifetime. This method of ascertaining cases and controls is also known as "broad", or "light-touch" phenotyping (Hyde *et al.*, 2016; Cai, Kendler and Flint, 2018; Davis, Cullen, *et al.*, 2019; Davies *et al.*, 2022). The accuracy with which this approach correctly defines cases and controls is reliant on the participant's ability to recall and willingness to report which diagnoses they have received (if any), as well as the accuracy of the diagnoses given by healthcare professionals.

Self-reported disorder/treatment seeking for related mental health problems

The shallowest phenotyping method is to categorise individuals as cases or controls based on whether they think they have ever had the psychiatric disorder of interest, or whether they have ever sought treatment for the disorder, or related mental health conditions (Howard *et al.*, 2018). As is the case with single-item self-reported clinical diagnoses, the validity of this phenotyping method relies on the recall of the participant, or the correct classification of their own mental health problems in the absence of a diagnosis from an expert. Focusing on treatment-seeking for related issues may mean that the specificity of case/control designation is lessened.

Figure 1.3. Four phenotyping approaches depicted as a swimming pool of varying depth. Self-reported diagnoses are considered the "shallowest" while diagnoses from structured or semi-structured clinical interviews are considered "deep".



The genetic basis of mood disorders

Before discussing the genetics of mood disorders, I will briefly describe the history of complex trait/disease genetics research.

A whistle-stop tour of the history of complex trait/disease genetics

Family members resemble each other in terms of their physical and behavioural characteristics and which diseases they suffer with. This observation has led researchers to ponder the question of nature and nurture in regard to human traits for centuries. In 1897, Francis Galton proposed that biological parents contribute one half of their genetics to offspring, with grandparents contributing one quarter (Galton, 1897). At this time, little was known about the specific mechanisms underlying familial inheritance.

Ronald A. Fisher was the first to propose a model of inheritance for complex human traits in a 1919 paper entitled *"The Correlation Between Relatives on the Supposition of Mendelian Inheritance"*. Here, Fisher suggested that quantitative traits (such as height) could be inherited in the same manner as Mendelian disorders (i.e., single gene disorders) if many genes acted upon the trait (i.e., the trait was polygenic) (Fisher, 1919). This model of inheritance has been developed since 1919 and is now known as the "infinitesimal model". Unlike the inheritance of single-gene disorders, also known as Mendelian disorders, the infinitesimal model assumed that quantitative traits are the outcome of a combination of genetic and non-genetic (environmental) influences. The model explained that the genetic component of offspring follows a normal distribution around the average genetic component of both parents. The variance of the genetic component is constant in a large out-crossing population (a large population with mating between unrelated individuals). However, in an in-breeding population, variance decreases in line with relatedness (Barton, Etheridge and Véber, 2017). The 'genetic component' comprises a number of different genetic factors which have small effects individually but, when taken in sum, produce a distribution of liability which underlies complex quantitative traits.

Twin and family studies, which were first developed and utilised in the 1900s (Luxenburger, 1928; Heston, 1966; Kendler and Zerbin-Rüdin, 1996), confirmed that all complex human traits are heritable to some extent (this is known as the "first law of behavioural genetics") (Turkheimer 2000). For instance, twin studies estimated the **broad-sense heritability** of major depressive disorder at 35-45% (Sullivan, Neale and Kendler, 2000; Polderman *et al.*, 2015) and of bipolar disorder at 65-75% (Polderman *et al.*, 2015). However, twin studies were unable to tell which specific genes or genetic variants actually caused the diseases in question. Therefore, with the knowledge of universal heritability, and that multiple **genomic loci** likely influenced liability for

complex human traits, discovering the locations of loci which increased risk or conferred protection to disease became a research priority.

Gene mapping by **family linkage studies** had been successful in Mendelian disorders which are caused by single genetic mutations with large effect sizes (i.e., high **penetrance** disorders) but unsuccessful in polygenic traits such as psychiatric disorders. This is because, for polygenic traits that involve both genetic and non-genetic influences, the effect sizes of associated genetic variants on their own are so tiny (i.e., low penetrance traits/disorders) they cannot be detected by cosegregation in **pedigrees** (Visscher *et al.*, 2012).

In 1996, Risch and Merikangas published a landmark paper entitled *"The Future of Genetic Studies of Complex Human Diseases"*. They demonstrated that performing an association study of one million genetic variants in the genomes of unrelated individuals could be more powerful than performing linkage analysis with a few hundred genetic markers in related individuals (Risch and Merikangas, 1996). In the following decade, the ability to map the genomes of hundreds of thousands of individuals was facilitated by profound advances in genomic technologies (Wang *et al.*, 2005). As a result, the human genome was sequenced and the haplotype structure of the genome was characterised (both completed in 2003). This enabled the discovery of structural variation comprised in the human genome, such as existence of millions of **common genetic variants**, such as **single nucleotide polymorphisms (SNPs; pronounced "snips")** and **insertion-deletion variants** of one or more nucleotides (indels), and the quantification of their correlation structure (known as **linkage disequilibrium [LD]**).

These advancements were indicators of a new era for genomics and biomedical research in the early 2000s (Hofker, Fu and Wijmenga, 2014). Following this, the introduction of genome-wide association studies (GWAS) was facilitated by the development of fairly inexpensive SNP microarrays (Visscher *et al.*, 2017) and the ability to impute millions of genetic variants from only a few thousand on these arrays (Coleman *et al.*, 2016). As a result, Risch and Merikangas's theoretical position was put into practice within a decade (Visscher *et al.*, 2012) and genome-wide association studies substantially altered the field's perception of the contributions of common genetic variation to complex traits (Hofker, Fu and Wijmenga, 2014).

Disclaimer. History of eugenics in complex trait/disease genetics

This introductory chapter mentions the seminal contributions of Ronald A. Fisher and Francis Galton to our understanding of statistical genetics. These academics, among others such as Karl Pearson, were pivotal in progressing the fields of statistics and statistical genetics through their thinking and research. They were also proponents of eugenics: the study of how to artificially engineer human reproduction to maximise the occurrence of heritable characteristics and traits that are regarded, by some in society, as desirable. It is important to recognise that the modern genetic principles that are integral to psychiatric genetics research, including the studies presented in this thesis, rest upon foundations that were built by individuals with prejudiced, ableist, and racist views and intentions. The scientific community is slowly grappling with this reality and steps are being put in place to address it (e.g., by removing the names of eugenicists from institutions and acknowledging this dark side of scientific history in research outputs) (Weiss and Lambert, 2011; Bodmer *et al.*, 2021).

Glossary of key terms 3:

Broad-sense heritability: The proportion of phenotypic variance (Vp) that is explained by genetic variance (Vg) in a particular population. Note that "heritability" and "broad-sense heritability" are often used interchangeably (Knopik, Neiderhiser, DeFries & Plomin, 2017).

Genomic locus: A specific position on a chromosome where a gene or genetic variant is located.

Family linkage studies: Children inherit DNA from both their mother and their father. DNA is inherited in large segments which means that genetic variants that are physically close to each other on the chromosome are expected to be inherited together more often than random chance. This is why the term "linkage" (i.e., the two genetic variants are linked) is used. In linkage analyses, a pedigree is used to measure the existence of genetic markers within a family which may co-occur with a particular disease in a hypothesis-free manner. If they are found to co-occur, then it is assumed that the genetic variant causing the disease must be near (i.e., linked) to the genetic marker being studied.

Penetrance: The penetrance of a genotype is the likelihood that the host to the genotype will develop clinical manifestations. Thus, penetrance is a term used to convey the proportion of individuals with a certain genotype who express the associated phenotype (i.e., show signs or symptoms of the disease or trait). Penetrance of 100% would mean that all individuals with the genotype go on to develop the disease. This is known as "complete penetrance". In reality, complete penetrance is relatively rare because most diseases and traits are not caused by genes alone. An example of a disease with complete penetrance is Huntington's disease (Zlotogora, 2003).

Pedigrees: A graphical representation of the inheritance of a trait through different generations of a family (National Human Genome Research Institute, 2022).

Common genetic variant: A common genetic variant is one which has a minor allele that is relatively frequent in the gene pool of a population. Common variants have a minor allele frequency of >0.05. Common variants contrast to rare variants, which have a minor allele frequency of <0.01. Rare variants tend to have larger phenotypic effects compared to common variants (Goswami, Chattopadhyay and Chuang, 2021).

Single Nucleotide Polymorphisms (SNPs): A SNP is a genomic locus at which there is a change of one base for another. SNPs are the most prevalent form of genetic variation in the human genome. A SNP is present in one in every 300 bases (equating to around ten million SNPs) in the human genome (Kruglyak and Nickerson, 2001; LaFramboise, 2009).

Insertion deletions (indels): Insertion deletions are one or more insertions or deletions in one or more nucleotides in the DNA sequence. Indels are the second most prevalent form of genetic variation in the human genome (SNPs are the most prevalent). Indels account for 15-21% of polymorphisms in the human genome (Mullaney *et al.*, 2010; Lin *et al.*, 2017).

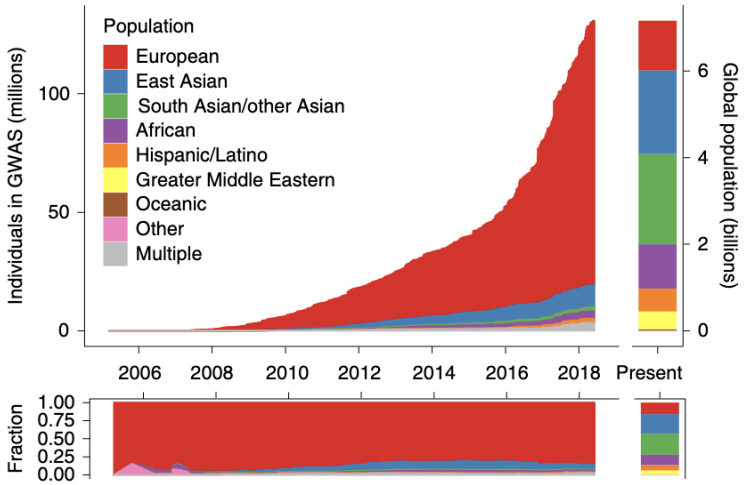
Linkage disequilibrium (LD): the non-random association of DNA variants at different loci (Slatkin, 2008). Due to recombination events during meiosis, genetic variants which are physically close together on a chromosome are more likely to be passed on from parent to offspring together. This means that certain genetic variants are found to be correlated with each other. The phenomenon of LD exists due to a number of factors including previous historical evolutionary forces, mutation, recombination rate, and natural selection. There are multiple ways to measure LD including D, D', and r^2 which are dependent on allele frequency.

An introduction to the genome-wide association study (GWAS)

A GWAS is an experimental design aimed at identifying associations between traits/diseases and polymorphisms at hundreds of thousands (sometimes millions) of loci in samples from populations (Visscher *et al.*, 2017). In brief, a GWAS performs a linear regression of a trait on each SNP (based on how many copies of a risk allele a participant carries [0, 1, 2] at each SNP being tested). A genotype is associated with the trait/disease if carrying a certain allelic variant is associated with a higher value of a continuous trait (e.g., height or depressive symptoms) or a higher likelihood of being a case than a control (e.g., if individuals affected by bipolar disorder carry the allelic variant at a higher frequency than individuals unaffected by bipolar disorder).

Disclaimer. Genome-wide association study (GWAS) samples disproportionately contain individuals of European ancestries.
Analysing the genomes of individuals from all over the world has demonstrated the genetic diversity of populations. Research proves that genetic differences are far greater *within* than *between* human populations (Rosenberg *et al.*, 2002) and an individual's ancestral background cannot be determined from their DNA alone. Nonetheless, at a population-level, ancestry is associated with subtle genetic differences. This is known as "population stratification" and refers to the process whereby structure in mating patterns leads to structures on genomic variation. This structure is therefore related to geographic location (Choi, Mak and O'Reilly, 2020). Differences can include levels of genetic diversity, the number of genetic variants in the genome, allele frequencies at certain variants, and the number and length of LD blocks. As a result, there are methodological issues that must be considered prior to conducting a GWAS. Genetic ancestry is calculated from genotype data and allows researchers to identify geographic groupings at a population level (Peterson *et al.*, 2019). Currently, individuals comprising a sample intended for a GWAS must all have the same (or similar) genetic ancestral background. Among the published GWAS literature, there is a lack of ancestral diversity (Haga, 2010). Despite representing 16% of the global population, White Europeans make up 79% of all GWAS samples (Martin *et al.*, 2019) (**figure 1.4**).

Figure 1.4. Individuals of European ancestries constitute a disproportionate amount of genome-wide association study (GWAS) samples.
The plot shows the relative proportions of published ancestry-specific GWAS between 2006-2018. This plot is from Martin et al., (2019).



The primary aim of the GWAS design is to discover the locations of polymorphisms which may increase risk or confer protection to the trait/disease being studied. The hope is that, in doing this, we will understand more about the cause(s) of diseases and this knowledge will aid the development of preventative strategies such as treatments and therapies. However, the value of any finding from GWAS is dependent on whether it can be replicated in an independent sample. One way to achieve this is to minimise the possibility of type I error (Coleman *et al.*, 2016). Type I error denotes a situation where the null hypothesis is rejected when it, in fact, was true (denoted as α). An example of a type I error in the context of a GWAS is the "discovery" of a statistically significant association between the trait and a specific polymorphism when the polymorphism is *not* actually causally related to the trait. A crucial step for minimising the chance of type I error is to apply strict quality control (QC) to the genetic data being used in the GWAS (Coleman *et al.*, 2016). QC refers to pre-analytical steps which ensure that the genetic data does not contain errors which could potentially lead to incorrect or biased results. Part of the QC process involves applying thresholds to various aspects of the genetic data and the individuals in the sample to ensure that any discoveries are due to true associations rather than unique characteristics of the data.

Another way to minimise the possibility of false positives is to correct for multiple testing. In order to account for the many millions of SNPs being tested for an association with the trait, GWASs use a strict threshold for significance of an association as $p<5x10^{-8}$ (Horwitz *et al.*, 2019). This p-value threshold was first introduced by Risch and Merikangas in 1996 (in the same paper that is cited above) (Risch and Merikangas, 1996). This was based upon the low replication rate of

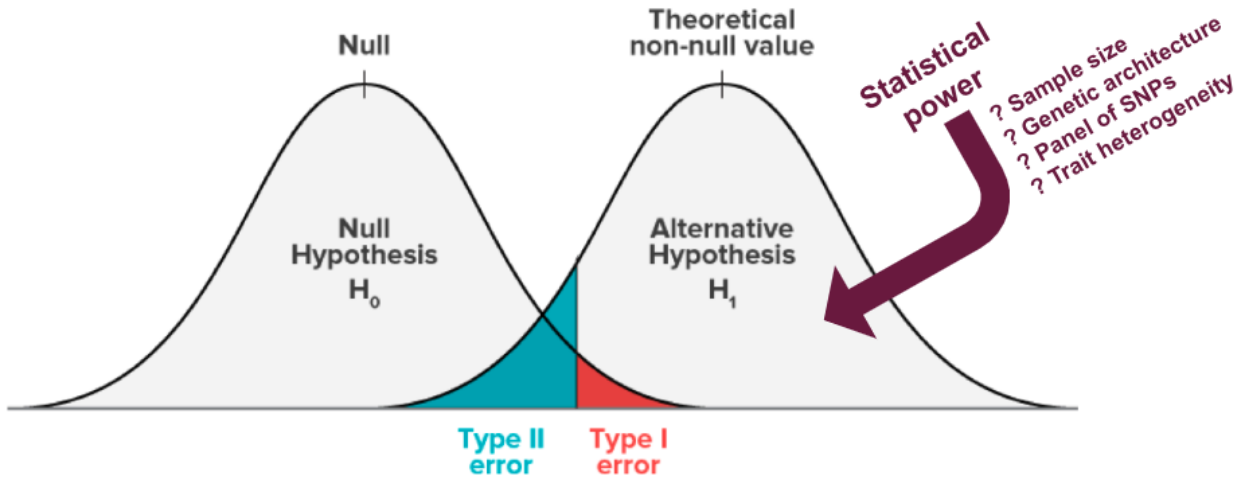
findings from **candidate gene studies** (Chen *et al.*, 2021). Subsequently, the International HapMap Project, which initially **genotyped** 1 million (later 3.1 million) polymorphisms the genomes of in 269 individuals of European, Han Chinese, Japanese, and Yoruban ancestry, proposed a near identical threshold for common genetic variants tested in GWASs (International HapMap Consortium, 2005). The widespread use of this threshold proved hugely successful in minimising false positives and, accordingly, increasing reproducible discoveries, the latter of which had been scarce in the field of complex trait genetics until the advent of the GWAS (Chen *et al.*, 2021).

The possibility of type II error also impacts the value of the results of a GWAS. Type II error refers to a situation where the null hypothesis is rejected when, in fact, the alternative is true (denoted as β). In the context of a GWAS, not detecting a true association between a polymorphism and the trait of interest is an example of a type II error (i.e., an association is not statistically significant when it should be).

Statistical power influences the likelihood of type II error (denoted as 1-β). In a GWAS, the statistical power to identify genetic variants that are associated with a trait or disease is dependent on multiple factors. Visscher et al., (2017) outline six considerations for statistical power (Visscher *et al.*, 2017):

1) Sample size
2) Distribution of effect sizes of causal genetic variants in study population
3) Frequency of those variants in study population
4) LD between observed, genotyped variants and unknown causal variants
5) Panel of genome-wide variants used to genotype individuals in sample
6) Heterogeneity of the trait/disease being studied (which relates to the clinical manifestations of the trait/disease and the way it is measured).

Figure 1.5. Type I error, type II error, and statistical power.
Multiple factors influence statistical power in a genome-wide association study (GWAS), including sample size, heterogeneity and genetic architecture of the trait being studied, and the panel of SNPs used for genotyping.



The first GWAS to be published was in 2005 (Klein *et al.*, 2005). However, the year 2007 marks the first well-designed GWAS with a sufficient sample size for discovery. This was performed by the Wellcome Trust Case Control Consortium (WTCCC) on seven complex traits, one of which was bipolar disorder (Wellcome Trust Case Control Consortium, 2007). This was followed by similar studies of multiple human traits and diseases, including quantitative traits such as body mass index (BMI), physical health conditions such as autoimmune disorders, sociobehavioural traits such as intelligence, educational attainment, and psychiatric disorders (Goldstein, 2009).

Despite identifying thousands of trait/disease-associated loci, the initial successes of the GWAS method were questioned for many reasons. First, the phenotypic variance that was attributable to common variants, sometimes known as **SNP-based heritability (h^2_{SNP})**, fell far below broad-sense estimates. This phenomenon was termed “missing heritability” (Maher, 2008). Second, the majority of significantly associated SNPs had little biological relevance or clinical utility for the trait/disease being studied (many of them were, in fact, not even located within genes). Third, the miniscule effect sizes of individual variants led some researchers to question whether they were truly associated with the trait/disease or simply a product of population stratification within the GWAS sample. As a result of these disappointments, some within the scientific field challenged the notion that common diseases were caused (in part) by common genetic variation at all (known as the common-disease-common-variant hypothesis) (Visscher *et al.*, 2012).

Since then, with vast increases in sample size and international collaboration efforts, the field of psychiatric genetics has seen an explosion in the discovery of robustly associated variants which replicate both within and between populations. A 2019 review paper of psychiatric GWASs

reported that, out of 514 studies, 1,123 SNPs had reached genome-wide significance ($p<5x10^{-8}$) and, of these, 453 had been reproduced in some capacity (Horwitz *et al.*, 2019). These successes have largely been due to international collaborative efforts to share and analyse data sets. While there are, of course, limitations to this method and (still) many unanswered questions, GWASs have led to a renewed understanding of the genetics behind complex traits and diseases, including those in the mood spectrum.

Lessons learnt from genome-wide association studies

Below, I outline **three** major conclusions drawn from the GWAS method with specific reference to recent research into disorders of the mood spectrum (Visscher *et al.*, 2017).

1. Complex traits are highly polygenic

Candidate gene studies of both major depressive disorder (Otte *et al.*, 2016; Border *et al.*, 2019; Norkeviciene *et al.*, 2022) and bipolar disorder (Jones and Craddock, 2001; Edvardsen *et al.*, 2008; Otte *et al.*, 2016; Border *et al.*, 2019; Oraki Kohshour *et al.*, 2022) suffered from inconsistent and non-replicable results. We now understand that these mood disorders, like all complex traits, are multifactorial (i.e., influenced by a combination of environmental factors and the small effects of many genetic variants; **figure 1.6**). Rather than focusing on single genes, GWASs scan the genome looking for associations at hundreds of thousands to millions of individual SNPs.

Major depressive disorder GWASs of the early 2010s yielded no positive findings (e.g., no genome-wide significant loci). This led to debates within the field about whether the main issue was the phenotyping approaches taken or the small samples. The first GWAS to identify and replicate any genetic variants associated with major depressive disorder was in a sample of Han Chinese women with recurrent major depressive disorder and 5,337 unaffected women by the China, Oxford, and Virginia Commonwealth University Experimental Research on Genetic Epidemiology (CONVERGE) consortium in 2015 (N_{total}=10,640) (CONVERGE consortium, 2015). This research group favoured homogeneous phenotyping over sample size and were able to identify and replicate two associations on chromosome 10. Although there is imperfect alignment between shallow and deeper measures (e.g., algorithmically-defined diagnoses) (Davis, Cullen, *et al.*, 2019; Davies *et al.*, 2022), genetics studies suggest that single-item self-reported diagnoses, as well as self-reported disorders and treatment-seeking, can be useful for increasing

statistical power in GWASs of depression because they allow for the accumulation of enormous samples. One year after the publication of the CONVERGE consortium's GWAS, Hyde et al., (2016) published a GWAS of self-reported data from the direct-to-consumer genetic testing company 23andMe (N_{cases}=75,607, $N_{controls}$=231,747, N_{total}=307,354). Here, cases were individuals of European ancestries who had self-reported a clinical diagnosis of depression, or thinking that they have had depression, in a single-item question. This was a much "shallower" and more heterogeneous phenotype than that studied by the CONVERGE consortium but their large sample size meant that they were able to replicate three genome-wide significant SNPs in a separate sample (Hyde *et al.*, 2016).

A rule learnt from decades of GWASs of polygenic traits, not just mood disorders, is that the number of discovered variants is initially low, but dramatically increases after sample sizes reach an inflection point (Levinson *et al.*, 2014). In 2018, the **Psychiatric Genomics Consortium (PGC)** published a GWAS of major depressive disorder including 135,458 cases and 344,901 controls of European ancestries (N_{total}=480,359) from seven cohorts. This involved international collaboration to achieve a larger sample size than ever before. The seven cohorts used various methods to identify cases, including diagnoses from electronic health records, structured interviews, and self-reported clinical diagnoses. The GWAS identified 44 loci which reached genome-wide significance. Of these, 30 were novel findings and 14 had been genome-wide significant in prior GWASs (Wray *et al.*, 2018).

In recent years, research groups that have collaborated to meta-analyse **GWAS summary statistics** have seen rapid progressions in the ability to identify genome-wide significant loci. In 2019, Howard et al., meta-analysed different measures of depression: a broad phenotype based on treatment seeking for *"nerves, anxiety or depression"* that showed high genetic overlap with clinically defined major depressive disorder (Howard *et al.*, 2018), the Hyde et al., (2016) GWAS, and Wray et al., (2018) GWAS. They discovered 102 genome-wide significant SNPs (82 of which replicated in an independent cohort) with a total sample size of 807,553 (N_{cases}=246,363 and $N_{controls}$=561,190) (Howard *et al.*, 2019). Last year, Levey et al., published an enormous meta-analysed GWAS which included many of the previously published depression studies and data from the Million Veteran Program in the US. With a sample size of 1,154,267 of individuals of European ancestries (N_{cases}=340,591), they replicated 99% of the 211 genome-wide significant variants in an independent data set. Levey at al. (2021) also performed a GWAS of depression in 59,600 individuals of African ancestries (N_{cases}=25,843) but found no genome-wide significant loci.

However, they did find that 61% of the genome-wide significant SNPs from the European GWAS had the same direction of effect (i.e., risk-increasing or risk-decreasing) in the African sample (Levey *et al.*, 2021)

The largest GWAS of bipolar disorder to date was published by the PGC in 2021. Genetic data from a total of 41,917 individuals affected by bipolar disorder (cases) and 371,549 unaffected individuals (controls) from Europe, North America, and Australia were included in the study. This represented an increase of 21,565 cases and 340,191 controls since their previous GWAS (Stahl *et al.*, 2019). These individuals had been recruited into one of 57 research cohorts who contributed data to the study. A total of 52 had contributed individual-level data which was then meta-analysed and five cohorts (iPSYCH30, deCODE genetics, Estonian Biobank, HUNT and UK Biobank) who contributed GWAS summary statistics. This GWAS identified 64 loci which were significantly associated with bipolar disorder (33 of which were new findings) thus reaffirming that bipolar disorder is a polygenic disorder (Mullins *et al.*, 2021). Unlike many of the major GWASs in major depressive disorder, this PGC GWAS did not utilise self-reported data on clinical diagnoses, self-reported disorder or treatment-seeking. All cases were required to meet DSM-IV or ICD-9/10 criteria at some point in their lifetime, made via structured diagnostic instruments from assessments by trained interviewers, clinician-administered checklists, or medical records.

In sum, mood disorders are polygenic. This means that each individual carries a number of risk alleles, and even the genomes of healthy individuals will be hosts to risk-increasing alleles (Gibson, 2012). Individual genetic variants confer a very small effect on overall risk (usually OR<1.1). But, when aggregated into a **polygenic risk score (PRS)**, they explain a much larger proportion of the variance (Smoller *et al.*, 2019). Whether a mood disorder manifests depends on the number of risk alleles a person carries and their unique life experiences. Therefore, inherited genetics cannot fully explain why an individual develops a mood disorder or not. **Figure 1.6** presents a visual aid for understanding how genetic and environmental factors act in tandem to influence liability to a psychiatric disorder. A PRS is based on the results of a GWAS (i.e., the summary statistics) and the variance that it can explain is limited by the SNP-based heritability of the trait being examined. Currently, the variance explained by PRS falls far below SNP-based heritability estimates (Choi, Mak and O'Reilly, 2020). But, the hope is that better powered GWASs and more sophisticated methods will enable PRSs to have far greater predictive power in the future.

2. The joint effect of common variants explains a substantial proportion of phenotypic variance

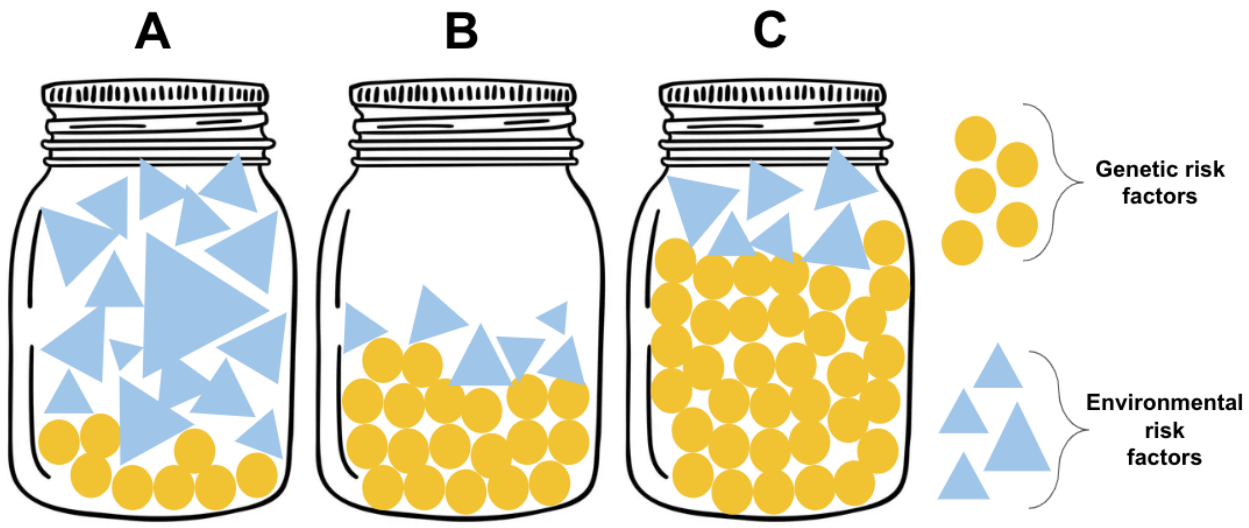
The proportion of phenotypic variance explained by additive genetic variation is known as **narrow-sense heritability (h^2)** (Visscher, Hill and Wray, 2008). One type of narrow-sense heritability is SNP-based heritability (h^2_{SNP}), which can be computed with a software called Linkage Disequilibrium Score Regression (LDSC) and GWAS summary statistics (Bulik-Sullivan *et al.*, 2015). These estimates fall below twin heritability estimates (Sullivan, Neale and Kendler, 2000; Polderman *et al.*, 2015). The most recent GWAS of major depressive disorder from the PGC reported a SNP-based heritability of 8.7% on the **liability scale** (assuming a population prevalence at 15%) (Wray *et al.*, 2018). Both smaller and larger SNP-based heritability have been estimated previously depending on the subtype of depression or the way it was measured (Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.*, 2013; Hyde *et al.*, 2016; Howard *et al.*, 2018; Coleman *et al.*, 2020).

The most recent GWAS of bipolar disorder from the PGC reported a SNP-based heritability of 15.6-18.6% on the liability scale (assuming a population prevalence at 1-2%) (Mullins *et al.*, 2021). These estimates were fractionally lower than those from the previous PGC GWAS of bipolar disorder (17-23% assuming a population prevalence of 0.5-2%) (Stahl *et al.*, 2019). When divided into subtypes, type I had a significantly larger SNP-based heritability than type II (20.9% versus 11.6%) which supported previous findings from family-based research (Song *et al.*, 2018). The two subtypes shared a high genetic correlation which was significantly different from one, thus demonstrating imperfect overlap in their genetic risk factors (Mullins *et al.*, 2021).

Another important point to bear in mind is that genetic risk does not explain all of the variance in mood disorders, as is the case with all complex, polygenic traits and diseases. Environmental exposures also contribute to overall risk (**figure 1.6**). While GWASs have been pivotal in confirming that common genetic variation contributes to the overall heritability of mood disorders, the challenge of closing the "missing heritability gap" remains. The issue of missing heritability has, in the past, been attributed to the assumption that GWASs can discover the effect of only a fraction of associated loci (Gibson, 2012). The hope is that, as sample sizes continue to grow and phenotypes become more refined, we will start to explain a greater percentage of disease risk within the population by common genetic variants. However, balancing these two requirements is challenging.

Figure 1.6. The "mental health jar" analogy developed by Professor Jehannine Austin (Austin, 2020)**.**

Mental health can be thought of as a jar. We are all born with varying levels of polygenic risk (yellow circles). Throughout our lives, we may experience environmental risk factors (blue triangles). These environmental risk factors, and our responses to them, are unique to each of us (demonstrated by the variable size of the blue triangles). Depending on our inherited genetic predisposition, we all require different amounts of environmental adversity to "fill up" our jar and experience an episode of mental illness. Person A has minimal genetic risk, which means they require a large amount of environmental risk before their jar fills up. By contrast, person C has a large amount of genetic risk. This means that they do not need to experience as much environmental adversity to have a full jar. Healthy behaviours such as, such as accessing effective treatment, finding effective ways to manage stress, exercise, and social support can help to make our jar taller and protect us from experiencing mental ill health.



3. Complex traits/diseases share genetic risk factors

An area of interest in psychiatric genetics is whether comorbidity between disorders is partially rooted in overlapping genetic risk factors via **pleiotropy**. Genetic overlap between psychiatric disorders was originally uncovered by family and twin studies (Smoller *et al.*, 2019). The advent of GWASs meant that DNA-level overlap could be researched, and this confirmed that pleiotropy is widespread in the human genome (Visscher and Yang, 2016). Using GWAS summary statistics,

genetic correlations (r_g) can be computed between polygenic traits and diseases in different study samples. A genetic correlation is a quantitative estimate, between -1 and 1, which represents the overlap in their additive genetic basis (based on a reference panel of SNPs) (van Rheenen *et al.*, 2019). This is based on the assumption that the trait is polygenic and pleiotropy is directional (Visscher and Yang, 2016). A positive genetic correlation between trait A and trait B suggests that, on average, the common genetic variants which increase risk for trait A also increase risk for trait B in the population. A negative genetic correlation between trait A and trait B suggests that, on average, the common variants which increase risk for trait A *decrease* risk for trait B (Abdellaoui and Verweij, 2021).

GWASs have demonstrated that the same genetic variants which exert influence on mood disorders also influence other traits, including 1) **psychiatric and neurodevelopmental disorders**, 2) **personality traits**, 3) **lifestyle factors and behaviours**, and 4) **anthropometric traits**. These findings have pointed towards additive genetics as a driver of comorbidity. There are similarities and differences between bipolar disorder and major depressive disorder.

Psychiatric and neurodevelopmental disorders. Mullins et al., (2021) showed that bipolar disorder's additive genetics were most closely related to schizophrenia (r_g=0.69), followed by depression-related phenotypes (major depressive disorder [r_g=0.48], major depression [r_g=0.44] and depressive symptoms [r_g=0.375]). Other significant associations in a positive direction were found with autism spectrum disorder (r_g=0.21) and attention deficit hyperactivity disorder (r_g=0.21). Wray et al., (2018) showed that major depressive disorder's greatest genetic overlap was with, unsurprisingly, depressive symptoms (r_g=0.91-0.98). Other positive genetic correlations included anxiety disorders (r_g=0.80), attention deficit hyperactivity disorder (r_g=0.42), autism spectrum disorders (r_g=0.44), schizophrenia (r_g=0.34), and anorexia nervosa (r_g=0.13).

Personality traits. Both bipolar disorder and major depressive disorder were positively genetically correlated with neuroticism. Major depressive disorder showed an association which was over 3x stronger (r_g= 0.70) than bipolar disorder's association (r_g=0.22).

Lifestyle factors and behaviours. Bipolar disorder was positively correlated with daytime sleepiness (r_g=0.13), sleep duration (r_g=0.12), insomnia (r_g=0.12), and smoking behaviours (r_g=0.13-0.15) (Mullins *et al.*, 2021). Major depressive disorder was similarly correlated with these sleep-related phenotypes (r_g=0.19-0.67) and smoking (r_g=0.29) (Wray *et al.*, 2018).

Anthropometric traits. Major depressive disorder was positively genetically correlated with body fat, body mass index (BMI), and obesity (r_g=0.09-0.20) (Wray *et al.*, 2018) while bipolar disorder showed no significant associations.

A revolutionary finding from the field of psychiatric genetics is that seemingly environmental exposures, such as negative life events, are heritable (Power *et al.*, 2013; Dalvie *et al.*, 2020). GWASs of trauma-related phenotypes have led to the discovery that they share much of their genetic risk with depression-related phenotypes. For instance, self-reported childhood trauma has been shown to share genetic risk with both major depressive disorder (r_g=0.71) and depressive symptoms (r_g=0.70) (Dalvie *et al.*, 2020). In terms of response to trauma, the PTSD working group of the PGC showed that PTSD was most strongly genetically associated with depressive symptoms (r_g=0.80), while major depressive disorder was significantly genetically correlated with PTSD at 0.62 (Nievergelt *et al.*, 2019). Such results add weight to the diathesis-stress model because it suggests that variability in sensitivity to traumatic events may be partly driven by genetic differences (which may form part of someone's diathesis to stressors).

Another intriguing discovery from the literature concerns the relationship between the genetic basis of mood disorders and the genetic basis of intelligence. Bipolar disorder has consistently been proven to have a positive genetic association with years of schooling (also known as educational attainment) but not with intelligence (Stahl *et al.*, 2019; Coleman *et al.*, 2020; Mullins *et al.*, 2021). By comparison, major depressive disorder has significant genetic correlations with both but in a negative direction (Wray *et al.*, 2018).

These findings are evidence for pleiotropic effects across different dimensions of psychopathology, including disorders of the mood spectrum. The next challenge involves using these findings to progress our understanding of how shared genetic risk contributes to the high prevalence of comorbidity in individuals affected by mood disorders.

Glossary of key terms 4:

Candidate gene study: A study which selects a gene based on prior putative evidence that it may be related to the development of the trait/disease of interest. Polymorphisms within the gene are then assessed for associations with the trait by observing its frequency in cases and controls. Its functional mechanisms are then followed up if an association is discovered.

Genotyping: A process used to measure polymorphisms at different loci (i.e., at different SNPs) in the human genome. The first step of genotyping is known as "hybridisation". Here, a genotyping microarray, which contains oligonucleotide probes at specific positions in the human genome where variants/polymorphisms occur, is washed with the DNA. If the genotyped individual has an identical genetic variant to the variant on the genotyping chip, the nucleotide will hybridise to its respective probe. After hybridisation, the probe will signal a fluorescent colour depending on which allele the participant has. These fluorescent probes can then be read by a computer and assigned to a specific SNP in the human genome. The result of genotyping is known as "raw intensity data" based on the intensity of the fluorescent signal of each probe. To identify the participant's genotype at each SNP, the raw data must then go through a process known as "calling" based on the clustering patterns of the fluorescent signal. This is usually done by a computer or (where the identity of the genotype is not clear) manually by a bioinformatician. This process yields genotype data which contains information about whether the participant is homozygous or heterozygous at various genetic variants across the genome (Coleman *et al.*, 2016). An important point is that genotyping is different to sequencing.

SNP-based heritability (h^2_{SNP}): The proportion of phenotypic variance (Vp) in a given trait that is explained by additive genetic variation of SNPs (Yang *et al.*, 2017).

Psychiatric Genomics Consortium (PGC): A collaborative organisation of international researchers aimed at understanding more about the genetic basis for psychiatric disorders. The PGC is divided into working groups for individual disorders, as well as a cross-disorder working group. Currently, the PGC involves over 800 researchers from 36 countries.

GWAS summary statistics: After conducting a GWAS, the results are presented in a tabulated file. Each row corresponds to each SNP. The columns include the effect size (beta or odds ratio), standard error, and p-value of every SNP regressed on the trait/disease of interest. Other columns contain further information about the two possible alleles at that SNP, the allele frequency, the chromosome it is found on, the base pair position, and the sample size of the GWAS (MacArthur *et al.*, 2021).

Polygenic risk score (PRS): A PRS is a weighted sum of risk alleles (0, 1 or 2) that an individual carries for a particular trait. The weights are derived from the SNP effect sizes in a discovery GWAS. PRSs can then be used in a separate sample for association testing or risk prediction (Maher, 2015).

Narrow-sense heritability: The proportion of phenotypic variance explained by additive genetic variation (Visscher, Hill and Wray, 2008).

Liability scale: A linear transformation applied to SNP-based heritability estimates based on an analyst-specified prevalence of the disorder in the population (Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.*, 2013). Conversion to the liability scale is performed in GWASs of binary traits when the ratio of cases:controls in the GWAS sample diverges from the true ratio of affected:unaffected individuals in the population.

Pleiotropy: Pleiotropy is a phenomenon whereby a polymorphism at a position in the DNA sequence has an effect on more than one outcome (i.e., trait or disease) (Visscher and Yang, 2016).

Mood disorders and rare variants

Although common variants in the human genome are the focus of this thesis, it is important to mention emerging evidence of the role that rare genetic variation may play in mood disorders. The effects of rare variants, which are found in <1% of the population, would not be identified in GWASs of mood disorders, since GWASs rely on the common variants on a SNP chip or microarray. One type of rare genetic mutation relevant to psychiatric disorders is known as a copy number variant (CNV). CNVs are segments of a chromosome that are greater than 1kb in length and vary in number from person-to-person in the form of deletions, duplications, insertions, and inversions (Malhotra and Sebat, 2012; Gordovez and McMahon, 2020). CNVs are found in all regions of the genome; some have no phenotypic effects while others increase a person's risk of developing certain disorders. The role of CNVs in some psychiatric and neurodevelopmental disorders, such as schizophrenia, autism spectrum disorders, and intellectual disability, is relatively well documented (Levy *et al.*, 2011; Sanders *et al.*, 2011; Kirov *et al.*, 2014; Marshall *et al.*, 2017; Viñas-Jornet *et al.*, 2018), while their impact on major depressive disorder and bipolar disorder is less well understood.

Nevertheless, there have been some key developments over the last decade in mood disorder research. For example, multiple studies have found that a CNV on chromosome 16p11.2 is more frequently observed in bipolar disorder cases compared to individuals unaffected by bipolar disorder (McCarthy *et al.*, 2009; Malhotra *et al.*, 2011; Green *et al.*, 2016). This CNV is also known to be associated with autism and schizophrenia (McCarthy *et al.*, 2009; Malhotra *et al.*, 2011; Green *et al.*, 2016). Additionally, two studies have pinpointed a CNV on chromosome 3q29 as potentially important (Quintero-Rivera, Sharifi-Hannauer and Martinez-Agosto, 2010; Green *et al.*, 2016). However, one of these was a case study (Quintero-Rivera, Sharifi-Hannauer and Martinez-Agosto, 2010), and the other found an association with bipolar disorder that was not significant (Green *et al.*, 2016), which means conclusions about 3q29 are currently tentative.

With regards to major depressive disorder, two large studies of CNVs were published in 2019 (Kendall *et al.*, 2019; Zhang *et al.*, 2019). Both found that CNVs were more common in individuals affected by major depressive disorder compared to unaffected individuals. Specifically, in Kendall et al., (2019), CNVs on chromosome 1q21.1, 15q11-13 (the CNV associated with an inherited condition called Prader-Willi syndrome), and 16p11.2 (the same CNV mentioned above in relation to bipolar disorder) were associated with major depressive disorder. As whole genome

sequencing becomes increasingly feasible for large-scale studies of mood disorders, we will likely discover more CNVs associated with mood disorders. Furthermore, we will be able to explore the way these rarer mutations interact with polygenic risk and investigate whether high rates of comorbidity in psychiatry can be partly explained by rare variants that affect multiple domains of psychopathology and neurodevelopment.

Measuring the mood spectrum: challenges in genetics research

The overarching aim of GWASs of mood disorders is to clarify their biological causes. Knowledge of this could improve therapeutic options by identifying novel "druggable" targets and/or providing personalised medicine, inform clinical processes through risk prediction, understand more about the epidemiology, and educate affected individuals about the cause(s) of their illness. The way mood disorders are phenotyped has implications for the conclusions drawn from genetic studies about them for two reasons. First, the phenotyping approach dictates both the sample size that can be obtained as well as the trait's heterogeneity, both of which influence statistical power and the likelihood of statistically significant discoveries (**figure 1.5**). Better powered GWASs are crucial for many post-GWAS analyses, such as increasing the predictive power of PRSs. Second, the validity of the phenotype is crucial for ensuring that any discoveries actually reflect true disorder aetiology. Third, definitional variance between GWASs of, supposedly, the same trait could lead to low replication rates which hinders scientific progress.

For ascertaining whether someone is a case or a control for mood disorders, clinical interviews are considered the gold-standard as they yield the most information and of the highest detail and accuracy. But, since GWASs require large samples to study the small effect sizes of common genetic variants, holding interviews with participants is neither cost-effective nor feasible (Davies *et al.*, 2022). An alternative is to collect data on diagnoses through linked medical records, which is becoming increasingly common (e.g., iPSYCH). As a result, in GWASs over the years, sample size has often been prioritised over clinical precision through the use of self-reported data via an online survey. GWASs already mentioned in this chapter demonstrate the utility of single-item self-reported diagnosis, at least in depression (Hyde *et al.*, 2016; Howard *et al.*, 2018, 2019).

The concept of an "inflection point" in sample size for GWASs of psychiatric disorders was introduced in this chapter. This refers to a critical point in sample size where the number of

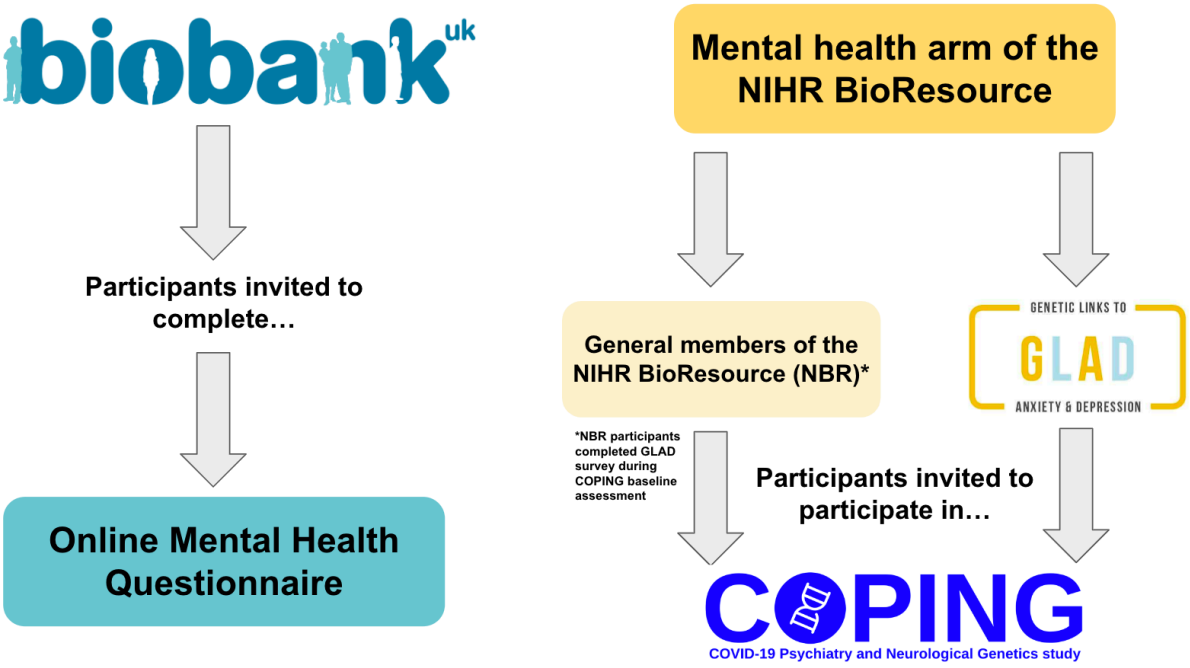
discoveries arising from genetic studies vastly increases (Levinson *et al.*, 2014). The UK Biobank, the Genetic Links to Anxiety and Depression (GLAD) study, and the COVID-19 Psychiatry and Neurological Genetics (COPING) study all offer avenues of ascertaining cases and controls in very large samples for genetic analyses, whether this be on their own or combined with other cohorts (as in the PGC). Beyond simply increasing sample size, discoveries can be accelerated by improving assessment of phenotypes, as reducing phenotypic heterogeneity is another avenue for increasing power (see **figure 1.5**). There is a diversity of phenotyping approaches that can be applied to mood disorders. The richness of mental health data collected by these three UK-based studies therefore represent an opportunity to explore alternative phenotyping strategies.

This thesis incorporates data from the UK Biobank and two studies from the mental health arm of the National Institute of Health and Care Research (NIHR) BioResource in the UK: the GLAD study and the COPING study. Between 2006-2010, the UK Biobank collected data on nearly half a million individuals aged 40-69 with the aim of providing a large research cohort for the study of multiple diseases and their risk factors. In 2016, the UK Biobank invited participants to complete a web-based Mental Health Questionnaire (MHQ) which offered the opportunity to characterise psychiatric disorders and their symptoms in detail. A total of 157,366 participants took part (Davis *et al.*, 2020). The GLAD study began in 2018 with the aim of recruiting 40,000 individuals aged 16+ with lifetime depression and/or an anxiety disorder (Davies *et al.*, 2019). The COPING study involved repeated assessment of mental health over the course of the COVID-19 pandemic (2021-2022) and was available to GLAD participants, as well as participants from the Eating Disorders Genetics Initiative (EDGI), and other members of the NIHR BioResource (NBR) (although this thesis only analyses COPING data from GLAD and NBR participants due to the unavailability of genetic data for EDGI participants). As a baseline assessment in the COPING study, all NBR participants completed the GLAD study survey, which allowed for data from these two independent sets of participants to be combined for analyses.

All three research studies have collected a huge variety of self-reported data on participants' personal experience of mental health. These data include both contemporaneous and lifetime measures, symptoms from screeners and psychometric tools, and DSM-5 diagnostic criteria via the CIDI Short Form (CIDI-SF) (Davies *et al.*, 2019; Davis, Coleman, *et al.*, 2019). These self-reported data can be leveraged to construct phenotypes of disorders and their constituent symptoms. Participants in all cohorts supplied genetic data through blood or saliva samples. Each of these studies used highly similar phenotype data collection and all were genotyped on the

same DNA microarray. For clarity, a diagram is presented below which demonstrates the sources of the data that form the basis of the research conducted in this thesis (**figure 1.7**).

Figure 1.7. United Kingdom (UK)-based studies included in the thesis.
This thesis incorporates self-reported data from the UK Biobank and two studies from the mental health arm of the National Institute of Health and Care Research (NIHR) BioResource in the UK: the GLAD study and the COPING study. The UK Biobank participants were invited to complete the online Mental Health Questionnaire (MHQ) in 2016. The COPING study was conducted over the course of the COVID-19 pandemic, and GLAD participants and general NIHR BioResource (NBR) participants were invited to take part (invites were sent via email between April 2020-January 2021). As a baseline assessment in the COPING study, all NBR participants completed the GLAD study survey, which allowed their self-reported data to be combined for the analyses presented in this thesis.



There is a need to evaluate mood disorder phenotyping approaches to determine whether the resulting phenotypes accurately reflect the mood spectrum and whether their implementation in genetic studies can contribute to our understanding of its underlying biology. **In chapters 2-4**,

three phenotyping approaches are applied to self-reported data in the UK Biobank MHQ, the GLAD, and COPING studies. They are:

1) Diagnostic subtypes
2) Continuous measures
3) Symptom-level analyses (including symptom subgroups individual symptoms)

Aims and structure of the thesis

The aim of the thesis is to explore whether various phenotyping approaches applied to self-reported data can improve our understanding of the mood spectrum's genetic basis. Overall, this thesis asks: how can we study the genetics of the mood spectrum with self-reported data? Are phenotypes derived from self-reported data valid for the mood spectrum? Can they improve our knowledge of the mood spectrum's genetic basis?

Chapter 2: There is high comorbidity (~50%) between major depressive disorder and PTSD (Kessler *et al.*, 1995; Rytwinski *et al.*, 2013). Clarifying the cause of this is challenging due to the complex aetiology of both the disorders. Major depressive disorder and PTSD are heritable and share genetic risk factors to some extent (Koenen *et al.*, 2008; Wolf *et al.*, 2010; Nievergelt *et al.*, 2019). Exposure to trauma is required for a diagnosis of PTSD, and is a strong risk factor for major depressive disorder (Kessler, Davis and Kendler, 1997; Green *et al.*, 2010; Nanni, Uher and Danese, 2012; Otte *et al.*, 2016). Not only this, but the two disorders have a high degree of overlap in terms of their diagnostic criteria. For instance, they are both characterised by the symptoms of anhedonia, sleep disturbances, irritability, and concentration difficulties (American Psychiatric Association, 2013). Previously, comorbidity between major depressive disorder and PTSD has been attributed to overlapping symptoms in those who have been exposed to trauma (Flory and Yehuda, 2015). However, several studies demonstrate that this is not an adequate explanation (Afzali *et al.*, 2017). An alternative source of PTSD-major depressive disorder comorbidity may be a shared genetic basis for trauma. In **chapter 2**, statistical genetics methods are applied to **diagnostic subtypes** of major depressive disorder to test this hypothesis and to further unravel the relationship between major depressive disorder and PTSD.

Chapter 3: Due to the delays in correctly diagnosing bipolar disorder and the high rate of misdiagnosis of bipolar disorder as unipolar depression (Hirschfeld, Lewis and Vornik, 2003), effective screening for possible hypomania/mania is hugely important. Many tools have been developed for this purpose. A well-known and widely used example is the Mood Disorder Questionnaire (MDQ) (Hirschfeld *et al.*, 2000). Based on clinical recommendations, the MDQ was included in the GLAD and COPING surveys as a means of assessing lifetime experience of thirteen manic symptoms. The hope was that the MDQ could provide an avenue for answering a number of research questions regarding the genetic basis of hypomania/mania in bipolar disorder. In **chapter 3**, the validity of this screening tool for genetic studies is tested. The participants' answers to questions in the MDQ are used to construct a **continuous measure** of hypomanic/manic symptoms. In addition to this, the latent factor structure of the MDQ is examined to identify **symptom subgroups**. GWASs of the **continuous measure** and **subgroups** are performed, and genetic correlations are calculated with bipolar disorder overall, type I, and type II.

Chapter 4: Major depressive disorder is the most common psychiatric disorder worldwide and, as discussed in this introductory chapter, is one of the leading causes of disability (GBD 2019 Mental Disorders Collaborators, 2022). Thus, developing effective treatments is crucial for bettering the lives of hundreds of millions of people around the world. There are numerous antidepressant medications available, and while they prove efficacious in some individuals (Cipriani et al., 2018), a sizable proportion fail to show improvement in symptoms after their second course of antidepressants (of sufficient dose and duration). These individuals have a **diagnostic subtype** of major depressive disorder known as treatment-resistant depression (Fava, 2003). Two studies have proposed anhedonia as a risk factor for poor response to antidepressants (Uher et al., 2012; McMakin et al., 2012). Anhedonia is a core symptom of major depressive disorder, but is also present transdiagnostically. Only two large, well-powered GWASs of anhedonia have been published to date, which confirmed that anhedonia has a heritable basis. But, these GWASs came with limitations which hampered our understanding of the genetic basis for this **specific symptom** of major depressive disorder (Ward *et al.*, 2019; Thorp *et al.*, 2020). To address some of these limitations, in **chapter 4**, a **continuous measure** of anhedonic symptoms is constructed from COPING study data. We perform a GWAS of this phenotype and compute genetic correlations with two measures of treatment-resistant depression and two measures of antidepressant response, and other external traits.

Chapter 5: The thesis concludes with a general discussion of lessons learnt from the three empirical studies, limitations of the study designs, and directions for future research of the genetic basis of the mood spectrum.

References

Abdellaoui, A. and Verweij, K.J.H. (2021) 'Dissecting polygenic signals from genome-wide association studies on human behaviour', *Nature human behaviour* [Preprint]. Available at: https://doi.org/10.1038/s41562-021-01110-y.

Afzali, M.H. *et al.* (2017) 'A network approach to the comorbidity between posttraumatic stress disorder and major depressive disorder: The role of overlapping symptoms', *Journal of affective disorders*, 208, pp. 490–496.

Albert, P.R. (2015) 'Why is depression more prevalent in women?', *Journal of psychiatry & neuroscience: JPN*, 40(4), pp. 219–221.

Alonso, J. *et al.* (2004) 'Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project', *Acta psychiatrica Scandinavica. Supplementum*, (420), pp. 21–27.

Altman, E.G. *et al.* (1997) 'The Altman Self-Rating Mania Scale', *Biological psychiatry*, 42(10), pp. 948–955.

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

Andrade, L. *et al.* (2003) 'The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys', *International journal of methods in psychiatric research*, 12(1), pp. 3–21.

Angst, F. *et al.* (2002) 'Mortality of patients with mood disorders: follow-up over 34-38 years', *Journal of affective disorders*, 68(2-3), pp. 167–181.

Angst, J. and Marneros, A. (2001) 'Bipolarity from ancient to modern times: conception, birth and rebirth', *Journal of affective disorders*, 67(1-3), pp. 3–19.

Arnold, L.M. (2003) 'Gender differences in bipolar disorder', *The Psychiatric clinics of North America*, 26(3), pp. 595–620.

Austin, J.C. (2020) 'Evidence-Based Genetic Counseling for Psychiatric Disorders: A Road Map', *Cold Spring Harbor perspectives in medicine*, 10(6). Available at: https://doi.org/10.1101/cshperspect.a036608.

Baldessarini, R.J. *et al.* (2010) 'Onset-age of bipolar disorders at six international sites', *Journal of affective disorders*, 121(1-2), pp. 143–146.

Baldessarini, R.J., Vázquez, G.H. and Tondo, L. (2020) 'Bipolar depression: a major unsolved challenge', *International journal of bipolar disorders*, 8(1), p. 1.

Bandelow, B. and Michaelis, S. (2015) ‘Epidemiology of anxiety disorders in the 21st century’, *Dialogues in clinical neuroscience*, 17(3), pp. 327–335.

Barton, N.H., Etheridge, A.M. and Véber, A. (2017) ‘The infinitesimal model: Definition, derivation, and implications’, *Theoretical population biology*, 118, pp. 50–73.

Baryshnikov, I. *et al.* (2020) ‘Diagnostic conversion from unipolar depression to bipolar disorder, schizophrenia, or schizoaffective disorder: A nationwide prospective 15-year register study on 43 495 inpatients’, *Bipolar disorders*, 22(6), pp. 582–592.

Beck, A.T. *et al.* (1961) ‘An inventory for measuring depression’, *Archives of general psychiatry*, 4, pp. 561–571.

Blanchard, E.B. *et al.* (1998) ‘Posttraumatic stress disorder and comorbid major depression: is the correlation an illusion?’, *Journal of anxiety disorders*, 12(1), pp. 21–37.

Blazer, D.G. *et al.* (1994) ‘The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey’, *The American journal of psychiatry*, 151(7), pp. 979–986.

Bodmer, W. *et al.* (2021) ‘The outstanding scientist, R.A. Fisher: his views on eugenics and race’, *Heredity*, 126(4), pp. 565–576.

Bond, D.J. *et al.* (2008) ‘Antidepressant-associated mood elevations in bipolar II disorder compared with bipolar I disorder and major depressive disorder: a systematic review and meta-analysis’, *The Journal of clinical psychiatry*, 69(10), pp. 1589–1601.

Border, R. *et al.* (2019) ‘No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples’, *The American journal of psychiatry*, 176(5), pp. 376–387.

Borrell-Carrió, F., Suchman, A.L. and Epstein, R.M. (2004) ‘The biopsychosocial model 25 years later: principles, practice, and scientific inquiry’, *Annals of family medicine*, 2(6), pp. 576–582.

Breslau, N. *et al.* (1997) ‘Psychiatric sequelae of posttraumatic stress disorder in women’, *Archives of general psychiatry*, 54(1), pp. 81–87.

Bromet, E. *et al.* (2011) ‘Cross-national epidemiology of DSM-IV major depressive episode’, *BMC medicine*, 9, p. 90.

Bulik-Sullivan, B. *et al.* (2015) ‘An atlas of genetic correlations across human diseases and traits’, *Nature genetics*, 47(11), pp. 1236–1241.

Burcusa, S.L. and Iacono, W.G. (2007) ‘Risk for recurrence in depression’, *Clinical psychology review*, 27(8), pp. 959–985.

Burton, R. (1989) *The Anatomy of Melancholy*. Clarendon Press.

Cai, N., Kendler, K.S. and Flint, J. (2018) ‘Minimal phenotyping yields GWAS hits of low specificity for major depression’, *bioRxiv*. Available at: https://doi.org/10.1101/440735.

Campbell, D.G. *et al.* (2007) ‘Prevalence of depression-PTSD comorbidity: implications for

clinical practice guidelines and primary care-based interventions', *Journal of general internal medicine*, 22(6), pp. 711–718.

Carvalho, A.F. *et al.* (2014) 'Rapid cycling in bipolar disorder: a systematic review', *The Journal of clinical psychiatry*, 75(6), pp. e578–86.

Cassano, G.B. *et al.* (2004) 'The mood spectrum in unipolar and bipolar disorder: arguments for a unitary approach', *The American journal of psychiatry*, 161(7), pp. 1264–1269.

Chen, Z. *et al.* (2021) 'Revisiting the genome-wide significance threshold for common variant GWAS', *G3* , 11(2). Available at: https://doi.org/10.1093/g3journal/jkaa056.

Chesney, E., Goodwin, G.M. and Fazel, S. (2014) 'Risks of all-cause and suicide mortality in mental disorders: a meta-review', *World psychiatry: official journal of the World Psychiatric Association* , 13(2), pp. 153–160.

Choi, S.W., Mak, T.S.-H. and O'Reilly, P.F. (2020) 'Tutorial: a guide to performing polygenic risk score analyses', *Nature protocols*, 15(9), pp. 2759–2772.

Clark, L.A. *et al.* (2017) 'Three Approaches to Understanding and Classifying Mental Disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC)', *Psychological science in the public interest: a journal of the American Psychological Society*, 18(2), pp. 72–145.

Coleman, J.R.I. *et al.* (2016) 'Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray', *Briefings in functional genomics*, 15(4), pp. 298–304.

Coleman, J.R.I., Peyrot, W.J., *et al.* (2020) 'Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank', *Molecular psychiatry* [Preprint]. Available at: https://doi.org/10.1038/s41380-019-0546-6.

Coleman, J.R.I., Gaspar, H.A., *et al.* (2020) 'The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls', *Biological psychiatry*, 88(2), pp. 169–184.

Conley, R.R. *et al.* (2007) 'The burden of depressive symptoms in the long-term treatment of patients with schizophrenia', *Schizophrenia research*, 90(1-3), pp. 186–197.

CONVERGE consortium (2015) 'Sparse whole-genome sequencing identifies two loci for major depressive disorder', *Nature*, 523(7562), pp. 588–591.

Coombes, B.J. *et al.* (2020) 'Dissecting clinical heterogeneity of bipolar disorder using multiple polygenic risk scores', *Translational psychiatry*, 10(1), p. 314.

Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.* (2013) 'Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs', *Nature genetics*, 45(9), pp. 984–994.

Crump, C. *et al.* (2013) 'Comorbidities and mortality in bipolar disorder: a Swedish national cohort study', *JAMA psychiatry* , 70(9), pp. 931–939.

Dahl, S.K. *et al.* (2017) 'Early adversity and risk for moderate to severe unipolar depressive

disorder in adolescence and adulthood: A register-based study of 978,647 individuals', *Journal of affective disorders*, 214, pp. 122–129.

Dalvie, S. *et al.* (2020) 'Genomic influences on self-reported childhood maltreatment', *Translational psychiatry*, 10(1), p. 38.

Daly, I. (1997) 'Mania', *The Lancet*, 349(9059), pp. 1157–1160.

Davies, M.R. *et al.* (2019) 'The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety', *Behaviour research and therapy*, 123, p. 103503.

Davies, M.R. *et al.* (2022) 'Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts', *Journal of anxiety disorders*, 85, p. 102491.

Davis, K.A.S., Cullen, B., *et al.* (2019) 'Indicators of mental disorders in UK Biobank-A comparison of approaches', *International journal of methods in psychiatric research*, 28(3), p. e1796.

Davis, K.A.S., Coleman, J.R.I., *et al.* (2019) 'Mental Health in UK Biobank Revised', *Epidemiology*. medRxiv. Available at: https://doi.org/10.1101/19001214.

Davis, K.A.S. *et al.* (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), p. e18.

De Aquino, J.P. and Ross, D.A. (2017) 'Kraepelin's Crumbling Twin Pillars: Using Biology to Reconstruct Psychiatric Nosology From the Bottom Up', *Biological psychiatry*, pp. e71–e74.

Dell'Osso, B., Cafaro, R. and Ketter, T.A. (2021) 'Has Bipolar Disorder become a predominantly female gender related condition? Analysis of recently published large sample studies', *International journal of bipolar disorders*, 9(1), p. 3.

Depue, R.A. *et al.* (1981) 'A behavioral paradigm for identifying persons at risk for bipolar depressive disorder: a conceptual framework and five validation studies', *Journal of abnormal psychology*, 90(5), pp. 381–437.

Drancourt, N. *et al.* (2013) 'Duration of untreated bipolar disorder: missed opportunities on the long road to optimal treatment', *Acta psychiatrica Scandinavica*, 127(2), pp. 136–144.

Dunner, D.L. (2003) 'Clinical consequences of under-recognized bipolar spectrum disorder', *Bipolar disorders*, 5(6), pp. 456–463.

Eaton, W.W. *et al.* (1997) 'Natural history of Diagnostic Interview Schedule/DSM-IV major depression. The Baltimore Epidemiologic Catchment Area follow-up', *Archives of general psychiatry*, 54(11), pp. 993–999.

Ebert, A. and Bär, K.-J. (2010) 'Emil Kraepelin: A pioneer of scientific understanding of psychiatry and psychopharmacology', *Indian journal of psychiatry*, 52(2), pp. 191–192.

Edvardsen, J. *et al.* (2008) 'Heritability of bipolar spectrum disorders. Unity or heterogeneity?', *Journal of affective disorders*, 106(3), pp. 229–240.

Endicott, J. and Spitzer, R.L. (1978) ‘A Diagnostic Interview: The Schedule for Affective Disorders and Schizophrenia’, *Archives of general psychiatry*, 35(7), pp. 837–844.

Fava, M. (2003) ‘Diagnosis and definition of treatment-resistant depression’, *Biological psychiatry*, 53(8), pp. 649–659.

Fekadu, A. *et al.* (2009) ‘A multidimensional tool to quantify treatment resistance in depression: the Maudsley staging method’, *The Journal of clinical psychiatry*, 70(2), pp. 177–184.

Fekadu, A., Donocik, J.G. and Cleare, A.J. (2018) ‘Standardisation framework for the Maudsley staging method for treatment resistance in depression’, *BMC psychiatry*, 18(1), p. 100.

Fiedorowicz, J.G. *et al.* (2011) ‘Subthreshold hypomanic symptoms in progression from unipolar major depression to bipolar disorder’, *The American journal of psychiatry*, 168(1), pp. 40–48.

First *et al.* (2015) ‘Structured clinical interview for DSM-5–Research version (SCID-5-RV, Version 1.0. 0)’, *Arlington: American Psychiatric* [Preprint].

Fisher, R.A. (1919) ‘XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance’, *Earth and environmental science transactions of the Royal Society of Edinburgh*, 52(2), pp. 399–433.

Flory, J.D. and Yehuda, R. (2015) ‘Comorbidity between post-traumatic stress disorder and major depressive disorder: alternative explanations and treatment considerations’, *Dialogues in clinical neuroscience*, 17(2), pp. 141–150.

Fountoulakis, K.N. (2012) ‘Refractoriness in bipolar disorder: definitions and evidence-based treatment’, *CNS neuroscience & therapeutics*, 18(3), pp. 227–237.

Frank, E. *et al.* (1991) ‘Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence’, *Archives of general psychiatry*, 48(9), pp. 851–855.

Friborg, O. *et al.* (2014) ‘Comorbidity of personality disorders in mood disorders: a meta-analytic review of 122 studies from 1988 to 2010’, *Journal of affective disorders*, 152-154, pp. 1–11.

Fried, E.I. and Nesse, R.M. (2015) ‘Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study’, *Journal of affective disorders*, 172, pp. 96–102.

Galton, F. (1897) ‘The average contribution of each several ancestor to the total heritage of the offspring’, *Proceedings of the Royal Society of London*, 61(369-377), pp. 401–413.

GBD 2019 Mental Disorders Collaborators (2022) ‘Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019’, *The lancet. Psychiatry*, 9(2), pp. 137–150.

Ghaemi, S.N., Lenox, M.S. and Baldessarini, R.J. (2001) ‘Effectiveness and safety of long-term antidepressant treatment in bipolar disorder’, *The Journal of clinical psychiatry*, 62(7), pp. 565–569.

Gibson, G. (2012) ‘Rare and common variants: twenty arguments’, *Nature reviews. Genetics*, 13(2), pp. 135–145.

Gitlin, M. (2006) ‘Treatment-resistant bipolar disorder’, *Molecular psychiatry*, 11(3), pp. 227–240.

Gold, S.M. *et al.* (2020) ‘Comorbid depression in medical diseases’, *Nature reviews. Disease primers*, 6(1), p. 69.

Goldstein, B.I. *et al.* (2015) ‘Major Depressive Disorder and Bipolar Disorder Predispose Youth to Accelerated Atherosclerosis and Early Cardiovascular Disease: A Scientific Statement From the American Heart Association’, *Circulation*, 132(10), pp. 965–986.

Goldstein, D.B. (2009) ‘Common genetic variation and human traits’, *The New England journal of medicine*, 360(17), pp. 1696–1698.

Gordovez, F.J.A. and McMahon, F.J. (2020) ‘The genetics of bipolar disorder’, *Molecular psychiatry*, 25(3), pp. 544–559.

Goswami, C., Chattopadhyay, A. and Chuang, E.Y. (2021) ‘Rare variants: data types and analysis strategies’, *Annals of translational medicine*, 9(12), p. 961.

de Graaf, R. *et al.* (2002) ‘Risk factors for 12-month comorbidity of mood, anxiety, and substance use disorders: findings from the Netherlands Mental Health Survey and Incidence Study’, *The American journal of psychiatry*, 159(4), pp. 620–629.

Green, E.K. *et al.* (2016) ‘Copy number variation in bipolar disorder’, *Molecular psychiatry*, 21(1), pp. 89–93.

Green, J.G. *et al.* (2010) ‘Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: associations with first onset of DSM-IV disorders’, *Archives of general psychiatry*, 67(2), pp. 113–123.

Haga, S.B. (2010) ‘Impact of limited population diversity of genome-wide association studies’, *Genetics in medicine: official journal of the American College of Medical Genetics*, 12(2), pp. 81–84.

Hamilton, M. (1960) ‘A rating scale for depression’, *Journal of neurology, neurosurgery, and psychiatry*, 23, pp. 56–62.

Harden, K.P. and Klump, K.L. (2015) ‘Introduction to the special issue on gene-hormone interplay’, *Behavior genetics*, 45(3), pp. 263–267.

Hardeveld, F. *et al.* (2010) ‘Prevalence and predictors of recurrence of major depressive disorder in the adult population’, *Acta psychiatrica Scandinavica*, 122(3), pp. 184–191.

Haro, J.M. *et al.* (2006) ‘Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys’, *International journal of methods in psychiatric research*, 15(4), pp. 167–180.

Heston, L.L. (1966) ‘Psychiatric disorders in foster home reared children of schizophrenic mothers’, *The British journal of psychiatry: the journal of mental science*, 112(489), pp. 819–825.

Hirschfeld, R.M. *et al.* (2000) ‘Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire’, *The American journal of psychiatry*,

157(11), pp. 1873–1875.

Hirschfeld, R.M.A. (2001) ‘The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care’, *Primary care companion to the Journal of clinical psychiatry*, 3(6), pp. 244–254.

Hirschfeld, R.M.A., Lewis, L. and Vornik, L.A. (2003) ‘Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder’, *The Journal of clinical psychiatry*, 64(2), pp. 161–174.

Hoch, A. and MacCurdy, J.T. (1922) ‘THE PROGNOSIS OF INVOLUTION MELANCHOLIA’, *Archives of Neurology & Psychiatry*, 7(1), pp. 1–37.

Hofker, M.H., Fu, J. and Wijmenga, C. (2014) ‘The genome revolution and its role in understanding complex diseases’, *Biochimica et biophysica acta*, 1842(10), pp. 1889–1895.

Horwitz, T. *et al.* (2019) ‘A decade in psychiatric GWAS research’, *Molecular psychiatry*, 24(3), pp. 378–389.

Howard, D.M. *et al.* (2018) ‘Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways’, *Nature communications*, 9(1), p. 1470.

Howard, D.M. *et al.* (2019) ‘Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions’, *Nature neuroscience*, 22(3), pp. 343–352.

Huxley, N. and Baldessarini, R.J. (2007) ‘Disability and its treatment in bipolar disorder patients’, *Bipolar disorders*, 9(1-2), pp. 183–196.

Hyde, C.L. *et al.* (2016) ‘Identification of 15 genetic loci associated with risk of major depression in individuals of European descent’, *Nature genetics*, 48(9), pp. 1031–1036.

International HapMap Consortium (2005) ‘A haplotype map of the human genome’, *Nature*, 437(7063), pp. 1299–1320.

Jablensky, A. (2016) ‘Psychiatric classifications: validity and utility’, *World psychiatry: official journal of the World Psychiatric Association* , 15(1), pp. 26–31.

James, A. *et al.* (2015) ‘Conversion from depression to bipolar disorder in a cohort of young people in England, 1999-2011: A national record linkage study’, *Journal of affective disorders*, 185, pp. 123–128.

Jansson, Å. (2011) ‘Mood disorders and the brain: depression, melancholia, and the historiography of psychiatry’, *Medical history*, 55(3), pp. 393–399.

Johnson, S.L. and Kizer, A. (2002) ‘Bipolar and unipolar depression: A comparison of clinical phenomenology and psychosocial predictors’, in I.H. Gotlib (ed.) *Handbook of depression , (pp*. New York, NY, US: The Guilford Press, xiii, pp. 141–165.

Jones, I. and Craddock, N. (2001) ‘Candidate gene studies of bipolar disorder’, *Annals of medicine*, 33(4), pp. 248–256.

Jouanna, J. and Allies, N. (2012) ‘THE THEORY OF THE FOUR HUMOURS’, in P. van der Eijk (ed.) *Greek Medicine from Hippocrates to Galen*. Brill (Selected Papers), pp. 335–360.

Judd, L.L. *et al.* (2000) ‘Does incomplete recovery from first lifetime major depressive episode herald a chronic course of illness?’, *The American journal of psychiatry*, 157(9), pp. 1501–1504.

‘Karl Ludwig Kahlbaum, M.D. 1828–1899’ (1999) *American Journal of Psychiatry*, 156(7), pp. 989–989.

Keck, P.E., Jr *et al.* (1998) ‘12-month outcome of patients with bipolar disorder following hospitalization for a manic or mixed episode’, *The American journal of psychiatry*, 155(5), pp. 646–652.

Keller, M.B. *et al.* (1992) ‘Time to recovery, chronicity, and levels of psychopathology in major depression. A 5-year prospective follow-up of 431 subjects’, *Archives of general psychiatry*, 49(10), pp. 809–816.

Kemp, D.E. *et al.* (2010) ‘Medical comorbidity in bipolar disorder: relationship between illnesses of the endocrine/metabolic system and treatment outcome’, *Bipolar disorders*, 12(4), pp. 404–413.

Kendall, K.M. *et al.* (2019) ‘Association of Rare Copy Number Variants With Risk of Depression’, *JAMA psychiatry* , 76(8), pp. 818–825.

Kendler, K.S. and Zerbin-Rüdin, E. (1996) ‘Abstract and review of “Studien Uber Vererbung und Entstehung Geistiger Störungen. I. Zur Vererbung und Neuentstehung der Dementia praecox.” (Studies on the inheritance and origin of mental illness: I. To the problem of the inheritance and primary origin of dementia praecox). 1916’, *American journal of medical genetics*, 67(4), pp. 338–342.

Kessing, L.V. *et al.* (2017) ‘Rate and predictors of conversion from unipolar to bipolar disorder: A systematic review and meta-analysis’, *Bipolar disorders*, 19(5), pp. 324–335.

Kessler, R.C. *et al.* (1995) ‘Posttraumatic stress disorder in the National Comorbidity Survey’, *Archives of general psychiatry*, 52(12), pp. 1048–1060.

Kessler, R.C. *et al.* (2003) ‘The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R)’, *JAMA: the journal of the American Medical Association*, 289(23), pp. 3095–3105.

Kessler, R.C. *et al.* (2005) ‘Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication’, *Archives of general psychiatry*, 62(6), pp. 593–602.

Kessler, R.C. *et al.* (2015) ‘Anxious and non-anxious major depressive disorder in the World Health Organization World Mental Health Surveys’, *Epidemiology and psychiatric sciences*, 24(3), pp. 210–226.

Kessler, R.C. and Bromet, E.J. (2013) ‘The epidemiology of depression across cultures’, *Annual review of public health*, 34, pp. 119–138.

Kessler, R.C., Davis, C.G. and Kendler, K.S. (1997) ‘Childhood adversity and adult psychiatric disorder in the US National Comorbidity Survey’, *Psychological medicine*, 27(5), pp. 1101–1119.

Kessler, R.C., Merikangas, K.R. and Wang, P.S. (2007) 'Prevalence, comorbidity, and service utilization for mood disorders in the United States at the beginning of the twenty-first century', *Annual review of clinical psychology*, 3, pp. 137–158.

Kirov, G. *et al.* (2014) 'The penetrance of copy number variations for schizophrenia and developmental delay', *Biological psychiatry*, 75(5), pp. 378–385.

Klein, R.J. *et al.* (2005) 'Complement factor H polymorphism in age-related macular degeneration', *Science*, 308(5720), pp. 385–389.

Koenen, K.C. *et al.* (2008) 'Common genetic liability to major depression and posttraumatic stress disorder in men', *Journal of affective disorders*, 105(1-3), pp. 109–115.

Kogan, J.N. *et al.* (2004) 'Demographic and diagnostic characteristics of the first 1000 patients enrolled in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD)', *Bipolar disorders*, 6(6), pp. 460–469.

Kotsopoulos, S. (1986) 'Aretaeus the Cappadocian on mental illness', *Comprehensive psychiatry*, 27(2), pp. 171–179.

Kroenke, K., Spitzer, R.L. and Williams, J.B. (2001) 'The PHQ-9: validity of a brief depression severity measure', *Journal of general internal medicine*, 16(9), pp. 606–613.

Kruglyak, L. and Nickerson, D.A. (2001) 'Variation is the spice of life', *Nature genetics*, 27(3), pp. 234–236.

Kuehner, C. (2003) 'Gender differences in unipolar depression: an update of epidemiological findings and possible explanations', *Acta psychiatrica Scandinavica*, 108(3), pp. 163–174.

Kuehner, C. (2017) 'Why is depression more common among women than among men?', *The lancet. Psychiatry*, 4(2), pp. 146–158.

Kupfer, D.J. *et al.* (2002) 'Demographic and clinical characteristics of individuals in a bipolar disorder case registry', *The Journal of clinical psychiatry*, 63(2), pp. 120–125.

LaFramboise, T. (2009) 'Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances', *Nucleic acids research*, 37(13), pp. 4181–4193.

Lai, H.M.X. *et al.* (2015) 'Prevalence of comorbid substance use, anxiety and mood disorders in epidemiological surveys, 1990-2014: A systematic review and meta-analysis', *Drug and alcohol dependence*, 154, pp. 1–13.

Lamers, F. *et al.* (2011) 'Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA)', *The Journal of clinical psychiatry*, 72(3), pp. 341–348.

Levey, D.F. *et al.* (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), pp. 954–963.

Levy, D. *et al.* (2011) 'Rare de novo and transmitted copy-number variation in autistic spectrum disorders', *Neuron*, 70(5), pp. 886–897.

Levinson, D.F. *et al.* (2014) 'Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it?', *Biological psychiatry*, 76(7), pp. 510–512.

Lin, M. *et al.* (2017) 'Effects of short indels on protein structure and function in human genomes', *Scientific reports*, 7(1), p. 9313.

Lish, J.D. *et al.* (1994) 'The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members', *Journal of affective disorders*, 31(4), pp. 281–294.

Luxenburger, H. (1928) 'Vorläufiger Bericht über psychiatrische Serienuntersuchungen an Zwillingen', *Archiv fur Psychiatrie und Nervenkrankheiten, vereinigt mit Zeitschrift fur die gesamte Neurologie und Psychiatrie*, 116(1), pp. 297–326.

MacArthur, J.A.L. *et al.* (2021) 'Workshop proceedings: GWAS summary statistics standards and sharing', *Cell genomics*, 1(1). Available at: https://doi.org/10.1016/j.xgen.2021.100004.

Maher, B. (2008) *Personal genomes: The case of the missing heritability*, *Nature Publishing Group UK*. Available at: https://doi.org/10.1038/456018a.

Maher, B.S. (2015) 'Polygenic Scores in Epidemiology: Risk Prediction, Etiology, and Clinical Utility', *Current epidemiology reports*, 2(4), pp. 239–244.

Malhi, G.S. and Mann, J.J. (2018) 'Depression', *The Lancet*, 392(10161), pp. 2299–2312.

Malhotra, D. *et al.* (2011) 'High frequencies of de novo CNVs in bipolar disorder and schizophrenia', *Neuron*, 72(6), pp. 951–963.

Malhotra, D. and Sebat, J. (2012) 'CNVs: harbingers of a rare variant revolution in psychiatric genetics', *Cell*, 148(6), pp. 1223–1241.

Mantere, O. *et al.* (2004) 'The clinical characteristics of DSM-IV bipolar I and II disorders: baseline findings from the Jorvi Bipolar Study (JoBS)', *Bipolar disorders*, 6(5), pp. 395–405.

Marshall, C.R. *et al.* (2017) 'Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects', *Nature genetics*, 49(1), pp. 27–35.

Martin, A.R. *et al.* (2019) 'Clinical use of current polygenic risk scores may exacerbate health disparities', *Nature genetics*, 51(4), pp. 584–591.

Maughan, B., Collishaw, S. and Stringaris, A. (2013) 'Depression in childhood and adolescence', *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent*, 22(1), pp. 35–40.

McCarthy, S.E. *et al.* (2009) 'Microduplications of 16p11.2 are associated with schizophrenia', *Nature genetics*, 41(11), pp. 1223–1227.

Merikangas, K.R. *et al.* (2007) 'Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication', *Archives of general psychiatry*, 64(5), pp. 543–552.

Merikangas, K.R. *et al.* (2011) 'Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative', *Archives of general psychiatry*, 68(3), pp. 241–251.

Miller, C.J., Johnson, S.L. and Eisner, L. (2009) 'Assessment Tools for Adult Bipolar Disorder', *Clinical psychology: a publication of the Division of Clinical Psychology of the American Psychological Association*, 16(2), pp. 188–201.

Miller, J.N. and Black, D.W. (2020) 'Bipolar Disorder and Suicide: a Review', *Current psychiatry reports*, 22(2), p. 6.

Mondimore, F.M. (2005) 'Kraepelin and manic-depressive insanity: an historical perspective', *International review of psychiatry* , 17(1), pp. 49–52.

Monroe, S.M. and Simons, A.D. (1991) 'Diathesis-stress theories in the context of life stress research: implications for the depressive disorders', *Psychological bulletin*, 110(3), pp. 406–425.

Montgomery, S.A. and Asberg, M. (1979) 'A new depression scale designed to be sensitive to change', *The British journal of psychiatry: the journal of mental science*, 134, pp. 382–389.

Moussavi, S. *et al.* (2007) 'Depression, chronic diseases, and decrements in health: results from the World Health Surveys', *The Lancet*, 370(9590), pp. 851–858.

Mullaney, J.M. *et al.* (2010) 'Small insertions and deletions (INDELs) in human genomes', *Human molecular genetics*, 19(R2), pp. R131–6.

Mullins, N. *et al.* (2021) 'Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology', *Nature genetics*, 53(6), pp. 817–829.

Muneer, A. (2017) 'Mixed States in Bipolar Disorder: Etiology, Pathogenesis and Treatment', *Chonnam medical journal*, 53(1), pp. 1–13.

Musliner, K.L. and Østergaard, S.D. (2018) 'Patterns and predictors of conversion to bipolar disorder in 91 587 individuals diagnosed with unipolar depression', *Acta psychiatrica Scandinavica*, 137(5), pp. 422–432.

Nanni, V., Uher, R. and Danese, A. (2012) 'Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis', *The American journal of psychiatry*, 169(2), pp. 141–151.

National Human Genome Research Institute (2022), *Pedigree*, link: https://www.genome.gov/genetics-glossary/Pedigree

Nievergelt, C.M. *et al.* (2019) 'International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci', *Nature communications*, 10(1), p. 4558.

Nordentoft, M. *et al.* (2013) 'Excess mortality, causes of death and life expectancy in 270,770 patients with recent onset of mental disorders in Denmark, Finland and Sweden', *PloS one*, 8(1), p. e55176.

Norkeviciene, A. *et al.* (2022) 'A Systematic Review of Candidate Genes for Major Depression', *Medicina* , 58(2). Available at: https://doi.org/10.3390/medicina58020285.

Olfson, M. *et al.* (1997) 'Mental disorders and disability among patients in a primary care group practice', *The American journal of psychiatry*, 154(12), pp. 1734–1740.

Oraki Kohshour, M. *et al.* (2022) ‘Genomic and neuroimaging approaches to bipolar disorder’, *BJPsych open*, 8(2), p. e36.

Ormel, J. *et al.* (1993) ‘Outcome of depression and anxiety in primary care. A three-wave 3 1/2-year study of psychopathology and disability’, *Archives of general psychiatry*, 50(10), pp. 759–766.

Osby, U. *et al.* (2001) ‘Excess mortality in bipolar and unipolar disorder in Sweden’, *Archives of general psychiatry*, 58(9), pp. 844–850.

Ostergaard, S.D., Jensen, S.O.W. and Bech, P. (2011) ‘The heterogeneity of the depressive syndrome: when numbers get serious’, *Acta psychiatrica Scandinavica*, 124(6), pp. 495–496.

Otte, C. *et al.* (2016) ‘Major depressive disorder’, *Nature reviews. Disease primers*, 2, p. 16065.

Palk, A.C. *et al.* (2019) ‘Potential use of clinical polygenic risk scores in psychiatry - ethical implications and communicating high polygenic risk’, *Philosophy, ethics, and humanities in medicine: PEHM*, 14(1), p. 4.

Park, S.-C. *et al.* (2014) ‘Does age at onset of first major depressive episode indicate the subtype of major depressive disorder?: the clinical research center for depression study’, *Yonsei medical journal*, 55(6), pp. 1712–1720.

Pavlova, B. *et al.* (2015) ‘Lifetime prevalence of anxiety disorders in people with bipolar disorder: a systematic review and meta-analysis’, *The Lancet Psychiatry*, 2(8), pp. 710–717.

Paykel, E.S. (2008) ‘Basic concepts of depression’, *Dialogues in clinical neuroscience*, 10(3), pp. 279–289.

Penninx, B.W.J.H. *et al.* (2011) ‘Two-year course of depressive and anxiety disorders: results from the Netherlands Study of Depression and Anxiety (NESDA)’, *Journal of affective disorders*, 133(1-2), pp. 76–85.

Perlis, R.H. *et al.* (2006) ‘Predictors of recurrence in bipolar disorder: primary outcomes from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD)’, *The American journal of psychiatry*, 163(2), pp. 217–224.

Peterson, R.E. *et al.* (2019) ‘Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations’, *Cell*, 179(3), pp. 589–603.

Plana-Ripoll, O. *et al.* (2019) ‘Exploring Comorbidity Within Mental Disorders Among a Danish National Population’, *JAMA psychiatry* , 76(3), pp. 259–270.

Plana-Ripoll, O. *et al.* (2020) ‘Nature and prevalence of combinations of mental disorders and their association with excess mortality in a population-based cohort study’, *World psychiatry: official journal of the World Psychiatric Association* , 19(3), pp. 339–349.

Polderman, T.J.C. *et al.* (2015) ‘Meta-analysis of the heritability of human traits based on fifty years of twin studies’, *Nature genetics*, 47(7), pp. 702–709.

Post, R.M. *et al.* (2010) ‘Early-onset bipolar disorder and treatment delay are risk factors for poor outcome in adulthood’, *The Journal of clinical psychiatry*, 71(7), pp. 864–872.

Power, R.A. *et al.* (2013) 'Estimating the heritability of reporting stressful life events captured by common genetic variants', *Psychological medicine*, 43(9), pp. 1965–1971.

Quintero-Rivera, F., Sharifi-Hannauer, P. and Martinez-Agosto, J.A. (2010) 'Autistic and psychiatric findings associated with the 3q29 microdeletion syndrome: case report and review', *American journal of medical genetics. Part A*, 152A(10), pp. 2459–2467.

Ramsawh, H.J. *et al.* (2014) 'Risk for suicidal behaviors associated with PTSD, depression, and their comorbidity in the U.S. Army', *Journal of affective disorders*, 161, pp. 116–122.

Ratheesh, A. *et al.* (2017) 'A systematic review and meta-analysis of prospective transition from major depression to bipolar disorder', *Acta psychiatrica Scandinavica*, 135(4), pp. 273–284.

Reddy, M.S. (2010) 'Depression: the disorder and the burden', *Indian journal of psychological medicine*, 32(1), pp. 1–2.

van Rheenen, W. *et al.* (2019) 'Genetic correlations of polygenic disease traits: from theory to practice', *Nature reviews. Genetics* [Preprint]. Available at: https://doi.org/10.1038/s41576-019-0137-z.

Risch, N. and Merikangas, K. (1996) 'The future of genetic studies of complex human diseases', *Science*, 273(5281), pp. 1516–1517.

Rizvi, S.J. *et al.* (2014) 'Treatment-resistant depression in primary care across Canada', *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, 59(7), pp. 349–357.

Robins, L.N. *et al.* (1988) 'The Composite International Diagnostic Interview. An epidemiologic Instrument suitable for use in conjunction with different diagnostic systems and in different cultures', *Archives of general psychiatry*, 45(12), pp. 1069–1077.

Rosenberg, N.A. *et al.* (2002) 'Genetic structure of human populations', *Science*, 298(5602), pp. 2381–2385.

Rosenthal, D. (ed.) (1963). New York, NY, US: Basic Books.

Roshanaei-Moghaddam, B. and Katon, W. (2009) 'Premature mortality from general medical illnesses among persons with bipolar disorder: a review', *Psychiatric services* , 60(2), pp. 147–156.

Rössler, W. (1992) '[Wilhelm Griesinger and community care]', *Der Nervenarzt*, 63(5), pp. 257–261.

Rössler, W., Riecher-Rössler, A. and Meise, U. (1994) 'Wilhelm Griesinger and the concept of community care in 19th-century Germany', *Hospital & community psychiatry*, 45(8), pp. 818–822.

Rowland, T.A. and Marwaha, S. (2018) 'Epidemiology and risk factors for bipolar disorder', *Therapeutic advances in psychopharmacology*, 8(9), pp. 251–269.

Rush, A.J. *et al.* (2006) 'Report by the ACNP Task Force on response and remission in major depressive disorder', *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 31(9), pp. 1841–1853.

Rytwinski, N.K. *et al.* (2013) 'The co-occurrence of major depressive disorder among individuals with posttraumatic stress disorder: a meta-analysis', *Journal of Traumatic Stress*, 26(3), pp. 299–309.

Sachs, G.S. and Rush, A.J. (2003) 'Response, remission, and recovery in bipolar disorders: what are the realistic treatment goals?', *The Journal of clinical psychiatry*, 64 Suppl 6, pp. 18–22; discussion 28.

Sanders, S.J. *et al.* (2011) 'Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism', *Neuron*, 70(5), pp. 863–885.

Sartorius, N. *et al.* (1996) 'Depression comorbid with anxiety: results from the WHO study on psychological disorders in primary health care', *The British journal of psychiatry. Supplement*, (30), pp. 38–43.

Schindel-Allon, I. *et al.* (2010) 'Longitudinal associations between post-traumatic distress and depressive symptoms following a traumatic event: a test of three models', *Psychological medicine*, 40(10), pp. 1669–1678.

Scott, K.M. *et al.* (2006) 'Mental disorder comorbidity in Te Rau Hinengaro: the New Zealand Mental Health Survey', *The Australian and New Zealand journal of psychiatry*, 40(10), pp. 875–881.

Seedat, S. *et al.* (2009) 'Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys', *Archives of general psychiatry*, 66(7), pp. 785–795.

Sforzini, L. *et al.* (2021) 'A Delphi-method-based consensus guideline for definition of treatment-resistant depression for clinical trials', *Molecular psychiatry* [Preprint]. Available at: https://doi.org/10.1038/s41380-021-01381-x.

Shih, R.A., Belmonte, P.L. and Zandi, P.P. (2004) 'A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders', *International review of psychiatry* , 16(4), pp. 260–283.

Shugar, G. *et al.* (1992) 'Development, use, and factor analysis of a self-report inventory for mania', *Comprehensive psychiatry*, 33(5), pp. 325–331.

Siris, S.G. (2001) 'Suicide and schizophrenia', *Journal of psychopharmacology* , 15(2), pp. 127–135.

Slatkin, M. (2008) 'Linkage disequilibrium--understanding the evolutionary past and mapping the medical future', *Nature reviews. Genetics*, 9(6), pp. 477–485.

Smoller, J.W. *et al.* (2019) 'Psychiatric genetics and the structure of psychopathology', *Molecular psychiatry*, 24(3), pp. 409–420.

Solmi, M. *et al.* (2022) 'Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies', *Molecular psychiatry*, 27(1), pp. 281–295.

Song, J. *et al.* (2018) 'Specificity in etiology of subtypes of bipolar disorder: Evidence from a Swedish population-based family study', *Biological psychiatry*, 84(11), pp. 810–816.

Soreca, I. *et al.* (2008) 'Relationship of general medical burden, duration of illness and age in patients with bipolar I disorder', *Journal of psychiatric research*, 42(11), pp. 956–961.

Spijker, J. *et al.* (2002) 'Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS)', *The British journal of psychiatry: the journal of mental science*, 181, pp. 208–213.

Sportiche, S. *et al.* (2017) 'Clinical factors associated with lithium response in bipolar disorders', *The Australian and New Zealand journal of psychiatry*, 51(5), pp. 524–530.

Stahl, E.A. *et al.* (2019) 'Genome-wide association study identifies 30 loci associated with bipolar disorder', *Nature genetics*, 51(5), pp. 793–803.

Steel, Z. *et al.* (2014) 'The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013', *International journal of epidemiology*, 43(2), pp. 476–493.

Steffen, A. *et al.* (2020) 'Mental and somatic comorbidity of depression: a comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data', *BMC psychiatry*, 20(1), p. 142.

Sullivan, P.F., Neale, M.C. and Kendler, K.S. (2000) 'Genetic epidemiology of major depression: review and meta-analysis', *The American journal of psychiatry*, 157(10), pp. 1552–1562.

Teesson, M., Slade, T. and Mills, K. (2009) 'Comorbidity in Australia: findings of the 2007 National Survey of Mental Health and Wellbeing', *The Australian and New Zealand journal of psychiatry*, 43(7), pp. 606–614.

Thayer, R.E. (1990) *The Biopsychology of Mood and Arousal*. Oxford University Press.

Thorp, J.G. *et al.* (2020) 'Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9', *Psychological medicine*, 50(14), pp. 2385–2396.

Tohen, M. *et al.* (2003) 'The McLean-Harvard First-Episode Mania Study: prediction of recovery and first recurrence', *The American journal of psychiatry*, 160(12), pp. 2099–2107.

Tondo, L., Vázquez, G. and Baldessarini, R.J. (2010) 'Mania associated with antidepressant treatment: comprehensive meta-analytic review', *Acta psychiatrica Scandinavica*, 121(6), pp. 404–414.

Tondo, L., Vázquez, G.H. and Baldessarini, R.J. (2017) 'Depression and Mania in Bipolar Disorder', *Current neuropharmacology*, 15(3), pp. 353–358.

Trivedi, M.H. *et al.* (2006) 'Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice', *The American journal of psychiatry*, 163(1), pp. 28–40.

Undurraga, J. *et al.* (2012) 'Bipolar depression: clinical correlates of receiving antidepressants', *Journal of affective disorders*, 139(1), pp. 89–93.

Viñas-Jornet, M. *et al.* (2018) 'High Incidence of Copy Number Variants in Adults with Intellectual Disability and Co-morbid Psychiatric Disorders', *Behavior genetics*, 48(4), pp. 323–336.

Visscher, P.M. *et al.* (2012) ‘Five years of GWAS discovery’, *American journal of human genetics*, 90(1), pp. 7–24.

Visscher, P.M. *et al.* (2017) ‘10 Years of GWAS Discovery: Biology, Function, and Translation’, *American journal of human genetics*, 101(1), pp. 5–22.

Visscher, P.M. *et al.* (2021) ‘Discovery and implications of polygenicity of common diseases’, *Science*, 373(6562), pp. 1468–1473.

Visscher, P.M., Hill, W.G. and Wray, N.R. (2008) ‘Heritability in the genomics era--concepts and misconceptions’, *Nature reviews. Genetics*, 9(4), pp. 255–266.

Visscher, P.M. and Yang, J. (2016) ‘A plethora of pleiotropy across complex traits’, *Nature genetics*, pp. 707–708.

Walker, E.R., McGee, R.E. and Druss, B.G. (2015) ‘Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis’, *JAMA psychiatry* , 72(4), pp. 334–341.

Wang, W.Y.S. *et al.* (2005) ‘Genome-wide association studies: theoretical and practical concerns’, *Nature reviews. Genetics*, 6(2), pp. 109–118.

Ward, J. *et al.* (2019) ‘Novel genome-wide associations for anhedonia, genetic correlation with psychiatric disorders, and polygenic association with brain structure’, *Translational psychiatry*, 9(1), p. 327.

Weiss, K.M. and Lambert, B.W. (2011) ‘When the time seems ripe: eugenics, the annals, and the subtle persistence of typological thinking’, *Annals of human genetics*, 75(3), pp. 334–343.

Weissman, M.M. *et al.* (1996) ‘Cross-national epidemiology of major depression and bipolar disorder’, *JAMA: the journal of the American Medical Association*, 276(4), pp. 293–299.

Wellcome Trust Case Control Consortium (2007) ‘Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls’, *Nature*, 447(7145), pp. 661–678.

Whiteford, H.A. *et al.* (2013) ‘Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010’, *The Lancet*, 382(9904), pp. 1575–1586.

Wigmore, E.M. *et al.* (2020) ‘Genome-wide association study of antidepressant treatment resistance in a population-based cohort using health service prescription data and meta-analysis with GENDEP’, *The pharmacogenomics journal*, 20(2), pp. 329–341.

Wilson, S. *et al.* (2015) ‘Age of onset and course of major depressive disorder: associations with psychosocial functioning outcomes in adulthood’, *Psychological medicine*, 45(3), pp. 505–514.

Wingo, A.P. *et al.* (2010) ‘Factors associated with functional recovery in bipolar disorder patients’, *Bipolar disorders*, 12(3), pp. 319–326.

Wolf, E.J. *et al.* (2010) ‘Posttraumatic stress disorder and the genetic structure of comorbidity’, *Journal of abnormal psychology*, 119(2), pp. 320–330.

Wray, N.R. *et al.* (2018) ‘Genome-wide association analyses identify 44 risk variants and refine

the genetic architecture of major depression', *Nature genetics*, 50(5), pp. 668–681.

Yang, J. *et al.* (2017) 'Concepts, estimation and interpretation of SNP-based heritability', *Nature genetics*, 49(9), pp. 1304–1310.

Young, R.C. *et al.* (1978) 'A rating scale for mania: reliability, validity and sensitivity', *The British journal of psychiatry: the journal of mental science*, 133, pp. 429–435.

Yousofpour, M. *et al.* (2015) 'Role of Heart and its Diseases in the Etiology of Depression According to Avicenna's Point of View and its Comparison with Views of Classic Medicine', *International journal of preventive medicine*, 6, p. 49.

Zhang, X. *et al.* (2019) 'Genome-wide Burden of Rare Short Deletions Is Enriched in Major Depressive Disorder in Four Cohorts', *Biological psychiatry*, 85(12), pp. 1065–1073.

Zimmerman, M. *et al.* (2009) 'Performance of the mood disorders questionnaire in a psychiatric outpatient setting', *Bipolar disorders*, 11(7), pp. 759–765.

Zimmerman, M., Chelminski, I. and McDermut, W. (2002) 'Major depressive disorder and axis I diagnostic comorbidity', *The Journal of clinical psychiatry*, 63(3), pp. 187–193.

Zlotogora, J. (2003) 'Penetrance and expressivity in the molecular age', *Genetics in medicine: official journal of the American College of Medical Genetics*, 5(5), pp. 347–352.

Chapter 2. Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder

This chapter is an exact copy of a peer-reviewed publication in *Psychological Medicine*.

Mundy, J., Hübel, C., Gelernter, J., Levey, D., Murray, R. M., Skelton, M., Stein, M. B., Vassos, E., Breen, G. and Coleman, J. R. I. (2021). Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder. *Psychological Medicine*. Cambridge University Press, pp. 1–10. doi: 10.1017/S0033291721000830.

Supplementary material is included in **appendix 2**.

Psychological Medicine

cambridge.org/psm

Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder

Jessica Mundy[1,2], Christopher Hübel[1,2,3], Joel Gelernter[4,5,6], Daniel Levey[4,5], Robin M. Murray[2,7], Megan Skelton[1,2], Murray B. Stein[8,9], The Million Veteran Program, Post Traumatic Stress Disorder Working Group of the Psychiatric Genomics Consortium, Evangelos Vassos[1,2], Gerome Breen[1,2] and Jonathan R. I. Coleman[1,2]

[1]Social, Genetic and Developmental Psychiatry Centre; Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [2]UK National Institute for Health Research (NIHR) Biomedical Research Centre, South London and Maudsley National Health Service (NHS) Trust, London, UK; [3]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [4]Division of Human Genetics, Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut, USA; [5]Department of Psychiatry, Veterans Affairs Connecticut Healthcare Center, West Haven, Connecticut, USA; [6]Departments of Genetics and Neuroscience, Yale University School of Medicine, New Haven, Connecticut, USA; [7]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [8]Psychiatry Service, VA San Diego Healthcare System, San Diego, California, USA and [9]Departments of Psychiatry and Family Medicine & Public Health, University of California San Diego, La Jolla, California, USA

Original Article

Cite this article: Mundy J *et al* (2021). Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder. *Psychological Medicine* 1–10. https://doi.org/10.1017/S0033291721000830

Received: 29 October 2020
Revised: 19 February 2021
Accepted: 24 February 2021

Key words:
Posttraumatic stress disorder; major depressive disorder; psychological trauma; genetics; genetic correlations; polygenic risk scores

Author for correspondence:
Jonathan R. I. Coleman,
E-mail: jonathan.coleman@kcl.ac.uk

Abstract

Background. Posttraumatic stress disorder (PTSD) and major depressive disorder (MDD) are commonly reported co-occurring mental health consequences of psychological trauma exposure. The disorders have high genetic overlap. Trauma is a complex phenotype but research suggests that trauma sensitivity has a heritable basis. We investigated whether sensitivity to trauma in those with MDD reflects a similar genetic component in those with PTSD.

Methods. Genetic correlations between PTSD and MDD in individuals reporting trauma and MDD in individuals not reporting trauma were estimated, as well as with recurrent MDD and single-episode MDD, using genome-wide association study (GWAS) summary statistics. Genetic correlations were replicated using PTSD data from the Psychiatric Genomics Consortium and the Million Veteran Program. Polygenic risk scores were generated in UK Biobank participants who met the criteria for lifetime MDD ($N = 29\,471$). We investigated whether genetic loading for PTSD was associated with reporting trauma in these individuals.

Results. Genetic loading for PTSD was significantly associated with reporting trauma in individuals with MDD [OR 1.04 (95% CI 1.01–1.07), Empirical-$p = 0.02$]. PTSD was significantly more genetically correlated with recurrent MDD than with MDD in individuals not reporting trauma (r_g differences = ~0.2, $p < 0.008$). Participants who had experienced recurrent MDD reported significantly higher rates of trauma than participants who had experienced single-episode MDD ($\chi^2 > 166$, $p < 0.001$)

Conclusions. Our findings point towards the existence of genetic variants associated with trauma sensitivity that might be shared between PTSD and MDD, although replication with better powered GWAS is needed. Our findings corroborate previous research highlighting trauma exposure as a key risk factor for recurrent MDD.

Introduction

Symptoms of posttraumatic stress disorder (PTSD) and major depressive disorder (MDD) are the most commonly described co-occurring problems following exposure to psychological trauma (Ben Barnes, Hayes, Contractor, Nash, & Litz, 2018). Across epidemiological samples, approximately 50% of individuals with PTSD have a comorbid diagnosis of MDD (Breslau, Davis, Peterson, & Schultz, 1997; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Rytwinski, Scur, Feeny, & Youngstrom, 2013). Similar or occasionally higher estimates are observed in primary care settings (Alim et al., 2006; Stein, McQuaid, Pedrelli, Lenox, & McCahill, 2000). Previously, high comorbidity rates were attributed to the classification of shared symptoms into the two diagnostic categories (Flory & Yehuda, 2015), such as negative mood, sleep disturbances, irritability and concentration difficulties (American Psychiatric Association, 2013). However, several studies demonstrate that comorbidity rates do not diminish after excluding shared symptoms from clinical diagnoses (Elhai et al., 2011; Grubaugh,

CAMBRIDGE UNIVERSITY PRESS

Long, Elhai, Frueh, & Magruder, 2010), suggesting that symptom overlap does adequately explain comorbidity. An alternative explanation might be the genetic overlap between the disorders (Sartor et al., 2012). Twin studies have previously indicated that PTSD shares genetic influences with MDD ($r = 0.77$) and related conditions (Koenen et al., 2008; Wolf et al., 2010). More recently, methods based on genome-wide association studies (GWAS) have been used to explore genetic correlations (r_g), a quantitative measure of the genetic relationship between two polygenic traits (van Rheenen, Peyrot, Schork, Lee, & Wray, 2019). Research from the Psychiatric Genomics Consortium (PGC) reported strong, positive genetic correlations of PTSD with depressive symptoms ($r_g = 0.80$) and with MDD ($r_g = 0.62$) (Nievergelt et al., 2019), thus supporting results from twin studies.

As well as shared genetics, a potential factor involved in PTSD-MDD comorbidity is exposure to trauma. There is a complex relationship between trauma exposure and mental health sequelae. Exposure to trauma is common (Breslau, Davis, Andreski, & Peterson, 1991; Kessler et al., 1995). In total, 50–90% of people will experience a traumatic event in their lifetime but only 8–12% will go on to develop PTSD (Shah, Shah, & Links, 2012), suggesting that certain individuals are at a greater risk of developing PTSD than others following exposure (Auxéméry, 2012; Duncan et al., 2018b; Nievergelt et al., 2019). Similarly, stressful and traumatic events are significant risk factors for MDD (Horesh, Klomek, & Apter, 2008; Hovens, 2015 ; Shapero et al., 2014), but the majority of people who are exposed do not develop the disorder (Kessler, 1997). Therefore, similar to PTSD, the effects of these events on the risk of developing MDD may be moderated by individual liability or sensitivity to trauma.

Reporting adverse or traumatic life events is heritable (Dalvie et al., 2020; Jay Schulz-Heik et al., 2009; Plomin, Lichtenstein, Pedersen, McClearn, & Nesselroade, 1990; Power et al., 2013). However, this alone does not provide direct evidence for a genetic basis for trauma sensitivity. There are a number of additional factors that could contribute to trauma's overall heritability, including whether or not the event is controllable and whether the individual plays an active or passive role in the event (Kendler, Karkowski, & Prescott, 1999; Plomin et al., 1990). Heritable personality traits are also important since these influence both the likelihood of exposure and willingness to report it (Sartor et al., 2012). We emphasise that the presence of these characteristics should not be interpreted as placing any blame on individuals who have experienced trauma. These heritable characteristics are difficult to disentangle from genetic influences on trauma sensitivity. However, studying mental health across individuals who have experienced trauma offers the opportunity to assess trauma sensitivity more specifically. Furthermore, with information on both life events and mental health in genotyped individuals, the extent that gene–environment interaction influences the development of psychopathology (such as the internalising symptoms in PTSD and MDD) can be investigated.

Aims

The heritability of PTSD (Duncan et al., 2018b; Nievergelt et al., 2019; Stein et al., 2021), which by definition requires trauma exposure, indicates that variance in sensitivity to such events may be partially genetically influenced and interacts with environmental factors to influence individual risk for developing symptoms. Trauma is a key risk factor for MDD. Recent research found that in UK Biobank participants with MDD who reported traumatic life events, MDD had higher SNP-based heritability compared to MDD in participants not reporting trauma (24% *v.* 12% respectively), suggesting that trauma sensitivity also has a heritable basis in MDD (Coleman et al., 2020). PTSD and MDD co-occur often among trauma-exposed individuals and the disorders have substantial genetic overlap. In our study, we aimed to understand whether trauma sensitivity in those with MDD reflects a similar genetic component in those with PTSD.

We addressed our research question in two parts. First, we used GWAS summary statistics to calculate genetic correlations between PTSD and (1) MDD with reported trauma and (2) MDD without reported trauma. Given the evidence from clinical studies, we hypothesised that PTSD and MDD with reported trauma would demonstrate higher genetic overlap compared to PTSD and MDD without reported trauma. In the original GWAS of MDD with reported trauma, both cases and controls were trauma-exposed (Coleman et al., 2020). This means the summary statistics specifically capture genetic variants associated with MDD in individuals who report trauma. A higher genetic correlation with PTSD, which requires trauma exposure, would therefore reflect a shared genetically driven component associated with trauma sensitivity. We also used summary statistics to calculate genetic correlations between PTSD and (3) recurrent MDD and (4) single-episode MDD. Research has shown that the type, frequency and severity of traumatic events are associated with the frequency and severity of subsequent depressive episodes (Hovens, Giltay, Spinhoven, van Hemert, & Penninx, 2015; Nanni, Uher, & Danese, 2012; Otte et al., 2016), with childhood maltreatment being particularly associated with recurrence (Danese, 2020). Accordingly, our second hypothesis was that PTSD would show greater genetic overlap with recurrent MDD compared to single-episode MDD, under the assumption that trauma exposure is more likely to be reported by individuals with recurrent MDD. We tested the validity of this assumption in the participants.

Secondly, to address our research question further, we computed PTSD polygenic risk scores (PRS) in 29 471 UK Biobank participants who met criteria for lifetime MDD and tested their association with reporting trauma and MDD recurrence. Following the logic of our previous hypotheses, we expected that individuals with MDD with a higher genetic risk for PTSD would be more likely to report trauma and would be more likely to have experienced recurrent episodes than those with a lower genetic risk for PTSD.

Methods

Major depressive disorder

In the first part of our study, GWAS summary statistics for the four MDD categories were obtained from pre-existing studies (Table 1). In the second part, we analysed phenotypic and genomic data from 29 471 UK Biobank participants who met the criteria for lifetime MDD and had been included in the previous GWAS of the MDD categories. Four participants had withdrawn from participating in the UK Biobank since the GWAS were published and were therefore not included in any individual-level analyses in our study. Participants were categorised as either having MDD with reported trauma or MDD without reported trauma. Subsequently, participants were categorised as having either recurrent or single-episode MDD. There is some degree of overlap between the categories (Fig. 1).

https://doi.org/10.1017/S0033291721000830 Published online by Cambridge University Press

Table 1. Information about the four posttraumatic stress disorder (PTSD) and four major depressive disorder (MDD) genome-wide association study (GWAS) summary statistics, including the original publication, characteristics of the sample, number (*N*) of cases and controls in original GWAS, liability scale SNP-based heritability (h^2_{SNP}) and standard error (S.E.) from High Definition Likelihood

Phenotype	Paper containing original GWAS	Sample characteristics	*N* cases	*N* controls	h^2_{SNP} (liability scale)	S.E.
UK Biobank PTSD	Nievergelt et al. (2019)	Probable PTSD phenotype was defined in UK Biobank participants based on self-report answers to PTSD Checklist (PCL) 6 (Civilian version) in the Mental Health Questionnaire	10 389	115 799	0.20	0.009
Psychiatric Genomics Consortium 1.5 PTSD	Nievergelt et al. (2019)	Sample comprises 59 studies of PTSD. Most cases were clinically ascertained through telephone or face-to-face interviews	12 823	35 648	0.06	0.011
Million Veteran Program PTSD	Stein et al. (2021)	Sample comprises US veterans. PTSD was algorithmically defined based on electronic health records Confirmed war- and combat exposure: 27.5% No exposure: 29.3% Unknown exposure: 43.1%	36 301	178 107	0.06	0.015
MDD with reported trauma	Coleman et al. (2020)	Phenotype was defined in UK Biobank participants. Sample comprises cases who met criteria for MDD and controls who did not meet criteria for MDD based on answers to the Composite International Diagnostic Interview Short Form (CIDI-SF). Cases and controls reported at least two traumatic life events (Table 3) in the Mental Health Questionnaire	13 393	10 701	0.24	0.017
MDD without reported trauma	Coleman et al. (2020)	Phenotype was defined in UK Biobank participants. Sample comprises cases who met criteria for MDD and controls who did not meet criteria for MDD based on answers to the Composite International Diagnostic Interview Short Form (CIDI-SF). Cases and controls reported no traumatic life events (Table 3) in the Mental Health Questionnaire	9487	39 677	0.15	0.020
Recurrent MDD	Coleman et al. (2019)	Phenotype was defined in UK Biobank participants. Cases met criteria for MDD based on answers to the Composite International Diagnostic Interview Short Form (CIDI-SF) and reported more than one depressive episode in the Mental Health Questionnaire. Controls did not meet criteria for MDD	17 451	63 482	0.22	0.009
Single-episode MDD	Coleman et al. (2019)	Phenotype was defined in UK Biobank participants. Cases met criteria for MDD based on answers to the Composite International Diagnostic Interview Short Form (CIDI-SF) and reported one depressive episode in the Mental Health Questionnaire. Controls did not meet criteria for MDD	12 024	63 482	0.10	0.008

Further details of the phenotypes and how observed scale h^2_{SNP} estimates were converted to the liability scale are presented in online Supplementary Material

Posttraumatic stress disorder

PTSD phenotypes can reflect sample characteristics and data collection methods. We used three sets of PTSD GWAS summary statistics to examine whether our findings were consistent across differing PTSD phenotypes. First, we calculated genetic correlations using summary statistics from a GWAS of probable PTSD in the UK Biobank which was based on self-reported answers in the Mental Health Questionnaire (MHQ) (UKB-PTSD). Secondly, we used summary statistics a GWAS of 59 mainly clinical PTSD samples from the PGC (PGC1.5-PTSD) (Nievergelt et al., 2019). Finally, we used summary statistics from a PTSD GWAS based on electronic health records of US veterans by the Million Veteran Program (MVP-PTSD) (Stein et al., 2021). We generated PRS using this MVP-PTSD phenotype.

The number of cases and controls and the SNP-based heritability (liability scale) of each GWAS can be found in Table 1. All summary statistics were produced from GWAS on individuals of European ancestries. Details of the contributing studies and phenotype definitions can be found in online Supplementary Methods.

Reported trauma in individuals with recurrent and single-episode major depressive disorder

We tested the assumption behind our second hypothesis: we expected the rates of trauma exposure to be higher among individuals who have experienced recurrent compared to single-episode MDD. The UK Biobank participants were categorised as having experienced either recurrent or single-episode MDD, as defined

https://doi.org/10.1017/S0033291721000830 Published online by Cambridge University Press

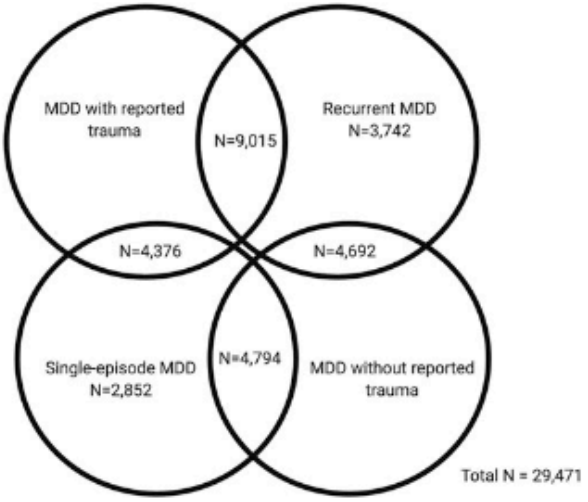


Fig. 1. Venn diagram showing participant overlap between the four major depressive disorder (MDD) categories in UK Biobank Mental Health Questionnaire (MHQ) respondents who met criteria for lifetime MDD ($N = 29\,471$). The MDD categories include MDD with reported trauma ($N = 13\,391$), MDD without reported trauma ($N = 9486$), recurrent MDD ($N = 17\,449$) and single-episode MDD ($N = 12\,022$).

by Coleman et al. (2019). Seven traumatic life events were included in the Coleman et al. (2020) definition of 'reported trauma exposure' due to them having a >2.5 odds ratio (OR) with MDD (Table 3). We performed χ^2 tests in *R* to establish whether there were differences in trauma reporting rates between individuals with recurrent and single-episode MDD. χ^2 statistics were considered significant if they reached or surpassed the Bonferroni-corrected α ($0.05/7 = 0.007$; to correct for the seven tests performed).

Throughout this paper, any mention of trauma exposure in UK Biobank participants refers specifically to retrospective, self-reported traumatic events due to the nature of data collection via the online MHQ. The events being reported may have occurred before, after or concurrently with MDD episodes.

Genetic correlations

GWAS summary statistics were used to calculate genetic correlations based on single nucleotide polymorphisms (SNP-based r_g) using High Definition Likelihood (HDL) and the 1 029 876 quality-controlled UK Biobank imputed HapMap3 SNPs reference panel. This reference panel is based on genotypes in the UK Biobank, which were imputed to HRC and UK10K + 1000 Genomes (Ning, Pawitan, & Shen, 2020).

First, we calculated genetic correlations between PTSD and (1) MDD with reported trauma, (2) MDD without reported trauma, (3) recurrent MDD and (4) single-episode MDD within UKB-PTSD. We repeated these genetic correlations using PGC1.5-PTSD and MVP-PTSD. Genetic correlations were tested for a significant difference from 0 (default in HDL) and from 1 (in Microsoft Excel, converting r_g to a χ^2 as $[(r_g - 1)/\text{se}]^2$). An explanation of HDL can be found in Ning et al. (2020). Genetic correlations were considered significantly different to 0 or to 1 if they surpassed the Bonferroni-corrected α ($0.05/4 = 0.0125$; to correct for the four tests per PTSD phenotype).

To test the significance of the differences between the genetic correlations, we performed a block-jackknife, which uses resampling to recalculate standard errors for the differences between two r_g estimates. We compared r_g estimates in a pairwise fashion, where each correlation pair was compared with all other correlation pairs within its group. This resulted in six different block-jackknife tests per PTSD phenotype. Differences between genetic correlations were considered statistically significant if they surpassed the Bonferroni-corrected α ($0.05/6 = 0.0083$; to correct for the six tests).

To maximise power, the PGC combined PGC1.5-PTSD and UKB-PTSD. We repeated these analyses with these summary statistics (PGC2-PTSD). Results can be found in online Supplementary Tables S1 and S5. Online Supplementary Methods contain further details of the HDL analysis, including the percentage overlap between the summary statistics and the HapMap3 reference panel. Online Supplementary Tables S3 and S4 contain the genetic correlations between the PTSD phenotypes and MDD categories, respectively.

We also ran these analyses using Linkage Disequilibrium Score Regression (LDSC), another command line tool for estimating heritability and genetic correlations from GWAS summary statistics (Bulik-Sullivan et al., 2015). In our study, we favoured HDL for estimating genetic correlations. Unlike LDSC, HDL uses a full likelihood-based method to estimate genetic correlations that fully accounts for linkage disequilibrium (LD) across the genome. When compared to LDSC, HDL reduces the variance of the genetic correlation by approximately 60% (Ning et al., 2020). Consequently, HDL is better powered to detect significant differences between correlations, which was a central aim of our study. The LDSC results and an explanation of any differences from HDL are presented in online Supplementary Results.

Polygenic risk scores

We computed PTSD PRS using PRSice v2.3.1. We controlled for the first six principal components, genotyping batch, assessment centre and current depression severity assessed by the Patient Health Questionnaire 9 (PHQ9). A significant difference in PHQ9 severity was found between the MDD with and without reported trauma group ($W = 45\,762\,133$, $p < 2.2 \times 10^{-16}$) and the recurrent and single-episode MDD group ($W = 70\,764\,216$, $p < 2.2 \times 10^{-16}$). Medians and IQRs are presented in Table 2. Therefore, PHQ9 severity was included as a covariate to control for negative mood at the time of recall influencing the reporting of traumatic events (Reuben et al., 2016).

PRS were calculated at 11 *p* value thresholds (5×10^{-8}, 1×10^{-5}, 1×10^{-3}, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1). Phenotype permutations were used to produce an empirical *p* value for the association at the best-fitting PRS, which accounts for testing at multiple thresholds (Euesden, Lewis, & O'Reilly, 2015). Once the best-fitting PRS had been calculated, we performed logistic regressions to examine whether genetic risk for PTSD showed a greater association with MDD with reported trauma or MDD without reported trauma, and with recurrent or single-episode MDD. The standardised β coefficients were converted to OR and 95% confidence intervals were calculated. The full six pairwise comparisons, as in the block-jackknife analysis, were not possible due to overlapping MDD categories (Fig. 1). Therefore, we limit the PRS analysis to two comparisons.

We performed power calculations using the Additive Variance Explained and Number of Genetic Effects Method of Estimation (AVENGEME) programme (Dudbridge, 2013). Details are presented in online Supplementary Methods. MVP-PTSD summary statistics were chosen based on their power and no overlap with the

https://doi.org/10.1017/S0033291721000830 Published online by Cambridge University Press

Table 2. Descriptive statistics for age, sex and current depression severity of the study sample of individuals who met criteria for lifetime major depressive disorder (MDD) in UK Biobank Mental Health Questionnaire (MHQ) respondents

	Sample overall	MDD with reported trauma	MDD without reported trauma	Recurrent MDD	Single-episode MDD
N	29 471	13 391	9486	17 449	12 022
Age					
Min	46	47	46	47	46
Max	80	80	80	80	80
Mean	62.36	61.66	63.24	61.83	63.14
Std	7.52	7.43	7.57	7.47	7.53
Sex					
Females (%)	20 323 (69%)	10 248 (77%)	5782 (61%)	12 295 (70%)	8028 (67%)
Males (%)	9148 (31%)	3143 (23%)	3704 (39%)	5154 (30%)	3994 (33%)
PHQ9 severity					
Min	0	0	0	0	0
Max	27	27	27	27	27
Median	4.76	5.82	3.49	5.90	3.12
IQR	6	6	5	7	4

Descriptive statistics and sample size (*N*) are given for the study sample as a whole and for the four MDD categories individually. Age was measured in years and refers to the age when the participant completed the MHQ. Sex was acquired from the central registry at recruitment, but in some cases was updated by the participant. Current depression severity was assessed using the Patient Health Questionnaire 9 (PHQ9). Min refers to minimum value, max refers to maximum value, mean refers to the mean value and std refers to standard deviation. PHQ9 severity was not normally distributed so the median and interquartile range (IQR) are provided instead of the mean and standard deviation

Table 3. Difference in reporting rates of traumatic life events between individuals with recurrent and single-episode major depressive disorder (MDD) in UK Biobank Mental Health Questionnaire (MHQ) respondents ($N = 29\,471$)

Trauma category	Traumatic event	Endorsement in single-episode MDD (%)	Endorsement in recurrent MDD (%)	χ^2 statistic	*p* value
Childhood emotional abuse	Felt hated by a family member as a child	2352 (20%)	5238 (30%)	405	$\mathbf{4.41 \times 10^{-90}}$
Childhood emotional neglect	Did not feel loved as a child	3121 (26%)	6702 (39%)	498	$\mathbf{2.71 \times 10^{-110}}$
Childhood sexual abuse	Was sexually molested as a child	1217 (10%)	2690 (16%)	175	$\mathbf{6.46 \times 10^{-40}}$
Adulthood emotional abuse	Was belittled by a partner or ex-partner	3887 (32%)	7590 (44%)	370	$\mathbf{1.57 \times 10^{-82}}$
Adulthood physical abuse	Was physically abused by a partner or ex-partner	2005 (17%)	3987 (23%)	166	$\mathbf{4.34 \times 10^{-38}}$
Adulthood sexual abuse	Was forced to have sex against my will by a partner or ex-partner	890 (7%)	2283 (13%)	240	$\mathbf{3.29 \times 10^{-54}}$
PTSD-related: sexual assault	Ever been a victim of sexual assault	2247 (19%)	4764 (28%)	293	$\mathbf{1.07 \times 10^{-65}}$

Traumatic events include three childhood events, three adulthood events and one posttraumatic stress disorder (PTSD)-related event. Differences were considered significant if they surpassed the Bonferroni-adjusted α ($p < 0.007$). Significant *p* values are shown in bold.

UK Biobank, as overlap between the training and target samples can lead to overfitting. Bonferroni adjustment was used to correct for the two tests, giving a final threshold of $p < 0.025$.

Results

Sample characteristics

Table 2 contains descriptive statistics, for age at MHQ, sex and current depression severity (assessed by the PHQ9) for the study sample as a whole and for the four MDD categories.

Reported trauma in individuals with recurrent and single-episode major depressive disorder

Each of the seven life events comprising the definition of 'reported trauma exposure' in Coleman et al. (2020) was significantly more commonly reported by participants who reported recurrent depressive episodes than single-episode MDD (Table 3). Nine further traumatic life events were also assessed by the UK Biobank MHQ. The difference in reporting rates between recurrent and single-episode MDD and the results of the χ^2 tests can be found in online Supplementary Table S2.

https://doi.org/10.1017/S0033291721000830 Published online by Cambridge University Press

Table 4. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (S.E.) and 95% confidence intervals (lower and upper CI) of (1) UK Biobank posttraumatic stress disorder (UKB-PTSD), (2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD) and (3) Million Veteran Program PTSD (MVP-PTSD) with the four major depressive disorder (MDD) categories

PTSD phenotype	MDD phenotype	r_g	S.E.	Lower CI	Upper CI	p (diff 0)	p (diff 1)
UKB-PTSD	MDD with reported trauma	0.6040	0.0550	0.4962	0.7118	$\mathbf{4.92 \times 10^{-28}}$	$\mathbf{6.02 \times 10^{-13}}$
UKB-PTSD	MDD without reported trauma	0.4701	0.0742	0.3247	0.6155	$\mathbf{2.43 \times 10^{-10}}$	$\mathbf{9.23 \times 10^{-13}}$
UKB-PTSD	Recurrent MDD	0.7134	0.0481	0.6191	0.8077	$\mathbf{1.03 \times 10^{-49}}$	$\mathbf{2.55 \times 10^{-9}}$
UKB-PTSD	Single-episode MDD	0.6466	0.0691	0.5112	0.7820	$\mathbf{8.35 \times 10^{-21}}$	$\mathbf{3.15 \times 10^{-7}}$
PGC1.5-PTSD	MDD with reported trauma	0.5520	0.0746	0.4058	0.6982	$\mathbf{1.35 \times 10^{-13}}$	$\mathbf{1.91 \times 10^{-9}}$
PGC1.5-PTSD	MDD without reported trauma	0.4841	0.1107	0.2671	0.7011	$\mathbf{1.22 \times 10^{-5}}$	$\mathbf{3.16 \times 10^{-6}}$
PGC1.5-PTSD	Recurrent MDD	0.6937	0.0821	0.5328	0.8546	$\mathbf{2.94 \times 10^{-17}}$	$\mathbf{1.91 \times 10^{-4}}$
PGC1.5-PTSD	Single-episode MDD	0.7560	0.1403	0.4810	1.0310	$\mathbf{7.09 \times 10^{-8}}$	0.08
MVP-PTSD	MDD with reported trauma	0.5397	0.0938	0.3559	0.7235	$\mathbf{8.77 \times 10^{-9}}$	$\mathbf{9.24 \times 10^{-7}}$
MVP-PTSD	MDD without reported trauma	0.4859	0.0871	0.3152	0.6566	$\mathbf{2.41 \times 10^{-8}}$	$\mathbf{3.58 \times 10^{-9}}$
MVP-PTSD	Recurrent MDD	0.5600	0.0532	0.4557	0.6643	$\mathbf{6.57 \times 10^{-26}}$	$\mathbf{1.33 \times 10^{-16}}$
MVP-PTSD	Single-episode MDD	0.6291	0.1113	0.4110	0.8472	$\mathbf{1.59 \times 10^{-8}}$	$\mathbf{8.61 \times 10^{-4}}$

p (diff 0) refers to p value to test whether the r_g differs from 0. p (diff 1) refers to p value to test whether the r_g differs from 1. Genetic correlations were considered significant if they surpassed the Bonferroni-adjusted α ($p < 0.0125$). Significant p values are shown in bold.

Genetic correlations

All genetic correlations were significantly different to 0. The genetic correlation between PGC1.5-PTSD and single-episode MDD did not differ significantly from 1, although this is likely due to the large standard errors of the r_g estimates, reflecting low power. All other genetic correlations were significantly different to 1 (Table 4).

Differences between genetic correlations

The genetic correlation between PTSD and recurrent MDD was significantly greater than that between PTSD and MDD without reported trauma when using UKB-PTSD and PGC1.5-PTSD (and PGC2-PTSD, which is presented in online Supplementary Results). All other genetic correlations were not significantly different from each other (online Supplementary Table S5). Genetic correlation estimates of PTSD with MDD with reported trauma were consistently higher than those with MDD without reported trauma, albeit not significant ($p = 0.14–0.65$). By contrast, no consistent pattern was observed between PTSD and recurrent *v.* single-episode MDD (Table 4, online Supplementary Table S5). These results were also observed when using PGC2-PTSD (online Supplementary Tables S1 and S5).

Polygenic risk scores

In individuals with MDD in the UK Biobank, genetic loading for PTSD was significantly associated with an increased likelihood of reporting trauma [OR 1.04 (95% CI 1.01–1.07), Empirical-p = 0.02]. In contrast, those with a higher genetic loading for PTSD were more likely to have experienced a single depressive episode rather than recurrent episodes, but this was not significant [OR 0.97 (95% CI 0.95–0.99), Empirical-$p = 0.08$]. The variance explained by the PRS ranged from 0.03% to 0.06% based on varying the population prevalence of the target phenotype. See online Supplementary Results for details of this analysis, including the number of SNPs in each PRS and Nagelkerke's R^2 for a range of population prevalences (online Supplementary Table S8).

Discussion

We investigated whether PTSD and MDD share a genetic component related to being exposed to traumatic events and experiencing internalising symptomatology. We addressed this by measuring the genetic overlap between PTSD and MDD with reported trauma and compared this to the genetic overlap between PTSD and MDD without reported trauma. We aimed to discover whether the genetic variants associated with MDD in individuals reporting trauma were shared with PTSD, which requires trauma exposure by definition. Additionally, we investigated whether genetic risk for PTSD was associated with reporting trauma in UK Biobank participants with MDD. Across all PTSD phenotypes, as hypothesised, genetic correlations with MDD in individuals reporting trauma were greater than genetic correlations with MDD in individuals not reporting trauma. However, the differences were not significant so strong conclusions cannot be drawn from this analysis alone. By contrast, the PRS analysis showed that genetic loading for PTSD in individuals with MDD was associated with a higher likelihood of reporting trauma. This result appears to be robust to recall bias since we controlled for depression severity at the time of reporting.

A potential explanation for this finding involves G × E. Certain individuals may be particularly sensitive to adverse life events due to their inherited genetics, and therefore have a propensity to develop psychopathology. This appears to be the case for PTSD symptoms: experiencing trauma is common, but only a minority of those who experience trauma develop PTSD (Auxéméry, 2012; Duncan, Cooper, & Shen, 2018a; Nievergelt et al., 2019). In our study, the PRS generated from the MVP-PTSD summary statistics capture the risk from common additive genetic variants for a persistent, negative response to traumatic events. Accordingly, the PTSD PRS represent a genetic component of trauma sensitivity in UK Biobank participants with MDD. We found a significant

https://doi.org/10.1017/S0033291721000830 Published online by Cambridge University Press

association between the PTSD PRS and reporting trauma in these individuals which suggests that there are genetic variants associated with PTSD that also influence an individual's sensitivity to trauma in MDD. This can be viewed as a form of G × E in line with the diathesis-stress approach, where the experience of certain life events increases the likelihood of developing psychopathology by activating a genetically driven vulnerability (Colodro-Conde et al., 2018; Meehl, 1962; Monroe & Simons, 1991). Research into G × E between adverse life events and the risk for MDD has previously yielded inconsistent results (Coleman et al., 2020; Colodro-Conde et al., 2018; Mullins et al., 2016; Peyrot et al., 2014). Our study adds to the literature suggesting that genetic risk for PTSD could reflect an underlying dimension of sensitivity to psychologically distressing events. Accordingly, having a higher genetic loading for PTSD may be associated with an increased risk of experiencing internalising symptoms in individuals who have experienced trauma.

The PRS finding is interesting in light of our hypothesis that PTSD would show higher genetic overlap with MDD in individuals reporting trauma compared to MDD in individuals not reporting trauma. Although the genetic correlation analysis yielded no conclusive results, the findings from the PRS analysis provide tentative evidence for an association between the genetics of PTSD and reported trauma in MDD. In the genetic correlation analysis, all PTSD phenotypes showed greater genetic overlap with the MDD with reported trauma phenotype compared to the MDD without reported trauma phenotype. However, the differences between the correlations were not significant so it is difficult to fully answer our research question. However, the lack of significance may be due to the limited power of the GWAS from which the summary statistics originated. Given the PRS results, it is possible that the greater genetic correlation between PTSD and MDD with reported trauma, compared to MDD without reported trauma, might be significant if the MDD summary statistics had been produced from better powered GWAS. The Genetic Links to Anxiety and Depression (GLAD) study, which aims to recruit 40 000 participants, will provide an opportunity to replicate these analyses with sufficient power to understand whether the differences were due to chance.

We note that UK Biobank participants who met the criteria for recurrent MDD reported significantly higher rates of trauma exposure in comparison to individuals who met the criteria for single-episode MDD. This corroborates previous psychiatric research that pinpoints exposure to stressful or traumatic events as a key risk factor for subsequent recurrent MDD (Hovens et al., 2015; Nanni et al., 2012; Otte et al., 2016). We expected PTSD to show a greater genetic correlation with the recurrent MDD phenotype compared to the single-episode MDD phenotype but found no evidence of this in the genetic correlation analysis. Findings from the PRS analysis show that genetic risk for PTSD showed stronger associations with single-episode MDD although the effect was small and not significant.

An interesting finding, which was consistent across UKB-PTSD and PGC1.5-PTSD (and PGC2-PTSD presented in online Supplementary Results), was the significantly higher genetic correlation between PTSD and recurrent MDD compared to PTSD and MDD without reported trauma. This might reflect similarities between PTSD and recurrent MDD. It is known that exposure to trauma, especially in childhood, is related to MDD that is severe and treatment-resistant, as well as recurrent, in later life (Danese, 2020; Nanni et al., 2012). Potentially, in terms of MDD subtypes, MDD without reported trauma may capture participants with milder symptoms, while the recurrent MDD subtype may capture severer symptoms due to the higher reported trauma in this category. Like recurrent MDD, PTSD is a severe psychiatric disorder where full, clinically significant symptoms may present many years after exposure (Kessler et al., 2017). Taking this into consideration, PTSD might share genetic variants associated with symptom severity and persistence with recurrent MDD, which may be shared to a lesser extent with MDD in individuals not reporting trauma. This could explain the significant difference between the genetic correlations. However, if this was the case, we would expect the finding to have replicated with MVP-PTSD. This is because war- and combat-related PTSD tends to be more severe and long-lasting compared to PTSD from other traumas (Kessler et al., 2017). Furthermore, war- and combat-related PTSD has been shown to correlate with heightened symptom severity (Guina, Nahhas, Sutton, & Farnsworth, 2018). As shown in Table 1, at least a quarter of the MVP-PTSD sample had been exposed to combat (Stein et al., 2021).

Unique features of the MVP-PTSD sample may explain why this significant finding did not replicate. First, unlike the UK Biobank and PGC samples (which have a more balanced sex division), MVP-PTSD overrepresents males (94.4%). Previous GWAS findings suggest that PTSD's heritability differs between men and women (Nievergelt et al., 2019). Secondly, although PTSD is often severe and disabling, it is not a homogeneous disorder (Smith, Summers, Dillon, & Cougle, 2016). It is well known that the nature of trauma(s) can affect subsequent clinical presentation (Kelley, Weathers, McDevitt-Murphy, Eakin, & Flood, 2009). War- and combat-related trauma has been found to be particularly associated with intrusive symptoms and arousal, such as excessive startle and physical reactivity (Guina et al., 2018). It is possible that the type of PTSD measured by the MVP in war veterans differs from that measured in civilians (such as the participants in UKB-PTSD and some participants in the PGC samples), which may alter its genetic sharing with internalising disorders such as MDD.

Merits and limitations

We were able to use a variety of PTSD definitions and data from the largest PTSD GWAS to date. These samples recruited participants who had experienced different types of trauma, exhibited varying levels of severity and were recruited in distinct ways. To participate in the UK Biobank, individuals visited recruitment centres for a number of hours to undergo physical assessments, provide data and a DNA sample (Sudlow et al., 2015). This level of investment may mean that people who were experiencing severe emotional and functional impairment were unlikely to participate. Contrastingly, the majority of the PGC1.5-PTSD participants were recruited directly from clinical studies of PTSD, using telephone diagnostic interviews and face-to-face clinical assessments (Nievergelt et al., 2019). Consequently, it is reasonable to assume that, on average, the participants comprising this sample report more severe symptoms than individuals drawn from the population without specific ascertainment for mental illness (as is the case with the UK Biobank). In contrast to the UK Biobank and PGC, the MVP sample was limited to US veterans (Stein et al., 2021). The benefit of using varying PTSD phenotypes was that it allowed us to examine whether the extent that trauma sensitivity is shared between MDD and PTSD depends on sample-specific characteristics. We saw that the significantly

https://doi.org/10.1017/S0033291721000830 Published online by Cambridge University Press

higher genetic correlation between PTSD and recurrent MDD compared to PTSD and MDD without reported trauma replicated when using PGC1.5-PTSD, suggesting this result is not only applicable to UK Biobank participants with probable PTSD but also to clinically defined PTSD. Likewise, the general pattern of the genetic correlation results was consistent across all PTSD phenotypes.

We note that the MDD phenotypes were defined in UK Biobank participants who show a 'volunteer selection bias' (Fry et al., 2017). This refers to the tendency of research participants to be more health-conscious and have a higher level of social capital than non-participants (Manolio et al., 2012). Therefore, although the UK Biobank offers the opportunity to amalgamate genetic and phenotypic in a large, homogenous, single-population cohort, its demographic features mean the MDD and trauma-related phenotypes cannot fully represent the experiences of other populations.

The interpretation of the results in this study is also affected by the fact that trauma exposure was measured retrospectively. This, and the older age of UK Biobank participants, may lead to inaccurate reporting of events (Colman et al., 2016). Secondly, the lack of temporal information regarding the onset of MDD in relation to traumatic experiences means we cannot infer causality between them. To overcome this, we could have limited the definition of trauma to the three childhood items which would have allowed a more robust measurement of the influence of trauma exposure on the later development of MDD. However, Coleman et al. (2020) reported that limiting the GWAS of MDD with reported trauma to only the events in the Childhood Trauma Screener did not significantly alter the SNP-based heritability of MDD, suggesting that the inclusion of the adulthood events is valid when investigating the relationship between reported trauma and MDD (Coleman et al., 2020). Overall, although this method of measuring trauma exposure is not ideal, it is the only feasible method for collecting large amounts of data required for genomic analyses.

Lastly, our results may not generalise to non-European populations. This limitation, which means the experiences of non-European individuals fail to be accounted for in genetics research, is increasingly being acknowledged. A recent PTSD GWAS from the MVP included individuals of African ancestries (Stein et al., 2021). Sample sizes are currently small but will hopefully grow as the field responds to the need for inclusivity and diversity in its research.

Overall, our findings tentatively point toward the existence of genetic variants which may interact with life events and influence the risk of experiencing internalising symptoms, although replication with better powered GWAS is needed. Our paper makes a step towards understanding the nature of trauma sensitivity in individuals with MDD, whether this has a genetic basis and whether this is shared with PTSD.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291721000830

Data. Genome-wide association study (GWAS) summary statistics for the posttraumatic stress disorder (PTSD) phenotypes were obtained from the PTSD working group of the Psychiatric Genomics Consortium and the Million Veteran Program. GWAS summary statistics for the major depressive disorder (MDD) phenotypes were obtained from the corresponding author (J.R.I.C.) at King's College London. Individual-level UK Biobank data are available to bona fide researchers with an approved application.

Acknowledgements. We are indebted to the scientists involved in the construction of the UK Biobank, the Psychiatric Genomics Consortium and the Million Veteran Program, and to the thousands of participants who have shared their life experiences. This research used data derived from the UK Biobank Resource under application 16577 (Professor Gerome Breen). The authors acknowledge the use of the research computing facility at King's College London, *Rosalind* (https://rosalind.kcl.ac.uk), which is delivered in partnership with the National Institute for Health Research (NIHR) Biomedical Research Centres at South London & Maudsley and Guy's & St. Thomas' NHS Foundation Trusts, and part-funded by capital equipment grants from the Maudsley Charity (award 980) and Guy's & St. Thomas' Charity (TR 130505). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, King's College London or the Department of Health and Social Care. Figure 1 was created with BioRender.com.

List of consortium authors. Adam X. Maihofer, Caroline M. Nievergelt, Dewlen G. Baker, Victoria B. Risborough, Joseph R. Calabrese, Sandro Galea, Dan J. Stein, Nastassja Koen, Shareefa Dalvie, Allison E. Aiello, Andrea L. Roberts, Koenen K. C., Nadia Solovieff, Henry R. Kranzler, Hongyu Zhao, Joel Gelernter, Lindsay A. Farrer, Eric Otto Johnson, John P. Rice, Laura J. Bierut, Nancy L. Saccone, Alexander McFarlane, David Forbes, Derrick Silove, Meaghan O'Donnell, Richard A. Bryant, Miranda van Hooff, Scott R. Sponheim, Seth G. Disner, Robert H. Pietrzak, Chia-Yen Chen, Jordan W. Smoller, Murray B. Stein, Robert J. Ursano, Ronald C. Kessler, Angela G. Junglen, Douglas L. Delahanty, Ananda B. Amstadter, Christina M. Sheerin, Ken Ruggiero, Katie A. McLaughlin, Matthew Peverill, Caldas-de-Almeida JM, S. Bryn Austin, Bizu Gelaye, Michelle A. Williams, Sixto E. Sanchez, Carol E. Franz, Matthew S. Panizzon, Michael J. Lyons, William S. Kremen, Ole A. Andreassen, Anders M. Dale, Bart P. F. Rutten, Christiaan Vinkers, Dick Schijven, Elbert Geuze, Eric Vermetten, Jurjen J. Luykx, Marco P. Boks, Allison E. Ashley-Koch, Jean C. Beckham, Melanie E. Garrett, Michael A. Hauser, Michelle F. Dennis, Nathan A. Kimbrel, Xue-Jun Qin, Karen-Inge Karstoft, Soren B. Andersen, Anders D. Borglum, David Michael Hougaard, Jonas Bybjerg-Grauholm, Laramie E. Duncan, Marie Bμkvad-Hansen, Merete Nordentoft, Ole Mors, P. B. Mortensen, Thomas Werge, Wesley K. Thompson, Yunpeng Wang, Andrew C. Heath, Elliot C. Nelson, Nicholas G Martin, Scott D. Gordon, Erika J. Wolf, Mark W. Logue, Mark W. Miller, Regina E. McGlinchey, William Milberg, Christopher R. Erbes, Melissa A. Polusny, Paul A. Arbisi, Alan L. Peterson, Douglas E. Williamson, Keith A. Young, Maurer D., Alex O. Rothbaum, Lori A. Zoellner, Matig R. Mavissakalian, Norah C. Feeny, P. Roy-Byrne, Barbara O. Rothbaum, Jessica Maples-Keller, Alaptagin Khan, Jonathan D. Wolff, Lauren A. M. Lebois, Martin H. Teicher, Milissa L. Kaufman, Sherry Winternitz, Elizabeth A. Bolger, Antonia V. Seligowski, Holly K. Orcutt, Alexandra Evans, Bozo Lugonja, Catrin Lewis, Ian Jones, Jonathan I. Bisson, Samuel A. McLean, Sarah D. Linnstaedt, Edward S. Peters, Ariane Rung, Edward Trapido, Megan Brashear, Adriana Lori, Alicia K. Smith, Bekh Bradley, Charles Gillespie, Guia Guffanti, Kerry J. Ressler, Lynn M. Almli, Tanja Jovanovic, Angela C. Bustamante, Monica Uddin, Leigh Luella van den Heuvel, S. Seedat, S. M. J Hemmings, Aferdita Goci Uka, Alma Dzubur Kulenovic, Christiane Wolf, Dragan Babic, Esmina Avdibegovic, Jürgen Deckert, Katharina Domschke, Miro Jakovljevic, Charles Marmar, Janine D. Flory, Marti Jett, Rachel Yehuda, Rasha Hammamieh, Anthony P. King, Israel Liberzon, Julia S. Seng, Charles Phillip Morris, Divya Mehta, Joanne Voisey, Lawford B. R., Ross McD. Young, Sarah McLeay, Nikolaos P. Daskalakis, Sonya B. Norman, Zhewu Wang, Gerome Breen, Jonathan R.I. Coleman, Torsten Klengel, Katy Torres, Karmel W. Choi, Daniel F Levey, Renato Polimanti, Elizabeth G. Atkinson, Alicia R. Martin, Benjamin M. Neale, Mark J. Daly, Stephan Ripke, Andrew Ratanatharathorn, Allison C. Provost, Magali Haas, Jennifer S. Stevens, Rajendra A. Morey, Supriya Harnal, Preben B. Mortensen, Soraya Seedat, Jennifer A. Sumner.

Author contributions. J.M., J.R.I.C. and G.B. were responsible for study conception and design. J.R.I.C., G.B., M.B.S., The Million Veteran Program and the PTSD working group of the Psychiatric Genomics Consortium were responsible for the acquisition of the data. J.M., J.R.I.C. and M.S. were responsible for data analysis. All authors were involved in the interpretation of the data. J.M. was responsible for the drafting of the paper, under the close supervision of J.R.I.C. and G.B. All authors read, edited and approved the final manuscript before submission. All authors agree to be accountable for all

aspects of the work, and in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated.

Financial support. This study represents independent research funded partly by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London and partly by the Lord Leverhulme Charitable Grant (Professor Robin M. Murray). High-performance computing facilities were funded with capital equipment grants from the Guy's & St. Thomas' Charity (TR 130505) and Maudsley Charity (award 980).

Conflict of interest. None.

Ethical standards. This study has been completed under UK Biobank approved study application 16577 (Professor Gerome Breen). The UK Biobank is approved by the North West Multi-centre Research Committee. All procedures performed in studies involving human participants were in accordance with the ethical standards of this committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethics standards. All participants provided written informed consent to participate in the study.

References

Alim, T. N., Graves, E., Mellman, T. A., Aigbogun, N., Gray, E., Lawson, W., & Charney, D. S. (2006). Trauma exposure, posttraumatic stress disorder and depression in an African-American primary care population. *Journal of the National Medical Association*, *98*, 1630–1636.

American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders* (Vol. *5*). Arlington, VA: American Psychiatric Association.

Auxéméry, Y. (2012). [Posttraumatic stress disorder (PTSD) as a consequence of the interaction between an individual genetic susceptibility, a traumatogenic event and a social context]. *L'Encephale*, *38*, 373–380.

Ben Barnes, J., Hayes, A. M., Contractor, A. A., Nash, W. P., & Litz, B. T. (2018). The structure of co-occurring PTSD and depression symptoms in a cohort of Marines pre- and post-deployment. *Psychiatry Research*, *259*, 442–449.

Breslau, N., Davis, G. C., Andreski, P., & Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry*, *48*, 216–222.

Breslau, N., Davis, G. C., Peterson, E. L., & Schultz, L. (1997). Psychiatric sequelae of posttraumatic stress disorder in women. *Archives of General Psychiatry*, *54*, 81–87.

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, … Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, *47*, 291–295.

Coleman JRI, Gaspar HA, Bryois J, Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, & Breen G (2019) The genetics of the mood disorder spectrum: Genome-wide association analyses of more than 185000 cases and 439000 controls. *Biological Psychiatry*, *88*, 169–184.

Coleman, J. R. I., Peyrot, W. J., Purves, K. L., Davis, K. A. S., Rayner, C., Choi, S. W., … Breen, G. (2020). Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Molecular Psychiatry*, *25*, 1430–1446.

Colman, I., Kingsbury, M., Garad, Y., Zeng, Y., Naicker, K., Patten, S., … Thompson, A. H. (2016). Consistency in adult reporting of adverse childhood experiences. *Psychological Medicine*, *46*, 543–549.

Colodro-Conde, L., Couvy-Duchesne, B., Zhu, G., Coventry, W. L., Byrne, E. M., Gordon, S., … Martin, N. G. (2018). A direct test of the diathesis-stress model for depression. *Molecular Psychiatry*, *23*, 1590–1596.

Dalvie, S., Maihofer, A. X., Coleman, J. R. I., Bradley, B., Breen, G., Brick, L. A., … Nievergelt, C. M. (2020). Genomic influences on self-reported childhood maltreatment. *Translational Psychiatry*, *10*, 38.

Danese, A. (2020). Annual research review: Rethinking childhood trauma-new research directions for measurement, study design and analytical strategies. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *61*, 236–250.

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, *9*, e1003348.

Duncan, L. E., Cooper, B. N., & Shen, H. (2018a). Robust findings from 25 years of PTSD genetics research. *Current Psychiatry Reports*, *20*, 115.

Duncan, L. E., Ratanatharathorn, A., Aiello, A. E., Almli, L. M., Amstadter, A. B., Ashley-Koch, A. E., … Koenen, K. C. (2018b). Largest GWAS of PTSD (N = 20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Molecular Psychiatry*, *23*, 666–673.

Elhai, J. D., de Francisco Carvalho, L., Miguel, F. K., Palmieri, P. A., Primi, R., & Christopher Frueh, B. (2011). Testing whether posttraumatic stress disorder and major depressive disorder are similar or unique constructs. *Journal of Anxiety Disorders*, *25*, 404–410.

Euesden, J., Lewis, C. M., & O'Reilly, P. F. (2015). PRSice: Polygenic risk score software. *Bioinformatics (Oxford, England)*, *31*, 1466–1468.

Flory, J. D., & Yehuda, R. (2015). Comorbidity between post-traumatic stress disorder and major depressive disorder: Alternative explanations and treatment considerations. *Dialogues in Clinical Neuroscience*, *17*, 141–150.

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., … Allen, N. E. (2017). Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. *American Journal of Epidemiology*, *186*, 1026–1034.

Grubaugh, A. L., Long, M. E., Elhai, J. D., Frueh, B. C., & Magruder, K. M. (2010). An examination of the construct validity of posttraumatic stress disorder with veterans using a revised criterion set. *Behaviour Research and Therapy*, *48*, 909–914.

Guina, J., Nahhas, R. W., Sutton, P., & Farnsworth, S. (2018). The influence of trauma type and timing on PTSD symptoms. *The Journal of Nervous and Mental Disease*, *206*, 72–76.

Horesh, N., Klomek, A. B., & Apter, A. (2008). Stressful life events and major depressive disorders. *Psychiatry Research*, *160*, 192–199.

Hovens, J. G. (2015) Emotional scars: Impact of childhood trauma on depressive and anxiety disorders, Universiteit Leiden, Belgium.

Hovens, J. G. F. M., Giltay, E. J., Spinhoven, P., van Hemert, A. M., & Penninx, B. W. J. H. (2015). Impact of childhood life events and childhood trauma on the onset and recurrence of depressive and anxiety disorders. *The Journal of Clinical Psychiatry*, *76*, 931–938.

Jay Schulz-Heik, R., Rhee, S. H., Silvern, L., Lessem, J. M., Haberstick, B. C., Hopfer, C., & Hewitt, J. K. (2009). Investigation of genetically mediated child effects on maltreatment. *Behavior Genetics*, *39*, 265–276.

Kelley, L. P., Weathers, F. W., McDevitt-Murphy, M. E., Eakin, D. E., & Flood, A. M. (2009). A comparison of PTSD symptom patterns in three types of civilian trauma. *Journal of Traumatic Stress*, *22*, 227–235.

Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). The assessment of dependence in the study of stressful life events: Validation using a twin design. *Psychological Medicine*, *29*, 1455–1460.

Kessler, R. C. (1997). The effects of stressful life events on depression. *Annual Review of Psychology*, *48*, 191–214.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Benjet, C., Bromet, E. J., Cardoso, G., … Koenen, K. C. (2017). Trauma and PTSD in the WHO world mental health surveys. *European Journal of Psychotraumatology*, *8*, 1353383.

Kessler RC, Sonnega A, Bromet E, Hughes M, & Nelson CB (1995) Posttraumatic stress disorder in the national comorbidity survey. jamanetwork.com *Archives of General Psychiatry 52*, 1048–1060.

Koenen, K. C., Fu, Q. J., Ertel, K., Lyons, M. J., Eisen, S. A., True, W. R., … Tsuang, M. T. (2008). Common genetic liability to major depression and posttraumatic stress disorder in men. *Journal of Affective Disorders*, *105*, 109–115.

Manolio, T. A., Weis, B. K., Cowie, C. C., Hoover, R. N., Hudson, K., Kramer, B. S., … Collins, F. S. (2012). New models for large prospective studies: Is there a better way? *American Journal of Epidemiology*, *175*, 859–866.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *The American Psychologist*, *17*, 827–838.

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, *110*, 406–425.

Mullins, N., Power, R. A., Fisher, H. L., Hanscombe, K. B., Euesden, J., Iniesta, R., … Lewis, C. M. (2016). Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychological Medicine*, *46*, 759–770.

Nanni, V., Uher, R., & Danese, A. (2012). Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: A meta-analysis. *The American Journal of Psychiatry, 169*, 141–151.

Nievergelt, C. M., Maihofer, A. X., Klengel, T., Atkinson, E. G., Chen, C.-Y., Choi, K. W., … Koenen, K. C. (2019). International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nature Communications, 10*, 4558.

Ning, Z., Pawitan, Y., & Shen, X. (2020). High-definition likelihood inference of genetic correlations across human complex traits. *Nature Genetics, 52*, 859–864.

Otte, C., Gold, S. M., Penninx, B. W., Pariante, C. M., Etkin, A., Fava, M., … Schatzberg, A. F. (2016). Major depressive disorder. *Nature Reviews. Disease Primers, 2*, 16065.

Peyrot, W. J., Milaneschi, Y., Abdellaoui, A., Sullivan, P. F., Hottenga, J. J., Boomsma, D. I., & Penninx, B. W. J. H. (2014). Effect of polygenic risk scores on depression in childhood trauma. *The British Journal of Psychiatry: the Journal of Mental Science, 205*, 113–119.

Plomin, R., Lichtenstein, P., Pedersen, N. L., McClearn, G. E., & Nesselroade, J. R. (1990). Genetic influence on life events during the last half of the life span. *Psychology and Aging, 5*, 25–30.

Power, R. A., Wingenbach, T., Cohen-Woods, S., Uher, R., Ng, M. Y., Butler, A. W., … McGuffin, P. (2013). Estimating the heritability of reporting stressful life events captured by common genetic variants. *Psychological Medicine, 43*, 1965–1971.

Reuben, A., Moffitt, T. E., Caspi, A., Belsky, D. W., Harrington, H., Schroeder, F., … Danese, A. (2016). Lest we forget: Comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 57*, 1103–1112.

Rytwinski, N. K., Scur, M. D., Feeny, N. C., & Youngstrom, E. A. (2013). The co-occurrence of major depressive disorder among individuals with posttraumatic stress disorder: A meta-analysis. *Journal of Traumatic Stress, 26*, 299–309.

Sartor, C. E., Grant, J. D., Lynskey, M. T., McCutcheon, V. V., Waldron, M., Statham, D. J., … Nelson, E. C. (2012). Common heritable contributions to low-risk trauma, high-risk trauma, posttraumatic stress disorder, and major depression. *Archives of General Psychiatry, 69*, 293–299.

Shah, R., Shah, A., & Links, P. (2012). Post-traumatic stress disorder and depression comorbidity: Severity across different populations. *Neuropsychiatry, 2*, 521–529.

Shapero, B. G., Black, S. K., Liu, R. T., Klugman, J., Bender, R. E., Abramson, L. Y., & Alloy, L. B. (2014). Stressful life events and depression symptoms: The effect of childhood emotional abuse on stress reactivity. *Journal of Clinical Psychology, 70*, 209–223.

Smith, H. L., Summers, B. J., Dillon, K. H., & Cougle, J. R. (2016). Is worst-event trauma type related to PTSD symptom presentation and associated features? *Journal of Anxiety Disorders, 38*, 55–61.

Stein, M. B., Levey, D. F., Cheng, Z., Wendt, F. R., Harrington, K., Cho, K., … Gelernter, J. (2021). Genome-wide association analyses of post-traumatic stress disorder and its symptom subdomains in the Million Veteran Program. *Nature Genetics, 53*, 174–184.

Stein, M. B., McQuaid, J. R., Pedrelli, P., Lenox, R., & McCahill, M. E. (2000). Posttraumatic stress disorder in the primary care medical setting. *General Hospital Psychiatry, 22*, 261–269.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., … 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 Medicine, 12*, e1001779.

van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H., & Wray, N. R. (2019). Genetic correlations of polygenic disease traits: From theory to practice. *Nature Reviews. Genetics, 20*, 567–581.

Wolf, E. J., Miller, M. W., Krueger, R. F., Lyons, M. J., Tsuang, M. T., & Koenen, K. C. (2010). Posttraumatic stress disorder and the genetic structure of comorbidity. *Journal of Abnormal Psychology, 119*, 320–330.

https://doi.org/10.1017/S0033291721000830 Published online by Cambridge University Press

Chapter 3. Genetic examination of the Mood Disorder Questionnaire and its relationship with bipolar disorder

This chapter is an exact copy of an article accepted pending minor revisions in *Neuropsychiatric Genetics Part B, of the American Medical Journal of Medical Genetics*.

Supplementary material is included in **appendix 3**.

Authors:

Jessica Mundy[1,2], Christopher Hübel[1,2,3], Brett N. Adey[1,2], Helena L. Davies[1,2], Molly R. Davies[1,2], Jonathan R. I. Coleman[1,2], Matthew Hotopf[1,2,4], Gursharan Kalsi[1,2], Sang Hyuck Lee[1,2], Andrew M. McIntosh[6], Henry C. Rogers[1,2], Thalia C. Eley[1,2], Robin M. Murray[1], Evangelos Vassos[1,2], Gerome Breen[1,2]

Author affiliations:

1. Institute of Psychiatry, Psychology and Neuroscience, King's College London, Denmark Hill, Camberwell, London, United Kingdom
2. UK National Institute for Health and Care Research (NIHR) Biomedical Research Centre, South London and Maudsley Hospital, London, United Kingdom
3. National Centre for Register-based Research, Aarhus Business and Social Sciences, Aarhus University, Aarhus, Denmark
4. South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital, Monks Orchard Road, Beckenham, Kent, United Kingdom
5. MRC Centre for Neuropsychiatric Genetics and Genomics, National Centre for Mental Health, Cardiff University, Cardiff, United Kingdom
6. Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom

Abstract

Background: The Mood Disorder Questionnaire (MDQ) is a common screening tool for bipolar disorder, assessing manic symptoms. Its utility for genetic studies of mania or bipolar traits has not been fully examined.

Methods: We psychometrically compared the MDQ to self-reported bipolar disorder in participants from the United Kingdom National Institute of Health and Care Research Mental Health BioResource. We conducted genome-wide association studies of manic symptom quantitative traits and symptom subgroups, derived from the MDQ items (N=11,568-19,859). We calculated genetic correlations with bipolar disorder and other psychiatric and behavioural traits.

Results: The MDQ screener showed low positive predictive value (0.29) for self-reported bipolar disorder. Neither concurrent nor lifetime manic symptoms were genetically correlated with bipolar disorder. Lifetime manic symptoms had a highest genetic correlation (r_g=1.0) with posttraumatic stress disorder although this was not confirmed by within-cohort phenotypic correlations (r_p=0.41). Other significant genetic correlations included attention deficit hyperactivity disorder (r_g=0.69), insomnia (r_g=0.55), and major depressive disorder (r_g=0.42).

Conclusion: Our study adds to existing literature questioning the MDQ's validity and suggests it may capture symptoms of general distress or psychopathology, rather than hypomania/mania specifically, in at-risk populations.

Introduction

Mania involves periods of elevated, expansive, or irritable mood. Symptoms may include feeling energetic or hyperactive, having unusually inflated self-confidence, requiring little or no sleep, or engaging in behaviour that some might consider impulsive or risky. Hypomania involves these symptoms but to a milder degree (American Psychiatric Association, 2013). Mania and hypomania generally alternate with episodes of depressed mood in bipolar disorder type I and type II, respectively. The lifetime prevalences of these psychiatric disorders are 0.6% and 0.4%, respectively (Merikangas *et al.*, 2011).

Individuals with bipolar disorder usually present with other psychiatric symptoms in the first instance, particularly depression (Musliner and Østergaard, 2018) and the time between initial presentation to services and receiving the correct diagnosis can be over ten years (Lish *et al.*, 1994; Hirschfeld, Lewis and Vornik, 2003). Therefore, some individuals with psychiatric problems, especially depression, anxiety, and substance use disorders, may later develop hypomania or mania (Zimmerman *et al.*, 2009; Kessing *et al.*, 2017; Baryshnikov *et al.*, 2020). Identifying "hidden" bipolar disorder patients can aid clinicians in diagnosis and earlier prescribing of correct medication (e.g., mood stabilisers) which reduces risk of antidepressant-induced mania, rapid cycling, and the costs associated with delayed treatment (Hirschfeld, 2010; Zimmerman, 2012).

Several screening tools have been developed to assess possible hypomania/mania in at-risk individuals. One of the most widely used is the Mood Disorder Questionnaire (MDQ) which involves questions about thirteen aspects of mania, symptom duration, functional impairment, and family history of bipolar disorder (Hirschfeld, 2010). The MDQ has been translated into 16 languages and is used globally (Zimmerman *et al.*, 2009; Hirschfeld, 2010). Bipolar disorder is heritable (Stahl *et al.*, 2019; Mullins *et al.*, 2021), and the MDQ has been applied in genetic studies of mania with mixed results. A twin study showed that the heritable basis of MDQ-assessed hypomania was moderately correlated with the heritable basis for bipolar disorder (r_g=0.40) in a non-clinical youth sample, which mirrored the phenotypic correlation (r_{ph}=0.39). However, hypomania did not show a significant correlation with bipolar disorder polygenic risk scores (PRSs) based on common genetic variants (Hosang *et al.*, 2021). Thus, the genetic basis of the symptoms assessed by the MDQ, and their relationship with bipolar disorder, warrant further examination.

The Genetic Links to Anxiety and Depression (GLAD) study, a nationwide resource of participants with a lifetime occurrence of depression and anxiety disorders, used the MDQ to assess for lifetime presence of hypomanic/manic symptoms. Participants were also asked about whether they had received a diagnosis of bipolar disorder by a professional. This large cohort of participants with mental health disorders and genetic data available offers an opportunity to examine the validity of the MDQ among individuals who are at increased risk of developing bipolar disorder.

Here, we investigated the validity of the MDQ. First, we assessed the psychometric properties of the MDQ as a screener in our sample, based on self-reported diagnoses of bipolar disorder. Since self-reported diagnoses of mental health disorders may be inaccurate (Davies *et al.*, 2022), we performed an extra step to validate the MDQ using genomic methods. We calculated the genetic correlation between the number of manic symptoms that participants reported in the MDQ and the largest available GWAS of bipolar disorder from the Psychiatric Genomics Consortium (PGC) (Mullins *et al.*, 2021). All cases included in the PGC GWAS (N=41,917) met DSM-IV or ICD-9/10 criteria for bipolar disorder, made from diagnostic interviews, clinician-administered checklists, or medical records. Thus, genetic correlations with this GWAS can be used to assess the external validity of the MDQ as a method of assessing hypomania/mania.

We assessed the MDQ items as a quantitative score in two ways: a) concurrent symptoms during one time period and b) cumulative symptoms across the lifetime (not specifying co-occurrence). Since nine of the MDQ items equate to the diagnostic criteria for hypomania/mania in the DSM-5, we hypothesised that both quantitative measures would show significant positive genetic correlations, a measure of the relationship between two polygenic phenotypes (van Rheenen *et al.*, 2019), with bipolar disorder. We expected the genetic correlation to be greater with the measure of concurrent symptoms because experiencing multiple symptoms within four days or one week is a requirement for a hypomania and mania diagnosis respectively (American Psychiatric Association, 2013) and a positive screen in the MDQ is made on the basis of concurrent symptoms (Hirschfeld *et al.*, 2000). In addition to bipolar disorder, we calculated genetic correlations with 34 other psychiatric and behavioural traits. We expected MDQ-assessed manic symptoms to have a higher genetic correlation with bipolar disorder compared to the other traits tested.

Since symptoms which collectively underlie a quantitative trait may vary in terms of their biology (Nagel *et al.*, 2018; Thorp *et al.*, 2020), we hypothesised that genetic risk for symptom subgroups of the MDQ, identified from factor analyses, would show genetic heterogeneity. We made no *a priori* predictions about the direction or strength of the overlap.

Methods

All code is available on GitHub (https://github.com/tnggroup/genetics_MDQ). This study was pre-registered on the Open Science Framework.

Study design

Data were examined from participants from the mental health arm of the National Institute of Health and Care Research (NIHR) BioResource in the UK. The largest group of participants were recruited via the Genetic Links to Anxiety and Depression (GLAD) Study (https://gladstudy.org.uk/), an online research platform for individuals with lifetime anxiety and/or major depressive disorder (MDD) (Davies *et al.*, 2019). Recruitment into GLAD began in September 2018 and was conducted via social media campaigns and NHS sites. Other participants were from the COVID-19 Psychiatry and Neurological Genetics (COPING) Study (https://gladstudy.org.uk/all-projects/current-projects/coping-study/). These individuals were initially recruited into the NIHR BioResource from various cohorts via several means (listed in **table S3.18a and S3.18b)**. They were secondarily invited into the COPING study (henceforth referred to as "COPING NBR participants").

Individuals were eligible to participate if they were aged 16+ and lived in the UK. GLAD participants were additionally required to have experienced MDD or an anxiety disorder in their lifetime. All participants provided demographic information, mental health histories, and some provided a saliva or blood sample. The COPING baseline survey comprised many of the same questionnaires from the GLAD sign-up survey which allowed for parallel assessments. All questionnaires were acquired using Qualtrics survey software (Qualtrics, Provo, UT). We analysed data from participants who completed the GLAD study sign-up survey or COPING baseline survey between 17th September 2018 and 3rd September 2021.

Study sample

We analysed data from individuals with experience of MDD and/or anxiety. COPING NBR participants who met symptom-based diagnostic criteria for MDD and/or any anxiety disorder were combined with GLAD participants to create a cohort who had been affected by these

common mental health disorders (**figure S3.28**). To compare the distribution of manic symptoms between those affected and unaffected by MDD and/or anxiety, we also measured lifetime MDQ items in participants with no history of these disorders. COPING NBR participants who did not meet criteria for MDD or any anxiety disorder were categorised as “unaffected participants”. COPING NBR participants without the data required to determine MDD or anxiety diagnosis were excluded (details on symptom-based diagnostic criteria in supplementary methods).

Ethics

Full informed consent was obtained from all participants. Ethical approval for the GLAD study was granted by the London-Fulham Research Ethics Committee (REC reference: 18/LO/1218) and for COPING, by the NHS Health Research Authority, South West - Central Bristol Research Ethics Committee (20/SW/0078).

Measures

The Mood Disorder Questionnaire

Lifetime experience of 13 hypomanic/manic symptoms were assessed via the MDQ (**table 3.1**) (henceforth “lifetime manic symptoms”). Participants who endorsed any of the lifetime manic symptoms were then presented with the question *“You ticked 'yes' to more than one of the previous symptoms - have several of these ever happened during the same period of time?”*. The participants who answered *“Yes"* were subsequently presented with a list of their previously endorsed symptoms and were asked to *“select all that occurred during the same period of time”* (henceforth “concurrent manic symptoms”).

For analyses of lifetime manic symptoms, GLAD and COPING NBR participants with complete data on all MDQ items were retained for analyses. For analyses of concurrent manic symptoms, only GLAD participants who reported more than one concurrent symptom and had complete data on all items were included in analyses. Data on concurrent manic symptoms were not available in COPING NBR.

See the supplementary methods for more information on how the MDQ screener was constructed.

Quantitative manic symptom phenotypes

The number of concurrent manic symptoms endorsed by participants affected by MDD and/or an anxiety disorder were summed. This quantitative phenotype represents the total number of MDQ items that a participant self-reported having experienced during one time period. Additionally, the number of endorsed lifetime manic symptoms were summed. This quantitative phenotype represents the total number of MDQ items that a participant self-reported having experienced during their lifetime.

Table 3.1. Hypomanic/manic symptoms assessed by the Mood Disorder Questionnaire (MDQ).

Question in MDQ	Abbreviated name	Endorsement (concurrent in affected participants) *N=30,342*	Endorsement (lifetime in affected participants) *N=47,787*	Endorsement (lifetime in unaffected participants) *N=6,308*
...you felt so good or so hyper that other people thought you were not your normal self or you were so hyper that you got into trouble?	Hyperactivity	32.3%	36.2%	2.4%
...you were so irritable that you shouted at people or started fights or arguments?	Irritability	61.8%	71.3%	17.5%
...you felt much more self-confident than usual?	More self-confidence	34.3%	38.0%	8.4%
...you got much less sleep than usual and found you didn't really miss it?	Decreased sleep	53.9%	40.7%	11.9%
...you were much more talkative or spoke much faster than usual?	More talkative	42.1%	43.2%	5.0%

...thoughts raced through your head or you couldn't slow your mind down?	Racing thoughts	77.6%	73.9%	15%
...you were so easily distracted by things around you that you had trouble concentrating or staying on track?	Concentration difficulties	70.8%	71.8%	14.3%
...you had much more energy than usual?	More energy	32.6%	35.7%	8.3%
...you were much more active or did many more things than usual?	More active	30.0%	37.3%	11.7%
...you were much more social or outgoing than usual, for example, you telephoned friends in the middle of the night?	More sociable	20.7%	21.7%	2.0%
...you were much more interested in sex than usual?	Higher libido	28.7%	33.9%	7.4%
...you did things that were unusual for you or that other people might have thought were excessive, foolish, or risky?	Risky behaviour	36.7%	36.8%	3.2%
...spending money got you or your family into trouble?	Reckless spending	28.3%	29.3%	1.9%

Note. "Affected" and "unaffected" refers to participants affected and unaffected by major depressive disorder (MDD) and/or an anxiety disorder. Each lifetime item is preceded with the question "Has there ever been a period of time when you were not your usual self and…". Endorsement refers to the % of participants in the three analytical groups who endorsed the item. Participants who endorsed any of the lifetime MDQ items were then presented with the question "You ticked 'yes' to more than one of the previous symptoms - have several of these ever happened during the same period of time?". The participants who answered "Yes" were subsequently presented with a list of their previously endorsed MDQ items and were asked to "select all that occurred during the same period of time". Note the difference in N between concurrent and lifetime groups (participants were excluded from the concurrent item analysis if they reported that fewer than two MDQ items occurred in the same time period). The reduction in overall N for this analytical group means that some of the concurrent symptoms appear to have a higher endorsement than the lifetime symptoms.

Manic symptom subgroups

To identify symptom subgroups in the MDQ, we performed factor analyses of the concurrent and lifetime MDQ items (details in supplementary methods).

We performed exploratory factor analysis (EFA) on 70% of each sample and selected the model with the best fit statistics. We then performed confirmatory factor analysis (CFA) on the remaining 30%. The CFA model was predefined to that identified by EFA which provided a more stringent test of model fit compared to EFA. Factor scores were computed for each factor in the best-fitting model in the whole sample. Factor scores were transformed using a rank-based inverse normal transformation and then standardised (mean=0, standard deviation [SD]=1). All factor analyses were performed in R (details in supplementary methods). We also performed factor analysis of lifetime MDQ items in COPING NBR participants who were unaffected by MDD and/or an anxiety disorder (supplementary material).

Validation of the Mood Disorder Questionnaire

Phenotypic validation

First, to assess the validity of the MDQ in our study sample, we calculated its sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) based on the participants' self-reported diagnoses of bipolar disorder by a professional. Note that these could only be calculated in GLAD participants. Second, based on the number of reported manic symptoms, we compared the mean reported items between participants who self-reported a diagnosis of bipolar disorder and those who self-reported no diagnosis. See supplementary methods for details on self-reported bipolar disorder diagnosis and the MDQ as a screener.

Genetic validation

We calculated the genetic correlation between the quantitative manic symptom phenotypes assessed via the MDQ with the largest GWAS of bipolar disorder from the Psychiatric Genomics Consortium (PGC) (N_{cases}=41,917, $N_{controls}$=371,549) (Mullins *et al.*, 2021). In addition to bipolar disorder, we calculated genetic correlations with 34 other psychiatric and behavioural traits (**table S3.11**). To do this, we first had to perform genome-wide association studies (GWASs) of the quantitative manic symptom phenotypes and the symptom subgroups.

Genotyping, imputation, and quality control

All data from GLAD and COPING NBR were genotyped by ThermoFisher on the Affymetrix UK Biobank Axiom Array v1 and v2 across numerous genotyping batches. Genetic data for GLAD and COPING NBR cohorts were separately subjected to quality control (QC) using the same pipeline (supplementary methods). For our specific analyses, additional QC was carried out. This included removing genotyped SNPs if missingness >5%, minor allele frequency (MAF)<0.01, or Hardy Weinberg Equilibrium $p<10^{-10}$. SNPs imputed with low confidence (INFO<0.3) were also excluded. Individuals with missingness >5%, a mismatch between their self-reported assigned sex at birth and genetic sex, or whose genetic sex could not be determined were excluded. Lastly, one participant of each pair of duplicated participants between the GLAD and COPING NBR cohorts was excluded.

Genome-wide association studies (GWASs)

GWASs were conducted with a mixed linear model using REGENIE, which controls for between-subject relatedness using whole-genome regression (Mbatchou *et al.*, 2021). We included the first ten ancestry principal components and genotyping batch as covariates (principal component analysis plots in **figure S3.18**). We performed GWASs of the quantitative manic symptom phenotypes in participants who were affected by MDD and/or an anxiety disorder. First, we performed GWASs of the total number of concurrent manic symptoms and of the factor scores for each subgroup identified by the factor analysis in participants of European ancestry. Second, we performed GWASs of lifetime manic symptoms and of the factor scores for each subgroup identified by the FA in participants of European ancestry.

SNP-based heritability and genetic correlations

Linkage Disequilibrium Score Regression (LDSC) (Bulik-Sullivan *et al.*, 2015) was used to estimate the SNP-based heritability (h^2_{SNP}) of each manic symptom phenotype. The SNP-based heritability estimates were statistically significant if their p-value surpassed the Bonferroni-corrected alpha of 0.006 ($\alpha = \frac{\mathbf{0.05}}{\mathbf{8}}$) which adjusted for the eight heritability estimates. LDSC was then used to calculate genetic correlations (r_g) between each of the manic symptom phenotypes and 37 GWAS summary statistics of psychiatric and behavioural traits (**table S3.11**) using the

extended 1000 Genomes linkage disequilibrium LD scores. These traits were selected from our internal GWAS summary statistics database. We only included traits that were sufficiently powered (a heritability z-score>4 and a mean chi-square>1.02). To reduce the multiple testing burden, we selected the most well-powered GWASs when more than one trait, or similar traits, were available. Summary statistics were munged using LDSC and a list of SNPs from the extended 1000 Genomes phase three reference panel (1000 Genomes Project Consortium *et al.*, 2015)

Eight sets of 37 genetic correlations were computed for each manic symptom phenotype. Within each set, the alpha value was adjusted to correct for multiple testing using the Bonferroni method giving an alpha value of 0.001 ($\alpha = \frac{\mathbf{0.05}}{\mathbf{37}}$).

We also calculated inter-genetic correlations between the manic symptom phenotypes. Genetic correlations were significant if the p-value surpassed the Bonferroni-adjusted alpha of 0.008 ($\alpha = \frac{\mathbf{0.05}}{\mathbf{6}}$) to correct for six sets of inter-genetic correlations.

Differences between genetic correlations

We tested whether the genetics of the symptom subgroups were differentially associated with the genetics of other traits. For traits that were significantly genetically correlated with more than one subgroup and the overall sum score, we used a block-jackknife to calculate the standard error of the difference between pairs of genetic correlations. In a pairwise fashion, we first compared each trait's genetic correlation with the overall sum score to that same trait's genetic correlation with each of the subgroups. Second, we compared each trait's genetic correlation with a particular subgroup to that same trait's genetic correlation with another subgroup. Genetic correlations were significantly different to each other if the block-jackknife p-value surpassed the Bonferroni-adjusted alpha ($\alpha = \frac{\mathbf{0.05}}{\mathbf{15}}$). The block-jackknife method applied to genetic correlations has been described elsewhere (Mundy *et al.*, 2021).

Results

Study sample

A total of 52,108 GLAD and COPING NBR participants met criteria for MDD or any anxiety disorder. A total of 6,308 COPING NBR participants did not meet criteria for MDD and any anxiety disorder and 13,195 were excluded from analyses for not having complete data needed to determine lifetime MDD or anxiety disorder diagnoses.

Manic symptoms

Quantitative phenotypes were derived from answers to questions in the MDQ about the symptoms alone. A total of 47,787 participants with mood or anxiety disorders (N excluded=4,321) and 6,119 unaffected participants (N excluded=189) had complete data for all 13 lifetime MDQ items. A total of 30,342 GLAD participants had complete data for all 13 concurrent MDQ items and endorsed more than one (flow-chart in **figure S3.28**). After inspecting a correlation matrix, the item “more active” had a correlation >0.8 with the item “more energy” (**figures S3.1-S3.3**). The item “more active” was removed due to the problems associated with including highly collinear items in factor analysis (Flora, Labrish and Chalmers, 2012). After the removal of this item, the total number of concurrent manic symptoms ranged 2-12 and the total number of lifetime manic symptoms ranged 0-12. The N included in the concurrent items analysis dropped to 29,899 after we removed anyone whose sum score was equal to one after the removal of “more active”.

The mean number of concurrent manic symptoms in participants affected by MDD and/or an anxiety disorder was 5.21 (SD=2.70). The mean number of lifetime manic symptoms was 5.30 (SD=3.50) in participants affected by MDD and/or an anxiety disorder and 0.97 (SD=1.63) in unaffected participants. Demographic information for the three study samples for the quantitative manic symptom phenotypes and factor analysis are presented in **table 3.2**.

Manic symptoms and self-reported bipolar disorder diagnosis

Among 27,751 participants with complete data on concurrent MDQ items (range 2-12) and data on bipolar disorder diagnosis status, 2,464 (9%) self-reported a diagnosis. The mean number of

concurrent symptoms reported by participants with a diagnosis was 8.35 (SD=2.86). The mean number reported by participants without a diagnosis was 4.82 (SD=2.44) (t=-59.155, $p<2.2x10^{-16}$).

Among 34,653 participants with complete data on lifetime MDQ items (range 0-12) and data on bipolar disorder diagnosis status, 2,614 (8%) self-reported a diagnosis. The mean number of lifetime symptoms reported by participants with a diagnosis was 10.62 (SD=2.83). The mean number reported by participants without a diagnosis was 6.35 (SD=3.28) (t=-73.101, p-value$<2.2x10^{-16}$). Descriptive statistics for the quantitative MDQ phenotypes in participants with and without a self-reported diagnosis of bipolar disorder are presented in **table S3.1**.

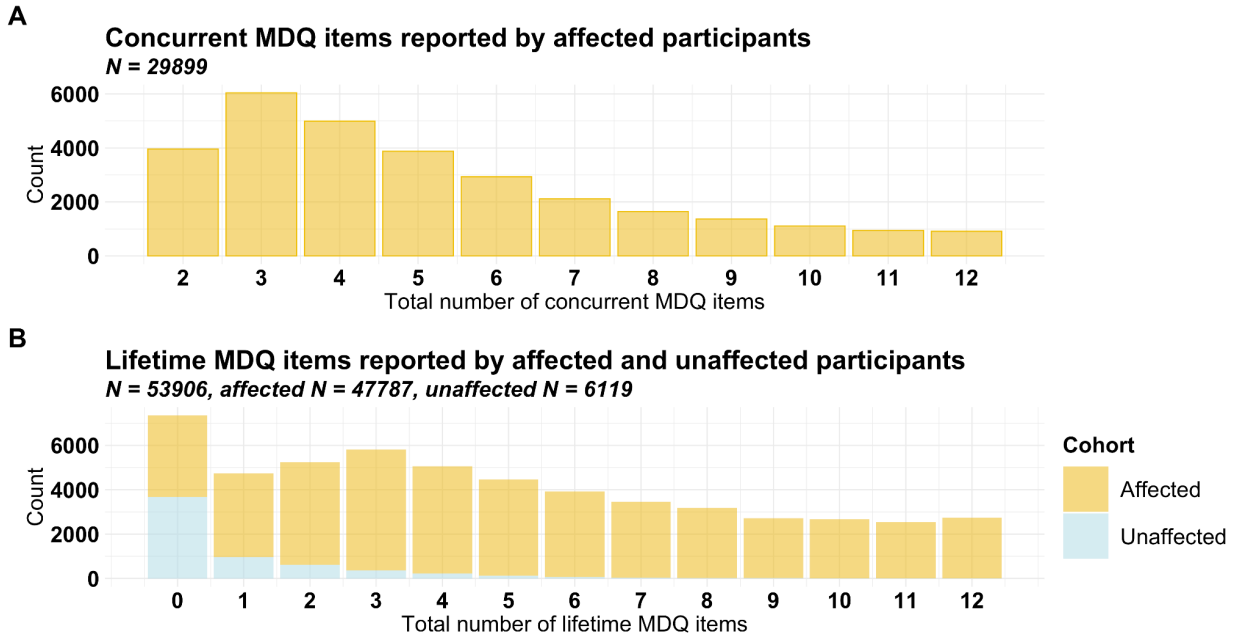
Table 3.2. Demographic information for participants included in the three sets of analyses of the Mood Disorder Questionnaire (MDQ).

From left to right these are: concurrent assessed manic symptoms in participants affected by major depressive disorder (MDD) and/or an anxiety disorder [range 2-12], middle), lifetime manic symptoms in participants affected by MDD and/or an anxiety disorder [range 0-12], and right), and lifetime manic symptoms in participants unaffected by MDD and/or an anxiety disorder [range 0-12].

	Concurrent MDQ items in affected participants	Lifetime MDQ items in affected participants	Lifetime MDQ items in unaffected participants
Variable	N = 29,899[1]	N = 47,787[1]	N = 6,119[1]
MDQ sum score	5.21 (2.70)	5.3 (3.5)	0.97 (1.63)
Age [years]	37 (14)	39 (15)	58 (13)
Sex			
Male	6,117 (20%)	10,143 (21%)	3,049 (50%)
Female	23,782 (80%)	37,644 (79%)	3,070 (50%)
Ethnicity			
White	28,106 (94%)	44,900 (95%)	5,773 (99%)
Mixed	826 (2.8%)	1,119 (2.4%)	16 (0.3%)
Asian or Asian British	427 (1.4%)	668 (1.4%)	45 (0.8%)
Black or Black British	156 (0.5%)	240 (0.5%)	14 (0.2%)
Arab	24 (<0.1%)	43 (<0.1%)	
Other	289 (1.0%)	429 (0.9%)	
Missing	71	388	271
Education			
No university degree	15,036 (51%)	22,049 (47%)	2,817 (47%)
University degree	14,479 (49%)	25,046 (53%)	3,160 (53%)
Missing	384	692	142

[1] Mean (SD); n (%)

Figure 3.1. Distribution of Mood Disorder Questionnaire scores.
A: distribution of the total number of concurrent manic symptoms reported by participants affected by major depressive disorder (MDD) and/or and anxiety disorder in the Mood Disorder Questionnaire (MDQ). B: distribution of the total number of lifetime manic symptoms reported by participants affected (yellow) and unaffected (blue) by major depressive disorder (MDD) and/or and anxiety disorder in the MDQ. Binary MDQ items (1="Yes", 0="No") were summed to create a quantitative sum score. The data originally included 13 items but one of a pair of highly correlated items was removed. Therefore, concurrent items ranged 2-12 and the lifetime items ranged 0-12.



Manic symptom subgroups

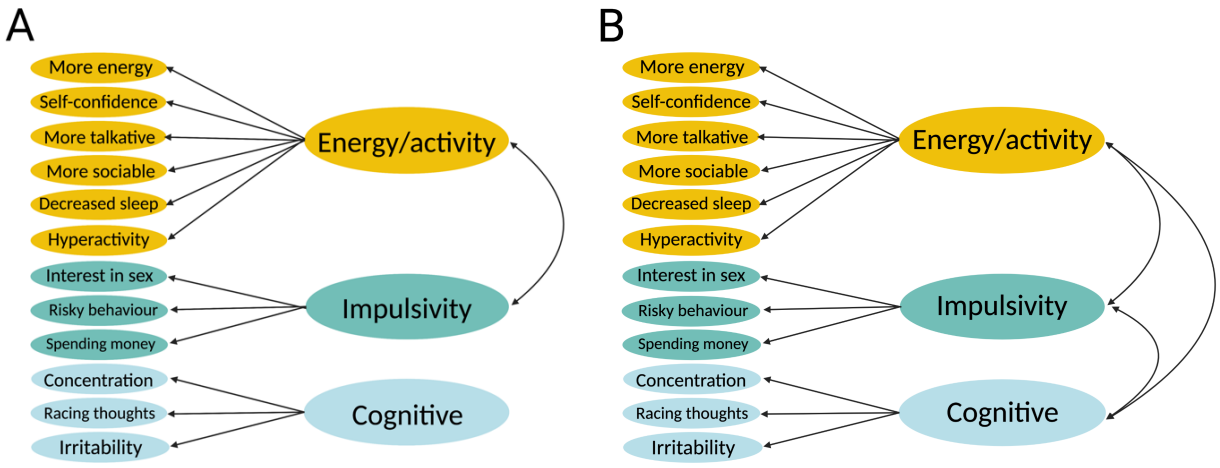
Manic symptom subgroups were identified with factor analyses (details and fit statistics in supplementary results).

Concurrent manic symptom subgroups

A total of 29,899 participants affected by MDD and/or an anxiety disorder were included in the factor analysis of concurrent MDQ items. Despite the scale's Cronbach's alpha being sufficient (**table S3.2**), results showed considerable evidence that the MDQ items, when measured concurrently, lacked internal consistency. Notably, there was a distinct pattern in the item-level correlations indicating that the items "concentration difficulties", "racing thoughts", and "irritability" did not correlate with other items (**figure S3.2**) (details in supplementary results). The 12

concurrent symptoms loaded onto three factors: *energy/activity*, *impulsivity*, and *cognitive*. The *energy/activity* and *impulsivity* factors correlated (r=0.54), but neither correlated with the *cognitive* factor (**figure 3.2; table S3.4b**). The model was confirmed in CFA on the remaining 30% of the sample (N=8,970) and showed good fit statistics (**table S3.5**) (details in supplementary results).

Figure 3.2. Factor analysis models of the Mood Disorder Questionnaire.
A: Simplified diagram of the best-fitting model identified by the exploratory factor analysis (EFA) of 12 concurrent manic symptoms reported by participants by major depressive disorder (MDD) and/or an anxiety disorder in the Mood Disorder Questionnaire (MDQ). The model was confirmed with confirmatory factor analysis (CFA). EFA N=20,929, CFA N=8,970. B: Best-fitting model identified by EFA of 12 lifetime manic symptoms reported by participants affected by MDD and/or an anxiety disorder in the MDQ. Model was confirmed with CFA. EFA N=33,450, CFA N=14,337. Item loadings are presented in figure S3.2 and figure S3.5 and the correlations between the factors are presented in table S3.4b and table S3.7b.



Lifetime manic symptom subgroups

A total of 47,787 participants affected by MDD and/or an anxiety disorder were included in the factor analysis of lifetime MDQ items. After performing EFA in 70% of the sample (N=33,450) on 12 items, a three-factor solution was selected as the final model because it had good fit statistics whilst retaining at least three items per factor (**table S3.7**).

These three lifetime factors perfectly mirrored those identified in the concurrent analysis and were named accordingly (*energy/activity, cognitive, and impulsivity*). However, unlike the concurrent symptoms, the three subgroups correlated with each other (r≥0.55) (**figure 3.2; figure S3.10; table S3.7b**). The model was confirmed in CFA on the remaining 30% of the sample (N=14,337) and showed good fit statistics (**table S3.8**) (details in supplementary results).

These results confirmed that the removal of "more active" from each analysis was justified. From inspecting the correlation matrices, it is clear that "more active" would have loaded onto the *energy/activity*, the largest factor, if it has been included. Therefore, even with this item being removed, energy/activity levels are well assessed by the MDQ.

Validation of the Mood Disorder Questionnaire

Phenotypic validation

A total of 34,479 GLAD participants had complete data on their bipolar disorder diagnosis status and the MDQ screener. Using a cut-off at ≥7 concurrent manic symptoms, the sensitivity of the MDQ screener was 0.58, the specificity was 0.89, the PPV was 0.29, and the NPV was 0.96.

Genetic validation

For the GWASs of concurrent MDQ-assessed manic symptoms, a total of 11,568 participants of European ancestries had genetic data available which passed the genotype and imputed data QC. The SNP-based heritability estimates from LDSC for the four concurrent manic symptom phenotypes ranged 3.8-6.8% but none were significantly different to zero (**table S3.16**).

For the GWASs of lifetime MDQ-assessed manic symptoms, a total of 19,859 participants of European ancestries had available genetic data that passed the genotype and imputed data QC. The SNP-based heritability estimates for the four lifetime manic symptom phenotypes ranged 5.1-7.6% and all were significantly different to zero ($p<0.006$) (**table S3.16**). Manhattan plots and quantile-quantile (QQ) plots, produced by the functional annotation and mapping software FUMA can be found in **figures S3.19-S3.26** (Watanabe *et al.*, 2017).

Genetic correlations with bipolar disorder

Against our hypothesis, we found weak genetic correlations between both of the quantitative manic symptom phenotypes (concurrent or lifetime) and bipolar disorder overall, type I, and type

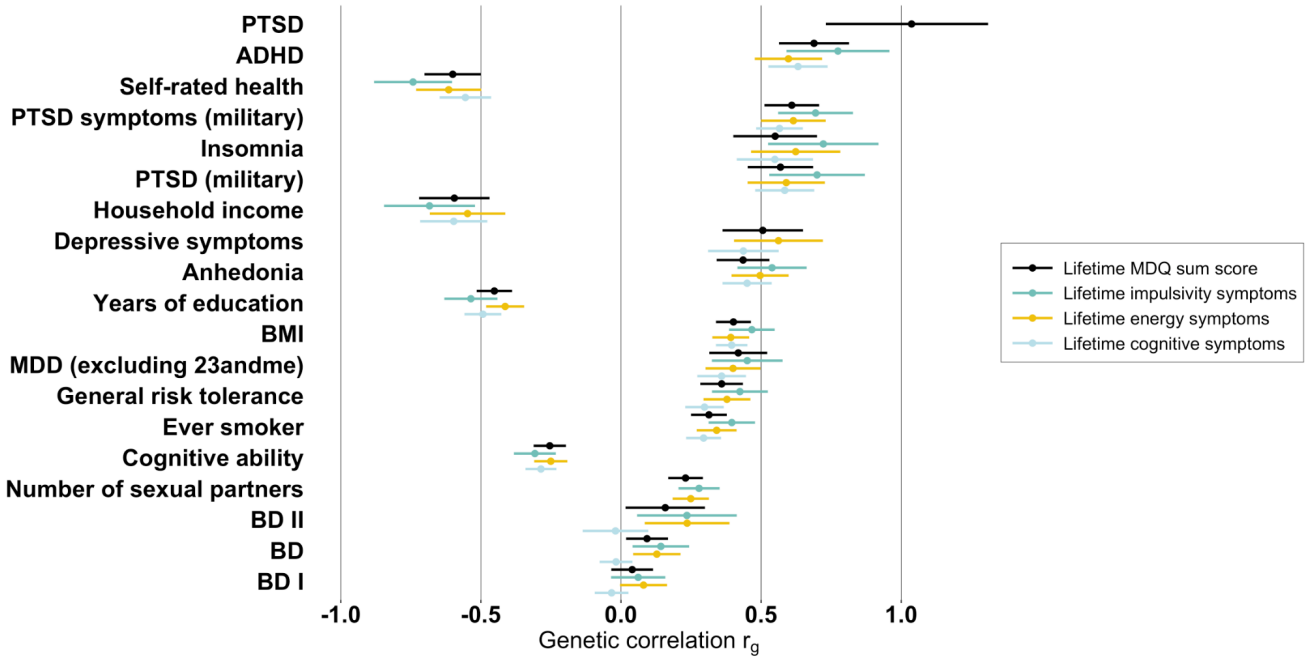
II. None of these genetic correlations were significantly different to zero (**table S3.14** and **table S3.15**).

Genetic correlations with other traits

As expected by the SNP-based heritability estimates (**table S3.16**), there was far stronger evidence of genetic influences on the lifetime symptoms compared to the concurrent symptoms. The concurrent manic symptom sum score and its three symptom subgroups were not genetically correlated with any of the psychiatric or behavioural traits (**table S3.14**).

Contrastingly, we found significant genetic correlations between the lifetime manic symptom phenotypes and 16 psychiatric and behavioural traits (**figure 3.3)**. The highest genetic correlation was between the overall sum score and PTSD (r_g=1.04, p=0.0007). The symptom subgroups were not significantly correlated with PTSD (although their point estimates were similar to the sum score with p-values just below significance) [**table S3.14]**).

Figure 3.3. Significant genetic correlations.
Genetic correlations were computed by Linkage Disequilibrium Score Regression (LDSC; see methods). Genetic correlations, indicated by dots with standard errors indicated by the lines either side of each estimate, were calculated between genome-wide association study (GWAS) summary statistics of lifetime Mood Disorder Questionnaire (MDQ) phenotypes and GWAS summary statistics of psychiatric and behavioural traits. All genetic correlations presented here, apart from bipolar disorder overall, type I, and type II, were significant after correcting for multiple testing (p<0.001) (bipolar disorder is included for comparison only).



Note: ADHD=attention deficit hyperactivity disorder, PTSD=posttraumatic stress disorder, BMI=body mass index, MDD=major depressive disorder, BD=bipolar disorder, BD I=bipolar disorder type 1, BD II=bipolar disorder type II. PTSD (military) and PTSD symptoms (military) refer to two GWASs of United States military subjects from the Million Veteran Program (MVP). PTSD refers to a GWAS of PTSD from the Psychiatric Genomics Consortium; PGC2). MDD refers to a GWAS of MDD from the PGC; PGC2 excluding 23andMe. Information about all the summary statistics used in our analysis, including the original publication and N, can be found in Table S3.11.

Differences between genetic correlations

Of the 16 traits which had a significant genetic correlation (p<0.001) with at least one of the lifetime manic symptom phenotypes, 15 were significantly genetically correlated with more than one symptom subgroup as well as the overall sum score. These were PTSD (military), PTSD symptoms (military), self-rated health, ADHD, insomnia, household income, depressive symptoms, ever smoker, years of education, anhedonia, MDD, BMI, general risk tolerance, number of sexual partners, and cognitive ability (**table S3.15**). We carried forward these 15 traits to a block-jackknife to test for significant differences a) between each subgroup and the overall sum score and b) between the subgroups themselves.

Genetic correlations were significantly different to each other if the block-jackknife p-value surpassed the Bonferroni-adjusted alpha of 0.003 ($\alpha = \frac{0.05}{15}$). When compared to the traits' genetic

correlations with the overall manic symptom sum score, none of the genetic correlations with the symptom subgroups differed significantly. Likewise, when compared to the traits' genetic correlations with other symptom subgroups, none of the genetic correlations differed significantly (**table S3.17**).

Genetic correlations between manic symptom phenotypes

The concurrent manic symptom sum score was significantly genetically correlated with the *concurrent energy/activity factor* (r_g=0.93, p=4.20x10^{-29}) and the *concurrent impulsivity factor* (r_g=0.89, p=2.47x10^{-18}) but not with the *concurrent cognitive factor* (r_g=-0.60, p=0.21).

Mirroring the phenotypic correlations between the symptom subgroups (**table S3.4b**), the *concurrent energy/activity factor* and *concurrent impulsivity factor* were significantly genetically correlated (r_g=0.89, p=1.68x10^{-18}), but the *concurrent cognitive factor* was not significantly genetically correlated with the *concurrent energy/activity factor* (r_g=-0.66, p=0.04) or *concurrent impulsivity factor* (r_g=-0.40, p=0.42) (**table S3.12**).

The lifetime manic symptoms sum score was significantly genetically correlated with all three of its symptom subgroups (with *energy/activity* r_g=1.00*,* p<0.008; with *cognitive* r_g=1.01*,* p<0.008; with *impulsivity* r_g=1.02*,* p<0.008) (**table S3.13**).

Reflected by the phenotypic correlations between the symptom subgroups (**figure S3.4**), the lifetime symptom subgroups were all significantly genetically correlated with each other in a positive direction (*energy/activity* with *cognitive* r_g=0.97*,* p=7.64x10^{-223}; *energy/activity* with *impulsivity* r_g=1.02*,* p<0.008; *cognitive* with *impulsivity* r_g=1.03*,* p=1.30x10^{-175}) (**table S3.13**).

Discussion

We assessed the validity of the MDQ as a screening tool for bipolar disorder in a large sample of individuals affected by mental health problems, which is the population that the MDQ was designed for use in. Taking into account the number, co-occurrence, severity, and duration of symptoms, the MDQ screener (using a cut-off of ≥7 items, as suggested by the MDQ developers (Hirschfeld *et al.*, 2000, 2003)) showed mediocre sensitivity (0.58) and high specificity (0.89). The

PPV was very poor (0.29). Our results showed that the MDQ may comprise three factors: *energy/activity, impulsivity*, and *cognitive*. When measured as concurrent symptoms, the MDQ items showed poor internal consistency with the *cognitive factor* not correlating with the *energy/activity* or the *impulsivity factor*. In our genetic analyses, the quantitative concurrent MDQ items were not heritable. When examining lifetime experience of the MDQ items (i.e., not specifying concurrence), the quantitative score and the three factors showed weak but significant SNP-based heritability. Lifetime MDQ items were genetically correlated with 16 other phenotypes, with the strongest correlation being with PTSD. An unexpected finding was the absence of significant genetic correlation with bipolar disorder overall, type I, and type II.

Very few studies have investigated the latent factor structure of the MDQ (Martino, Valerio and Parker, 2020). One previous study found two latent factors: *energised-activity* and *irritability-racing thoughts* (Benazzi and Akiskal, 2003). While they reported a dual factor structure to the MDQ (although they only included six of the items), we found that both the concurrent and lifetime comprised three factors: *energy/activity, cognitive,* and *impulsivity*. In terms of similarities, the items which loaded onto their *irritability-racing thoughts* factor were the same as the items which loaded onto our *cognitive* factor ("irritability", "racing thoughts", "concentration difficulties"). Likewise, their *energised-activity* factor contained three items ("more active", "more energy", and "decreased sleep") which loaded onto our *energy/activity* factor. However, we also found that four additional items loaded onto this factor ("more sociable", "more self-confidence", "hyperactivity", and "more talkative").

The concurrent items showed poor internal consistency. An unexpected observation was that the items in the *concurrent cognitive factor* ("irritability", "concentration difficulties", and "racing thoughts") did not correlate with the other two factors (**figure 3.2**). This separation from the other items was also found when we performed a one-factor EFA to check that all the items represented a unified latent construct. Here, the items in the *cognitive* factor did not load onto the single factor along with the other items (**table S3.3; figure S3.4**). This was also reflected in the genetic results. This is troubling given that these items constitute part of a validated scale (Hirschfeld *et al.*, 2000).

One explanation for this finding is that irritable mood, racing thoughts, and problems concentrating do not coincide with the other manic symptoms assessed in the MDQ. An alternative and more probable explanation, given that it is unusual to have items in a psychometric scale that are not correlated with each other, concerns recall and memory bias. Potentially, the participants were

able to accurately recall that they have experienced these types of psychiatric problems at some point during their lifetime but failed to recognise that these symptoms occurred at the same time as the other MDQ items. This would explain why these three concurrent MDQ items were not correlated with all remaining items and branched as their own independent subgroup (**figure 3.2**). The characteristics of the study sample may also play a role. The experiences asked about in the "irritability", "racing thoughts", and "concentration difficulties" items are common features of both anxiety and depression (Faravelli *et al.*, 2012; Vidal-Ribas *et al.*, 2016). Of note is the fact that they were the most endorsed symptoms (**table 3.1**). It is likely that these symptoms are not specific to mania in this study sample. This may explain why they were generally not reported alongside other symptoms.

We found that the genetics of the items in the MDQ, measured continuously either as concurrent or cumulative lifetime symptoms, were not genetically correlated with bipolar disorder overall, type I, or type II. This was contrary to our hypothesis that all of the quantitative MDQ-assessed manic symptom phenotypes would show significant, positive genetic correlations with bipolar disorder. We anticipated that the effect size would be larger for concurrent items since co-occurrence of hypomanic or manic symptoms within the same four days or week, respectively, is a requirement for DSM-5 diagnosis (American Psychiatric Association, 2013) and a positive screen in the MDQ (Hirschfeld *et al.*, 2000). There are several possible explanations for this unexpected finding.

First, the quantitative phenotypes were made by summing the number of MDQ items that a participant reported (answer options were "Yes" or "No"). Therefore, the composite scores simply reflect the number of manic symptoms a person has experienced and do not capture any information about the severity or duration of the symptoms (these are separate questions in the MDQ). Comparing quantitative scores between those who do and do not self-report a diagnosis of bipolar (**table S3.1**) does seem to suggest that the number of reported MDQ items is relevant for bipolar disorder. However, in the absence of information about duration and severity, it is possible that these quantitative phenotypes do not reflect hypomania/mania experienced in bipolar disorder. The bipolar disorder GWASs used for genetic correlations were from the most recent PGC analysis, with cases that had clinically diagnosed bipolar disorder. The DSM-5 stipulates that hypomanic or manic symptoms must be present for four days or one week for a diagnosis of bipolar type II or type I, respectively, to be given. For bipolar disorder type I, a diagnosis can only be made when the *"mood disturbance is sufficiently severe to cause marked impairment in social or occupational functioning"* (American Psychiatric Association, 2013).

Therefore, the phenotype of the PGC GWAS relates not just to the number of symptoms but also their duration and associated impairment. By contrast, the GWASs performed in our study only captured the *number* of symptoms a participant had experienced. These phenotypes therefore tell us nothing about whether the participant has experienced no clinically relevant symptoms, subthreshold symptoms, hypomania, or mania.

In support of this conclusion is the genetic correlation of 0.38 with depressive symptoms in the most recent PGC bipolar disorder GWAS (Mullins *et al.*, 2020). Since bipolar disorder involves *both* depressive and manic episodes, we expected bipolar disorder to show a similarly high genetic correlation with manic symptoms in our study. A crucial difference between our quantitative mania scores and the depressive symptom score was that information about severity and duration were included in the latter; depressive symptoms were assessed with two items from the nine item Patient Health Questionnaire (PHQ9) and the answer options were “Not at all”, “Several days”, “More than half the days”, “Nearly everyday” (Okbay *et al.*, 2016). By contrast, the individual MDQ items can be answered with “Yes” or “No”, while two separate questions measure duration and severity. Therefore, it is not straightforward to construct a quantitative hypomania/mania phenotype with the MDQ without applying the same severity and duration to all endorsed items.

The second possible explanation for our lack of genetic correlation with bipolar disorder relates to the type of genomic methodology that we applied to the MDQ. GWASs are only able to capture additive genetic risk from the SNPs in the genotyping or imputation panel. Similar to our results, the study of the MDQ by Hosang et al. (2022) found that MDQ-assessed hypomania was not significantly genetically correlated with bipolar disorder PRS, but they did find a positive and significant twin-based genetic correlation (Hosang *et al.*, 2021). A study by Mistry et al. (2019) reported a similar result. Hypomania, assessed via the Hypomania Checklist 32, was not significantly associated with bipolar disorder PRS (Mistry *et al.*, 2019). Taken together with the results of our study, it may be the case that common variant influences on hypomania/mania are not the same as those influencing bipolar disorder. Other sources of genetic variation, such as rare variants, could drive shared genetic influences between the two but these would not be captured by GWAS or PRS methods.

The final possible explanation for the lack of positive genetic correlation with bipolar disorder is that the MDQ is not a valid measure of hypomania/mania in our study sample. Most studies

reporting high sensitivity and specificity of the MDQ screener include participants with well-established mood disorder diagnoses, who are stabilised, or undergoing treatment. Consequently, their insight into the clinical utility of the MDQ in classifying bipolar disorder among outpatients, or those presenting with a variety of psychiatric complaints and unknown diagnoses, is limited (Zimmerman *et al.*, 2009). The few studies that have investigated this report sensitivity values of 46-64% and specificity values of 65-83% (Hardoy *et al.*, 2005; Konuk *et al.*, 2007; Gervasoni *et al.*, 2009; Zimmerman *et al.*, 2009). Therefore, nearly half of individuals with bipolar disorder, in the population that the MDQ was designed for, could screen negatively, and a significant proportion who do not have bipolar disorder could screen positively. In our study, the sensitivity of the MDQ as a screener was similar to these previous studies at 0.58, and the specificity was good at 0.89. The PPV was low at 0.29. This suggests that, while the MDQ performed well at identifying participants without bipolar disorder, it falsely identified lots of participants as having bipolar disorder when they actually do not.

The MDQ has poorer accuracy in identifying bipolar type II compared to type I (Hirschfeld *et al.*, 2000; Hardoy *et al.*, 2005; Gervasoni *et al.*, 2009; Zimmerman *et al.*, 2009). This, combined with the fact that the MDQ performs more poorly in community samples compared to clinical samples (Miller, Johnson and Eisner, 2009), suggests that symptom severity is an important factor dictating the psychometric properties of the MDQ. Due to the characteristics of our study sample, some of the MDQ items were very highly endorsed (**table 3.1**). This could be because, among individuals with MDD and/or anxiety, they ask about relatively common experiences rather than symptoms of hypomania/mania. This is supported by the finding that MDQ overestimated the prevalence of bipolar disorder in our study sample (PPV of 0.29). Overall, it is possible that the items in the MDQ are not capturing hypomanic/manic symptoms with much precision. This may also be a factor influencing our genetic correlation results; lack of specificity to hypomania/mania could have led to noise into our phenotypes which, as a result, may have diluted the MDQ's genetic sharing with bipolar disorder.

Viewing our genetic correlation results overall, there is no obvious pattern. It appears that the MDQ items are indexing many traits. Significant genetic correlations were found with MDD, depressive symptoms, insomnia, anhedonia, and PTSD symptoms, as well as with risk-taking and smoking. Although the genetic correlation between lifetime MDQ items with PTSD was high, notably this was not confirmed by its phenotypic correlation with PTSD symptoms measured in the same participants (supplementary material). In **table 3.1**, there is a notable difference in

endorsement of the MDQ items between participants affected and unaffected by MDD and anxiety. One explanation for this is that a proportion of the affected participants have undiagnosed bipolar disorder and therefore report more of the MDQ items. However, given the psychometric properties of the MDQ screener in our study, a more likely explanation is that the MDQ items are capturing non-specific aspects of mental illness. Combined with the genetic correlation results, this suggests that, among individuals with MDD and/or anxiety, the MDQ captures symptoms of general distress or psychopathology rather than mania specifically.

A strength of our study was that we were able to measure manic symptoms quantitatively within individuals with experience of MDD or anxiety. Previous GWASs that have isolated mania for genetic analyses have dealt with the phenotype as a binary variable which often incurs a loss of statistical power (Greenwood, Bipolar Genome Study (BiGS) Consortium and Kelsoe, 2013; Lee *et al.*, 2013). Another strength was the measurement of two different types of manic symptoms: concurrent and lifetime symptoms. Our findings suggest that the internal consistency of concurrent MDQ items is compromised when applied to at-risk populations, possibly due to recall and memory biases. The process of screening via the MDQ involves participants self-reporting their symptoms. Our findings suggest that this may be inadequate, especially for individuals who are presenting with symptoms of anxiety and/or depression, due to poor recall of their experience of “irritability”, “racing thoughts”, and “concentration difficulties”. This may have implications for the application of the MDQ as a bipolar disorder screener to these individuals.

Our conclusions should be considered in light of several limitations. First is the relatively modest size of our study sample for genetic analyses compared to modern GWAS standards. This may have impacted statistical power, especially in the GWASs of concurrent symptoms (which had an already attenuated sample size). Second, the criteria for the MDQ screener was based upon symptoms causing “moderate” or “severe” functional impairment (Hirschfeld *et al.*, 2000, 2003). Therefore, it is possible that some individuals with bipolar disorder type II may have been missed. However, even with possible under recognition of type II, the MDQ screener showed an overestimation of bipolar disorder cases in our study sample. Third, the GLAD study is a cohort with generally severe symptomatology, and we cannot generalise to individuals with milder forms of MDD and anxiety (Davies *et al.*, 2019). These sample characteristics also meant that the SNP-based heritability estimates were difficult to interpret, as they depend on the population in which the phenotype is measured. Since we did not use a general population sample, it is difficult to gauge what the SNP-based heritability represents in our study. Another way to measure

hypomania/mania's SNP-based heritability would be to perform GWAS of bipolar disorder type I/type II vs. MDD. However, some eventually convert to bipolar disorder from MDD (Angst *et al.*, 2005; Baryshnikov *et al.*, 2020) which means that we may inadvertently include "hidden" bipolar disorder cases in the MDD comparison group which would dilute the genetic signal of mania. This point relates to a final limitation that our assessment of bipolar disorder was based on self-reports. Calculations of sensitivity, specificity, PPV, and NPV should be made against a "gold-standard" reference (e.g., a diagnosis from a clinical interview). Given that bipolar disorder can go undiagnosed for up to ten years (Hirschfeld, Lewis and Vornik, 2003; Mantere *et al.*, 2004; Drancourt *et al.*, 2013) and unipolar depression is the most likely misdiagnosis (Hirschfeld, Lewis and Vornik, 2003) the eligibility criteria of the GLAD study means that it is highly probable that a proportion of the participants have undiagnosed bipolar disorder. This may have contributed to the seemingly low PPV (0.29).

Overall, our study adds to existing literature questioning the MDQ's validity by showing that, among individuals with MDD and/or anxiety, the items alone capture dimensions of general psychopathology rather than hypomania/mania. Furthermore, our results question the concurrent items' internal consistency. Researchers using the MDQ to measure bipolar disorder in epidemiological studies or biobanks should be cautious of its ability to accurately index symptoms of hypomania/mania and should consider ways to incorporate symptom severity and duration into their phenotyping method.

Data availability

GLAD and COPING study data is available via a data request application to the NIHR BioResource (https://bioresource.nihr.ac.uk/using-our-bioresource/academic-and-clinical-researchers/apply-for-bioresource-data/). The data are not publicly available due to restrictions outlined in the study protocol and specified to participants during the consent process. A specific data freeze is available including the variables for the analyses described in this paper; email gladstudy@kcl.ac.uk for details. The summary statistics from the GWASs performed in this study are available by contacting Professor Gerome Breen.

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Author contributions

Data curation: Molly R. Davies, Gursharan Kalsi, Henry C. Rogers, Brett N. Adey, Sang Hyuck Lee, Jonathan R. I. Coleman, Andrew M. McIntosh, Matthew Hotopf, Thalia C. Eley, Gerome Breen
Data cleaning: Jessica Mundy, Helena L. Davies, Christopher Hübel, Molly R. Davies, Jonathan R. I. Coleman
Formal analysis: Jessica Mundy, Christopher Hübel, Evangelos Vassos, Gerome Breen
Funding acquisition: Robin M. Murray, Gursharan Kalsi, Matthew Hotopf, Thalia C. Eley, Gerome Breen
Investigation: Jessica Mundy, Christopher Hübel, Jonathan R. I. Coleman, Evangelos Vassos, Gerome Breen
Methodology: Jessica Mundy, Christopher Hübel, Jonathan R. I. Coleman, Evangelos Vassos, Gerome Breen
Project administration: Gursharan Kalsi, Christopher Hübel, Henry C. Rogers, Molly R. Davies, Gerome Breen
Resources: Molly R. Davies, Gursharan Kalsi, Henry C. Rogers, Gerome Breen
Software: Christopher Hübel, Brett N. Adey, Jonathan R. I. Coleman, Sang Hyuck Lee
Supervision: Robin Murray, Evangelos Vassos, Gerome Breen
Visualisation: Jessica Mundy
Writing – original draft: Jessica Mundy
Writing – review & editing: Jessica Mundy, Helena L. Davies, Molly R. Davies, Brett N. Adey, Christopher Hübel, Matthew Hotopf, Thalia C. Eley, Robin M. Murray, Evangelos Vassos, Gerome Breen
Preparation: Jessica Mundy, Robin Murray, Evangelos Vassos, Gerome Breen

Funding

This work was supported by the National Institute for Health and Care Research (NIHR) BioResource [RG94028, RG85445], NIHR Biomedical Research Centre [IS-BRC-1215-20018],

HSC R&D Division, Public Health Agency [COM/5516/18], MRC Mental Health Data Pathfinder Award (MC_PC_17,217), and the National Centre for Mental Health funding through Health and Care Research Wales. Jessica Mundy acknowledges funding from the Lord Leverhulme Charitable Grant. Helena L Davies acknowledges funding from the Economic and Social Research Council (ESRC) as part of a PhD studentship. Brett N Adey acknowledges funding through a pre-doctoral Fellowship from the NIHR (NIHR301067). Dr Christopher Hübel acknowledges funding from Lundbeckfonden (R276-2018-4581).

Conflicts of interest

GB has received honoraria, research or conference grants and consulting fees from Illumina, Otsuka, and COMPASS Pathfinder Ltd. The remaining authors have nothing to disclose.

Acknowledgements

We thank all GLAD study and NIHR BioResource volunteers for their participation, and gratefully acknowledge the NIHR BioResource, NIHR BioResource centres, NHS Trusts, and staff for their contribution. We also acknowledge King's College London, South London and Maudsley NHS Trust and King's Health Partners. We thank the National Institute for Health and Care Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Programme. We acknowledge use of the research computing facility at King's College London, *Rosalind* (https://rosalind.kcl.ac.uk), which is delivered in partnership with the NIHR Biomedical Research Centres at South London & Maudsley and Guy's & St. Thomas' NHS Foundation Trusts, and part-funded by capital equipment grants from the Maudsley Charity (award 980) and Guy's & St. Thomas' Charity (TR130505). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, King's College London, or the Department of Health and Social Care.

References

1000 Genomes Project Consortium *et al.* (2015) 'A global reference for human genetic variation', *Nature*, 526(7571), pp. 68–74.

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

Angst, J. *et al.* (2005) 'Diagnostic conversion from depression to bipolar disorders: results of a long-term prospective study of hospital admissions', *Journal of affective disorders*, 84(2-3), pp. 149–157.

Baryshnikov, I. *et al.* (2020) 'Diagnostic conversion from unipolar depression to bipolar disorder, schizophrenia, or schizoaffective disorder: A nationwide prospective 15-year register study on 43 495 inpatients', *Bipolar disorders*, 22(6), pp. 582–592.

Benazzi, F. and Akiskal, H.S. (2003) 'The dual factor structure of self-rated MDQ hypomania: energized-activity versus irritable-thought racing', *Journal of affective disorders*, 73(1-2), pp. 59–64.

Bulik-Sullivan, B.K. *et al.* (2015) 'LD Score regression distinguishes confounding from polygenicity in genome-wide association studies', *Nature genetics*, 47(3), pp. 291–295.

Davies, M.R. *et al.* (2019) 'The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety', *Behaviour research and therapy*, 123, p. 103503.

Davies, M.R. *et al.* (2022) 'Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts', *Journal of anxiety disorders*, 85, p. 102491.

Drancourt, N. *et al.* (2013) 'Duration of untreated bipolar disorder: missed opportunities on the long road to optimal treatment', *Acta psychiatrica Scandinavica*, 127(2), pp. 136–144.

Faravelli, C. *et al.* (2012) 'Generalized anxiety disorder: is there any specific symptom?', *Comprehensive psychiatry*, 53(8), pp. 1056–1062.

Flora, D.B., Labrish, C. and Chalmers, R.P. (2012) 'Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis', *Frontiers in psychology*, 3, p. 55.

Gervasoni, N. *et al.* (2009) 'Performance of the Mood Disorder Questionnaire (MDQ) according to bipolar subtype and symptom severity', *European psychiatry: the journal of the Association of European Psychiatrists*, 24(5), pp. 341–344.

Greenwood, T.A., Bipolar Genome Study (BiGS) Consortium and Kelsoe, J.R. (2013) 'Genome-wide association study of irritable vs. elated mania suggests genetic differences between clinical subtypes of bipolar disorder', *PloS one*, 8(1), p. e53804.

Hardoy, M.C. *et al.* (2005) 'Validation of the Italian version of the "Mood Disorder Questionnaire" for the screening of bipolar disorders', *Clinical practice and epidemiology in mental health: CP & EMH*, 1, p. 8.

Hirschfeld, R.M. *et al.* (2000) 'Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire', *The American journal of psychiatry*, 157(11), pp. 1873–1875.

Hirschfeld, R.M.A. *et al.* (2003) 'Validity of the mood disorder questionnaire: a general

population study', *The American journal of psychiatry*, 160(1), pp. 178–180.

Hirschfeld, R.M.A. (2010) 'The mood disorder questionnaire: its impact on the field [corrected]', *Depression and anxiety*, 27(7), pp. 627–630.

Hirschfeld, R.M.A., Lewis, L. and Vornik, L.A. (2003) 'Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder', *The Journal of clinical psychiatry*, 64(2), pp. 161–174.

Hosang, G.M. *et al.* (2021) 'Association of Etiological Factors for Hypomanic Symptoms, Bipolar Disorder, and Other Severe Mental Illnesses', *JAMA psychiatry* [Preprint]. Available at: https://doi.org/10.1001/jamapsychiatry.2021.3654.

Kessing, L.V. *et al.* (2017) 'Rate and predictors of conversion from unipolar to bipolar disorder: A systematic review and meta-analysis', *Bipolar disorders*, 19(5), pp. 324–335.

Konuk, N. *et al.* (2007) '[Validation of the Turkish version of the mood disorder questionnaire for screening bipolar disorders]', *Turk psikiyatri dergisi = Turkish journal of psychiatry*, 18(2), pp. 147–154.

Lee, H.-J. *et al.* (2013) 'A genome-wide association study of seasonal pattern mania identifies NF1A as a possible susceptibility gene for bipolar disorder', *Journal of affective disorders*, 145(2), pp. 200–207.

Lish, J.D. *et al.* (1994) 'The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members', *Journal of affective disorders*, 31(4), pp. 281–294.

Mantere, O. *et al.* (2004) 'The clinical characteristics of DSM-IV bipolar I and II disorders: baseline findings from the Jorvi Bipolar Study (JoBS)', *Bipolar disorders*, 6(5), pp. 395–405.

Martino, D.J., Valerio, M.P. and Parker, G. (2020) 'The structure of mania: An overview of factorial analysis studies', *European psychiatry: the journal of the Association of European Psychiatrists*, 63(1), p. e10.

Mbatchou, J. *et al.* (2021) 'Computationally efficient whole-genome regression for quantitative and binary traits', *Nature genetics*, 53(7), pp. 1097–1103.

Merikangas, K.R. *et al.* (2011) 'Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative', *Archives of general psychiatry*, 68(3), pp. 241–251.

Miller, C.J., Johnson, S.L. and Eisner, L. (2009) 'Assessment Tools for Adult Bipolar Disorder', *Clinical psychology: a publication of the Division of Clinical Psychology of the American Psychological Association*, 16(2), pp. 188–201.

Mistry, S. *et al.* (2019) 'Genetic risk for bipolar disorder and psychopathology from childhood to early adulthood', *Journal of affective disorders*, 246, pp. 633–639.

Mullins, N. *et al.* (2020) 'Genome-wide association study of over 40,000 bipolar disorder cases provides novel biological insights', *medRxiv*, p. 2020.09.17.20187054.

Mullins, N. *et al.* (2021) 'Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology', *Nature genetics*, 53(6), pp. 817–829.

Mundy, J. *et al.* (2021) ‘Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder’, *Psychological medicine*, pp. 1–10.

Musliner, K.L. and Østergaard, S.D. (2018) ‘Patterns and predictors of conversion to bipolar disorder in 91 587 individuals diagnosed with unipolar depression’, *Acta psychiatrica Scandinavica*, 137(5), pp. 422–432.

Nagel, M. *et al.* (2018) ‘Item-level analyses reveal genetic heterogeneity in neuroticism’, *Nature communications*, 9(1), p. 905.

Okbay, A. *et al.* (2016) ‘Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses’, *Nature genetics*, 48(6), pp. 624–633.

van Rheenen, W. *et al.* (2019) ‘Genetic correlations of polygenic disease traits: from theory to practice’, *Nature reviews. Genetics* [Preprint]. Available at: https://doi.org/10.1038/s41576-019-0137-z.

Stahl, E.A. *et al.* (2019) ‘Genome-wide association study identifies 30 loci associated with bipolar disorder’, *Nature genetics*, 51(5), pp. 793–803.

Thorp, J.G. *et al.* (2020) ‘Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9’, *Psychological medicine*, 50(14), pp. 2385–2396.

Vidal-Ribas, P. *et al.* (2016) ‘The Status of Irritability in Psychiatry: A Conceptual and Quantitative Review’, *Journal of the American Academy of Child and Adolescent Psychiatry*, 55(7), pp. 556–570.

Watanabe, K. *et al.* (2017) ‘Functional mapping and annotation of genetic associations with FUMA’, *Nature communications*, 8(1), p. 1826.

Zimmerman, M. *et al.* (2009) ‘Performance of the mood disorders questionnaire in a psychiatric outpatient setting’, *Bipolar disorders*, 11(7), pp. 759–765.

Zimmerman, M. (2012) ‘Misuse of the Mood Disorders Questionnaire as a case-finding measure and a critique of the concept of using a screening scale for bipolar disorder in psychiatric practice’, *Bipolar disorders*, 14(2), pp. 127–134.

Chapter 4. Exploring the genetics of anhedonia and its relationship with treatment response in major depressive disorder

This chapter is a manuscript that is in preparation for peer-review.

Supplementary material is included in **appendix 4**.

Authors:

Jessica Mundy[1,2], Kirstin L. Purves[1,2], Katherine S. Young[1,2], Brett N. Adey[1,2], Cherie Armour[3], Molly R. Davies[1,2], Jonathan R. I. Coleman[1,2], Matthew Hotopf[1,2,5], Christopher Hübel[1,2,6], Ian R. Jones[8], Gursharan Kalsi[1,2], Andrew M. McIntosh[9], Alicia J. Peel[1], Henry Rogers[1,2], Daniel J. Smith[10], Abigail ter Kuile[1,2], James T.R. Walters[8], Evangelos Vassos[1,2], Robin M. Murray[1], Thalia C. Eley[1,2], Gerome Breen[1,2]

Author affiliations:

1. Institute of Psychiatry, Psychology and Neuroscience, King's College London, Denmark Hill, Camberwell, London, United Kingdom
2. UK National Institute for Health Research (NIHR) Biomedical Research Centre, South London and Maudsley Hospital, London, United Kingdom
3. Research Centre for Stress, Trauma, and Related Conditions (STARC), School of Psychology, Queen's University Belfast (QUB), Belfast, Northern Ireland, United Kingdom
4. Research Department of Clinical, Educational and Health Psychology, University College London, Gower Street, London, United Kingdom
5. South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital, Monks Orchard Road, Beckenham, Kent, United Kingdom

6. National Centre for Register-based Research, Aarhus Business and Social Sciences, Aarhus University, Aarhus, Denmark
7. MRC Centre for Neuropsychiatric Genetics and Genomics, National Centre for Mental Health, Cardiff University, Cardiff, United Kingdom
8. National Centre for Mental Health, Division of Psychiatry and Clinical Neuroscience, Cardiff University, Cardiff, United Kingdom
9. Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom
10. Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

Abstract

Background: Anhedonia is a core symptom of major depressive disorder (MDD) but also presents transdiagnostically. Anhedonia has been proposed as a risk factor for poor response to antidepressants in those who have MDD. Anhedonia has a heritable basis.

Aims: We aimed to investigate whether anhedonic symptoms share genetic overlap with treatment-resistant depression, antidepressant response, and other traits including MDD.

Methods: We measured anhedonic symptoms in participants in the COVID-19 Psychiatric and Neurological Genetics (COPING) study. We used the anhedonic depression subscale of the 30-item Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30) to create a continuous score representing anhedonic symptoms (range 0-40). We performed a genome-wide association study and calculated genetic correlations with other psychiatric and behavioural traits, including measures of treatment resistance and treatment response in MDD. We compared our GWAS of anhedonic symptoms to a previous GWAS of the single anhedonia item from the nine item Patient Health Questionnaire (PHQ9).

Results: Anhedonic symptoms had a SNP-based heritability of 8% and a genetic correlation (r_g) of 0.45 with the previous single item PHQ9 anhedonia GWAS. The phenotypic correlation was higher at 0.59. Anhedonic symptoms did not have significant genetic correlations with treatment-resistant depression or antidepressant response, but had positive genetic correlations with

depressive symptoms (r_g=0.71), anxiety (r_g=0.50), and neuroticism (r_g=0.46), and negative genetic correlations with self-rated health (r_g=-0.42), and risk-taking behaviour (r_g=-0.28).

Conclusion: We found no evidence to suggest that anhedonic symptoms share genetic overlap with either treatment response in MDD. The AD-MASQ-D30 is a more detailed method of assessing anhedonia than the single PHQ9 item due to the inclusion of ten items each rated 0-4. The AD-MASQ-D30 may be better able to capture anhedonia distinct from MDD than the single anhedonia item from the PHQ9 (range 0-3).

Introduction

Anhedonia, a cardinal symptom of major depressive disorder (MDD), is defined in the the 5th edition of the Diagnostic Statistical Manual of Mental Disorders (DSM-5) as *"markedly diminished interest or pleasure in all, or almost all activities most of the day, nearly every day"* (American Psychiatric Association, 2013). Anhedonia was implicated as a risk factor for poor treatment outcomes, including longer time to remission and fewer depression-free days, in 12-18 year olds with MDD (McMakin *et al.*, 2012). In a study of 811 participants being treated for moderate to severe MDD, baseline interest-activity symptoms (created to represent anhedonic symptoms across multiple measures of depression) were associated with lower likelihood of remission irrespective of depression severity, type of prescribed antidepressant, and type of scale used to measure depressive symptoms (Uher *et al.*, 2012). Therefore, knowledge of a patient's propensity for experiencing anhedonia may be useful for predicting treatment response in MDD.

As well as MDD, anhedonia is observed in many other forms of psychopathology including schizophrenia (Horan, Kring and Blanchard, 2006), obsessive compulsive disorder (OCD) (Abramovitch *et al.*, 2014; Li *et al.*, 2019), and some anxiety disorders (Winer *et al.*, 2017). Anhedonia is also present in neurological disorders such as Parkinson's disease (Loas, Krystkowiak and Godefroy, 2012), and epilepsy (Roberts-West, Vivekananda and Baxendale, 2022), and is a risk factor for greater impairment in Alzheimer's disease (Natta *et al.*, 2013). Additionally, anhedonia is a symptom often reported by trauma-exposed individuals, such as those with posttraumatic stress disorder (PTSD) (American Psychiatric Association, 2013).

Anhedonia is heritable (Bogdan and Pizzagalli, 2009; Liu *et al.*, 2016) but few studies have explored its genetic basis. The largest study analysed data from the UK Biobank. Participants were asked to complete the nine item Patient Health Questionnaire (PHQ9) as part of the baseline assessment which assessed current depression symptoms (Kroenke, Spitzer and Williams, 2001). The PHQ9 included a single question aimed at capturing anhedonia: "*Over the past two weeks, how often have you had little interest or pleasure in doing things?*". Numeric answers (ranging 0-3) were then used as an ordinal phenotype in a GWAS of 375,275 participants of European ancestries. This phenotype had a SNP-based heritability of 5.6% and significant, positive genetic correlations with MDD, bipolar disorder, and schizophrenia. The GWAS identified 11 loci, some of which had been previously implicated in MDD (e.g., NCAM1 on chromosome 11), schizophrenia (e.g., PRKD1 on chromosome 14, and NRG4 on chromosome 15), mood instability, and suicidality (e.g., DCC on chromosome 18) (Ward *et al.*, 2019).

While a strength of this study was the large sample size, the authors recognised that relying on a single item to measure a trait as complex as anhedonia is insufficient. Research has shown that the experience of anhedonia, although a symptom of depression on its own, is complex and involves a number of different dimensions (Thomsen, 2015; Winer, Jordan and Collins, 2019). In addition, the PHQ9 was only assessed at one time point and external factors influencing participant answers (such as season or current health status) were not controlled for. Furthermore, genetic correlations were estimated only with MDD, bipolar disorder, schizophrenia, Parkinson's disease, and OCD. These psychiatric and neurological disorders were selected *a priori* because anhedonia is sometimes reported by individuals affected by them. Therefore, significant genetic associations with other traits may have gone unnoticed.

To overcome the limitations of Ward et al. (2019) and to better understand the genetic basis of anhedonia, there is a need to perform a well-powered GWAS with a detailed and valid measure from multiple time points. Additionally, an exploration into whether anhedonia shares genetic variants with a broader range of traits than has already been measured is needed, especially with MDD treatment response. To address this, we analysed mental health and genetic data from a large cohort of individuals affected and unaffected by MDD or anxiety disorders in the United Kingdom (UK). We measured anhedonic symptoms quantitatively with a validated ten-item scale, estimated its SNP-based heritability, and calculated genetic correlations with 42 traits, including two measures of treatment-resistant depression and two measures of antidepressant response.

We aimed to elucidate whether the relationship between anhedonia and treatment-response in MDD is partly rooted in shared genetic risk factors.

Methods

All code is available on GitHub (https://github.com/tnggroup/genetics_anhedonia). This study was pre-registered on the Open Science Framework (https://osf.io/pb769/).

Study sample

Anhedonic symptoms: Data were analysed from participants in the mental health arm of the National Institute of Health and Care Research (NIHR) BioResource in the UK who took part in the COVID-19 Psychiatry and Neurological Genetics (COPING) study. Collectively, this cohort comprises participants from three independent cohorts: the Genetic Links to Anxiety and Depression (GLAD) study, the Eating Disorders Genetics Initiative (EDGI), and participants from the NIHR BioResource (NBR). Recruitment for the GLAD study began in September 2018 and was conducted via social media campaigns and NHS sites. GLAD participants were recruited based on having lifetime experience of an anxiety or depressive disorder. EDGI participants were recruited from February 2020 based on having lifetime experience of an eating disorder. NBR participants were recruited via existing research studies (**table S4.10**) and included individuals affected and unaffected by mental illnesses. Henceforth, these individuals will be referred to as COPING participants.

Baseline invitations to participate in the COPING study were first sent to existing GLAD, EDGI, and NBR participants on 30th April 2020 with all invites sent by 11th May 2020. Invitations were sent to newly enrolled (i.e., not already enrolled on 30th April 2020) GLAD and EDGI participants periodically throughout the pandemic with the last round of invites sent on 19th January 2021. Individuals were eligible to participate if they were aged 16+ and lived in the UK. All participants provided demographic information, detailed mental health information, and a saliva or blood sample. Surveys were sent to participants every 14 days until the 28th July 2021, at which point they were sent every 28 days.

Treatment-resistant depression: Treatment-resistant depression was defined based on data collected when the participants completed the GLAD sign-up survey. We analysed data from participants who completed the GLAD study sign-up (for staged treatment-resistant depression) or COPING baseline survey (for anhedonic symptoms) between 17th September 2018 and 3rd September 2021 and had available genetic data.

Ethics

Informed consent was obtained from all participants. Ethical approval for the GLAD study was granted by the London-Fulham Research Ethics Committee (REC reference: 18/LO/1218). Ethical approval for COPING was granted by the NHS Health Research Authority, South West - Central Bristol Research Ethics Committee (20/SW/0078).

Main measures

Current anhedonic symptoms

As part of the repeated assessment of mental health in the COPING study, participants answered questions from a subscale in the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (MASQ-D30) every two weeks. The original MASQ was developed by Clark & Watson (1999) to represent their tripartite model of anxious, depressive, and somatic symptoms (Clark and Watson, 1991). The subscale included ten questions about *"feelings, sensations, problems, and experiences that people sometimes have"* over the past two weeks (**table 4.1**). Participants could answer with the following options: *"Not at all", "A little bit", "Moderately", "Quite a bit"* and *"Extremely"*. These statements refer to different aspects of positive affect (**table 4.1**). When these items are reverse scored, thus representing a *lack of* positive affect, they are a valid and internally consistent measure of symptoms of anhedonic depression (Wardenaar *et al.*, 2010). Henceforth, this subscale will be referred to as the "anhedonic depression subscale of the MASQ-D30" (AD-MASQ-D30).

At each completed COPING survey, we reverse coded the numeric answers for each participant [4-0] and summed them. This resulted in a sum score for each participant ranging 0-40 which represented the severity of anhedonic symptoms, with higher scores representing more severe levels of anhedonia. Note that participants with any item(s) missing were removed from the

creation of a sum score. They were therefore deemed as “missing” for that particular COPING survey.

Exclusions: We first excluded participants who had complete data on the AD-MASQ-D30 at fewer than three COPING survey time points. Following this, we excluded anyone who’s genetic data did not pass the standard genotype quality control (QC) (more details in supplementary methods). More detail about the rationale behind the exclusion criteria and genotype QC is presented in the supplementary methods.

The AD-MASQ-D30 was included at each COPING survey time point throughout the pandemic so the majority of participants had data about their anhedonic symptoms at more than one point in time. We therefore created three phenotypes for GWASs from the participants’ available data. These were scores representing each participant’s highest, lowest, and mean anhedonic symptoms. A flow-chart depicting each stage of exclusions can be found in table **S4.10**.

Treatment-resistant depression

Response to antidepressants is not a dichotomous phenotype (Sforzini *et al.*, 2021), although it is often referred to as so. To address this, tools have been designed to measure treatment resistance continuously, which takes into account variability in response to treatments. One such method is the Maudsley Staging Method (MSM) (Fekadu *et al.*, 2009). The MSM uses a points-based system in three domains: 1) severity of illness, 2) duration of presenting illness, and 3) treatment response. The MSM was included in the GLAD study baseline survey and participants were eligible to answer the scale if they were currently experiencing a depressive episode (i.e., a severity score of 1/“mild” or more on the PHQ9). The questions included in the MSM are presented in **table S4.1** and the supplementary methods. The first domain, severity of the presenting depressive episode, was assessed via the PHQ9 at the same time the MSM was completed (Kroenke, Spitzer and Williams, 2001).

Exclusions: We excluded anyone who had missing data on either the PHQ9 or the MSM. This included participants who were not currently depressed at the time of completing the GLAD survey (i.e., a PHQ9 severity score of 0/”None”) and therefore did not complete the MSM. It also included

participants who reported that they had not taken any antidepressants for six weeks or longer for their current or most recent depressive episode (these participants were not shown the next question about add-on/augmentation medications). Individuals who have experienced depression may appear treatment-resistant if they have undiagnosed bipolar disorder (McAllister-Williams *et al.*, 2018). We wanted to maximise the chance that all individuals in our GWAS sample were affected by major depressive disorder and not bipolar disorder. *(Individuals were eligible to complete the MSM if they were currently depressed at the time of completing the GLAD study survey. This means that individuals with bipolar disorder who happened to be in a depressive episode answered the MSM).* Therefore, we retained participants who self-reported that they had NOT received a diagnosis of bipolar disorder in BOTH of these questions. Following this, we excluded anyone who's genetic data did not pass the standard genotype QC. More detail about the rationale behind the exclusion criteria and genotype QC is presented in the supplementary methods. A flow-chart depicting each stage of exclusions can be found in table **S4.11**.

In those who completed all questions of the MSM, we summed their answers to create a score (ranging 2-10). This was then added to their PHQ9 severity score (ranging 1-4). Thus, their score (ranging 3-14) represented levels of treatment resistance which also took into account the severity of their presenting depressive episode. A more detailed explanation of how staged treatment-resistant depression was phenotyped in the GLAD study is presented in the supplementary methods.

Additional measures for phenotypic analyses

We calculated phenotypic correlations between participants' highest anhedonic symptoms and their concurrent depression and anxiety symptoms (i.e., that they were experiencing at the time of their highest anhedonia score). We also calculated the correlation between the participants' highest anhedonic symptoms measured with the AD-MASQ-D30 and their concurrent anhedonia measured by the single PHQ9 item: "*Over the past two weeks, how often have you had little interest or pleasure in doing things?*". This was the same phenotype used for the Ward et al. (2019) GWAS of anhedonia (Ward *et al.*, 2019). A correlation matrix of Pearson's product-moment correlations was computed using R.

Current depression symptoms

Current depression symptoms were measured continuously with the PHQ9 (Kroenke, Spitzer and Williams, 2001), which asked questions relating to mood and feelings that participants may have experienced over the past two weeks. Participants could answer with the following: "*Not at all*", "*Several days*", "*More than half the days*", and "*Nearly every day*" which were coded as 0, 1, 2, and 3 respectively. Numeric answers were summed to create a continuous measure of current depression symptoms ranging 0-27 at each COPING survey, with higher values reflecting more severe symptoms.

Current anxiety symptoms

Current generalised anxiety symptoms were measured continuously with the seven item Generalised Anxiety Disorder questionnaire (GAD7) (Spitzer *et al.*, 2006), which asked questions relating to *"problems"* that participants may have experienced over the past two weeks. Participants could answer with the following: "*Not at all*", "*Several days*", "*More than half the days*", and "*Nearly every day*" which were coded as 0, 1, 2, and 3 respectively. Numeric answers were summed to create a continuous measure of current anxiety ranging 0-21 at each COPING survey, with higher values reflecting more severe symptoms.

Similarly to anhedonic symptoms, participants with any item(s) missing at a COPING survey were removed from the creation of a symptom sum score. They were therefore deemed as "missing" for that particular COPING survey.

Genetic analyses

Genotyping, imputation and quality control

All data from GLAD and COPING NBR were genotyped by ThermoFisher on the Affymetrix UK Biobank Axiom Array v1 and v2 across numerous genotyping batches. Genetic data for the GLAD and NBR cohorts were separately subjected to QC with the same pipeline. For our specific analyses, additional QC was carried out (see supplementary methods for QC pipeline and analyses-specific QC).

Genome-wide association studies

First, we conducted GWASs of the COPING participants' lowest, highest, and mean anhedonic symptoms using a mixed linear model with REGENIE, which controls for between-subject relatedness using whole-genome regression (Mbatchou *et al.*, 2021). Second, we conducted a GWAS of staged treatment-resistant depression in GLAD participants, also using a mixed linear model with REGENIE. We included the first ten ancestry principal components (PCs) and genotyping batch as covariates.

A previous GWAS of treatment-resistant depression recommended combining cohorts to increase statistical power to detect and replicate genetic variants (Li *et al.*, 2020). Therefore, we meta-analysed our results with a previous GWAS of binary treatment resistant depression assessed from primary care records in the UK Biobank (treatment-resistant depression vs. non treatment-resistant depression; N_{cases}=2,165, $N_{controls}$=14,207, N_{total}=16,372) (Fabbri *et al.*, 2021). We used METAL for the meta-analysis with p-values and directions of SNP effects as inputs. We allowed for heterogeneity. We filtered the resulting summary statistics to keep SNPs which were present in both sets of input GWAS summary statistics.

The three resulting GWAS summary statistics were each inputted into the Functional Annotation and Mapping (FUMA) software (Watanabe *et al.*, 2017) to produce Manhattan plots, quantile-quantile (QQ) plots, and to investigate whether any gene-based associations reached statistical significance. Gene-based associations were computed by the Multi-marker Analysis of GenoMic Annotation (MAGMA) software.

SNP-based heritability

We estimated the SNP-based heritability with individual-level data using genomic relatedness–based restricted maximum likelihood (GREML) in the software Genome-wide Complex Trait Analysis (GCTA) (Yang *et al.*, 2011). GCTA uses all genotyped SNPs and a genetic-relatedness matrix (GRM) to estimate the variance explained by the genotyped SNPs regarding a particular complex trait while controlling for gross genetic similarity (i.e., relatedness) between participants. The GRMs were adjusted for incomplete tagging of causal SNPs and then pruned for relatedness (i.e., one of a pair of individuals with estimated relatedness larger than or equal to a pi-hat of 0.05 was removed). Since GCTA-GREML requires individual-level genotype data, we were not able to calculate GCTA-heritability of meta-analysed GLAD-UK Biobank treatment-resistant depression.

We also calculated SNP-based heritability estimates using *Linkage Disequilibrium Score Regression* (LDSC) (Bulik-Sullivan *et al.*, 2015). To maximise the power to detect significant associations with other traits, we selected the measure of anhedonic symptoms with the largest SNP-based heritability estimated by LDSC to calculate genetic correlations with external traits. LDSC uses GWAS summary statistics and linkage disequilibrium (LD) between SNPs based on a chosen reference panel.

The SNP-based heritability estimates were statistically significant if their p-value surpassed the Bonferroni-corrected alpha of p<0.017 ($\alpha = \frac{\mathbf{0.05}}{\mathbf{3}}$) to adjust for the three independent phenotypes (anhedonia, staged treatment-resistant depression, and meta-analysed GLAD-UK Biobank treatment-resistant depression). *Since the three correlated measures of anhedonic symptoms [highest, lowest, and mean] were calculated in the same participants, the estimations accounted for one independent statistical test (**figure S4.1**).*

Genetic correlations (GCTA bivariate-REML)

GCTA requires individual-level genotype data for genetic correlations. Since we had this for participants in COPING and GLAD participants, we calculated the genetic correlation between anhedonic symptoms and staged treatment-resistant depression using GCTA bivariate-REML. For this, we prioritised the measure of anhedonic symptoms with the largest heritability estimated by GCTA-GREML, which were the mean symptoms (**table 4.2** and **table S4.4**). We created a new GRM for this analysis which was based upon the merged COPING NBR and GLAD genotype data (details of the merge process are presented in the supplementary methods).

Genetic correlations (bivariate LDSC)

We also used LDSC to calculate genetic correlations between GWAS summary statistics of anhedonic symptoms and external psychiatric and behavioural traits. For this analysis, we prioritised the highest anhedonic symptoms, rather than the mean anhedonic symptoms, as the phenotype of interest. This is because LDSC computes genetic correlations based on the SNPs in an LD reference panel, rather than full genotype data. We wanted to maximise power to detect significant genetic correlations, so we selected the phenotype with the highest heritability as

estimated by LDSC (highest anhedonic symptoms). We calculated genetic correlations (r_g) between this phenotype and external psychiatric and behavioural traits (**table S4.5**) including staged treatment-resistant depression. Where possible, we included traits that had been analysed in the previous Ward et al. (2019) GWAS of anhedonia in the UK Biobank, and this GWAS itself for comparison of the two different approaches to measuring anhedonia. We also included other psychiatric and behavioural traits which were sufficiently powered (a heritability z-score>4 and a mean chi-square>1.02) from our internal GWAS summary statistics database (**table S4.7**). To reduce the multiple testing burden, we selected the most well-powered GWASs when more than one trait, or similar traits, were available.

In addition to our measure of staged treatment-resistant depression, we included three further traits representing different aspects of treatment resistance or response to antidepressants in depression. The first was the aforementioned GWAS of treatment-resistant depression in the UK Biobank (N_{cases}=2,165, $N_{controls}$=14,207, N_{total}=16,372) (Fabbri *et al.*, 2021). The second and third were the results of two GWASs of antidepressant response in 10 cohorts from Europe and the United States. This included a continuous measure (percentage improvement in symptoms after treatment) (N=5,218) and a binary measure (responders vs. non-responders) (N_{cases}=1,852, $N_{controls}$=3,299, N_{total}=5,151) (Pain *et al.*, 2022). Information about the external traits is presented in **table S4.7**.

Summary statistics were munged with the extended 1000 Genomes reference panel in LDSC. A total of 42 genetic correlations were computed. The alpha value for the genetic correlations was adjusted to correct for multiple testing with the Bonferroni method. This meant that genetic correlations were significantly different to zero if p<0.001 ($\alpha = \frac{0.05}{42}$).

We also calculated the LDSC genetic correlations between the Ward et al. (2019) GWAS of anhedonia in the UK Biobank and any traits which were significantly genetically correlated with AD-MASQ-D30-assessed anhedonic symptoms, plus the three measures of treatment-resistant depression and two measures of antidepressant response (Fabbri *et al.*, 2021; Pain *et al.*, 2022). This totalled 10 phenotypes. These genetic correlations were significantly different to zero if p<0.005 ($\alpha = \frac{0.05}{10}$).

Results

Study sample

Anhedonic symptoms

The study sample for our analysis of anhedonic symptoms included a total of 13,433 COPING participants who had complete data available for the AD-MASQ-D30 in at least three COPING surveys and had genetic data that passed the standard genotype QC. Of these participants, 5,590 were GLAD participants (42%) and 7,843 were NBR participants (58%) (no genetic data was available for EDGI participants). The mean age was 53 years (SD=15). The sample was mainly female (67%). Bearing in mind that individuals aged 16+ were eligible to take part in the COPING study and therefore some may not have yet completed university, 58% of the sample had a university degree.

Staged treatment-resistant depression

The study sample for our analysis of staged treatment-resistant depression included 8,165 participants who completed the MSM, had genetic data that passed the standard genotype QC, and self-reported that they had not received a diagnosis of bipolar disorder. The mean age was 53 (SD=15). The sample was overwhelmingly female (80%). Those with a university degree constituted 55%. Since we limited our analytical samples to those with available genetic data which passed the standard genotype QC, our sample is of European ancestry. Full descriptive statistics are presented in **table 4.1**. Note that 2,669 GLAD participants were in both the anhedonic symptoms and staged treatment-resistant depression study samples.

Table 4.1. Demographic descriptive statistics for the two analytical study samples (anhedonic symptoms and staged treatment-resistant depression).
Anhedonic symptoms were measured in participants from the COVID-19 Psychiatry and Neurological Genetics (COPING) study with the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30) [ranging 0-40]. Staged treatment-resistant depression was measured in participants from the Genetic Links to Anxiety and Depression (GLAD) study using two questionnaires: the Maudsley Staging Method (MSM) and the nine item Patient Health

Questionnaire (PHQ9) [ranging 3-14]. All participants included in analyses had genetic data available that passed the standard genotype quality control (QC).

	Anhedonic symptoms in COPING participants	Treatment-resistant depression in GLAD participants
Variable	**N = 13,433**[1]	**N = 8,165**[1]
Highest anhedonic symptoms	32 (7)	
Lowest anhedonic symptoms	19 (8)	
Mean anhedonic symptoms	25 (8)	
Sex		
Female	9,046 (67%)	6,497 (80%)
Male	4,387 (33%)	1,668 (20%)
Age [years]	53 (15)	39 (14)
Missing	5	
Self-reported ethnicity		
Mixed	33 (0.3%)	45 (0.6%)
Other	42 (0.3%)	40 (0.5%)
White	13,053 (99%)	8,067 (99%)
Missing	305	13
Education		
No university degree	5,499 (42%)	3,592 (45%)
University degree	7,638 (58%)	4,476 (55%)
Missing	296	97
TRD sum score (MSM)		5.22 (1.68)

[1] Mean (SD); n (%)

Note: anhedonic symptoms were measured at multiple timepoints throughout the COPING study (surveys were sent every two weeks at first, and then every month). This means that participants had data on anhedonic symptoms at multiple timepoints. We calculated three phenotypes, highest, lowest, and mean anhedonic symptoms, for each participant. The values shown in the table are the mean values and standard deviations of each of these three phenotypes. Plots of the three anhedonic symptoms and staged treatment-resistant depression are presented in ***figure 4.1****.*

Phenotype descriptive statistics

Anhedonic symptoms

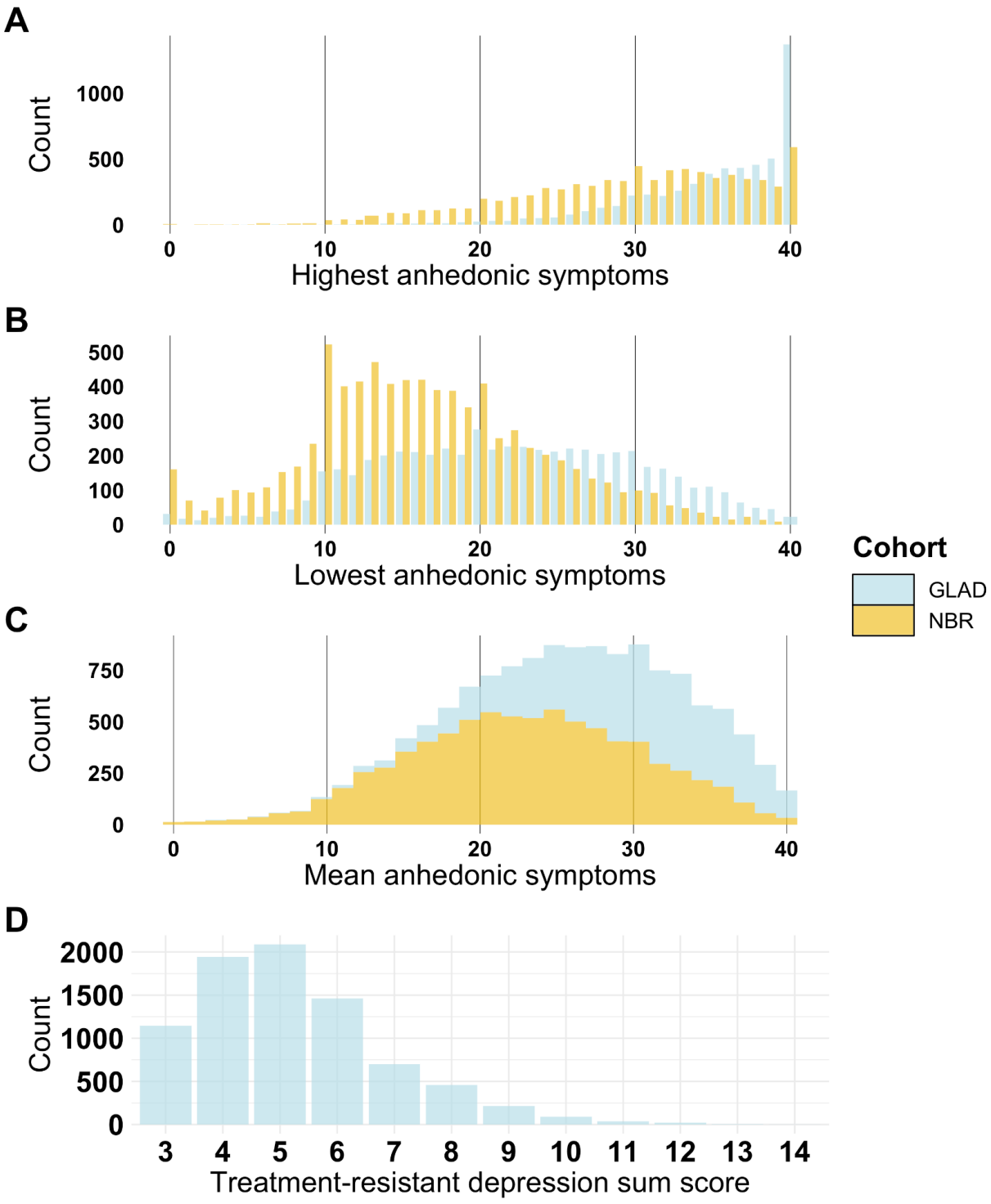
At each survey time point, we calculated sum scores from the participants' answers to the AD-MASQ-D30 which represented current anhedonic symptoms. The sum scores ranged 0-40 with higher scores representing higher levels of anhedonia. The average (mean) of the participants'

highest score (i.e., most severe anhedonia across all non-missing timepoints) was 32.0 (SD=7.2), the average of their lowest score (i.e., least severe anhedonia across all non-missing timepoints) was 18.1 (SD=8.5), and the average of their mean score (i.e., average anhedonia across all non-missing time points) was 25.8 (SD=6.7). Bar plots (highest and lowest) and a histogram (mean) of the symptoms are presented in **figure 4.1**.

Staged treatment-resistant depression

We calculated a sum score, representing stages of treatment-resistant depression, based on answers to the MSM and PHQ9 in the GLAD study sign-up questionnaire (ranging 3-14) which also takes into account the severity of their presenting depressive episode (based on severity in the PHQ9). A bar plot of the GLAD participants' scores is presented in **figure 4.1**. Note that participants can score three or more in the overall sum score (i.e., it is impossible to gain a score of one or two) (Fekadu *et al.*, 2009). An explanation for this is presented in the supplementary methods. The mean score was 5.22 (SD=1.68) and the modal score was 5.

Figure 4.1. A-C: Plots of the three measures of anhedonic symptoms calculated for participants in the COVID-19 Psychiatry and Neurological Genetics (COPING) study (N=13,334).
Anhedonic symptoms were measured with the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30). D: Plot of staged treatment-resistant depression in Genetic Links to Anxiety and Depression (GLAD) study participants (N=8,165). Treatment-resistant depression was measured as a continuous scale with the Maudsley Staging Method (MSM) and nine item Patient Health Questionnaire (PHQ9).


A
Count
1000
500
0
0
10
20
30
40
Highest anhedonic symptoms
B
Count
500
400
300
200
100
0
0
10
20
30
40
Lowest anhedonic symptoms
C
Count
750
500
250
0
0
10
20
30
40
Mean anhedonic symptoms
D
Count
2000
1500
1000
500
0
3
4
5
6
7
8
9
10
11
12
13
14
Treatment-resistant depression sum score
Cohort
GLAD
NBR

Phenotypic analyses

The correlations between the highest, lowest, and mean anhedonic symptoms among COPING participants, and the correlations between highest anhedonic symptoms and concurrent depression and anxiety symptoms, and the single anhedonia item from the PHQ9, are presented in the supplementary results. A key finding was that anhedonic symptoms had moderate correlations with concurrent depression (PHQ9) and anxiety (GAD7) symptoms.

Another key finding was that the single anhedonia item from the PHQ9 showed a much lower correlation with anhedonic symptoms measured by the AD-MASQ-D30 than its correlations with depression and anxiety symptoms (**figures S4.1-S4.2**). In participants who overlapped between the two analytical study samples (N=2,669), there were low correlations between their staged treatment-resistant depression (assessed when they completed the GLAD study) and their highest and mean anhedonic symptoms (assessed during the COPING study) (highest symptoms r_{ph}=0.23, SE=0.02, p=1.70x10^{-32}; mean symptoms r_{ph}=0.29, SE=0.02, p=2.12x10^{-54}).

Genetic analyses

For the estimation of heritability with GCTA-GREML, two separate GRMs were computed for 1) COPING participants and 2) GLAD participants. In each GRM, we included the same participants from the respective GWASs but some were removed by pruning for relatedness. For anhedonic symptoms in COPING participants, 13,271 were included in the estimation of the GRM (162 removed). For staged treatment-resistant depression in GLAD participants, 8,062 participants were included in the estimation of the GRM (184 removed).

SNP-based heritability

GCTA: GCTA estimated the participants' mean anhedonic symptoms as the most heritable and significantly different to zero (h^2_{SNP}=0.13, SE=0.04, p=0.001). GCTA estimated the SNP-based heritability of staged treatment-resistant depression at 0.03 (SE=0.06, p=0.35) but this was not significant (**table 4.2** and **table S4.4**).

Table 4.2. Single Nucleotide Polymorphism (SNP)-based heritability estimates (h^2_{SNP}).

SNP-based heritability estimates (h^2_{SNP}), standard errors (SE), z-scores, and p-values of participants' anhedonic symptoms and staged treatment-resistant depression.

Phenotype	Study sample	N	h^2_{SNP}	SE	Z-score	P-value (0)
Highest anhedonic symptoms	COPING participants	13,271	0.11	0.04	2.48	**3.83×10^{-3}**
Lowest anhedonic symptoms	COPING participants	13,271	0.09	0.05	2.08	**2.19×10^{-2}**
Mean anhedonic symptoms	COPING participants	13,271	0.13	0.04	2.81	**1.96×10^{-3}**
Staged treatment-resistant depression	GLAD participants	8,062	0.03	0.06	0.41	0.35

Note: anhedonic symptoms were measured in participants of the COVID-19 Psychiatry and Neurological Genetics (COPING) study. Staged treatment-resistant depression was measured using the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in participants of the Genetic Links to Anxiety and Depression (GLAD) study. The h^2_{SNP} estimates were calculated using Genome-wide Complex Trait Analysis (GCTA). "N" refers to the number of individuals included in the genetic-relatedness matrix (GRM) in GCTA. P-values were calculated by GCTA.

LDSC: LDSC estimated the participants' highest anhedonic symptoms as the most heritable (h^2_{SNP}=0.07, SE=0.03, p=0.006). For staged treatment-resistant depression, LDSC gave a non-significant SNP-based heritability estimate of 0.05 (SE=0.04, p=0.20) (**table 4.2**).

After meta-analysing our GWAS results with the results of the staged treatment-resistant depression in the UK Biobank, the summary statistics had 12,114,577 SNPs. We filtered the summary statistics for overlapping SNPs (i.e., those present in both the GLAD GWAS and UK Biobank GWAS). This removed 7,347,155 SNPs altogether and left 4,767,422 remaining. The SNP-based heritability of meta-analysed GLAD-UK Biobank treatment-resistant depression was 0.02 (SE=0.01, p=0.08) (**table 4.2** and **S5**).

LDSC suggested that there was no evidence of genomic inflation for all phenotypes (i.e., lambda GC<1.05) (**table S4.5**). Manhattan plots and QQ plots for SNP associations for highest anhedonic symptoms, staged treatment-resistant depression, and meta-analysed GLAD-UK Biobank treatment-resistant depression are presented in **figures S4.4, S4.6** and **S4.8** respectively).

SNP and gene-based associations (FUMA)

In genome-wide analysis of highest anhedonic symptoms, staged treatment-resistant depression, and meta-analysed GLAD-UK Biobank treatment-resistant depression, no SNPs were associated at genome-wide statistical significance ($p<5x10^{-8}$) and no genes were significant after Bonferroni correction (**figures S4.4-S4.9**).

Genetic correlations between anhedonic symptoms, treatment-resistant depression, and antidepressant response

GCTA: A total of 13,199 participants with data on mean anhedonic symptoms and 8,012 participants with data on staged treatment-resistant depression were included (N_{total}=21,211). The genetic correlation between the COPING participants' mean anhedonic symptoms and staged treatment-resistant depression was not significant (r_g=0.40, SE=0.62, p=0.15) (**table S4.6**).

LDSC: The COPING participants' highest anhedonic symptoms were not significantly correlated with staged treatment-resistant depression in GLAD (r_g=0.10, SE=0.34, p=0.76), treatment-resistant depression in the UK Biobank (r_g=0.32, SE=0.44, p=0.47), or the meta-analysis of the two (r_g=0.26, SE=0.32, p=0.42). Additionally, there were no significant genetic correlations with the binary measure of antidepressant response by Pain et al. (2022) (r_g=-0.08, SE=0.32, p=0.79). *The genetic correlation with the continuous measure of antidepressant response (percentage improvement) could not be computed because its LDSC-estimated SNP-based heritability was not significant (Pain et al., 2022)* (**table S4.8**).

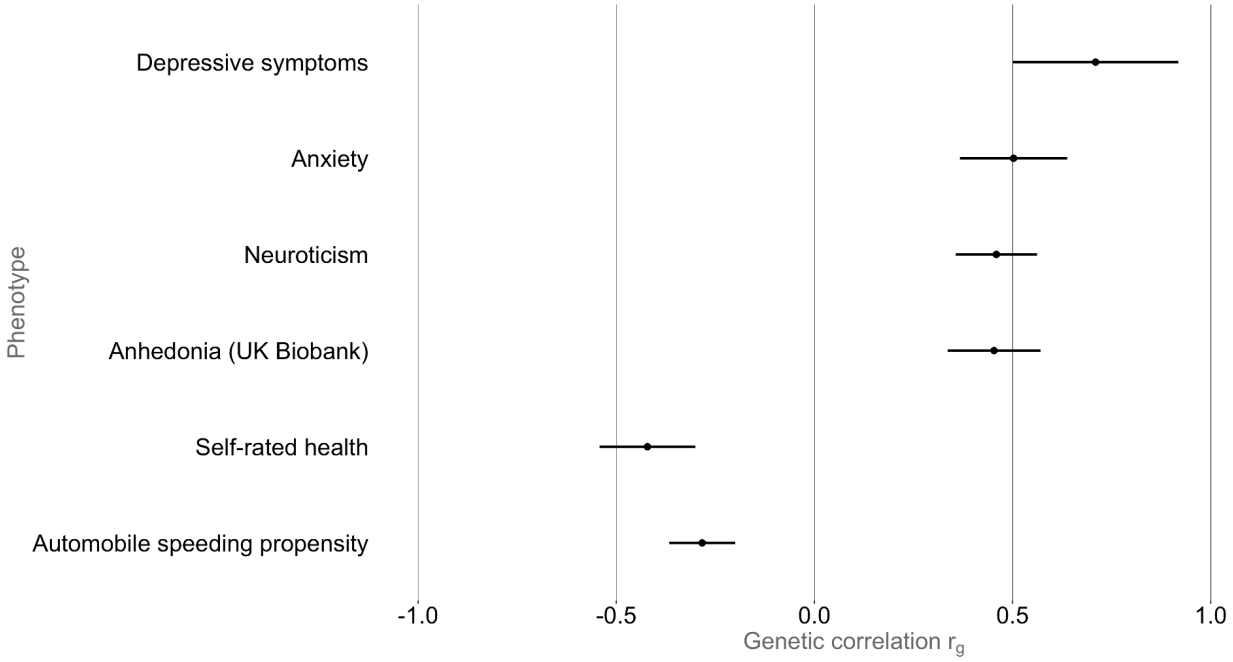
Genetic correlations between anhedonic symptoms and other traits

AD-MASQ-D30 anhedonic symptoms in COPING: Based on GWAS summary statistics, there were significant positive genetic correlations between the participants' highest anhedonic symptoms and four external traits, and significant negative genetic correlations with two traits.

The positive genetic correlations were with: depressive symptoms (r_g=0.71, SE=0.21, p=6.75x10^{-4}), anxiety (r_g=0.50, SE=0.14, p=2.10x10^{-4}), neuroticism (r_g=0.46, SE=0.10, p=7.60x10^{-6}), and anhedonia measured in the UK Biobank with the single PHQ9 item (r_g=0.45, SE=0.12, p=1.08x10^{-4}). The negative genetic correlations were with self-rated health (r_g=-0.42, SE=0.12, p=4.98x10^{-4}) and automobile speeding propensity (r_g=-0.28, SE=0.08, p=6.90x10^{-4}) (**figure 4.4**). All genetic correlations were significantly different to one except for depressive symptoms (p=0.16).

Figure 4.2. Significant genetic correlations between the participants' highest anhedonic symptoms and external traits.

Genetic correlations were estimated with Linkage Disequilibrium Score Regression (LDSC) and the extended 1000 Genomes Linkage Disequilibrium (LD) reference panel. Genetic correlations were significantly different to zero if they surpassed the Bonferroni-corrected alpha of 0.001 to correct for 42 tests. The bars represent the point estimates +/- their respective standard errors (SEs). Details of these external phenotypes are presented in table S4.7 and all LDSC genetic correlation results are presented in table S4.8.



UK Biobank GWAS of the single anhedonia item from the PHQ9: Similarly to anhedonic symptoms measured by the AD-MASQ-D30, the single item measure of anhedonia from the PHQ9 was not significantly correlated with staged treatment-resistant depression (r_g=0.23,

SE=0.17, p=0.18), binary treatment-resistant depression (r_g=0.57, SE=0.53, p=0.28), or a meta-analysis of the two (r_g=0.38, SE=0.15, p=0.01). Likewise, this phenotype was not significantly correlated with the binary measure of antidepressant response (r_g=0.006, SE=0.12, p=0.95). Regarding the traits that we found to be significantly genetically correlated with AD-MASQ-D30-assessed anhedonic symptoms, the single item measure of anhedonia from the PHQ9 was significantly genetically correlated with all of them. Each genetic correlation was in the same direction as its respective genetic correlation with AD-MASQ-D30-assessed anhedonic symptoms (**table S4.9**).

Discussion

We performed GWASs of continuous measures of anhedonic symptoms (N=13,433) and staged treatment-resistant depression (N=8,165) and investigated whether they were significantly genetically correlated. We also explored the relationship between anhedonic symptoms and antidepressant response using previously published GWAS summary statistics (Pain *et al.*, 2022). Anhedonia has been hypothesised as a potential factor influencing treatment resistance in individuals with MDD (McMakin *et al.*, 2012; Uher *et al.*, 2012). To our knowledge, our study is the first to investigate the genetic links between them. While prior research has shown that anhedonia and treatment-resistant depression are both heritable, our results suggest that the genetics underlying anhedonia, whether measured continuously via the AD-MASQ-D30 or the single item from the PHQ9, are not shared with either staged treatment-resistant depression or binary treatment-resistant depression (the latter of which was defined from electronic health records) (Fabbri *et al.*, 2021). Likewise, we did not find any significant genetic correlations with two measures of antidepressant response (Pain *et al.*, 2022). In phenotypic analyses, we found low but highly significant correlations between participants' anhedonic symptoms during the COPING study and their staged treatment-resistant depression when they completed the GLAD study survey (r=~0.2). Our results add evidence for a weak association between anhedonia and treatment resistance in MDD but do seem to suggest that this association is not primarily due to overlapping genetic risk factors.

The largest GWAS of anhedonia was performed by Ward et al. (2019) based on the single PHQ9 item, treated continuously, in UK Biobank participants. They reported a SNP-based heritability of

5.6% estimated by BOLT-LMM. We adopted a different approach in our study by collecting data on ten different aspects of anhedonia at multiple timepoints over the course of 16 months and collating them into a composite score which represented anhedonic symptoms. We performed a GWAS of the participants' highest, lowest, and mean anhedonic symptoms. In GCTA, which uses individual-level genotype data, the genome-wide common variant heritability (often termed SNP-based heritability) was 0.11 (SE=0.04, $p=3.83\times10^{-3}$). The participants' mean and lowest anhedonic symptoms also both had significant SNP-based heritability estimates (h^2_{SNP}=0.13 and h^2_{SNP}=0.09 respectively). While our GCTA estimates of heritability were significant for all phenotypes, only the participants' highest anhedonic symptoms was significantly heritable in LDSC (h^2_{SNP}=7.6%). Another UK Biobank GWAS of the single anhedonia item from the PHQ9 reported an LDSC SNP-based heritability of 3.8% when treated continuously and 8.1% when dichotomised (Thorp *et al.*, 2020). Since the highest anhedonic symptoms phenotype was the only one that was significantly heritable in LDSC, we used this phenotype to test for genetic correlations with other psychiatric and behavioural traits. We included the UK Biobank single item GWAS of anhedonia (Ward *et al.*, 2019) as one of these traits and the resulting genetic correlation was highly significant but moderate (r_g=0.45, SE=0.12, p=0.0001). This demonstrates that the additive genetic risk factors which contribute to variance in both traits are not completely overlapping.

Overall, our continuous anhedonia measure was more heritable than that estimated in either of the UK two Biobank single item analyses (Ward *et al.*, 2019; Thorp *et al.*, 2020). Considering our modest sample size compared to these two previous GWASs of continuous anhedonia (COPING N=13,433; Ward et al. (2019) N=375,275; Thorp et al. (2020) N=148,752), the magnitude and significance of our heritability estimates may owe to increased variance in the continuous measure of anhedonia through the use of a more detailed phenotyping approach. The moderate genetic correlation between two anhedonia phenotypes (COPING vs. Ward et al.) may also be due to the two different approaches taken to measure them. In our phenotypic analyses, we found only a moderate phenotypic correlation between anhedonic symptoms and this single item measure of anhedonia from the PHQ9 (r=0.59, SE=0.006, $p<2.22\times10^{-16}$). This adds further weight to the notion that the single anhedonia item from the PHQ9 and AD-MASQ-D30-assessed anhedonic symptoms, while related, only partially capture the same trait, even in the same individual at the same point in time.

A benchmark for evaluating these two differing modes of measuring anhedonia is to compare their respective associations with depression-related phenotypes. In terms of genetic results, we found that AD-MASQ-D30-assessed anhedonic symptoms were not significantly correlated with MDD, whereas the UK Biobank GWAS of the single anhedonia item from the PHQ9 was (r_g=0.77, p=1.34x10^{-139}) (Ward *et al.*, 2019). However, this is not wholly unsurprising given that the anhedonia symptom asked about in the PHQ9 is one of two core symptoms ∂for diagnosing MDD. We did find a high and significant genetic correlation between AD-MASQ-D30-assessed anhedonic symptoms and a previous meta-analysis GWAS of depressive symptoms (Okbay *et al.*, 2016) (r_g=0.71, SE=0.21, p=0.0007). When the genetic correlation of depressive symptoms (Okbay *et al.*, 2016) was calculated with the UK Biobank GWAS of the single anhedonia item from the PHQ9, we found an even stronger result (r_g=0.96, SE=0.03, p=1.8x10^{-213}) but neither of these estimates were significantly different to one (p<0.005). However, it should be noted that the depressive symptoms phenotype definition from Okbay et al. (2016) involved two questions, one of which is the single PHQ9 anhedonia item, which likely explains the high genetic correlation with the Ward et al. (2019) GWAS.

In phenotypic analyses of COPING participants, we found that depressive symptoms sum score (assessed via the PHQ9) was more strongly correlated with the single PHQ9 anhedonia item than the it was with AD-MASQ-D30-assessed anhedonic symptoms (PHQ9 anhedonia item r_{ph}=0.87 vs. AD-MASQ-D30=assessed anhedonic symptoms r_{ph}=0.59). This pattern remained even after we removed the anhedonia item from the overall PHQ9 sum score (PHQ9 anhedonia item r_{ph}=0.82 vs. AD-MASQ-D30 anhedonic symptoms r_{ph}=0.58). This finding, along with the fact that the UK Biobank single PHQ9 anhedonia item GWAS was highly genetically correlated with MDD, demonstrates that it is difficult to disentangle the single PHQ9 anhedonia item from MDD overall. Conversely, our measure of anhedonic symptoms was not significantly genetically correlated with MDD, suggesting they are biologically distinct traits. These findings have implications for phenotyping decisions in future studies of anhedonia. While the single PHQ9 item is a more efficient method to collect data on a large-scale it may not fully represent the complex, multi-faceted trait of anhedonia. Furthermore, using the single item measure may mean that genetic liability for MDD is captured because it is one of the core symptoms used to define a depression diagnosis, rather than transdiagnostic anhedonia which may be separate from depression.

Antidepressants are one of the most commonly prescribed medications but a substantial proportion do not show signs of improvement after their first course (Rush *et al.*, 2006). Treatment resistance is said to occur when an individual does not show clinically significant improvement in symptoms after two different types of antidepressant medications, each at an adequate dose and duration, and is a concerning issue facing public health due to associated distress, high rate of comorbidities, and health care costs (Fekadu, Donocik and Cleare, 2018). Previous studies of treatment resistance, using electronic health records or prescription data, have reported significant SNP-based heritability estimates of 8% (Li *et al.*, 2020; Fabbri *et al.*, 2021). In our study, we conducted a GWAS of staged treatment-resistant depression phenotypes in the GLAD study based on the MSM (Fekadu *et al.*, 2009), but did not have a significant SNP-based heritability based upon individual-level genotype data (GCTA; r_g=0.03, SE=0.06, p=0.35) or GWAS summary statistics (LDSC; LDSC r_g=0.02, SE=0.01, p=0.08). Post-hoc power analyses using the GCTA-GREML power calculator revealed that, with a sample size of 8,062, we were powered at 80% to detect a SNP-based heritability of 10.989%. Thus, a heritability of 2% would require a much larger sample to be significantly different to zero.

Other than the genetic correlation with the single anhedonia item from the PHQ9 and depressive symptoms, we found significant genetic correlations between the participants' highest anhedonic symptoms and four psychiatric and behavioural traits. Positive genetic correlations were with anxiety (r_g=0.50, SE=14, P=0.0002) and neuroticism (r_g=0.46, SE=0.10, p=7.6x10^{-6}). Negative genetic correlations were with self-rated health (r_g=-0.42, SE=0.12, p=0.0005) and automobile speeding propensity (r_g=-0.29, SE=0.08, 0.0007). We did not replicate significant associations found in Ward et al. (2019) with MDD, bipolar disorder, or schizophrenia. However, similar to findings from Ward et al. (2019), the participants' highest anhedonic symptoms were not genetically correlated with OCD (Ward *et al.*, 2019).

The relationship between anhedonia and anxiety, and anhedonia and neuroticism, is not well understood, but there has been some research proposing mechanistic links between them (Jacobson and Newman, 2014; Liao *et al.*, 2019; Winer, Jordan and Collins, 2019). The literature on anhedonia and risk-taking is mixed. In one study alone, different dimensions of anhedonia had both a positive and a negative relationship with risk-taking behaviours (Currin *et al.*, 2022), and another study found a link between some types of risk-taking behaviours but not others (Testa and Steinberg, 2010). The final significant genetic correlation discovered with anhedonic

symptoms was with self-rated health (r_g=-0.42, SE=0.12, p=5.0x10^{-4}) in a negative direction, as expected. Explanations for this include the fact that individuals with mental illnesses, especially MDD, are at a higher risk of comorbid physical and mental health disorders for a variety of reasons (Kessler, Merikangas and Wang, 2007; Robson and Gray, 2007; Gao *et al.*, 2013), and the possibility that individuals with more severe anhedonia are more likely to perceive their own health negatively.

The conclusions discussed here should be considered alongside some important limitations of our study design. First, we relied on self-reported data to phenotype anhedonic symptoms as well as treatment-resistant depression that may have suffered from recall bias. However, both measures use current symptoms, for which recall bias may be more limited (Young *et al.*, 2021). Nonetheless, parts of the MSM may have been retrospective for some participants (e.g., about how long their current/most recent depressive episode had lasted for, how many medications they had tried, and whether or not they had received ECT). The use of retrospective, self-reported data may have introduced noise into the staged treatment-resistant depression phenotype which may have contributed to low statistical power in addition to the small sample size. A further limitation is the fact that, because the MSM was only asked of individuals currently experiencing a depressive episode, people who may have once scored highly on the MSM, but have since shown symptomatic recovery, were not included. Even though these individuals may now be in remission, they would arguably be regarded as having treatment resistance. This may have reduced the variance in the phenotype which also could have contributed to low power.

Another limitation is the fact that the COPING data in our analyses was collected from May 2020 to September 2021. For some of this time period, the UK was in a national lockdown due to the COVID-19 pandemic. This unprecedented change in circumstance for individuals all over the world profoundly affected mental health (Xiong *et al.*, 2020). Thus, the COPING participants may have experienced better or worse mental health than usual which could have biased measures in either direction. A final limitation is the small sample sizes. GWASs of psychiatric disorders are now being regularly performed on samples well into the tens to hundreds of thousands (Pardiñas *et al.*, 2018; Wray *et al.*, 2018). Our sample sizes of ~13,000 and ~8,000 are potentially insufficient for discovery. The depth of our phenotyping approach to anhedonic symptoms may have compensated for low power, but our GWAS of treatment-resistant depression likely suffered from

both a small sample and a noisy phenotype. The GLAD study is continuing to recruit and genotype participants and therefore sample sizes available for these analyses will grow.

In conclusion, we found that anhedonic symptoms are heritable and share genetic associations with a number of other traits, including anxiety, neuroticism, and risk-taking. The use of a detailed phenotyping method and a continuous measure likely increased statistical power in our GWAS and suggests that this is an optimal way to study psychiatric traits in the future. We found no evidence for genetic links between anhedonia and treatment resistance or treatment-response in those with MDD. However, GWASs of treatment-response in MDD are currently underpowered. A priority for future research is the collection of data on a larger number of individuals to allow for the relationship between this important dimension of MDD and other traits to be studied effectively.

Funding

This work was supported by the National Institute for Health and Care Research (NIHR) BioResource [RG94028, RG85445], NIHR Biomedical Research Centre [IS-BRC-1215-20018], HSC R&D Division, Public Health Agency [COM/5516/18], MRC Mental Health Data Pathfinder Award (MC_PC_17,217), and the National Centre for Mental Health funding through Health and Care Research Wales. Jessica Mundy acknowledges funding from the Lord Leverhulme Charitable Grant. Helena L. Davies acknowledges funding from the Economic and Social Research Council (ESRC) as part of a PhD studentship. Brett N. Adey acknowledges funding through a pre-doctoral Fellowship from the NIHR (NIHR301067). Dr Christopher Hübel acknowledges funding from Lundbeckfonden (R276-2018-4581).

Conflicts of interest

Professor Gerome Breen has received honoraria, research or conference grants and consulting fees from Illumina, Otsuka, and COMPASS Pathfinder Ltd. Professor James Walters has received grant funding from Takeda for work unrelated to the GLAD study. The remaining authors have nothing to disclose.

Acknowledgements

We thank all GLAD study, EDGI, and NIHR BioResource volunteers for their participation, and gratefully acknowledge the National Institute for Health and Care Research (NIHR) BioResource, NIHR BioResource centres, NHS Trusts, and staff for their contribution. We also acknowledge King's College London, South London and Maudsley NHS Trust and King's Health Partners. We thank the National Institute for Health and Care Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Programme. We acknowledge use of the research computing facility at King's College London, *CREATE*. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, King's College London, or the Department of Health and Social Care.

References

Abramovitch, A. *et al.* (2014) 'Anhedonia in obsessive-compulsive disorder: beyond comorbid depression', *Psychiatry research*, 216(2), pp. 223–229.

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

Bogdan, R. and Pizzagalli, D.A. (2009) 'The heritability of hedonic capacity and perceived stress: a twin study evaluation of candidate depressive phenotypes', *Psychological medicine*, 39(2), pp. 211–218.

Bulik-Sullivan, B.K. *et al.* (2015) 'LD Score regression distinguishes confounding from polygenicity in genome-wide association studies', *Nature genetics*, 47(3), pp. 291–295.

Clark, L.A. and Watson, D. (1991) 'Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications', *Journal of abnormal psychology*, 100(3), pp. 316–336.

Currin, D.L. *et al.* (2022) 'The role of anhedonia in predicting risk-taking behavior in university students', *Journal of psychiatric research*, 155, pp. 451–457.

Fabbri, C. *et al.* (2021) 'Genetic and clinical characteristics of treatment-resistant depression using primary care records in two UK cohorts', *Molecular psychiatry*, 26(7), pp. 3363–3373.

Fekadu, A. *et al.* (2009) ‘A multidimensional tool to quantify treatment resistance in depression: the Maudsley staging method’, *The Journal of clinical psychiatry*, 70(2), pp. 177–184.

Fekadu, A., Donocik, J.G. and Cleare, A.J. (2018) ‘Standardisation framework for the Maudsley staging method for treatment resistance in depression’, *BMC psychiatry*, 18(1), p. 100.

Gao, K. *et al.* (2013) ‘Should an assessment of Axis I comorbidity be included in the initial diagnostic assessment of mood disorders? Role of QIDS-16-SR total score in predicting number of Axis I comorbidity’, *Journal of affective disorders*, 148(2-3), pp. 256–264.

Horan, W.P., Kring, A.M. and Blanchard, J.J. (2006) ‘Anhedonia in schizophrenia: a review of assessment strategies’, *Schizophrenia bulletin*, 32(2), pp. 259–273.

Jacobson, N.C. and Newman, M.G. (2014) ‘Avoidance mediates the relationship between anxiety and depression over a decade later’, *Journal of anxiety disorders*, 28(5), pp. 437–445.

Kessler, R.C., Merikangas, K.R. and Wang, P.S. (2007) ‘Prevalence, comorbidity, and service utilization for mood disorders in the United States at the beginning of the twenty-first century’, *Annual review of clinical psychology*, 3, pp. 137–158.

King's College London. (2022). King's Computational Research, Engineering and Technology Environment (CREATE). Retrieved March 2, 2022, from https://doi.org/10.18742/rnvf-m076

Kroenke, K., Spitzer, R.L. and Williams, J.B. (2001) ‘The PHQ-9: validity of a brief depression severity measure’, *Journal of general internal medicine*, 16(9), pp. 606–613.

Liao, A. *et al.* (2019) ‘Anxiety and anhedonia in depression: Associations with neuroticism and cognitive control’, *Journal of affective disorders*, 245, pp. 1070–1078.

Li, Q.S. *et al.* (2020) ‘Genome-wide association studies of antidepressant class response and treatment-resistant depression’, *Translational psychiatry*, 10(1), p. 360.

Li, S. *et al.* (2019) ‘Patients With Obsessive-Compulsive Disorder Exhibit Deficits in Consummatory but Not Anticipatory Pleasure’, *Frontiers in psychology*, 10, p. 1196.

Liu, W.-H. *et al.* (2016) ‘Anhedonia is associated with blunted reward sensitivity in first-degree relatives of patients with major depression’, *Journal of affective disorders*, 190, pp. 640–648.

Loas, G., Krystkowiak, P. and Godefroy, O. (2012) ‘Anhedonia in Parkinson’s disease: an overview’, *The Journal of neuropsychiatry and clinical neurosciences*, 24(4), pp. 444–451.

Mbatchou, J. *et al.* (2021) ‘Computationally efficient whole-genome regression for quantitative and binary traits’, *Nature genetics*, 53(7), pp. 1097–1103.

McAllister-Williams, R.H. *et al.* (2018) ‘Multiple-therapy-resistant major depressive disorder: a clinically important concept’, *The British journal of psychiatry: the journal of mental science*, 212(5), pp. 274–278.

McMakin, D.L. *et al.* (2012) ‘Anhedonia predicts poorer recovery among youth with selective serotonin reuptake inhibitor treatment-resistant depression’, *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(4), pp. 404–411.

Natta, L.E. *et al.* (2013) ‘An examination of anhedonia separate from depression in Alzheimer’s disease: a preliminary study’, *The American journal of geriatric psychiatry: official journal of the American Association for Geriatric Psychiatry*, 21(3), pp. S78–S79.

Okbay, A. *et al.* (2016) ‘Genome-wide association study identifies 74 loci associated with educational attainment’, *Nature*, 533(7604), pp. 539–542.

Pain, O. *et al.* (2022) ‘Identifying the Common Genetic Basis of Antidepressant Response’, *Biological psychiatry global open science*, 2(2), pp. 115–126.

Pardiñas, A.F. *et al.* (2018) ‘Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection’, *Nature genetics*, 50(3), pp. 381–389.

Roberts-West, L., Vivekananda, U. and Baxendale, S. (2022) ‘Anhedonia in epilepsy’, *Epilepsy & behavior: E&B*, 0(0). Available at: https://doi.org/10.1016/j.yebeh.2022.108966.

Robson, D. and Gray, R. (2007) ‘Serious mental illness and physical health problems: a discussion paper’, *International journal of nursing studies*, 44(3), pp. 457–466.

Rush, A.J. *et al.* (2006) ‘Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report’, *The American journal of psychiatry*, 163(11), pp. 1905–1917.

Sforzini, L. *et al.* (2021) ‘A Delphi-method-based consensus guideline for definition of treatment-

resistant depression for clinical trials', *Molecular psychiatry* [Preprint]. Available at: https://doi.org/10.1038/s41380-021-01381-x.

Spitzer, R.L. *et al.* (2006) 'A brief measure for assessing generalized anxiety disorder: the GAD-7', *Archives of internal medicine*, 166(10), pp. 1092–1097.

Testa, C.R. and Steinberg, L. (2010) 'Depressive symptoms and health-related risk-taking in adolescence', *Suicide & life-threatening behavior*, 40(3), pp. 298–305.

Thomsen, K.R. (2015) 'Measuring anhedonia: impaired ability to pursue, experience, and learn about reward', *Frontiers in psychology*, 6, p. 1409.

Thorp, J.G. *et al.* (2020) 'Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9', *Psychological medicine*, 50(14), pp. 2385–2396.

Uher, R. *et al.* (2012) 'Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms', *Psychological medicine*, 42(5), pp. 967–980.

Wardenaar, K.J. *et al.* (2010) 'Development and validation of a 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (MASQ)', *Psychiatry research*, 179(1), pp. 101–106.

Ward, J. *et al.* (2019) 'Novel genome-wide associations for anhedonia, genetic correlation with psychiatric disorders, and polygenic association with brain structure', *Translational psychiatry*, 9(1), p. 327.

Watanabe, K. *et al.* (2017) 'Functional mapping and annotation of genetic associations with FUMA', *Nature communications*, 8(1), p. 1826.

Winer, E.S. *et al.* (2017) 'Mapping the relationship between anxiety, anhedonia, and depression', *Journal of affective disorders*, 221, pp. 289–296.

Winer, E.S., Jordan, D.G. and Collins, A.C. (2019) 'Conceptualizing anhedonias and implications for depression treatments', *Psychology research and behavior management*, 12, pp. 325–335.

Wray, N.R. *et al.* (2018) 'Genome-wide association analyses identify 44 risk variants and refine

the genetic architecture of major depression', *Nature genetics*, 50(5), pp. 668–681.

Xiong, J. *et al.* (2020) 'Impact of COVID-19 pandemic on mental health in the general population: A systematic review', *Journal of affective disorders*, 277, pp. 55–64.

Yang, J. *et al.* (2011) 'GCTA: a tool for genome-wide complex trait analysis', *American journal of human genetics*, 88(1), pp. 76–82.

Young, K.S. *et al.* (2021) 'Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK'. Available at: https://doi.org/10.31234/osf.io/sf7b6.

Chapter 5. General discussion

Mood disorders represent a significant public health concern due to their high prevalence and associated disability. Family and twin studies, now supported by research from the field of statistical genetics, confirm that genetic variation is responsible for a substantial proportion of susceptibility for mood disorders in the population. Research into genetic influences on mood disorders aims to better understand their underlying biology, with the overarching hope of improving risk stratification, prevention, and treatment for the hundreds of millions of people affected by them. The purpose of this thesis was to explore various phenotyping approaches that can be applied to self-reported data about disorders and symptoms within the mood spectrum. Phenotyping approaches are of particular importance in mood disorder research due to their substantial heterogeneity. Many of the largest GWASs of mood disorders have employed a case/control design. The UK Biobank, the GLAD, and COPING study have collected rich self-reported data on the mood spectrum in samples large enough for genetic analyses. This thesis endeavoured to explore these self-reported data and apply three phenotyping approaches to them: diagnostic subtypes, continuous measures, and symptom-level analyses. The validity of mood disorder phenotypes for genetic studies is crucial for maximising the potential for discovery, decreasing the chance of type I and type II error, and increasing the likelihood of replication in other datasets. This final chapter draws together findings from the three empirical chapters. First, I will discuss **three lessons** that have been learnt from this body of research. These relate to phenotyping decisions that a researcher may encounter when studying the genetics of the mood spectrum. Second, I will consider general limitations relating to the study design. Last, I will discuss challenges that remain and directions for future research.

Three lessons learnt from this thesis

Lesson 1. Refined phenotypes offer insight into genetics as a driver of comorbidity

Researchers studying mood disorders are required to make a number of phenotyping decisions before conducting their study. One is the choice between taking a broad, sweeping view of a disorder category or zooming in on a **diagnostic subtype** or even a **specific**

symptoms/symptom domains. Within psychiatric genetics, sample size has traditionally been prioritised due to the challenges of studying the small effects of common genetic variants (Nishino *et al.*, 2018). As a result, many of the most well powered GWASs, such as those from the PGC, have focused on "umbrella" diagnostic categories (e.g., "depression", "major depressive disorder", "bipolar disorder", "schizophrenia", "anxiety disorders") (Howard *et al.*, 2018; Pardiñas *et al.*, 2018; Wray *et al.*, 2018; Purves *et al.*, 2019; Stahl *et al.*, 2019; Mullins *et al.*, 2021). This strategy has been undeniably successful for achieving the many aims of GWASs: discovering robustly-associated loci, increasing the variance explained by the additive effect of SNPs (SNP-based heritability), and calculating significant genetic correlations with other traits. But, this strategy has limited scope to explore differences in genetic architecture between disorder subtypes (termed "genetic heterogeneity") and the intricate relationships between them and other traits (Polimanti, 2022).

The heterogeneous nature of major depressive disorder was introduced in **chapter 1** (Ostergaard, Jensen and Bech, 2011; Fried and Nesse, 2015). This stems from variability in symptom presentation, recurrence, severity, duration, age of onset, sex, clinical specifiers, aetiology, comorbidity, and response to treatment, among other features (Nguyen *et al.*, 2022). Gender-specific subtypes have also been posited (Kuehner, 2017). Individuals classified as having the umbrella category of "major depressive disorder" may actually present with diverse disorder manifestations (Fried and Nesse, 2015). As a result, major depressive disorder can be separated into subtypes which can then be examined individually. In this thesis, **chapter 2** and **chapter 4** collectively studied five **diagnostic subtypes** of major depressive disorder: 1) recurrent depression, 2) single episode depression, 3) major depressive disorder with self-reported trauma, 4) major depressive disorder without self-reported trauma (**chapter 2**), and treatment-resistant depression (**chapter 4**).

In **chapter 2**, I explored whether recurrence and self-reported trauma were relevant factors in major depressive disorder's genetic overlap with PTSD. The rationale behind this was rooted in the high prevalence of comorbid major depressive disorder in individuals with PTSD, the fact that trauma is a key risk factor for both, and emerging evidence for a heritable basis for trauma response. I hypothesised that the heritability of trauma sensitivity is shared between major depressive disorder and PTSD, and this could be a factor driving their high levels of comorbidity among the trauma-exposed. I used **diagnostic subtypes** of major depressive disorder to test my hypothesis. The main conclusions from **chapter 2** were, firstly, the underlying additive genetic

basis of recurrent major depressive disorder was not significantly more correlated with that of PTSD compared to single-episode major depressive disorder. Likewise, at a genome-wide level, major depressive disorder in the presence of self-reported trauma was not more genetically correlated with PTSD than major depressive disorder in the absence of self-reported trauma. However, this lack of difference may have resulted from statistical power that was too minimal to detect small differences. But, by looking at the participants' individual genetic liability through PRSs, I found that individuals with major depressive disorder who self-reported trauma had a significantly higher genetic liability for PTSD than individuals with major depressive disorder who self-reported no trauma (Mundy *et al.*, 2021). This provided some evidence in favour of my hypothesis, although I recommended that future studies should attempt to replicate this finding in external cohorts.

All the UK Biobank participants included in the case groups in **chapter 2** could be categorised as having "major depressive disorder" but focusing on the subtypes was helpful for further unravelling the intricate relationship with PTSD. Therefore, a lesson from this thesis is that refined mood disorder phenotypes, such as **diagnostic subtypes**, hold promise for understanding more about whether shared genetic liability could drive comorbidity. Subtypes may have distinctive underlying biology, including their additive genetic risk factors, which contribute to their disease manifestations and would be missed when grouped together under one umbrella diagnostic category (Feczko *et al.*, 2019). In the Australian Genetics of Depression Study (AGDS), which runs parallel to the GLAD study, ubiquitous genetic heterogeneity among subtypes of major depressive disorder has been reported (Mitchell *et al.*, 2022). In the UK Biobank, a recent study showed similar findings based on sixteen major depressive disorder diagnostic subtypes. While the inter-genetic correlations ranged 0.55-0.86, indicating a degree of genetic sharing, there were some clear differences between the genetics of the individual subtypes. For instance, depression with mild impairment or later age of onset had lower SNP-based heritability estimates compared to "major depressive disorder" as a general diagnostic category. Also, clinically challenging subtypes (e.g., earlier age of onset, presence of suicidal ideation/self harm, recurrent episodes, and severe impairment) had higher genetic correlations with other psychiatric disorders, such as schizophrenia, compared to less clinically challenging/milder subtypes (Nguyen *et al.*, 2022). I also found evidence of genetic heterogeneity among the major depressive disorder subtypes in **chapter 2** (see **table S2.4**).

In **chapter 4**, I tested the hypothesis that additive genetic risk factors for treatment-resistant depression were partly shared with anhedonia, as there have been two studies which purported to find a link between them (McMakin *et al.*, 2012; Uher *et al.*, 2012). I found no evidence in support of this hypothesis. However, power issues (mentioned in **chapter 4**) could have been a cause of this. Also in **chapter 4**, I studied an even more refined phenotype than a **diagnostic category** by focusing on a **specific symptom domain** of major depressive disorder. Our analysis revealed that anhedonic symptoms had a SNP-based heritability of 9-13% and showed significant positive genetic correlations with depressive symptoms, anxiety, and neuroticism. Aside from depressive symptoms, there is a dearth of published literature on the relationship between anhedonia and these internalising phenotypes, but one 2017 study proposed that anhedonia could be a factor explaining the high rates of anxiety-depression comorbidity. In this study, anxiety and depression were positively associated in individuals showing anhedonic tendencies (e.g., avoidance of activities). They also found that anxiety led to anhedonia, which then encouraged the development of a depressive episode (Winer *et al.*, 2017). They proposed a pathway whereby anxiety leads to avoidance of pleasurable or rewarding experiences which, in turn, puts the individual at high risk of developing major depressive disorder (Jacobson and Newman, 2014). For example, social anxiety might make a person withdraw from their peer group, avoid meeting new people, or stop attending activities where there are (or even might be) large groups. As a result of missing out on these important, sociable, and pleasurable experiences, they may begin to experience anhedonia which, in turn, makes them more vulnerable to experiencing depression. Previous research shows that a lack of positive experiences is predictive of depression (Spinhoven *et al.*, 2011). This hypothesis could be easily tested with the data used in this thesis, since the COPING study has collected data longitudinally on depression, anxiety, and anhedonic symptoms on upwards of 20,000 UK-based participants, many of which have genetic data available.

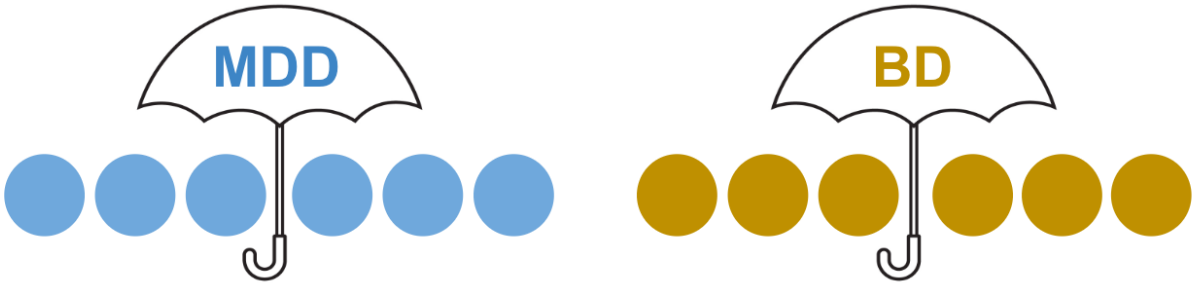
Focusing on **specific symptoms**, as opposed to disorder categories, is beginning to attract more attention in psychiatric genetics research. For example, Thorp et al., (2020) studied each of the nine symptoms in the PHQ9 and found widespread genetic heterogeneity, both in the individual symptom's heritability, and their genetic associations with psychiatric disorders and complex traits (Thorp *et al.*, 2020). They also found that this phenotyping approach led to the discovery for seven genome-wide significant loci which had not been identified in previous major depressive disorder GWASs. A similar approach was taken by Nagel et al., (2018) who conducted item-level GWASs of neuroticism items (Nagel *et al.*, 2018). Studies such as these highlight the importance of

examining the individual symptoms which make up a composite/sum score or a binary case/control phenotype, as grouping them together may hide important differences in genetic architecture and relationships with other traits. In **chapter 3** of this thesis, a similar approach was adopted by applying factor analysis to the items in the MDQ. The results showed that the MDQ may comprise three symptom subgroups but, contrary to expectations, there were no significant differences between the three **symptom subgroups** and their genetic correlations with other psychiatric and behavioural traits. This was likely due to low power to detect small differences, thus emphasising the need for large sample sizes when comparing refined phenotypes to one another.

Summary: There is value in refining phenotypes for genetic studies of mood disorders. This can be achieved by delving into **diagnostic subtypes** or going even further by focusing on **specific symptoms** or **symptom subgroups**. When measured with precision, refined phenotypes reduce heterogeneity and increase statistical power (see **figure 1.5**). Not only this, but differentiating between them may reveal specific genetic mechanisms which are lost when studying general "umbrella" diagnostic categories, including shared genetic liability with other disorders. The application of genetic methods to refined phenotypes in **chapter 2** and **chapter 4** revealed possible mechanisms explaining comorbidity. The results of the staged treatment-resistant depression GWAS (**chapter 4**) and the GWASs of the MDQ (**chapter 3**) emphasise that sufficient sample sizes and valid phenotyping instruments are paramount for studying the genetics of refined mood disorder phenotypes and looking for subtle differences between them. A key question which should be addressed in the future is whether diagnostic subtypes represent variations of one disorder or biologically distinct disorders which share symptoms. From the perspective of genetic aetiology, this question will only be answered with large sample sizes combined with precision phenotyping.

Figure 5.1. Mood disorders are heterogeneous.

General “umbrella” categories can be refined into diagnostic subtypes based upon episode recurrence, symptom severity, treatment resistance, aetiology, clinical specifiers, and other features (represented by the circles beneath the umbrellas).



Lesson 2. Continuous measures can increase statistical power in a GWAS if used appropriately

Statistical power is influenced not only by sample size, but by the way the phenotype is assessed (see **figure 1.5**). Phenotypes for GWASs can be measured continuously (e.g., regressing the number of copies of a risk allele a participant carries against their numeric score), or by categorising participants as cases and controls (e.g., comparing allele frequencies between cases and controls). Compared to a case-control design, **continuous measures** (e.g., a sum score representing varying levels in symptom severity) have greater statistical power because a smaller amount of information is lost to artificially categorising participants and phenotypic variance is maintained (van der Sluis *et al.*, 2013). For instance, binary categorisation of participants into cases (who are given a value of one) and controls (who are given a value of zero) means that all information on phenotypic differences between cases, or phenotypic differences between controls, is lost. Furthermore, individuals who are very mild cases may be genetically very similar to controls, which can reduce power to detect genetic differences between the two experimental groups. Despite the benefits of using a **continuous measure**, the largest published GWASs of mood disorders have opted for binary phenotypes (Wray *et al.*, 2018; Howard *et al.*, 2019; Levey *et al.*, 2021; Mullins *et al.*, 2021). This is because these large-scale GWASs required meta-analysis of different cohorts around the world and constructing a **continuous measure** would

require access to, and harmonisation of, the underlying symptom-based tools that defined cases and controls. This would be a hugely complex and time-consuming task.

If researchers *do* have access to uniform symptom-based data among participants, using a **continuous measure** for a GWAS may be an alternative to a case-control design. The UK Biobank, the GLAD study, and the COPING study included many questionnaires and scales which can be used to construct **continuous measures** of dimensions of the mood spectrum. In this thesis, I applied **continuous measures** to three mood disorder phenotypes: manic symptoms (**chapter 3**), anhedonic symptoms (**chapter 4**), and treatment resistant depression (**chapter 4**). When viewed collectively, the results of these chapters offer insight into the optimal way to construct **continuous measures** for genetic studies of psychiatric disorders.

First, when the phenotype is intended to represent a participant's collective experience of many symptoms, the individual items that make up the composite score must have adequate internal consistency. In **chapter 3**, the concurrent MDQ items had poor internal consistency: the items "irritability", "racing thoughts", and "concentration difficulties" were not correlated with the other items (**figure 3.1**). These three items loaded onto the factor which I named the *cognitive* factor, and this factor did not correlate with the other two factors (*energy/activity* and *impulsivity*). This was reflected in the genetic analysis: the *cognitive* factor was not genetically correlated with the other two factors (**table 12** in **appendix 3**). This lack of internal consistency introduced phenotypic heterogeneity which likely hampered power in the GWAS. As a result, the sum score, representing the number of concurrent manic symptoms a participant had experienced, was not heritable. Thus, when researchers are planning to perform a GWAS of symptoms in a continuous manner, they should first ensure that the items in the score are well correlated, as this suggests that they all capture different elements of the same underlying trait.

Second, **continuous measures** comprising detailed information (e.g., multiple scale items) have more power than continuous measures based on a single item from a scale. Part of the reason for this is that composite scores that are based on multiple scale items likely have larger ranges than single items, and large ranges do a comparatively better job at representing the phenotypic variation in the population. In **chapter 4**, GCTA-GREML detected a SNP-based heritability of 8% for anhedonic symptoms assessed by the AD-MASQ-D30 with a sample of 13,000 COPING participants. The Ward et al., (2019) GWAS detected a SNP-based heritability of 6% for anhedonia assessed by the single item from the PHQ9 with a sample of 375,000 UK Biobank

participants (Ward *et al.*, 2019). By using the AD-MASQ-D30, **chapter 4** utilised data on ten different facets of anhedonia, each scored 0-4. This contrasts with Ward et al., (2019)'s use of a single item from the PHQ9, scored 0-3. The inclusion of a larger amount of data into the phenotype increased the range (0-40 versus 0-3) and likely boosted statistical power to compensate for the fairly modest sample size. I also found that the single anhedonia item from the PHQ9 had a genetic correlation with major depressive disorder that was not significantly different to one. By contrast, the more detailed measure of anhedonic symptoms was not significantly genetically correlated with major depressive disorder. Anhedonia is a core symptom of major depressive disorder but is also present transdiagnostically. A more detailed phenotyping approach meant that I was able to create a continuous measure of anhedonia that is partially distinct from major depressive disorder. Compared to the option of using the single PHQ9 item, my phenotyping method with the AD-MASQ-D30 may be more useful for future studies wishing to understand the role genetics plays in the presence of anhedonia in other psychiatric and neurodevelopmental disorders compared to the simple item measure from the PHQ9, which likely only captures liability for major depressive disorder.

Third, when a **continuous measure** is intended to represent symptoms of a disorder, it is necessary to include information about severity and/or duration, rather than just a count of symptoms. In **chapter 3**, I posited that the lack of significant genetic correlation with bipolar disorder may have been due to the composite score not containing information about severity or duration. I compared this to a previous depressive symptoms GWAS which, being based on the PHQ9, included such information and found a significant positive genetic correlation with bipolar disorder (Okbay *et al.*, 2016; Mullins *et al.*, 2021). Overall, symptom counts alone are potentially not sufficiently informative for genetic studies.

Lastly, retrospective data may not be suitable for constructing **continuous measures**. In **chapter 3**, the MDQ assessed *lifetime* rather than *current* manic symptoms. This could have led to the issues with internal consistency for the concurrent items. Likewise, some of the questions in the MSM might have required some of the participants to recall information from a long time ago, especially those who had been in their current depressive episode for a very long period (e.g., two years or more). In a paper that I contributed to during my PhD, we showed that GLAD participants rated their depression and anxiety as worsening during the pandemic. But, when we looked at objective current measures, the participants had actually shown an overall improvement in their symptoms (Young *et al.*, 2022), thus demonstrating the inaccuracy of retrospective data.

In **chapter 4**, individuals who had been prescribed numerous different medications, or those who had been depressed for many years, may have given erroneous answers to the first three questions of the MSM. Exacerbating this is the fact that memory is sometimes impaired in individuals suffering from depression (Burt, Zembar and Niederehe, 1995; Rock *et al.*, 2014). This may be especially important in my interpretation of the staged treatment-resistant depression phenotype, since I limited the sample to those who were currently depressed.

Summary. In theory, **continuous measures** increase statistical power in GWAS over case-control studies and offer a promising avenue for increasing replicable discoveries and the detectable SNP-based heritability. However, there are some important caveats. The first is that, when constructing a composite score of symptoms, all contributing items should be adequately correlated to ensure good internal consistency. The second is that composite scores should utilise as much information as possible so that the numeric variance represents the phenotypic variance in the population as best as it can. Third, continuous symptom scores should try to include data on severity and duration. Lastly, retrospective data may compromise reliability and validity of **continuous measures**.

Lesson 3. Clinical tools are not always suitable for data collection via self-reports

In this thesis, I measured mood disorder symptomatology with self-reported data from a variety of scales and instruments. In **chapter 3**, a mania screener (the MDQ) was used. In chapter 4, a psychometric scale (the AD-MASQ-D30) and a clinical staging tool (the MSM) was used. The results of **chapter 4** suggested that the AD-MASQ-D30 captured anhedonia with a fair degree of accuracy. For instance, the symptoms in the AD-MASQ-D30 had a high genetic correlation with depressive symptoms, as well as moderate phenotypic correlations with the anhedonia item from the PHQ9 and depressive symptoms. By contrast, **chapter 3** suggested that self-reported answers to the MDQ, a clinical screening tool, did not reflect the type of hypomanic and manic symptoms experienced in bipolar disorder. **Chapter 4** revealed problems with collecting self-reported data via the MSM, which is also a clinical instrument. Thus, a lesson learnt from this thesis is that tools intended for use by mental health professionals are not always appropriate for collecting self-reported data via an online survey.

Treatment response is a hugely complex area of psychiatry. Indeed, a whole spectrum that involves different levels of "response" has been previously posited (**figure 5.2**) (McAllister-Williams *et al.*, 2018; Sforzini *et al.*, 2021). To gain an understanding of where a patient is positioned along this spectrum, information on length of depressive episode, medications, alternative treatments, and episode severity needs to be assessed, hence the existence of staging tools such as the MSM (Fekadu *et al.*, 2009). The staged treatment-resistant depression phenotype in **chapter 4** was not heritable. The fairly small sample size (N=8,154) could have been a cause of this. However, another GWAS of stages of treatment resistance, and a GWAS of a related phenotype (antidepressant response), both had significant non-zero SNP-based heritabilities with even smaller samples (N=3,452 and N=5,151 respectively) (Wigmore *et al.*, 2020; Pain *et al.*, 2022). These two previous GWASs assessed treatment resistance/antidepressant response objectively via prescription records/comparing depressive symptoms pre- and post-treatment respectively. By contrast, the assessment of treatment-resistant depression in the GLAD study relied on self-reported data. The MSM's developers state that it is intended for use by "a clinician with mental health training" (Fekadu, Donocik and Cleare, 2018) thus suggesting that the staging process benefits from interaction between patient and professional. Accuracy might significantly improve if the patient is known to the mental health professional and/or they have objective answers to the MSM's questions, including the exact number of antidepressant and/or augmentation medications from prescription records, and the length of current depressive episode. Furthermore, a group of experienced psychiatrists in the UK suggested that psychosocial factors, such as the patient's mindset and willingness to engage in antidepressant treatment, should be assessed alongside more obvious indicators of treatment resistance (McAllister-Williams *et al.*, 2018). Overall, treatment response and resistance requires an enormous complexity of information. Such information could be garnered from the patient in a one-to-one medical setting, but not via an online survey. For these reasons, my assessment of treatment resistance in the GLAD study might have lacked validity and this could have contributed to noise in the GWAS.

The same could be true for the MDQ which, although self-scored (Hirschfeld *et al.*, 2000), is intended for screening patients for possible bipolar disorder in a primary-care setting rather than data collection via an online survey. Bipolar disorder, despite being one of the most debilitating psychiatric disorders, is difficult to classify correctly (Phillips and Kupfer, 2013). As a result of this, unipolar depression is a common misdiagnosis (Hirschfeld, Lewis and Vornik, 2003). While mania is more pronounced, hypomania is milder and can often go unnoticed. Further complicating the

diagnosis of bipolar disorder is the fact that insight into symptoms is often impaired, even in euthymic periods, among individuals with bipolar disorder (Varga *et al.*, 2006). Not only this, but hypomania and mania show a high degree of phenotypic heterogeneity, and the symptoms involved are experienced in other disorders such as schizoaffective disorder (American Psychiatric Association, 2013). Thus, screening tools with good psychometric properties are essential for identifying individuals who potentially have undetected bipolar disorder and enabling access to effective treatments. Mentioned already in **chapter 3** is the attenuation of the MDQ's sensitivity and specificity when removed from clinical contexts (Miller, Johnson and Eisner, 2009). In **chapter 3**, I found that, based on the symptoms alone, the MDQ had no significant genetic correlation with bipolar disorder. Even with information on duration and severity of symptoms, the MDQ screener poorly identified individuals with bipolar disorder in the GLAD study.

A previous study showed that hypomania assessed by the MDQ was not significantly associated with bipolar disorder PRS constructed in a community sample (Hosang *et al.*, 2021). GLAD participants, although recruited based on lifetime experience of depression and/or anxiety, do not represent a clinical sample. Nonetheless, given that a higher proportion of individuals who have experienced depression develop bipolar disorder than those who have never experienced depression, I assumed that levels of hypomanic/manic symptoms would be more severe than a traditional community sample. However, the GLAD participants' subjective experience of the symptoms asked in the MDQ may not reflect those involved in hypomania/mania but another, unknown, dimension of psychopathology. This could explain why I observed significant genetic correlations with an array of psychiatric and behavioural traits but not with bipolar disorder. Despite its recommendation from clinical experts, the results of **chapter 3** suggest that the MDQ was not the optimal mode of assessing hypomanic/manic symptoms in the GLAD study.

During my PhD, I wrote diagnostic algorithms for bipolar disorder type I and type II using MDQ data in GLAD, EDGI, and NBR participants. Based on the results of **chapter 3**, we decided that the MDQ is not suitable for this purpose. For example, of the 2,428 participants identified with lifetime mania, only 43% self-reported a diagnosis of bipolar disorder by a healthcare professional. An important caveat here, however, is that bipolar disorder can go undiagnosed for up to ten years (Hirschfeld, Lewis and Vornik, 2003; Mantere *et al.*, 2004; Drancourt *et al.*, 2013), and unipolar depression is the most likely misdiagnosis (Hirschfeld, Lewis and Vornik, 2003). Given the eligibility criteria of the GLAD study, it is highly probable that a proportion of the participants have undiagnosed bipolar disorder. This may have contributed to the misalignment between the

diagnostic algorithm and the single-item self-report measure. Having said this, of the 2,760 participants who *did* self-report a diagnosis, 34% were categorised as having *never* experienced mania or hypomania by the algorithm, thus calling into question the utility of this MDQ-based algorithm to identify bipolar disorder cases and controls in future studies. By comparison, the case/control status designation from diagnostic algorithms of other major depressive disorder and anxiety disorders in GLAD and NBR participants' (derived from answers to the CIDI-SF), while showing variable overlap with self-reported diagnoses, was more often correct than incorrect. Algorithms for major depressive disorder had the highest accuracy (84% correct overlap) (Davies *et al.*, 2022). Questions from a structured interview, which are specifically designed for epidemiological assessments, might have been a better choice than the MDQ, although they would still suffer from issues of retrospective data collection.

Overcoming the difficulties of assessing hypomania/mania in research studies is not a simple task. Digital phenotyping methods, which collect data on moment-by-moment symptoms but also more objective indicators of mood (e.g., keyboard strokes, speech patterns, sleep, and activity levels) might offer a better route (McInnis, Gideon and Mower Provost, 2017; Zulueta *et al.*, 2018; Ebner-Priemer *et al.*, 2020; Orsolini, Fiorani and Volpe, 2020). However, this involves longitudinal assessments with technology equipped for digital phenotyping which is much more intensive and costly than administering self-reported online surveys. Digital phenotyping may yield richer and more precise measures of hypomanic/manic symptoms but the resulting sample sizes would likely be insufficient for genetic analyses.

Summary: Clinical tools, especially those intended for screening, may not be suitable for studying the mood spectrum when used to collect self-reported data for two reasons. First, they are designed for interaction between physician and patient which is lost during self-report surveys. Second, they are intended for patients who are presenting at primary care settings rather than large-scale epidemiological assessments. As a result, their ability to accurately capture specific mood disorder symptoms/dimensions may be lessened when taken out of this context. For particularly challenging mood disorder phenotypes, alternative approaches to data collection (e.g., medical records or digital phenotyping) may be preferred, although these come with limits on feasible sample sizes.

Figure 5.2. A proposed spectrum of treatment response in major depressive disorder (MDD) (McAllister-Williams *et al.*, 2018; Sforzini *et al.*, 2021).

In a paper by Sforzini et al., (2021) which draws upon a 2018 paper by McAllister-Williams et al., (2018), a spectrum of treatment response was proposed. This ranges from partially-responsive depression (a 25-50% reduction in depressive symptoms after treatment), to treatment-resistant depression (no clinically significant improvement in symptoms after two antidepressant treatments), to multi-therapy resistant MDD (no clinically significant improvement in symptoms after more than two interventions, including non-pharmacological treatments), and finally to refractory depression (absence of response to all currently available treatments).

Limitations

Within each empirical chapter, study-specific limitations have been raised. In this next section, broader challenges that have arisen from the research presented in this thesis are discussed.

The study samples lack generalisability

The first limitation is the lack of generalisability of the study samples included in this thesis. This stems from selection biases regarding demographic and health-related characteristics. Selection bias refers to a situation where the way in which research participants are recruited means that, collectively, they do not represent the population from which they were sampled (Holmberg and Andersen, 2022). Between 2006-2010, over half a million participants provided their data to the UK Biobank. This represented just 5% of the original number who were sent invitations. Compared to the UK population as a whole, those who chose to participate were more likely to be of a White ethnic background, be female, and come from a higher socioeconomic background (Fry *et al.*, 2017). In addition to this, the UK Biobank suffers from “healthy volunteer” selection bias. This refers to a phenomenon whereby individuals who participate in research show, on average, higher levels of health than those who tend not to participate (Manolio *et al.*, 2012). For instance, UK Biobank participants tended to have a lower BMI and a smaller waist circumference, report fewer diseases, and have lower rates of all-cause mortality and cancer than those in the

same age bracket from the general population (Fry *et al.*, 2017). This is not wholly unsurprising, participating in research involves dedication of time and effort, and those with poorer health may not have the capabilities or motivation to engage with the research.

The GLAD study suffers from similar concerns about demographic generalisability, including a lack of ethnic diversity (95% of participants are White), high levels of education (54.8% have a university degree), and primarily female participants (79.7%) (Davies *et al.*, 2019). An extra consideration is the GLAD study's overall picture of mental health. Unlike the UK Biobank, the GLAD study recruited individuals specifically on the basis of having lifetime experience of depression and/or an anxiety disorder. Accordingly, the health-related characteristics differ substantially from the UK Biobank. Initial analyses show that GLAD participants have high levels of symptom severity and comorbidity. Furthermore, trauma exposure and PTSD are frequently reported by the participants: as much as 63% self-report a traumatic event, and 60% meet criteria for PTSD in the last month (Davies *et al.*, 2019). This means that GLAD participants fall towards the severe end of the psychopathology spectrum compared to individuals with depression and/or anxiety in the general population.

Lack of generalisability is problematic for studies wishing to provide accurate prevalences of certain exposures and outcomes. But, this was not the goal of this thesis. In the body of research presented here, lack of generalisability may be a problem for the construction of valid phenotypes. For instance, in **chapter 2**, summary statistics from GWASs of major depressive disorder with and without trauma in UK Biobank MDQ respondents were assessed for a genetic association with PTSD. Here, trauma exposure was assessed via the CTS, ATS, and PCL6 and events with a >2.5 odds ratio with major depressive disorder were combined to create a single binary variable (Coleman *et al.*, 2020). Traumatic experiences among those of a higher socioeconomic status may have different consequences to those who were socioeconomically disadvantaged. Factors such as levels of social support and access to treatment, which are generally higher among more socioeconomically advantaged groups (Weyers *et al.*, 2008; McMaughan, Oloruntoba and Smith, 2020), may moderate post-trauma mental wellness and resilience (Ozbay *et al.*, 2007; Lee, 2019; Cho and Bulger, 2021), including whether or not they develop major depressive disorder. Therefore, the characteristics of the UK Biobank sample may have influenced which traumatic events were selected. By studying a narrow section of society, the research presented in **chapter 2** may not have captured the full range of experiences in this important risk factor for mood disorders. Likewise, the high rates of trauma and PTSD reported by GLAD participants could have

contributed to the lack of genetic validity of the MDQ items presented in **chapter 3**. For instance, I found a phenotypic correlation of 0.41 ($p<2x10^{-16}$) between the MDQ sum score and PTSD symptoms. This association may have artificially inflated the genetic correlation with PTSD (as seen in the genetic correlation results), thus calling into question the external validity of the findings.

The aims of this thesis were to explore phenotyping approaches that can be applied to self-reported data about the mood spectrum's disorders and symptoms. Research studies which expect to gather information on a more representative slice of the UK population may find different results to those presented in this thesis. Thus, the conclusions drawn about optimal methods for phenotyping mood disorders may not have applicability to more other study populations. An example is the upcoming study Our Future Health which hopes to recruit five million UK-based adults. Their aim is to ensure that this five million represents the UK population on a variety of demographic features. Unless increasing diversity is specifically prioritised with sufficient funding and time, as Our Future Health has done, the ascertainment biases discussed above are likely to persist in research studies of the future. Therefore, the conclusions presented here are still useful for researchers designing studies to study the mood spectrum. The likely biases of a study sample (demographic, health-related, or otherwise) should ideally be anticipated and given due consideration during initial deliberation about which scales and measures to include. For instance, forthcoming epidemiological or genetics-focused studies should take note of the MDQ's unexpected poor ability to assess symptoms specific to hypomania/mania. If these studies recruit participants by similar means as the GLAD study, and therefore expect similar biases, alternative hypomania/mania scales should be prioritised.

The GWASs presented in this thesis only included individuals of European ancestries

A second shortcoming, which is one faced by the field of psychiatric genetics and genetics more broadly (Haga, 2010; Sirugo, Williams and Tishkoff, 2019), is the lack of ancestral diversity in the GWAS samples. Currently, there is a disproportionate amount of published GWASs which contain only individuals of European genetic ancestries (Haga, 2010; Martin *et al.*, 2019; Mills and Rahal, 2019). **Chapter 1** already highlighted this grave issue. All the empirical chapters presented in this

thesis restricted their GWAS samples to individuals of European ancestries. The primary reason for this relates to sample size. As mentioned in **chapter 1**, the statistical power in genetics studies of polygenic traits and diseases is partly dependent on the number of individuals included in the study sample (see **figure 1.5**). Different traits and diseases require varying minimum sample sizes due to their unique genetic architectures (e.g., polygenicity and heritability). In the UK Biobank, the GLAD study, and the COPING study, individuals from racially minoritised groups constitute <5% of the overall study sample (Fry *et al.*, 2017; Davies *et al.*, 2019). Thus, after further limiting the sample to individuals with information on the phenotypes of interest for this thesis, they would have been insufficiently sized to conduct well-powered GWASs.

The challenge faced here is not unique. The lack of diverse participants in GWAS samples has persisted since the dawn of the GWAS method (see **figure 1.2**). At the root of the overabundance of published genetic studies involving only individuals of European ancestries is a complex web of *"logistical, systematic, and historical factors"* (Popejoy and Fullerton, 2016). Oft-cited logistical factors include the vast majority of researchers who publish scientific GWAS being based in North America or Europe, and the well documented challenges of engaging individuals from racially marginalised backgrounds in scientific research (Paskett *et al.*, 2008; Haga, 2010). Both issues relate to wider systematic factors, such as disparities in funding allocation, and historical factors, such as mistrust and suspicion of medical organisations. This mistrust stems from countless examples of horrific medical exploitation of racially minoritised groups throughout history. The lower uptake of the COVID-19 vaccine among such groups in the UK, as observed in many other countries too, serves as a pertinent reminder that regaining the confidence of minoritised groups in science and medicine is not straightforward (Dolby *et al.*, 2022). These historical factors are even more pertinent for research which asks participants to provide a DNA sample. Worries, or a lack of understanding, about the collection, storage, and use of DNA may contribute to individuals from racially minoritised backgrounds being even less likely to feel comfortable participating in genetics research (Moorman *et al.*, 2004; Catz *et al.*, 2005).

The absence of diversity in the genetic research presented in this thesis has implications for the wider conclusions drawn. The genetic results may generalise poorly to those of non-European ancestry. The scientific community has highlighted examples of ethnicity influencing association between biomarkers and outcomes. As a result, a “one-size-fits-all” approach is incorrect and some clinical tests need to be adjusted for application in different groups (Veeranna *et al.*, 2013; Rappoport *et al.*, 2018). Genetic risk is not exempt from this. Existing research already shows that findings from GWASs of White Europeans and, by extension, any findings produced from

methods which rely on them (e.g., PRS) have limited applicability to groups of different genetic ancestries (Martin *et al.*, 2019). An example of this low generalisability is the consistent observation that genome-wide significant variants often do not replicate with effects in the same direction, or at all, in non-European samples (Gudbjartsson *et al.*, 2007; Lewis *et al.*, 2008; Yamada *et al.*, 2009; Levey *et al.*, 2021). Much more pressing than the limitations of this thesis is the very real threat that the application of GWASs to human traits and diseases could exacerbate pre-existing health disparities between ethnic groups (Martin *et al.*, 2019). In doing so, the GWAS method would actively contradict its primary goal of improving health outcomes of humans. Personalised polygenic risk information is gradually being entered into mainstream healthcare. While this is an exciting time for genomic medicine, the prospect that not all members of society will benefit equally from these advancements is hugely concerning. This will be the situation as long the European bias in GWAS is not addressed.

Increasing participation of individuals from racially minoritised backgrounds is paramount. The principal investigators and research team behind the GLAD study are keenly aware of this and efforts have been made to try to access minoritised pockets of society and encourage them to enrol (Davies *et al.*, 2019). During my PhD, I attended the GLAD NIHR BioResource team meetings and learnt about the strategies that have been employed to achieve this, including targeted social media campaigns and increasing representation of non-White individuals in recruitment material. An acknowledgement of the issue from scientists in their various research outputs is also crucial. A key challenge here is ensuring that conversations about genetic ancestry do not inadvertently stir up false theories or beliefs about distinctions between “races” being rooted in biology (Carlson *et al.*, 2022). Part of this relates to human geneticists using correct terminology in discussions of genetic ancestry and ensuring that its complexity is well presented and explained. For instance, disseminated research should emphasise that individuals from different populations may be more genetically similar than two individuals from the same population (Witherspoon *et al.*, 2007), and that ancestral origins do not follow modern geopolitical borders or concepts of “race” but a continuum of genetic variation (Jorde and Wooding, 2004). Rather, a complex combination of geographical, topographical, socio-cultural, and historical factors, which impact migration, admixture, gene flow, and genetic drift, influence human genetic diversity within and between populations (Li *et al.*, 2008). Overall, in addition to striving for better representation of minoritised groups, scientists must also improve the communication of findings from genetic studies where population stratification and/or the participants’ genetic ancestral origins are mentioned.

Absence of replication in external data sets

A final limitation of the research presented in this thesis is the absence of replication in external data sets. As mentioned in **chapter 1**, replication is a crucial part of the GWAS method because it ensures that findings are not simply a product of the characteristics of the study sample (i.e., false positives). Robust findings that generalise to external study samples are important for ensuring that time, money, and effort is not wasted following up false results. The absence of replication in this thesis is due to the unavailability of suitable samples to achieve this. At the time of conducting analyses for **chapter 2**, no external study sample had data on major depressive disorder via the CIDI-SF, and the CTS, ATS, and PCL6. Now, this information is available in GLAD, EDGI, and NBR participants. Thus, replication of the UK Biobank results is possible (albeit with the caveat that UK Biobank and GLAD/NBR have dissimilar mental health-related characteristics). Likewise, for **chapter 3** and **chapter 4**, no external cohort had data on the MDQ, anhedonic symptoms, and staged treatment-resistant depression alongside genetic data. Now, the next phase of the Twins Early Development Study (TEDS26) has been launched. TEDS26 included the MDQ in its questionnaire. It will be interesting to see if the results of **chapter 3** can be replicated in this twin-based cohort, as the TEDS twins are not recruited based on lifetime history of depression or anxiety and are generally more demographically representative of England and Wales than the GLAD study is of the UK (Rimfeld *et al.*, 2019).

Remaining challenges and future directions

Medical record linkage in the GLAD study

An exciting prospect for the study of mood disorders in the GLAD study is the newly granted medical record linkage via the UK Longitudinal Linkage Collaboration (UK LLC). Medical record linkage will include data from cancer registers, General Practitioner (GP) records, Hospital Episode Statistics (HES), Improving Access to Psychological Therapies (IAPT), and Mental Health Services Data Set (MHSDS). Such wide ranging data will be pivotal in validating mood disorder phenotypes that have been derived from self-reported data in this thesis. For instance, the number of antidepressant and augmentation medications that an individual has taken will be accessible from prescription records. Furthermore, the length of an individual's depressive episodes could be estimated more accurately from GP records, IAPT, and MHSDS. Cross-referencing these

data with the participants' answers to the MSM could help to finetune the staged treatment-resistant depression phenotype in **chapter 4**. Following this, the GWAS could be repeated with a (hopefully) less noisy phenotype and, accordingly, more power. This would confirm whether the lack of a significant SNP-based heritability was due to issues regarding the retrospective self-reported nature of the MSM. Additionally, access to prescription records will open up alternative avenues to measure treatment resistance in MDD. For example, the GWAS by Wigmore et al., (2020), which used antidepressant-switching to derive a staged treatment-resistant phenotype, could be replicated with data from the GLAD study. Another future direction could involve investigating whether individuals with a diagnosis of bipolar disorder from GP records fall on the upper end of the distribution of MDQ items. Given the results of **chapter 3**, it is likely that this will not be the case.

Diagnostic categories versus spectrum of mental illness

As mentioned in **chapter 1**, psychiatric disorders are classified on the basis of operational criteria, such as those in the DSM. This multinomial taxonomic system was established based on consensus from experts about which signs and symptoms "belonged" to each category. This practice, which has been upheld for decades, has undeniably been useful clinically (being diagnosed with a psychiatric disorder is often the first step towards receiving treatment) and for progressing research (diagnostic categories offer operational criteria for researchers to define cases and controls in a reproducible manner). Yet, there is little evidence for the boundaries between disorders (Jablensky, 2016), as well as the boundary demarcating "affected" from "unaffected" within each disorder (see **figure 1.1**).

Epidemiological and psychiatric literature show that subthreshold symptoms are common among individuals who would be deemed unaffected by clinical cutoffs (Murphy *et al.*, 2012; Rodríguez *et al.*, 2012; Burstein *et al.*, 2014), including depressive and hypomania/mania symptoms. These individuals are more susceptible to developing later major depressive disorder and bipolar disorder respectively compared to individuals who have no symptoms (Fiedorowicz *et al.*, 2011; Hill *et al.*, 2014; Axelson *et al.*, 2015) thus adding weight to the argument that psychiatric disorders operate on a continuum. Furthermore, the co-occurrence of two or more forms of mental illness within the same individual is the rule rather than the exception (Clark *et al.*, 2017). Even in cases of non-comorbidity, an individual diagnosed with one psychiatric disorder may report subclinical symptoms of another. Therefore, while arguably useful for clinical purposes, the validity of discrete diagnostic categories, as in the DSM and ICD, is questionable.

Findings from genetics research similarly challenge the "DSM paradigm" (Craddock and Owen, 2010; Smoller *et al.*, 2019). This includes the clustering of nosologically separate psychiatric disorders within biological relatives, the fact that a single *de novo* mutation can contribute to risk for numerous psychiatric and neurodevelopmental outcomes, and genetic correlations across a web of psychiatric disorders (Malhotra and Sebat, 2012; Bulik-Sullivan *et al.*, 2015; Smoller *et al.*, 2019). Such research demonstrates that underlying genetics do not support the clear-cut boundaries between psychiatric disorders upheld by diagnostic systems. Nonetheless, many published scientific papers employ dichotomous phenotyping methods suggesting there is some value in doing so. Even within this thesis, dichotomous diagnostic subtypes were helpful in unravelling the relationship between major depressive disorder and PTSD (**chapter 2**). A challenge now faced by the field of psychiatric and statistical genetics is whether diagnostic categories should continue to be pursued, or whether research efforts could be better spent exploring the "continuum hypothesis" by prioritising cross-disorder, dimensional phenotypes.

In **chapters 3** and **4**, mood symptomatology was measured as a composite score, as opposed to a case/control study design. Assessing symptoms quantitatively outside the boundaries of a traditional diagnosis is consistent with the continuum hypothesis, which is backed up by research showing that subsyndromal symptoms confer risk for the later development of a psychiatric disorder. There is evidence to suggest that psychiatric disorders are the extreme end of a quantitative trait within a population (Plomin, Haworth and Davis, 2009). The concept of mental illness as a continuum from "unaffected" to "severely affected" is also backed up by genetics research showing that PRSs increase linearly with increasing odds of being a case versus a control (Mitchell *et al.*, 2022). The questionable validity of the MDQ and the biases of the GLAD study sample means the results of **chapter 3** are not particularly helpful here, but **chapter 4** emphasises the value of applying genomic methods to quantitative measures of psychopathology. Assessing a trait continuously, as opposed to a case/control design, captures more phenotypic variance and bolsters statistical power (van der Sluis *et al.*, 2013). Not only this but focusing on a **specific symptom** (anhedonia) revealed potential mechanisms that would otherwise have been missed when studying major depressive disorder, which contains a plethora of symptoms, as a general category. Future research should aim to combine these two approaches. Such an approach could be considered "transdiagnostic" for two reasons. First, **continuous measures** transcend the binary affected/unaffected paradigm. Second, **specific symptoms** often present across a range of psychiatric syndromes. Adopting a transdiagnostic perspective in this way could

contribute to our understanding of mood disorders as a spectrum, above and beyond what could be achieved by studying a general disorder as a dichotomous phenotype.

Is it possible to marry the two positions?

Methods which unveil the underlying genetic architecture of multiple complex traits, such as Genomic Structural Equation Modelling (Genomic SEM) (Grotzinger *et al.*, 2019), represent an opportunity to actually use the nosological categories to investigate the dimensional basis of mental health. Genomic SEM has unearthed clusters of psychiatric phenotypes which are influenced by joint sources of additive genetic liability. For instance, a study of schizophrenia, bipolar disorder, major depressive disorder, PTSD, and anxiety revealed that all diagnostic categories loaded onto one *"p"* factor (Grotzinger *et al.*, 2019). A study published earlier this year combined GWAS of five diagnostic categories (bipolar type I, bipolar type II, schizophrenia, schizoaffective disorder, and major depressive disorder) with GWAS of their cardinal symptoms (psychosis, depression, and mania) in Genomic SEM. Results showed two correlated but distinct factors with divergent genetic architectures (Mallard *et al.*, 2022). Genomic SEM provides the analytical tools to unpack the continuum on which diagnostic categories lie by drawing attention to psychiatric syndromes which share similar genetic bases and those which diverge. Thus, the application of a common factor model to discrete diagnostic categories can provide evidence of the blurred boundaries between their aetiologies (as paradoxical as that may seem).

A future direction for mood disorder genetics could involve models, such as those described above, with much more refined phenotypes. **Chapter 2** highlighted the value of focusing on **diagnostic subtypes** for exploring genetic correlations with other traits. Given the vast heterogeneity of mood disorders (Fried and Nesse, 2015; Coombes *et al.*, 2020), future research could apply Genomic SEM to **diagnostic subtypes** based on a variety of variables, for example, recurrence, treatment resistance, aetiology, and clinical specifiers. A previous study of UK Biobank data already hinted at the utility of this research design by highlighting the varying strength of pairwise genetic pathways between mood disorders. Recurrent, single-episode, and subthreshold depression correlated highly (r_g=0.9-0.94) and each genetic correlation did not differ significantly from one. The bipolar disorder type I and type II correlation was also not significantly different from one (r_g=0.87). Interestingly, bipolar disorder type II shared high genetic correlations with recurrent depression (r_g=0.69) and single-episode depression (r_g=0.61), while bipolar disorder type I did not. The authors concluded that this pattern represented evidence for a

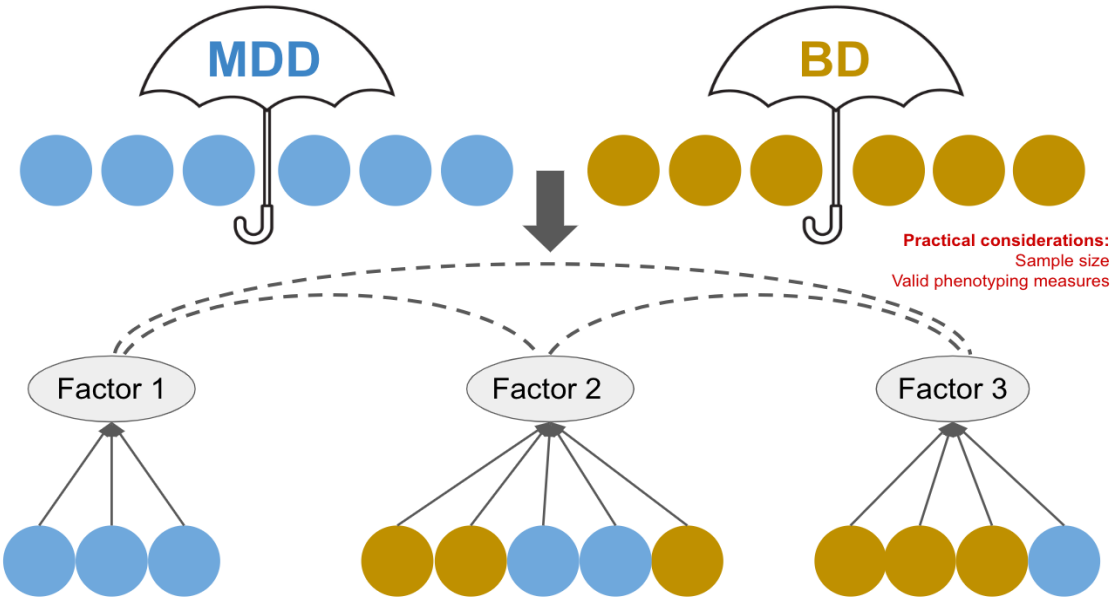
spectrum of genetic liability mirroring clinical observations, with major depressive disorder and bipolar disorder at either pole, and bipolar disorder type II acting as a bridge between them (Coleman *et al.*, 2020). To build on this research design, one could divide these disorder categories into more refined phenotypes, perform GWASs of them, and apply Genomic SEM to the summary statistics. Similar to findings from genetic studies of major depressive disorder subtypes (Mitchell *et al.*, 2022; Nguyen *et al.*, 2022), recent research has demonstrated that clinical heterogeneity in bipolar disorder, beyond type I and II, is reflected by genetic heterogeneity (Coombes *et al.*, 2020; Richards *et al.*, 2022). Thus, **diagnostic subtypes** *within* type I and type II could be explored in a Genomic SEM model along with major depressive disorder **diagnostic subtypes**. This type of research design, illustrated in **figure 5.3**, would allow a more fine-grained analysis of the genetic risk underlying the complexity of the mood spectrum by revealing which mood subtypes cluster together (suggesting common genetic risk) and which segregate (suggesting divergent genetic risk).

An extension to the research design in **figure 5.3** involves including disorders outside the mood spectrum. A PRS analysis demonstrated that bipolar disorder subtypes had unique genetic overlap with other psychiatric disorders such as schizophrenia, attention-deficit hyperactivity disorder (ADHD), and major depressive disorder. For instance, those with bipolar disorder and a history of psychosis had a higher PRS for schizophrenia, while those who had experienced rapid cycling had a higher PRS for ADHD (Coombes *et al.*, 2020). A similar approach was adopted in a recent PRS study of over two million Swedish adults to clarify which aspects of liability to non-mood psychopathology was shared and separate from subtypes of bipolar disorder. They found a high degree of overlap between the top two PRS deciles of major depressive disorder and bipolar disorder genetic risk (both influenced risk for nonpsychotic bipolar disorder) but they also found some key differences. High bipolar disorder PRS and low major depressive disorder PRS increased risk for psychotic bipolar disorder, non-psychotic bipolar disorder, and SAD, while low bipolar disorder PRS and high major depressive disorder PRS increased for non-psychotic major depressive disorder, non-psychotic bipolar disorder, and anxiety disorders (Kendler *et al.*, 2022). The symptom-level analyses of depressive symptoms by Thorp et al., (2021) showed that the nine symptoms in the DSM-5 criteria of major depressive disorder had differential genetic correlations with other psychiatric disorders. This led to the conclusion that studying symptom combinations could prove beneficial for understanding disease pathogenesis and, one day, tailoring treatments to the individual (Thorp *et al.*, 2021). Taken together, these studies hint at the

potential discoveries that could be made by integrating mood disorder **diagnostic subtypes** in a model with non-mood psychiatric disorders.

A practical consideration is sample size. Before the proposed research design in **figure 5.3** can be realised, the sample sizes of the GWASs of mood subtypes need to be sufficient, especially if relatively minor differences in their genetic liabilities are hoped to be discovered. This important issue was exemplified by the results of **chapter 2**: comparing GWAS summary statistics of the different major depressive disorder subtypes revealed no significant differences. But, when looking at individual-level genetic susceptibility through PRS, I did find that those who had self-reported having been exposed to trauma carried a higher genetic loading for PTSD than those who self-reported no trauma. In the discussion of **chapter 2**, I highlighted the need for larger sample sizes before small differences between phenotypes can be detected. The GLAD study, which continues to recruit participants with detailed mental health data and DNA samples, represents an opportunity to do this. Confidence in the refined phenotypes will grow once they have been validated against linked medical records. Sample size may be particularly limiting for bipolar disorder, which has a far lower lifetime prevalence than major depressive disorder. To overcome this, collaboration and data-sharing between different bipolar disorder-focused research studies will be crucial.

Figure 5.3. Hypothetical research study framework.
Diagnostic subtypes in the mood spectrum can be used in Genomic Structural Equation Modelling (SEM) to identify shared and divergent genetic risk factors.



Integration of multimodal data to investigate the mood spectrum

The chapters in this thesis combined participants' self-reported phenotypic data on mood disorders and their symptoms with their genotype data. While genetics and genomics research has been pivotal in advancing our understanding of the mood spectrum, the integration of other data types would accelerate discoveries and paint a more holistic picture of the underlying biology. Researchers from different fields could join forces to collate and analyse multiple data types from individuals with different mental illnesses. For instance, the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) Consortium collected various cognitive and neuroimaging data from individuals with psychosis who had been diagnosed with either bipolar disorder, schizoaffective disorder, or schizophrenia, their first degree relatives, and individuals with no history of psychosis. From this, three neurobiologically distinct groups were identified (which the authors called "biotypes"). These biotypes were agnostic to the taxonomy of the DSM as each one spanned the three diagnostic categories (Clementz *et al.*, 2016). Data on individual genetic risk (e.g., PRS) could very easily supplement a research design like this. The Danish High Risk and Resilience study (also known as the VIA study) is an example of a study which has collected wide-ranging data types to provide a comprehensive picture of the antecedents and outcomes of psychiatric illness. A total of 522 seven year olds born to a parent with bipolar disorder, schizophrenia, or two healthy parents have been followed up over the course of their childhood and adolescence. Data on neuro and social cognition, motor functioning, psychopathology, home life, attachment, socioeconomic indicators, brain structure (from age eleven), and genetics have been collected. Furthermore, by studying the children from before puberty, their developmental course will also be able to be assessed (Thorup *et al.*, 2018, 2022). The study of individuals based upon multiple facets of disease biology represents a transition towards pathology-based definitions and away from definitions based purely on signs and symptoms. Such an approach would also help to clarify the dimensional nature of mood disorders, by addressing whether they can be classified based on biomarkers rather than purely their emotional and/or behavioural manifestations. For this approach to be realised in the future, collaboration, skills-sharing, and data-sharing from experts across multiple disciplines, including molecular and statistical genetics, psychiatry and psychology, neuroimaging, and bioinformatics, will be paramount.

Public education about risk factors for psychiatric disorders

Personal genetic information is gradually becoming embedded into healthcare and support for genetic testing for psychiatric disorders is growing (Morosoli *et al.*, 2021). Although there are no formal genetic tests for psychiatric disorders, interested individuals can very easily access their PRSs through direct-to-consumer genetic testing companies (e.g., "23andMe") and websites such as "impute.me". Without the careful guidance of a genetic counsellor or healthcare professional trained in genetics, such information could easily be misinterpreted. This is hugely problematic because research shows that personal beliefs about the cause(s) of mental illness can influence emotional and behavioural outcomes, such as self-stigma, feelings of blame or guilt, or reproductive concerns (Austin, Smith and Honer, 2006; Meiser *et al.*, 2007; Austin, Hippman and Honer, 2012). Therefore, education about the joint contributions of genes and the environment to mental illness is paramount for ensuring that affected individuals, their families, and the general public base their perceptions in accurate and up-to-date science.

The task of communicating information about personal psychiatric risk is not straightforward for several reasons. First, while genetic literacy appears to be improving over time, it is typically poor among the general public, even in those who are well-educated (Chapman *et al.*, 2019; Little, Koehly and Gunter, 2022). Second, an individual's liability for a psychiatric disorder is complex and comes with a high degree of uncertainty. This uncertainty may be difficult to convey, especially if basic knowledge of genetics is low, or if individuals hold deterministic views about heritable conditions. Emerging research from the genetic counselling literature suggests that, even in the absence of a genetic test, individuals affected by mental health disorders can benefit from receiving explanations of the relative influences of inherited genetic factors and environmental exposures to their disorder(s) (Austin and Honer, 2005; Moldovan, Pintea and Austin, 2017; Michael *et al.*, 2020). This can be achieved with the "mental health jar" analogy presented in **figure 1.6** (Austin, 2020). One US study delivered fake genetic test results to 165 participants with depression and showed that those who were told that their result *"contains a gene that has been shown to significantly increase a person's risk of developing major depression"* exhibited lower levels of agency with regard to mood regulation than those who were told that they did not have the susceptibility gene. However, members of the "gene-present" group who were also shown a short, informative video about the non-deterministic nature of genetics in psychiatric disorders did not show this reduction in mood regulation agency, thus demonstrating

the importance of nuanced education when delivering personalised genetic information to individuals with depression (Lebowitz and Ahn, 2018).

Although the precise causes are still unknown, the scientific community understands much more about the contribution of genetics to psychiatric disorders compared to thirty years ago. Even without delivering personalised genetic risk information, it is important to extend this knowledge to the public in an understandable format. During my PhD, myself and another PhD student designed and launched the Perceptions of Psychiatric Risk (PerPsych) project to achieve this. The PerPsych project has two arms: 1) "lived experience" and 2) "mental health professionals, students, and trainees" which aim to learn about how individuals affected by mental ill health, and the professionals to interact with them, think about the possible cause(s) of mental illness.

PerPsych: lived experience

The "lived experience" arm recruits participants from the GLAD study and asks them to complete three surveys which are each sent two weeks apart. The primary aim is to explore the way in which the participants evaluate the cause(s) of their major depressive disorder/anxiety and how this impacts their emotional wellbeing and behaviour. We also adapted an animated video from Professor Danielle Dick's lab which used the "mental health jar" analogy (see **figure 1.6**) to explain that both genetic and environmental risk factors contribute to mental illness, that certain behaviours can protect against future episodes, and that feeling better is very much possible. We want to investigate whether watching this video, compared to a 'control' animated video which talks about mental health generally and does not mention genetic and environmental risk, can improve participants' understanding of risk factors for depression and anxiety and encourage healthy behaviours. We also ask participants if they have questions about their disorder that have not been adequately answered by healthcare services to identify gaps in knowledge that need addressing. Note that we are also running a parallel PerPsych study for EDGI UK participants. **Figure 5.4** contains a flow-chart summarising the study design for the "lived experience" arm.

Figure 5.4. Flow-chart of the study design for the ''lived experience'' arm of the Perceptions of Psychiatric Risk (PerPsych) project.
Participants from the Genetic Links to Anxiety and Depression (GLAD) study and Eating Disorders Genetics Initiative (EDGI UK) are invited to take part via email. Participants complete the part 1 survey (15-20 minutes) and are asked to watch either a 'test' video (which includes the "mental

health jar" analogy) or a 'control' video from Mind the mental health charity (which does not discuss genetic or environmental risk). They are then immediately asked follow-up questions. Two weeks and one month after the part 1 survey is completed, the participants are sent the part 2 and part 3 surveys respectively (5-10 minutes). An example question is presented in **figure 5.5**.

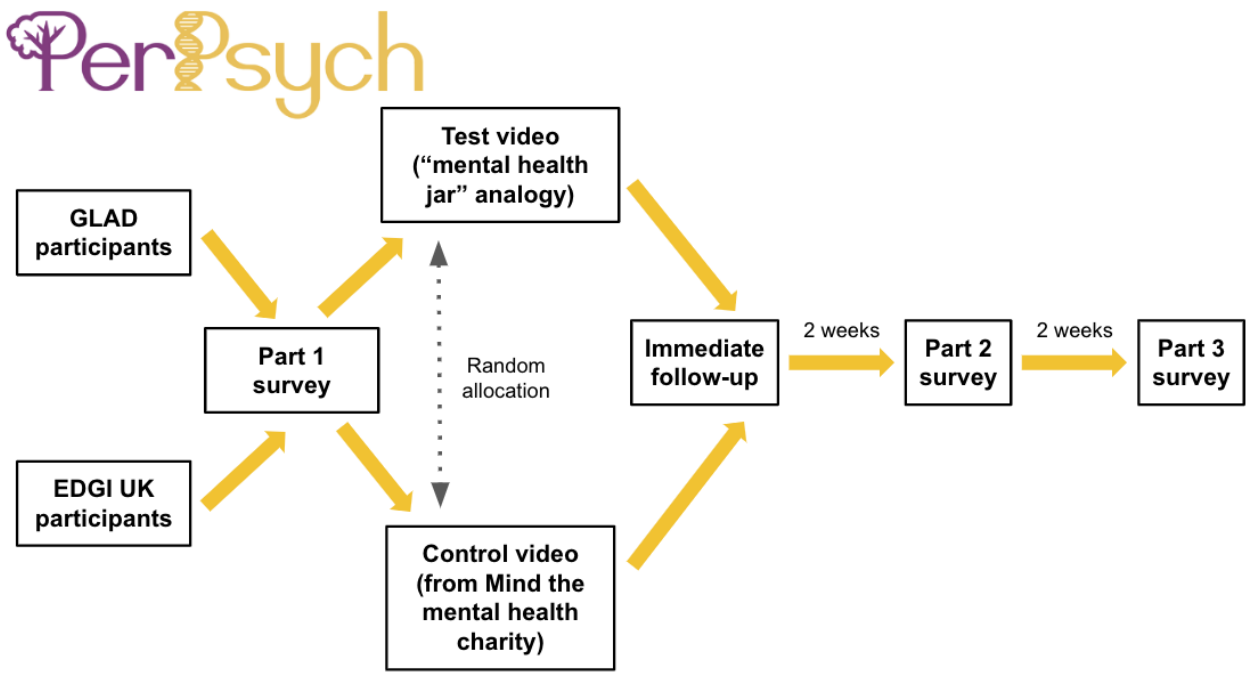


Figure 5.5. Example Qualtrics question from the "lived experience" arm of the Perceptions of Psychiatric Risk (PerPsych) survey.

This question aims to capture the perceptions of participants about the possible genetic and/or environmental causes of their disorder.

Please select an option according to how much you believe your **depression** was caused by **inherited genetic factors** and/or **environmental factors**.
You can think about **environmental factors** as your experiences or things that have happened to you throughout your life.

If using a mobile device, please scroll to view all answer options

	Only genetic factors	Almost only genetic factors	Mainly genetic factors	Slightly more genetic factors	Equal genetic and environmental factors	Slightly more environmental factors	Mainly environmental factors	Almost only environmental factors	Only environmental factors	Prefer not to answer
I believe my depression was caused by...	○	○	○	○	○	○	○	○	○	○

→

Mental health professionals are in a position to clarify risk factors for mental illness but their knowledge of the heritable basis of psychiatric disorders, and confidence discussing this, are lacking (Finn *et al.*, 2005; Martorell *et al.*, 2019). Through the PerPsych project, we are aiming to investigate the perceptions of professionals who work with mentally unwell individuals, and those in training for such careers, of genetic and environmental contributions to mental illness. We are interested in gaining a broad set of perspectives from UK-based clinical psychologists, psychiatrists, educational psychologists, assistant psychologists, general practitioners (GPs), social workers, psychological wellbeing practitioners (PWPs), mental health nurses, therapists, and counsellors (invited to take part via email or social media advertisement). The survey asks questions to ascertain their level of genetic knowledge and how confident they feel discussing genetic and environmental risk with their patients/clients. We also ask for their opinions on whether they believe the animated video (as in the "lived experience" arm) would be beneficial for their patients/clients.

Figure 5.6. Positive feedback about the test animated video in the "lived experience" arm of the Perceptions of Psychiatric Risk (PerPsych) project.

The test animated video used the "mental health jar" analogy (see **figure 1.6**). The full script of this video can be found in **appendix 5**. I received consent from participants to present their data from fee-text boxes.

We pre-registered the PerPsych project on the Open Science Framework (OSF) prior to its launch. This can be found in **appendix 5** including all measures and proposed research questions. At the time of submitting this thesis, we have recruited 2,598 GLAD and 979 EDGI UK participants, and 162 mental health professionals, students, and trainees. Once we have completed data collection, we will begin analysing the data with regard to the research questions posed in the OSF pre-registration (included in **appendix 5**). To continue to engage the participants in the research process, we plan to send out a newsletter with lay summaries of our results to participants to keep them updated about what we have found. Additionally, we plan to send the test video (with the "mental health jar" analogy) to all participants at the end of the study. This is to ensure that even those who were randomly allocated the control video can benefit from this information. Some positive feedback from participants about the 'test' animated video is presented in **figure 5.6**.

Conclusion

Mood psychopathology has existed throughout human history. Despite centuries of dedicated study, the precise causes of mood disorders are unclear. Genetic studies offer one avenue of investigation and, over the past two decades, mood disorder genetics research has been

expanding. There are multiple end goals of such research: identifying risk variants for druggable targets, stratifying individuals based on risk, exploring shared aetiology with other human traits and disorders, and using aetiological knowledge to inform and educate affected individuals. Accurate phenotyping is important for achieving all these goals. Research studies collecting both phenotypic and genetic data offer the opportunity to study the genetic basis of both major depressive disorder and bipolar disorder. However, such studies require very large sample sizes to study these polygenic conditions, and this means that self-reported data collection is routinely used. While offering many opportunities, self-reported data also brings with it challenges regarding the construction of phenotypes for analyses, including GWASs. In this thesis, I investigated the various ways that mood psychopathology can be phenotyped with self-reported data and applied statistical genetics methods to them. Phenotypic validity is essential for minimising the likelihood of false positive findings and increasing likelihood of replication, which are key goals of genetics research. The studies presented in this thesis emphasise the value of taking refined phenotypes (e.g., diagnostic subtypes) as well as continuous measures of mood. These strategies can be combined, as per **chapter 4**. However, researchers wishing to adopt these methods in future studies should be warned of the limitations of retrospective data as well as the application of clinical tools to self-reported data collection. To continue to unravel the perpetual debate about mental ill health as a spectrum, it is paramount that researchers adopt data-driven/bottom-up approaches to classifying mood disorders, which considers genetic risk as one of many aetiological indicators.

References

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

Austin, J.C. (2020) 'Evidence-Based Genetic Counseling for Psychiatric Disorders: A Road Map', *Cold Spring Harbor perspectives in medicine*, 10(6). Available at: https://doi.org/10.1101/cshperspect.a036608.

Austin, J.C., Hippman, C. and Honer, W.G. (2012) 'Descriptive and numeric estimation of risk for psychotic disorders among affected individuals and relatives: implications for clinical practice', *Psychiatry research*, 196(1), pp. 52–56.

Austin, J.C. and Honer, W.G. (2005) 'The potential impact of genetic counseling for mental illness', *Clinical genetics*, 67(2), pp. 134–142.

Austin, J.C., Smith, G.N. and Honer, W.G. (2006) 'The genomic era and perceptions of

psychotic disorders: genetic risk estimation, associations with reproductive decisions and views about predictive testing', *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*, 141B(8), pp. 926–928.

Axelson, D. *et al.* (2015) 'Diagnostic Precursors to Bipolar Disorder in Offspring of Parents With Bipolar Disorder: A Longitudinal Study', *The American journal of psychiatry*, 172(7), pp. 638–646.

Bulik-Sullivan, B. *et al.* (2015) 'An atlas of genetic correlations across human diseases and traits', *Nature genetics*, 47(11), pp. 1236–1241.

Burstein, M. *et al.* (2014) 'Threshold and subthreshold generalized anxiety disorder among US adolescents: prevalence, sociodemographic, and clinical characteristics', *Psychological medicine*, 44(11), pp. 2351–2362.

Burt, D.B., Zembar, M.J. and Niederehe, G. (1995) 'Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity', *Psychological bulletin*, 117(2), pp. 285–305.

Carlson, J. *et al.* (2022) *Counter the weaponization of genetics research by extremists*, *Nature Publishing Group UK*. Available at: https://doi.org/10.1038/d41586-022-03252-z.

Catz, D.S. *et al.* (2005) 'Attitudes about genetics in underserved, culturally diverse populations', *Community genetics*, 8(3), pp. 161–172.

Chapman, R. *et al.* (2019) 'New literacy challenge for the twenty-first century: genetic knowledge is poor even among well educated', *Journal of community genetics*, 10(1), pp. 73–84.

Cho, S. and Bulger, M. (2021) 'Social Support and Depressive Symptoms among Trauma-Impacted Older Adults', *Journal of evidence-based social work*, 18(4), pp. 371–378.

Clark, L.A. *et al.* (2017) 'Three Approaches to Understanding and Classifying Mental Disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC)', *Psychological science in the public interest: a journal of the American Psychological Society*, 18(2), pp. 72–145.

Clementz, B.A. *et al.* (2016) 'Identification of Distinct Psychosis Biotypes Using Brain-Based Biomarkers', *The American journal of psychiatry*, 173(4), pp. 373–384.

Coleman, J.R.I., *et al.* (2020) 'Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank', *Molecular psychiatry* [Preprint]. Available at: https://doi.org/10.1038/s41380-019-0546-6.

Coleman, J.R.I., *et al.* (2020) 'The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls', *Biological psychiatry*, 88(2), pp. 169–184.

Coombes, B.J. *et al.* (2020) 'Dissecting clinical heterogeneity of bipolar disorder using multiple polygenic risk scores', *Translational psychiatry*, 10(1), p. 314.

Craddock, N. and Owen, M.J. (2010) 'The Kraepelinian dichotomy - going, going... but still not

gone', *The British journal of psychiatry: the journal of mental science*, 196(2), pp. 92–95.

Davies, M.R. *et al.* (2019) 'The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety', *Behaviour research and therapy*, 123, p. 103503.

Davies, M.R. *et al.* (2022) 'Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts', *Journal of anxiety disorders*, 85, p. 102491.

Dolby, T. *et al.* (2022) 'Monitoring sociodemographic inequality in COVID-19 vaccination uptake in England: a national linked data study', *Journal of epidemiology and community health*, 76(7), pp. 646–652.

Drancourt, N. *et al.* (2013) 'Duration of untreated bipolar disorder: missed opportunities on the long road to optimal treatment', *Acta psychiatrica Scandinavica*, 127(2), pp. 136–144.

Ebner-Priemer, U.W. *et al.* (2020) 'Digital phenotyping: towards replicable findings with comprehensive assessments and integrative models in bipolar disorders', *International journal of bipolar disorders*, 8(1), p. 35.

Feczko, E. *et al.* (2019) 'The Heterogeneity Problem: Approaches to Identify Psychiatric Subtypes', *Trends in cognitive sciences*, 23(7), pp. 584–601.

Fekadu, A. *et al.* (2009) 'A multidimensional tool to quantify treatment resistance in depression: the Maudsley staging method', *The Journal of clinical psychiatry*, 70(2), pp. 177–184.

Fekadu, A., Donocik, J.G. and Cleare, A.J. (2018) 'Standardisation framework for the Maudsley staging method for treatment resistance in depression', *BMC psychiatry*, 18(1), p. 100.

Fiedorowicz, J.G. *et al.* (2011) 'Subthreshold hypomanic symptoms in progression from unipolar major depression to bipolar disorder', *The American journal of psychiatry*, 168(1), pp. 40–48.

Finn, C.T. *et al.* (2005) 'Psychiatric genetics: a survey of psychiatrists' knowledge, opinions, and practice patterns', *The Journal of clinical psychiatry*, 66(7), pp. 821–830.

Fried, E.I. and Nesse, R.M. (2015) 'Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study', *Journal of affective disorders*, 172, pp. 96–102.

Fry, A. *et al.* (2017) 'Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population', *American journal of epidemiology*, 186(9), pp. 1026–1034.

Grotzinger, A.D. *et al.* (2019) 'Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits', *Nature human behaviour*, 3(5), pp. 513–525.

Gudbjartsson, D.F. *et al.* (2007) 'Variants conferring risk of atrial fibrillation on chromosome 4q25', *Nature*, 448(7151), pp. 353–357.

Haga, S.B. (2010) 'Impact of limited population diversity of genome-wide association studies', *Genetics in medicine: official journal of the American College of Medical Genetics*, 12(2), pp. 81–84.

Hill, R.M. *et al.* (2014) 'Escalation to Major Depressive Disorder among adolescents with subthreshold depressive symptoms: evidence of distinct subgroups at risk', *Journal of affective disorders*, 158, pp. 133–138.

Hirschfeld, R.M. *et al.* (2000) 'Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire', *The American journal of psychiatry*, 157(11), pp. 1873–1875.

Hirschfeld, R.M.A., Lewis, L. and Vornik, L.A. (2003) 'Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder', *The Journal of clinical psychiatry*, 64(2), pp. 161–174.

Holmberg, M.J. and Andersen, L.W. (2022) 'Collider Bias', *JAMA: the journal of the American Medical Association*, 327(13), pp. 1282–1283.

Hosang, G.M. *et al.* (2021) 'Association of Etiological Factors for Hypomanic Symptoms, Bipolar Disorder, and Other Severe Mental Illnesses', *JAMA psychiatry* [Preprint]. Available at: https://doi.org/10.1001/jamapsychiatry.2021.3654.

Howard, D.M. *et al.* (2018) 'Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways', *Nature communications*, 9(1), p. 1470.

Howard, D.M. *et al.* (2019) 'Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions', *Nature neuroscience*, 22(3), pp. 343–352.

Jablensky, A. (2016) 'Psychiatric classifications: validity and utility', *World psychiatry: official journal of the World Psychiatric Association* , 15(1), pp. 26–31.

Jacobson, N.C. and Newman, M.G. (2014) 'Avoidance mediates the relationship between anxiety and depression over a decade later', *Journal of anxiety disorders*, 28(5), pp. 437–445.

Jorde, L.B. and Wooding, S.P. (2004) 'Genetic variation, classification and "race"', *Nature genetics*, 36(11 Suppl), pp. S28–33.

Kendler, K.S. *et al.* (2022) 'Risk for Mood, Anxiety, and Psychotic Disorders in Individuals at High and Low Genetic Liability for Bipolar Disorder and Major Depression', *JAMA psychiatry* [Preprint]. Available at: https://doi.org/10.1001/jamapsychiatry.2022.2873.

Kuehner, C. (2017) 'Why is depression more common among women than among men?', *The lancet. Psychiatry*, 4(2), pp. 146–158.

Lebowitz, M.S. and Ahn, W.-K. (2018) 'Blue Genes? Understanding and Mitigating Negative Consequences of Personalized Information about Genetic Risk for Depression', *Journal of genetic counseling*, 27(1), pp. 204–216.

Lee, J.-S. (2019) 'Perceived social support functions as a resilience in buffering the impact of trauma exposure on PTSD symptoms via intrusive rumination and entrapment in firefighters', *PloS one*, 14(8), p. e0220454.

Levey, D.F. *et al.* (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), pp. 954–963.

Lewis, J.P. *et al.* (2008) 'Association analysis in african americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies', *Diabetes*, 57(8), pp. 2220–2225.

Li, J.Z. *et al.* (2008) 'Worldwide human relationships inferred from genome-wide patterns of variation', *Science*, 319(5866), pp. 1100–1104.

Little, I.D., Koehly, L.M. and Gunter, C. (2022) 'Understanding changes in genetic literacy over time and in genetic research participants', *American journal of human genetics*, 0(0). Available at: https://doi.org/10.1016/j.ajhg.2022.11.005.

Malhotra, D. and Sebat, J. (2012) 'CNVs: harbingers of a rare variant revolution in psychiatric genetics', *Cell*, 148(6), pp. 1223–1241.

Mallard, T.T. *et al.* (2022) 'Multivariate GWAS of psychiatric disorders and their cardinal symptoms reveal two dimensions of cross-cutting genetic liabilities', *Cell genomics*, 2(6). Available at: https://doi.org/10.1016/j.xgen.2022.100140.

Manolio, T.A. *et al.* (2012) 'New models for large prospective studies: is there a better way?', *American journal of epidemiology*, 175(9), pp. 859–866.

Mantere, O. *et al.* (2004) 'The clinical characteristics of DSM-IV bipolar I and II disorders: baseline findings from the Jorvi Bipolar Study (JoBS)', *Bipolar disorders*, 6(5), pp. 395–405.

Martin, A.R. *et al.* (2019) 'Clinical use of current polygenic risk scores may exacerbate health disparities', *Nature genetics*, 51(4), pp. 584–591.

Martorell, L. *et al.* (2019) 'Genetics and genetic counseling in psychiatry: Results from an opinion survey of professionals and users', *Molecular genetics & genomic medicine*, 7(8), p. e830.

McAllister-Williams, R.H. *et al.* (2018) 'Multiple-therapy-resistant major depressive disorder: a clinically important concept', *The British journal of psychiatry: the journal of mental science*, 212(5), pp. 274–278.

McInnis, M., Gideon, J. and Mower Provost, E. (2017) 'T14 - Digital Phenotyping In Bipolar Disorder', *European neuropsychopharmacology: the journal of the European College of Neuropsychopharmacology*, 27, p. S440.

McMakin, D.L. *et al.* (2012) 'Anhedonia predicts poorer recovery among youth with selective serotonin reuptake inhibitor treatment-resistant depression', *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(4), pp. 404–411.

McMaughan, D.J., Oloruntoba, O. and Smith, M.L. (2020) 'Socioeconomic Status and Access to Healthcare: Interrelated Drivers for Healthy Aging', *Frontiers in public health*, 8, p. 231.

Meiser, B. *et al.* (2007) 'Attitudes towards childbearing, causal attributions for bipolar disorder and psychological distress: a study of families with multiple cases of bipolar disorder', *Psychological medicine*, 37(11), pp. 1601–1611.

Michael, J.E. *et al.* (2020) ‘Perceptions of genetic risk, testing, and counseling among individuals with eating disorders’, *The International journal of eating disorders* [Preprint]. Available at: https://doi.org/10.1002/eat.23333.

Miller, C.J., Johnson, S.L. and Eisner, L. (2009) ‘Assessment Tools for Adult Bipolar Disorder’, *Clinical psychology: a publication of the Division of Clinical Psychology of the American Psychological Association*, 16(2), pp. 188–201.

Mills, M.C. and Rahal, C. (2019) ‘A scientometric review of genome-wide association studies’, *Communications biology*, 2, p. 9.

Mitchell, B.L. *et al.* (2022) ‘The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes’, *Biological psychiatry*, 92(3), pp. 227–235.

Moldovan, R., Pintea, S. and Austin, J. (2017) ‘The efficacy of genetic counseling for psychiatric disorders: A meta-analysis’, *Journal of genetic counseling*, 26(6), pp. 1341–1347.

Moorman, P.G. *et al.* (2004) ‘Racial differences in enrolment in a cancer genetics registry’, *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 13(8), pp. 1349–1354.

Morosoli, J.J. *et al.* (2021) ‘Investigating perceived heritability of mental health disorders and attitudes toward genetic testing in the United States, United Kingdom, and Australia’, *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*, 186(6), pp. 341–352.

Mullins, N. *et al.* (2021) ‘Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology’, *Nature genetics*, 53(6), pp. 817–829.

Mundy, J. *et al.* (2021) ‘Psychological trauma and the genetic overlap between posttraumatic stress disorder and major depressive disorder’, *Psychological medicine*, pp. 1–10.

Murphy, J. *et al.* (2012) ‘A population based analysis of subclinical psychosis and help-seeking behavior’, *Schizophrenia bulletin*, 38(2), pp. 360–367.

Nagel, M. *et al.* (2018) ‘Item-level analyses reveal genetic heterogeneity in neuroticism’, *Nature communications*, 9(1), p. 905.

Nguyen, T.-D. *et al.* (2022) ‘Genetic heterogeneity and subtypes of major depression’, *Molecular psychiatry*, 27(3), pp. 1667–1675.

Nishino, J. *et al.* (2018) ‘Sample Size for Successful Genome-Wide Association Study of Major Depressive Disorder’, *Frontiers in genetics*, 9, p. 227.

Okbay, A. *et al.* (2016) ‘Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses’, *Nature genetics*, 48(6), pp. 624–633.

Orsolini, L., Fiorani, M. and Volpe, U. (2020) ‘Digital Phenotyping in Bipolar Disorder: Which Integration with Clinical Endophenotypes and Biomarkers?’, *International journal of molecular sciences*, 21(20). Available at: https://doi.org/10.3390/ijms21207684.

Ostergaard, S.D., Jensen, S.O.W. and Bech, P. (2011) 'The heterogeneity of the depressive syndrome: when numbers get serious', *Acta psychiatrica Scandinavica*, 124(6), pp. 495–496.

Ozbay, F. *et al.* (2007) 'Social support and resilience to stress: from neurobiology to clinical practice', *Psychiatry* , 4(5), pp. 35–40.

Pain, O. *et al.* (2022) 'Identifying the Common Genetic Basis of Antidepressant Response', *Biological psychiatry global open science*, 2(2), pp. 115–126.

Pardiñas, A.F. *et al.* (2018) 'Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection', *Nature genetics*, 50(3), pp. 381–389.

Paskett, E.D. *et al.* (2008) 'Recruitment of minority and underserved populations in the United States: the Centers for Population Health and Health Disparities experience', *Contemporary clinical trials*, 29(6), pp. 847–861.

Phillips, M.L. and Kupfer, D.J. (2013) 'Bipolar disorder diagnosis: challenges and future directions', *The Lancet*, 381(9878), pp. 1663–1671.

Plomin, R., Haworth, C.M.A. and Davis, O.S.P. (2009) 'Common disorders are quantitative traits', *Nature reviews. Genetics*, 10(12), pp. 872–878.

Polimanti, R. (2022) 'Not Only Gene Discovery: Genome-wide Association Studies and Polygenic Risk Scores as Tools to Dissect the Heterogeneity of Major Depressive Disorder', *Biological psychiatry*, pp. 177–178.

Popejoy, A.B. and Fullerton, S.M. (2016) *Genomics is failing on diversity*, *Nature Publishing Group UK*. Available at: https://doi.org/10.1038/538161a.

Purves, K.L. *et al.* (2019) 'A major role for common genetic variation in anxiety disorders', *Molecular psychiatry* [Preprint]. Available at: https://doi.org/10.1038/s41380-019-0559-1.

Rappoport, N. *et al.* (2018) 'Comparing Ethnicity-Specific Reference Intervals for Clinical Laboratory Tests from EHR Data', *The journal of applied laboratory medicine*, 3(3), pp. 366–377.

Richards, A.L. *et al.* (2022) 'Genetic Liabilities Differentiating Bipolar Disorder, Schizophrenia, and Major Depressive Disorder, and Phenotypic Heterogeneity in Bipolar Disorder', *JAMA psychiatry* , 79(10), pp. 1032–1039.

Rimfeld, K. *et al.* (2019) 'Twins Early Development Study: A Genetically Sensitive Investigation into Behavioral and Cognitive Development from Infancy to Emerging Adulthood', *Twin research and human genetics: the official journal of the International Society for Twin Studies*, 22(6), pp. 508–513.

Rock, P.L. *et al.* (2014) 'Cognitive impairment in depression: a systematic review and meta-analysis', *Psychological medicine*, 44(10), pp. 2029–2040.

Rodríguez, M.R. *et al.* (2012) 'Definitions and factors associated with subthreshold depressive conditions: a systematic review', *BMC psychiatry*, 12, p. 181.

Sforzini, L. *et al.* (2021) 'A Delphi-method-based consensus guideline for definition of treatment-resistant depression for clinical trials', *Molecular psychiatry* [Preprint]. Available at:

https://doi.org/10.1038/s41380-021-01381-x.

Sirugo, G., Williams, S.M. and Tishkoff, S.A. (2019) 'The Missing Diversity in Human Genetic Studies', *Cell*, 177(1), pp. 26–31.

van der Sluis, S. *et al.* (2013) 'Power in GWAS: lifting the curse of the clinical cut-off', *Molecular psychiatry*, 18(1), pp. 2–3.

Smoller, J.W. *et al.* (2019) 'Psychiatric genetics and the structure of psychopathology', *Molecular psychiatry*, 24(3), pp. 409–420.

Spinhoven, P. *et al.* (2011) 'Personality, Life Events and the Course of Anxiety and Depression', *European journal of personality*, 25(6), pp. 443–452.

Stahl, E.A. *et al.* (2019) 'Genome-wide association study identifies 30 loci associated with bipolar disorder', *Nature genetics*, 51(5), pp. 793–803.

Thorp, J.G. *et al.* (2020) 'Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9', *Psychological medicine*, 50(14), pp. 2385–2396.

Thorp, J.G. *et al.* (2021) 'Symptom-level modelling unravels the shared genetic architecture of anxiety and depression', *Nature human behaviour*, 5(10), pp. 1432–1442.

Thorup, A.A.E. *et al.* (2018) 'The Danish High Risk and Resilience Study-VIA 11: Study Protocol for the First Follow-Up of the VIA 7 Cohort -522 Children Born to Parents With Schizophrenia Spectrum Disorders or Bipolar Disorder and Controls Being Re-examined for the First Time at Age 11', *Frontiers in psychiatry / Frontiers Research Foundation*, 9, p. 661.

Thorup, A.A.E. *et al.* (2022) 'The Danish High-Risk and Resilience Study-VIA 15 - A Study Protocol for the Third Clinical Assessment of a Cohort of 522 Children Born to Parents Diagnosed With Schizophrenia or Bipolar Disorder and Population-Based Controls', *Frontiers in psychiatry / Frontiers Research Foundation*, 13, p. 809807.

Uher, R. *et al.* (2012) 'Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms', *Psychological medicine*, 42(5), pp. 967–980.

Varga, M. *et al.* (2006) 'Insight, symptoms and neurocognition in bipolar I patients', *Journal of affective disorders*, 91(1), pp. 1–9.

Veeranna, V. *et al.* (2013) 'Association of novel biomarkers with future cardiovascular events is influenced by ethnicity: results from a multi-ethnic cohort', *International journal of cardiology*, 166(2), pp. 487–493.

Ward, J. *et al.* (2019) 'Novel genome-wide associations for anhedonia, genetic correlation with psychiatric disorders, and polygenic association with brain structure', *Translational psychiatry*, 9(1), p. 327.

Weyers, S. *et al.* (2008) 'Low socio-economic position is associated with poor social networks and social support: results from the Heinz Nixdorf Recall Study', *International journal for equity in health*, 7, p. 13.

Wigmore, E.M. *et al.* (2020) 'Genome-wide association study of antidepressant treatment

resistance in a population-based cohort using health service prescription data and meta-analysis with GENDEP', *The pharmacogenomics journal*, 20(2), pp. 329–341.

Winer, E.S. *et al.* (2017) 'Mapping the relationship between anxiety, anhedonia, and depression', *Journal of affective disorders*, 221, pp. 289–296.

Witherspoon, D.J. *et al.* (2007) 'Genetic similarities within and between human populations', *Genetics*, 176(1), pp. 351–359.

Wray, N.R. *et al.* (2018) 'Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression', *Nature genetics*, 50(5), pp. 668–681.

Yamada, H. *et al.* (2009) 'Replication of prostate cancer risk loci in a Japanese case-control association study', *Journal of the National Cancer Institute*, 101(19), pp. 1330–1336.

Young, K.S. *et al.* (2022) 'Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK', *Psychological medicine*, pp. 1–14.

Zulueta, J. *et al.* (2018) 'Predicting Mood Disturbance Severity with Mobile Phone Keystroke Metadata: A BiAffect Digital Phenotyping Study', *Journal of medical Internet research*, 20(7), p. e241.

Appendix 1. Supplementary material for chapter 1

Supplementary material

Diagnostic criteria of a major depressive episode as outlined in the 5th edition of the Diagnostic Statistical Manual for Mental Disorders (DSM-5) (American Psychiatric Association, 2013)

- Five (or more) of the following symptoms have been present during the same two week period and represent a change from previous functioning; at least one of the symptoms is either **(1) depressed mood** or **(2) loss of interest or pleasure**.
    - *Note: Do not include symptoms that are clearly attributable to another medical condition.*

    1. **Depressed most of the day, nearly every day as indicated by subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful)**
    2. **Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by subjective account or observation).**
    3. Significant weight loss when not dieting or weight gain (e.g., change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day.
    4. Insomnia or hypersomnia nearly every day.
    5. Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).
    6. Fatigue or loss of energy nearly every day Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
    7. Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others)
    8. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide

- The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
- The episode is not attributable to the physiological effects of a substance or to another medical condition.
    - *The above criteria represent a major depressive episode.*
- The occurrence of the major depressive episode is not better explained by schizoaffective disorder, schizophrenia, schizophreniform disorder, delusional disorder, or other specified and unspecified schizophrenia spectrum and other psychotic disorders.
- There has never been a manic episode or a hypomanic episode.
    - *Note: This exclusion does not apply if all of the manic-like or hypomanic-like episodes are substance-induced or are attributable to the physiological effects of another medical condition.*

Diagnostic criteria of a manic and hypomanic episode as outlined in the 5th edition of the Diagnostic Statistical Manual for Mental Disorders (DSM-5) (American Psychiatric Association, 2013)

Key differences between mania and hypomania are highlighted in bold.

Manic episode:

- A distinct period of abnormally and persistently elevated, expansive, or irritable mood and abnormally and persistently increased goal-directed activity or energy, lasting **at least 1 week** and present most of the day, nearly every day (or any duration if hospitalization is necessary).
- During the period of mood disturbance and increased energy or activity, 3 (or more) of the following symptoms (4 if the mood is only irritable) are present to a significant degree and represent a noticeable change from usual behavior:
    - Inflated self-esteem or grandiosity
    - Decreased need for sleep (e.g., feels rested after only 3 hours of sleep)
    - More talkative than usual or pressure to keep talking
    - Flight of ideas or subjective experience that thoughts are racing
    - Distractibility (i.e., attention too easily drawn to unimportant or irrelevant external stimuli), as reported or observed

    - Increase in goal-directed activity (either socially, at work or school, or sexually) or psychomotor agitation (i.e., purposeless, non-goal-directed activity)
    - Excessive involvement in activities that have a high potential for painful consequences (e.g., engaging in unrestrained buying sprees, sexual indiscretions, or foolish business investments)
- The mood disturbance is **sufficiently severe** to cause marked impairment in social or occupational functioning, or to necessitate hospitalization to prevent harm to self or others, or there are psychotic features.
- The episode is not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication, or other treatment) or to another medical condition.
    - *Note: A full manic episode that emerges during antidepressant treatment [e.g., medication, electroconvulsive therapy (ECT)], but persists at a fully syndromal level beyond the physiological effect of treatment is sufficient evidence for a manic episode, and therefore, a bipolar I diagnosis*

Hypomanic episode:

- A distinct period of abnormally and persistently elevated, expansive, or irritable mood and abnormally and persistently increased activity or energy, lasting at least **4 consecutive days** and present most of the day, nearly every day.
- During the period of mood disturbance and increased energy and activity, 3 (or more) of the above (manic) symptoms (4 if the mood is only irritable) have persisted, represent a noticeable change from usual behavior, and have been present to a significant degree.
- The episode is associated with an unequivocal change in functioning that is uncharacteristic of the individual when not symptomatic.
- The disturbance in mood and the change in functioning are observable by others.
- The episode is **not severe enough** to cause marked impairment in social or occupational functioning or to necessitate hospitalization. If there are psychotic features, the episode is, by definition, manic.
- The episode is not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication, or other treatment).
    - *Note: A full hypomanic episode that emerges during antidepressant treatment (e.g., medication, ECT) but persists at a fully syndromal level beyond the*

physiological effect of that treatment is sufficient evidence for a hypomanic episode diagnosis. However, caution is indicated so that one or two symptoms (particularly increased irritability, edginess or agitation following antidepressant use) are not taken as sufficient for a diagnosis of a hypomanic episode nor necessarily indicative of a bipolar diathesis

Bipolar disorder type I and type II:

- A person can be diagnosed with **bipolar disorder type I** if they meet criteria for a manic episode. This manic episode **may have been preceded and/or followed by a hypomanic or major depressive episode.**
- A person can be diagnosed with **bipolar disorder type II** if they meet criteria for at least one hypomanic episode **and** at least one major depressive episode and there has **never** been a manic episode.

Supplementary references

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

Appendix 2. Supplementary material for chapter 2

Supplementary material

Supplementary methods

Study sample and phenotype definitions

UK Biobank

Between 2006 and 2010, the UK Biobank recruited over 500,000 individuals for a cohort study aimed at improving diagnosis, treatment and prevention of serious diseases(Allen *et al.*, 2014). Participants gave full informed consent, answered surveys and provided physical measurements including DNA samples at a baseline visit to one of 22 assessment centres across the UK (Sudlow *et al.*, 2015). The phenotypes assessed in this study were derived from the online follow-up Mental Health Questionnaire (MHQ), which received 157,366 responses. This online questionnaire comprises a number of adapted versions of clinically used psychiatric questionnaires to assess common mental health disorders. Case definitions based on responses to the psychiatric questionnaires were derived by the working committee who wrote the MHQ (Davis *et al.*, 2020). The individual-level analysis sample for the present study specifically focused on participants who had completed the MHQ and met criteria for lifetime major depressive disorder (MDD) (*N*=29,41).

Major depressive disorder (MDD)

Participants were considered cases for probable MDD based on their responses to questions derived from the Composite International Diagnostic Interview Short Form (CIDI-SF). Reporting on a period of depression lasting at least two weeks, cases endorsed at least one of the two core symptoms (“ever had prolonged feelings of sadness or depression” and “ever had prolonged loss of interest in normal activities”), at least five of the nine symptoms queried overall, and reported that they were affected almost every day, most days during the period, with more than a little impact on normal functioning. Controls did not meet case criteria and did not meet criteria for a current episode of depression. Participants who self-reported a diagnosis of schizophrenia, other psychoses, or bipolar disorder were excluded. Controls were excluded if they self-reported any

mental illness, taking any drug with an antidepressant indication, or had been hospitalised with a mood disorder or met previously defined criteria for a mood disorder.

Major depressive disorder with/without reported psychological trauma

Participants were classified as having either “MDD with reported exposure to psychological trauma” or “MDD without reported exposure psychological trauma”. Phenotype definitions for these are included in the Supplementary Material of Coleman et al. (2020). Questions relating to traumatic experiences in childhood were assessed on a five point scale (ranging from “never” to “often”) using the Childhood Trauma Screener (an adapted version of the Childhood Trauma Questionnaire (Bernstein *et al.*, 1994; Grabe *et al.*, 2012; Bellis *et al.*, 2014). An equivalent screener was constructed for traumatic events in adulthood. Only traumatic experiences with an odds ratio >2.5 with MDD were selected to obtain a single binary variable for trauma exposure. This included: three events in childhood (did not feel loved, felt hated by a family member, sexually abused); three events in adulthood (physical violence, belittlement, sexual interference); and one PTSD-related event (ever a victim of sexual assault). Participants were included in the “MDD with reported exposure to psychological trauma” analysis if they reported two or more of these events and met criteria for MDD. Participants were included in the “MDD without reported exposure to psychological trauma” analysis if they reported none of these events and met case criteria for MDD (Coleman *et al.*, 2020).

Recurrent and single-episode major depressive disorder

Participants were classified as having either recurrent MDD or single-episode MDD. The definitions of the recurrent and single-episode MDD phenotypes can be found in the Supplementary Material of Coleman et al. (2019). In brief, participants who met criteria for MDD were classified with recurrent MDD if they reported multiple depressed periods across their lifetime and single-episode MDD otherwise (Coleman *et al.*, 2019).

Posttraumatic stress disorder phenotypes

Posttraumatic stress disorder in the UK Biobank

The definition of posttraumatic stress disorder (PTSD) in the UK Biobank is included in the Supplementary Material of Nievergelt et al. (2019). In brief, the PTSD phenotype was derived from six questions asked in the follow-up online mental health questionnaire. These questions

were derived from the brief civilian version of the PTSD Checklist Screener (PCL-S) which measures PTSD symptoms experienced in the previous month: avoidance of activities; disturbing thoughts; and feeling upset; and two additional questions related to feeling distant and feeling irritable. Each item was scored on a five-point Likert item measuring the amount of concern caused by that symptom in the past month (1="Not at all" to 5="Extremely"). In addition, a “trouble concentrating” question from the Patient Health Questionnaire-9 (PHQ9) depression questionnaire was added to replace a similar item that would normally be included in the PCL-S. This item was scored on a four-point Likert item according to frequency of difficulties associated with trouble concentrating (1="Not at all" and 4="Nearly every day"). For each participant, all items were summed into a total score ranging 3-29. Participants were considered PTSD cases if they had an overall PCL-S score ≥ 13. Participants were considered PTSD controls if they responded to all of the initial three questions and had PCL-S score ≤ 12 (Nievergelt *et al.*, 2019).

Posttraumatic Stress Disorder working group of the Psychiatric Genomics Consortium

The PTSD working group of the Psychiatric Genomics Consortium (PGC) meta-analysed data from 59 studies of PTSD to perform a GWAS known as the PTSD Freeze 1.5 (PGC1.5-PTSD). This sample involved 12,823 cases and 35,648 controls. The PGC gathered data for PGC1.5-PTSD through a number of independent studies who used a wide range of methods, primarily telephone diagnostic interviews and face-to-face clinical assessments. Some of the participants included in this cohort are veterans who have been combat or war-zone exposed. Other traumatic events assessed by the PGC include serious car accidents, campus shootings, domestic violence, and childhood physical and sexual abuse. Participants were assessed for current and lifetime PTSD using various instruments and different versions of the DSM (Nievergelt *et al.*, 2019). Further details of the contributing studies and instruments used to assess PTSD are contained in the Supplementary Material of Nievergelt et al. (2019).

To maximise power in genome-wide analysis, the UK Biobank and PTSD Freeze 1.5 (PGC1.5-PTSD) were combined. This combined data set is known as the PTSD Freeze 2 (PGC2-PTSD). In this sample, the number of cases was 23,212 and the number of controls was 151,447. The SNP-based heritability (liability scale) of the PGC2-PTSD phenotype was 0.06 (SE=0.011). In the main paper, we present genetic correlations between the MDD categories and the PGC1.5-PTSD summary statistics. We also calculated genetic correlations with the PGC2-PTSD summary statistics which are presented in Supplementary Table 1.

The Million Veteran Program

The Million Veteran Program (MVP) is a longitudinal study of United States military veterans who provided a blood sample for biobanking and responses to various questionnaires (Gaziano *et al.*, 2016). While the MVP sample consists of United States veterans, only 27.5% have confirmed war- or combat- exposure, while 29.3% who had not been exposed. The remaining 43.1% had unknown war- and combat-exposure (Stein *et al.*, 2020). The MVP-PTSD GWAS was of a binary, algorithmically-defined probable PTSD phenotype based on the veterans' Electronic Health Records. Further details of how the MVP phenotype was defined can be found in Stein et al. (2020).

Computational Methods

High Definition Likelihood inference of genetic correlations

Genetic correlations were estimated using High Definition Likelihood (HDL). Firstly for each phenotype, GWAS summary statistics were used to estimate the proportion of variance explained by common genetic variants (h^2_{SNP}) using High Definition Likelihood (HDL). While this is not a specific aim of the study, this step is necessary for interpreting genetic correlations between the PTSD phenotypes and the four MDD categories. HDL rests upon the principle of linkage-disequilibrium and extends the regression formula used by Linkage Disequilibrium Score Regression (LDSC) (see below). Unlike LDSC, HDL uses a full, likelihood-based method to estimate genetic correlations (Ning *et al.*, 2020). Further details can be found in the original paper by Ning et al. (2020).

All summary statistics were wrangled using the built-in HDL function for data wrangling (https://github.com/zhenin/HDL/wiki/Format-of-summary-statistics). All GWAS summary statistics used in this analysis had at least 99% SNP overlap with the UK Biobank LD reference panel using HapMap3 variants, apart from the Million Veteran Program (MVP) GWAS which had an overlap of 94.23%. Through correspondence with the HDL authors, we are confident that this level of SNP overlap is acceptable since the mismatch is not due to differences in the ancestral population from which the samples were created (e.g European vs. non-European populations) since both the MVP and the UK Biobank samples comprise participants from European ancestries only. The authors of HDL have performed a simulation to test this and discovered that missing SNPs lead

to more conservative results but should not generate false positives (Z Ning, personal correspondence). Therefore, a missing rate around 5% is acceptable. Furthermore, we repeated the data wrangling using the smaller HapMap2 (as opposed to HapMap3) reference panel and this did not improve the SNP overlap appreciably (data not shown). The authors of HDL therefore advised us that we continue to use the HapMap3 reference panel for the wrangling of the MVP Case-Control GWAS summary statistics.

High Definition Likelihood block-jackknife

The block-jackknife method was used to compare genetic correlations for a statistically significant difference between them. Each genetic correlation was compared in a pairwise fashion with all other genetic correlations within each set (the sets being UKB-PTSD, PGC2-PTSD, and MVP-PTSD presented in the main paper and PGC1.5-PTSD presented in the Supplementary Material). The block-jackknife uses a resampling method to estimate standard errors for each genetic correlation, which is then used to determine whether the differences between the genetic correlations are significantly different to zero. Here we provide an explanation of the block-jackknife method (Z Ning, personal correspondence).

When estimating genetic correlations, r_g1 and r_g2, one may then want to test whether the difference between them, referred to as the "global difference", is significantly different to zero. This would mean the null hypothesis is: $r_g1 - r_g2 = 0$. For r_g1, by setting jackknife.df=TRUE when using the HDL tool, you can get jackknife estimates of $r_g1.1$ to $r_g1.61$. This creates a file with 61 jackknife estimates because the genome is split into 61 pieces during the resampling process. $r_g1.k$ represents the estimated r_g with piece k removed. Similarly, you have $r_g2.1$ to $r_g2.61$ for r_g2.

With these two files containing the values in for both correlations, one can create the block-jackknife estimates by doing $r_g1.1 - r_g2.1$ all the way up to $r_g1.61 - r_g2.61$. The block-jackknife standard error can be found using this formula:

$$\widehat{se_{jack}} = \sqrt{\frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \overline{\hat{\theta}}_{(.)})^2}$$

Where n = 61 (for each of the resampling estimates), $\hat{\theta}$ = the difference between each resampling estimate (e.g., r_g1.1 - r_g2.1 all the way up to r_g1.61 - r_g2.61), θ = the mean of these estimates.

Then, the block-jackknife standard error and the global difference between the two correlations can be used to perform a Wald test:

$$Z = \text{global difference} \div \text{block-jackknife standard error}$$

The Wald test will give a z-score, which can then be used to find the two-tailed p-value, which will tell you whether the global difference is significantly different to zero. In our study, we corrected for multiple testing by considering $p<0.008$ as the threshold for significance (0.05/6 = 0.008, to account for the 6 block-jackknife tests carried out for each PTSD phenotype). Supplementary Table 5 contains all results from the HDL block-jackknife analysis.

Converting observed scale heritability estimates to the liability scale

The heritability of each trait was estimated by the HDL programme irrespective of population prevalence. Therefore, these estimates were converted to the liability scale in R, using code from the Nievergelt Lab github: https://gist.github.com/nievergeltlab/fb8a20feded72030907a9b4e81d1c6ea. Standard errors were also converted to the liability scale using the same formula(Lee *et al.*, 2011). Population prevalence for PTSD was adopted from Nievergelt et al. (2019), from Coleman et al. (2020) for reported trauma in MDD and from Burcusa & Iacono (2007) for recurrence in MDD.

Linkage Disequilibrium Score Regression

In addition to using HDL, we calculated genetic correlations using Linkage Disequilibrium Score Regression (LDSC). LDSC is a command line tool for estimating heritability and genetic correlation from GWAS summary statistics. It rests upon the principle of linkage-disequilibrium (LD). LD describes the degree to which an allele of one SNP is inherited or correlated with an allele of another SNP within a population (Bush and Moore, 2012). In this method, an “LD Score” of a given SNP refers to “the sum of LD r^2 measured with all other SNPs”. LDSC works by performing regression analysis on the LD scores and the test statistic of each SNP included in the GWAS, including those that do not meet genome-wide significance (Bulik-Sullivan *et al.*, 2015a). LDSC relies on the fact that the GWAS effect-size estimate for a given SNP incorporates

the effect of all SNPs in LD with that particular SNP. For most complex human traits, which are polygenic, SNPs with high LD will have higher chi-square test statistics, on average, than SNPs with low LD (Bulik-Sullivan *et al.*, 2015b).

As in the HDL analysis, GWAS summary statistics were used to estimate the proportion of variance explained by common genetic variants (h^2_{SNP}) using LDSC. While this is not a specific aim of the study, this step is necessary for interpreting genetic correlations.

Linkage Disequilibrium Score Regression (LDSC) block-jackknife

The block-jackknife method was used to compare genetic correlations for a statistically significant difference between them. Each genetic correlation was compared in a pairwise fashion with all other genetic correlations within each set (the sets being UKB-PTSD, PGC1.5-PTSD, PGC2-PTSD, and MVP-PTSD). As described above for HDL, the block-jacknife works by repeated re-estimation of blocks of jack-knife estimates. In the case of LDSC, the number of blocks is set by the user (we used 200), but otherwise the calculation of significant differences follows the description provided for HDL.

Polygenic Risk Scores

PRS were calculated PRSice v2.3.1, a command line programme that uses GWAS summary statistics to calculate genetic risk of a base phenotype in individuals from an independent sample. A PRS refers to the summation of alleles across many genetic loci associated with a particular trait or disease. These alleles are typically weighted by effect sizes estimated from GWAS (Euesden *et al.*, 2015). In our study, the PRS represents the aggregated PTSD risk conferred by many DNA variants in participants of the UK Biobank who have MDD. Since even well-powered GWAS offer only tentative evidence for causally associated variants, PRS are calculated at a range of different *P*-value thresholds to provide the 'best-fit', or most predictive, PRS (Dudbridge, 2013). Once the best fitting PRS has been estimated, these are used as predictors of a target phenotype in individuals in an independent sample in a regression. PRS were calculated at 11 p-value thresholds ($5x10^{-8}$, $1x10^{-5}$, $1x10^{-3}$, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1). Phenotype permutations (implemented in PRSice) were used to produce an empirical p-value for the association at the best-fitting PRS, which accounts for testing multiple thresholds. We then performed a logistic regression to examine whether the risk scores are more strongly associated

with MDD with reported trauma or MDD without reported trauma, and recurrent or single-episode MDD.

Power calculations

We calculated the power of our PRS analyses using the Additive Variance Explained and Number of Genetic Effects Method of Estimation (AVENGEME) programme in R (Dudbridge, 2013). We were uncertain about the covariance between genetic effects in the target sample (MVP PTSD) and the two target samples. Therefore, based on the parameters required by the AVENGEME package (see below), we calculated that for the PRS analysis testing the risk score's association with MDD with and without reported trauma to be powered at least 80%, the covariance between the genetic effect sizes in the training and target samples would need to be at least 0.024. We calculated that for the PRS analysis testing the risk score's association with recurrent and single-episode MDD to be powered at least 80%, the covariance between the genetic effect sizes in the training and target samples would need to be at least 0.0305. Due to these minimum covariance estimates being so low, we were confident that the PRS analyses were powered by at least 80%.

Parameters required by the AVENGEME package:

- Number of SNPs included in PRS after clumping = 183,881
- Proportion of variance in PTSD explained by genetic effects = 0.03 (Stein *et al.*, 2020)
- Training sample prevalence of PTSD = 0.18 (Stein *et al.*, 2020)
- Population prevalence of PTSD = 0.30 (Nievergelt *et al.*, 2019)
- Target sample prevalence of reported trauma among MDD cases = 0.59
- Population prevalence of reported trauma among MDD cases = 0.52 (Coleman *et al.*, 2020)
- Target sample prevalence of recurrent MDD = 0.59
- Population prevalence of recurrent MDD among MDD cases = 0.5 (Burcusa and Iacono, 2007)

Supplementary results

High Definition Likelihood block-jackknife

The results of the HDL block-jackknife analysis can be found in Supplementary Table 5. Differences between genetic correlations were considered statistically significant if they

surpassed the Bonferroni corrected alpha (0.05/6 = 0.008; i.e. to correct for the six block-jackknife tests per PTSD phenotype).

Linkage Disequilibrium Score Regression genetic correlations

Genetic correlations (Supplementary Table 6) estimated by LDSC were considered significantly different to zero and to one if they reached or surpassed the Bonferroni-corrected alpha in each analysis (0.05/4 = 0.0125; i.e. to correct for the four tests).

Linkage Disequilibrium Score Regression block-jackknife

We tested the differences between the genetic correlations using a block-jackknife. Differences were considered statistically significant if they passed a Bonferroni-corrected alpha of 0.008 (0.05/6 = 0.008; i.e. to correct for the six block-jackknife tests per PTSD phenotype). Supplementary Table 7 contains the results of the LDSC block-jackknife.

Comparison of the two methods for estimating genetic correlations

In our study, we performed the genetic correlation and block-jackknife analyses in both HDL and LDSC. We have opted to only present the HDL results in the main text since it is the preferable method (for reasons discussed in the paper). Nonetheless, there are some key differences between the HDL and LDSC results which need to be discussed here.

An anticipated difference was that no genetic correlations were found to differ significantly from any other genetic correlations when using LDSC. The absence of statistically significant differences in genetic correlation must be interpreted in the context of the power of the original GWAS from which the summary statistics used in this study were created. As shown in Supplementary Table 6, the standard errors surrounding the LDSC point estimates are notably large, meaning that any differences between genetic correlation would need to be large to be detected as significantly different to zero in the block-jackknife analysis. The advantage of using HDL to estimate genetic correlations is reduction in variance of the point estimate (Ning *et al.*, 2020), which provides better power to observe differences between genetic correlations. This was the case in our study, where we found that PTSD is significantly more genetically correlated with recurrent MDD than it is genetically correlated with MDD without reported trauma when using

HDL. Similar to the HDL results, when using LDSC we find that all PTSD phenotypes have a higher genetic correlation with recurrent MDD than with MDD without reported trauma; however, unlike when using the HDL block-jackknife, this difference was not significantly different to zero in the results of LDSC block-jackknife.

Another difference between the HDL and LDSC results is the pattern of genetic correlations. Firstly, when using LDSC, we observe a clear pattern where all PTSD phenotypes are more genetically correlated with recurrent MDD than single-episode MDD. We do not see this pattern when using HDL. Secondly, the consistent pattern whereby all PTSD phenotypes are more genetically correlated with MDD with reported trauma compared MDD without reported trauma when using HDL only holds true for the PGC 1.5, PGC 2 and MVP PTSD phenotypes when using LDSC, whereas the UK Biobank PTSD phenotype appears to be slightly more genetically correlated with MDD with without reported trauma (r_g difference = 0.0039).

These differences are likely due to inconsistent SNP reference panels used by HDL and LDSC, which are needed to estimate the LD-scores. Therefore, a difference in the reference panel between the two methods is likely to lead to differing results. In HDL, the reference panels with imputed SNPs are based on genotypes in UK Biobank, which were imputed to HRC and UK10K + 1000 Genomes. Specifically in our study, we used the 1,029,876 Quality Controlled UK Biobank imputed HapMap3 SNPs reference panel (https://github.com/zhenin/HDL/wiki/Reference-panels). On the other hand, LDSC uses a reference panel from the 1000 Genomes Project which is not specific to the UK Biobank (Bulik-Sullivan *et al.*, 2015b). According to correspondence with the authors, HDL uses the UK Biobank reference panels instead of 1000 Genomes because the UK Biobank has a larger sample size. This therefore leads to more accurate estimates of LD. Further details can be found at the discussion section of Ning et al. (2020). However, this increase in the accuracy of LD estimation is likely to apply most strongly to summary statistics from the UK in general and UK Biobank in particular, and less so to samples descended from European ancestry populations from elsewhere in Europe. This being the case, it is feasible that the differences in genetic correlations we observe in our study partly reflect differences in the proportion of UK ancestry in the PTSD summary statistics (as all of the MDD summary statistics were drawn from UK Biobank). This inconsistency between LDSC and HDL, and its potential relationship to UK ancestry in the summary statistics assessed, is likely to have wider implications than our study alone and requires a detailed examination beyond the scope of this study.

Polygenic risk scores

Polygenic risk scores for PTSD were calculated PRSice v2.3.1. Two PRS analyses were performed: two regressions using PRS based on the MVP PTSD summary statistics. Full results of these analyses can be found in Supplementary Table 8. Beta coefficients were exponentiated in R to give odds ratios (OR) and 95% confidence intervals (CI).

Nagelkerke's R^2

The PRS were initially calculated without consideration of the prevalence of the target phenotype (reported psychological trauma in individuals with MDD and episode recurrence in individuals with MDD) within the population. Based on the results from PRSice, Nagelkerke's R^2 was calculated for the estimated population prevalence, +10% and -10% in R. We obtained an estimated population prevalence for trauma exposure among MDD cases of 52% from the Supplementary Material of Coleman et al. (2020). We obtained an estimated population prevalence for recurrence among MDD cases of 50% from Burcusa & Iacono (2007). For the regression testing the PTSD risk scores' association with MDD with reported trauma and MDD without reported trauma, the Nagelkerke's R^2 based on a population prevalence of 42%, 52% and 62% were 0.0563%, 0.0569% and 0.0555% respectively. For the regression testing the PTSD risk scores' association with recurrent MDD compared to single-episode MDD, the Nagelkerke's R^2 based on population prevalence 40%, 50% and 60% were 0.0296%, 0.0301% and 0.0296% respectively.

Supplementary references

Allen NE, Sudlow C, Peakman T, Collins R, UK Biobank (2014) UK biobank data: come and get it. *Science translational medicine* 6, 224ed4.

Bellis MA, Hughes K, Leckenby N, Perkins C, Lowey H (2014) National household survey of adverse childhood experiences and their relationship with resilience to health-harming behaviors in England. *BMC medicine* 12, 72.

Bernstein DP, Fink L, Handelsman L, Foote J, Lovejoy M, Wenzel K, Sapareto E, Ruggiero J (1994) Initial reliability and validity of a new retrospective measure of child abuse and neglect. *The American journal of psychiatry* 151, 1132–1136.

Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R…Neale BM (2015a) An atlas

of genetic correlations across human diseases and traits. *Nature genetics* 47, 1236–1241.

Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium…Neale BM (2015b) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* 47, 291–295.

Burcusa SL, Iacono WG (2007) Risk for recurrence in depression. *Clinical psychology review* 27, 959–985.

Bush WS, Moore JH (2012) Chapter 11: Genome-wide association studies. *PLoS computational biology* 8, e1002822.

Coleman JRI, Gaspar HA, Bryois J, Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Breen G (2019) The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biological psychiatry*.

Coleman JRI, Peyrot WJ, Purves KL, Davis KAS, Rayner C, Choi SW…Breen G (2020) Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Molecular psychiatry* 25, 1430-1446.

Davis KAS, Coleman JRI, Adams M, Allen N, Breen G, Cullen B…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, e18.

Dudbridge F (2013) Power and predictive accuracy of polygenic risk scores. *PLoS genetics* 9, e1003348.

Euesden J, Lewis CM, O'Reilly PF (2015) PRSice: Polygenic Risk Score software. *Bioinformatics* 31, 1466–1468.

Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J…O'Leary TJ (2016) Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *Journal of clinical epidemiology* 70, 214–223.

Grabe HJ, Schulz A, Schmidt CO, Appel K, Driessen M, Wingenfeld K…Freyberger HJ (2012) [A brief instrument for the assessment of childhood abuse and neglect: the childhood trauma screener (CTS)]. *Psychiatrische Praxis* 39, 109–115.

Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. *American journal of human genetics* 88, 294–305.

Nievergelt CM, Maihofer AX, Klengel T, Atkinson EG, Chen C-Y, Choi KW…Koenen KC (2019) International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nature communications* 10, 4558.

Ning Z, Pawitan Y, Shen X (2020) High-definition likelihood inference of genetic correlations across human complex traits. *Nature genetics* 52, 859–864.

Stein MB, Levey DF, Cheng Z, Wendt FR, Harrington K, Cho K…Gelernter J (2020) Genomic Characterization of Posttraumatic Stress Disorder in a Large US Military Veteran Sample. *BioRxiv*.

Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J…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 medicine* 12, e1001779.

Supplementary tables

Table S2.1. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (SE) and 95% confidence intervals (CI) of Psychiatric Genomics Consortium 2 PTSD (PGC2-PTSD) with the four major depressive disorder (MDD) categories. *P* (diff 0) refers to *p*-value to test whether the r_g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the r_g estimate differs from 1. Genetic correlations were considered significant if they reached or surpassed the Bonferroni adjusted threshold (*p*<0.0125). Significant *p*-values are shown in bold. The SNP-based heritability of the PGC2-PTSD phenotype was estimated by HDL to be 0.06 (SE = 0.006).

PTSD Phenotype	MDD phenotype	r_g	SE	Lower CI	Upper CI	*P* (diff 0)	*P* (diff 1)
PGC2-PTSD	MDD with reported trauma	0.6497	0.0825	0.4880	0.8114	**3.39×10^{-15}**	**2.18×10^{-5}**
PGC2-PTSD	MDD without reported trauma	0.5509	0.1107	0.3339	0.7679	**5.12×10^{-7}**	**4.24×10^{-5}**
PGC2-PTSD	Recurrent MDD	0.7915	0.0821	0.6306	0.9524	**8.45×10^{-22}**	**0.01**
PGC2-PTSD	Single-episode MDD	0.8147	0.1129	0.5934	1.0360	**5.27×10^{-13}**	0.1

Table S2.2. Difference in reporting rates of traumatic life events, assessed by the Mental Health Questionnaire (MHQ), between individuals with recurrent and single-episode major depressive disorder (MDD) in UK Biobank MHQ respondents who met criteria for lifetime MDD (*N*=29,471). Traumatic events include two childhood events, two adulthood events and five catastrophic/posttraumatic stress disorder (PTSD)-related events. Differences were considered significant if they surpassed the Bonferroni adjusted alpha (*p*<0.006) to correct for the nine chi-square tests. Significant differences are shown in bold.

Trauma category	Traumatic event	Endorsement in single-episode MDD (%)	Endorsement in recurrent MDD (%)	χ^2 statistic	*P*-value
Childhood physical abuse	Physically abused by family as a child	2,644 (22%)	4,846 (28%)	126	**3.82x10-29**
Childhood physical neglect	Someone to take to doctor when needed as a child	1,948 (16%)	3,556 (20%)	80	**3.80x10-19**
Adulthood emotional neglect	Been in a confiding relationship as an adult	3,758 (32%)	6,937 (40%)	205	**2.08x10-46**
Adulthood physical neglect	Able to pay rent/mortgage as an adult	1,963 (17%)	3,603 (21%)	87	**1.07x10-20**
PTSD-related: experience of war or combat	Been involved in combat or exposed to war-zone	400 (3%)	595 (3%)	0.13	0.72
PTSD-related: serious accident	Been in serious accident believed to be life-threatening	1,247 (10%)	2,313 (13%)	55	**9.47x10-14**
PTSD-related: life-threatening illness	Diagnosed with life-threatening illness	2,092 (17%)	3,211 (18%)	5	0.03
PTSD-related: physically violent crime	Victim of physically violent crime	2,332 (19%)	4,273 (25%)	106	**8.12x10-25**
PTSD-related: witnessed sudden violent death	Witnessed sudden violent death	1,778 (15%)	2,866 (16%)	14	**0.0002**

Table S2.3. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (SE), and 95% confidence intervals (CI) between the four posttraumatic stress disorder (PTSD) phenotypes: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD). *P* (diff 0) refers to *p*-value to test whether the r_g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the r_g estimate differs from 1. Genetic correlations were considered significant if they surpassed the Bonferroni adjusted threshold ($p<0.008$) to correct for the 6 tests. Significant *p*-values are shown in bold.

Phenotype 1	Phenotype 2	r_g	SE	Lower CI	Upper CI	*P* (diff 0)	*P* (diff 1)
UKB-PTSD	PGC1.5-PTSD	0.5738	0.1155	0.3474	0.8001	**$6.83Ex10^{-7}$**	**0.0002**
UKB-PTSD	PGC2-PTSD	0.8016	0.0703	0.6638	0.9394	**$3.80x10^{-30}$**	**0.005**
UKB-PTSD	MVP-PTSD	0.6702	0.0707	0.5316	0.8088	**$2.56x10^{-21}$**	**$3.10x10^{-6}$**
PGC1.5-PTSD	PGC2-PTSD	0.9853	0.1481	0.6950	1.2756	**$2.85x10^{-11}$**	0.92
PGC1.5-PTSD	MVP-PTSD	1.0053	0.1669	0.6782	1.3324	**$1.71x10^{-9}$**	0.97
PGC2-PTSD	MVP-PTSD	0.9810	0.1105	0.7644	1.1976	**$6.96x10^{-19}$**	0.86

Table S2.4. High Definition Likelihood (HDL) genetic correlation estimates (r_g), standard errors (SE), and 95% confidence intervals (CI) between the four major depressive disorder (MDD) categories: 1) MDD with reported trauma, 2) MDD without reported trauma, 3) recurrent MDD, 4) single-episode MDD. *P* (diff 0) refers to *p*-value to test whether the r_g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the r_g estimate differs from 1. Genetic correlations were considered significant if they surpassed the Bonferroni adjusted threshold ($p<0.008$) to correct for the 6 tests. Significant *p*-values are shown in bold.

Phenotype 1	Phenotype 2	r_g	SE	Lower CI	Upper CI	*P* (diff 0)	*P* (diff 1)
MDD with reported trauma	MDD without reported trauma	0.6068	0.0727	0.4643	0.7493	**7.11×10^{-17}**	**6.35×10^{-8}**
MDD with reported trauma	Recurrent MDD	0.8120	0.0438	0.7262	0.8978	**1.46×10^{-76}**	**1.77×10^{-5}**
MDD with reported trauma	Single-episode MDD	0.8508	0.0629	0.7275	0.9741	**1.13×10^{-41}**	0.02
MDD without reported trauma	Recurrent MDD	0.8686	0.0748	0.7222	1.0152	**3.31×10^{-31}**	0.08
MDD without reported trauma	Single-episode MDD	0.9175	0.1013	0.7190	1.1160	**1.35×10^{-19}**	0.42
Recurrent MDD	Single-episode MDD	0.9424	0.0596	0.8256	1.0592	**2.49×10^{-56}**	0.33

Table S2.5. Genetic correlation results from the High Definition Likelihood (HDL) block-jackknife analysis of posttraumatic stress disorder (PTSD) and the four major depressive disorder (MDD) categories. The four PTSD phenotypes include: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD). Each genetic correlation was compared in a pairwise fashion with all other genetic correlations in the set (i.e. for each PTSD phenotype). r_g difference refers to the difference between the two genetic correlation estimates, SE refers to the standard error and *P* (diff 0) refers to *p*-value to test whether the r_g difference differs significantly from 0. Differences between genetic correlations were considered statistically significant if they surpassed the Bonferroni adjusted threshold (p<0.008). Significant *p*-values are shown in bold.

HDL genetic correlation 1	HDL genetic correlation 2	r_g difference	SE	*P* (diff 0)
UKB-PTSD and MDD with reported trauma	UKB-PTSD and MDD without reported trauma	0.1339	0.0910	0.14
UKB-PTSD and MDD with reported trauma	UKB-PTSD and recurrent MDD	-0.1094	0.0572	0.06
UKB-PTSD and MDD with reported trauma	UKB-PTSD and single-episode MDD	-0.0426	0.0864	0.62
UKB-PTSD and MDD without reported trauma	UKB-PTSD and recurrent MDD	-0.2433	0.0696	**4.77×10^{-4}**
UKB-PTSD and MDD without reported trauma	UKB-PTSD and single-episode MDD	-0.1765	0.0773	0.02
UKB-PTSD and recurrent MDD	UKB-PTSD and single-episode MDD	0.0668	0.0752	0.37
PGC1.5-PTSD and MDD with reported trauma	PGC1.5-PTSD and MDD without reported trauma	0.0679	0.0988	0.49
PGC1.5-PTSD and MDD with reported trauma	PGC1.5-PTSD and recurrent MDD	-0.1417	0.0730	0.05
PGC1.5-PTSD and MDD with reported trauma	PGC1.5-PTSD and single-episode MDD	-0.2040	0.1413	0.15
PGC1.5-PTSD and MDD without reported trauma	PGC1.5-PTSD and recurrent MDD	-0.2096	0.0745	**4.90×10^{-3}**
PGC1.5-PTSD and MDD without reported trauma	PGC1.5-PTSD and single-episode MDD	-0.2719	0.1183	0.02

PGC1.5-PTSD and recurrent MDD	PGC1.5-PTSD and single-episode MDD	-0.0623	0.1228	0.61
PGC2-PTSD and MDD with reported trauma	PGC2-PTSD and MDD without reported trauma	0.0988	0.1261	0.43
PGC2-PTSD and MDD with reported trauma	PGC2-PTSD and recurrent MDD	-0.1418	0.0875	0.10
PGC2-PTSD and MDD with reported trauma	PGC2-PTSD and single-episode MDD	-0.1650	0.1047	0.12
PGC2-PTSD and MDD without reported trauma	PGC2-PTSD and recurrent MDD	-0.2406	0.0862	**5.28×10^{-3}**
PGC2-PTSD and MDD without reported trauma	PGC2-PTSD and single-episode MDD	-0.2638	0.1062	0.01
PGC2-PTSD and recurrent MDD	PGC2-PTSD and single-episode MDD	0.0232	0.1024	0.82
MVP-PTSD and MDD with reported trauma	MVP-PTSD and MDD without reported trauma	0.0538	0.1197	0.65
MVP-PTSD and MDD with reported trauma	MVP-PTSD and recurrent MDD	-0.0203	0.0722	0.78
MVP-PTSD and MDD with reported trauma	MVP-PTSD and single-episode MDD	-0.0894	0.1028	0.38
MVP-PTSD and MDD without reported trauma	MVP-PTSD and recurrent MDD	-0.0741	0.0880	0.40
MVP-PTSD and MDD without reported trauma	MVP-PTSD and single-episode MDD	-0.1432	0.1080	0.18
MVP-PTSD and recurrent MDD	MVP-PTSD and single-episode MDD	-0.0691	0.1130	0.54

Table S2.6. Linkage Disequilibrium Score Regression (LDSC) genetic correlation estimates (r_g) and standard errors (SE) and 95% confidence intervals (CI) of 1) UK Biobank posttraumatic stress disorder (PTSD), 2) Psychiatric Genomics Consortium (PGC) 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD) with the four major depressive disorder (MDD) categories. *P* (diff 0) refers to *p*-value to test whether the r_g estimate differs from 0. *P* (diff 1) refers to *p*-value to test whether the r_g estimate differs from 1. Genetic correlations were considered significant if they passed the Bonferroni adjusted threshold ($p<0.0125$). Significant *p*-values are shown in bold.

PTSD Phenotype	MDD phenotype	r_g	SE	Lower CI	Upper CI	*P* (diff 0)	*P* (diff 1)
UKB-PTSD	MDD with reported trauma	0.6499	0.0832	0.4868	0.8130	**5.70×10^{-15}**	**2.58×10^{-5}**
UKB-PTSD	MDD without reported trauma	0.6538	0.1365	0.3863	0.9213	**1.67×10^{-6}**	**0.01**
UKB-PTSD	Recurrent MDD	0.7815	0.0550	0.6737	0.8893	**7.55×10^{-46}**	**7.11×10^{-5}**
UKB-PTSD	Single-episode MDD	0.7158	0.1045	0.5110	0.9206	**7.29×10^{-12}**	**6.54×10^{-3}**
PGC1.5-PTSD	MDD with reported trauma	0.6976	0.2131	0.2799	1.1153	**1.06×10^{-3}**	0.16
PGC1.5-PTSD	MDD without reported trauma	0.4514	0.2471	-0.0329	0.9357	**6.77×10^{-2}**	0.03
PGC1.5-PTSD	Recurrent MDD	0.7438	0.1545	0.4410	1.0466	**1.48×10^{-6}**	0.10
PGC1.5-PTSD	Single-episode MDD	0.5743	0.1696	0.2418	0.9067	**7.07×10^{-4}**	**0.01**
PGC2-PTSD	MDD with reported trauma	0.6263	0.1160	0.3989	0.8537	**6.72×10^{-8}**	**1.27×10^{-3}**
PGC2-PTSD	MDD without reported trauma	0.5461	0.1680	0.2168	0.8754	**1.15×10^{-3}**	**6.90×10^{-3}**
PGC2-PTSD	Recurrent MDD	0.7592	0.0880	0.5867	0.9317	**6.47×10^{-18}**	**6.21×10^{-3}**
PGC2-PTSD	Single-episode MDD	0.6597	0.1161	0.4321	0.8873	**1.32×10^{-8}**	**3.38×10^{-3}**
MVP-PTSD	MDD with reported trauma	0.3906	0.0766	0.2405	0.5407	**3.40×10^{-7}**	**1.78×10^{-15}**
MVP-PTSD	MDD without reported trauma	0.3593	0.1035	0.1564	0.5622	**5.15×10^{-4}**	**6.00×10^{-10}**
MVP-PTSD	Recurrent MDD	0.4812	0.0624	0.3589	0.6035	**1.30×10^{-14}**	**9.24×10^{-17}**
MVP-PTSD	Single-episode MDD	0.4392	0.0885	0.2657	0.6127	**6.84×10^{-7}**	**2.35×10^{-10}**

Table S2.7. genetic correlation results from the Linkage Disequilibrium Score Regression (LDSC) block-jackknife analysis of posttraumatic stress disorder (PTSD) and the four major depressive disorder (MDD) categories. The four PTSD phenotypes include: 1) UK Biobank PTSD (UKB-PTSD), 2) Psychiatric Genomics Consortium 1.5 PTSD (PGC1.5-PTSD), 3) PGC 2 PTSD (PGC2-PTSD), 4) Million Veteran Program PTSD (MVP-PTSD). r_g difference refers to the difference between the two genetic correlation estimates, SE refers to the standard error and *P* (diff 0) refers to *p*-value to test whether the r_g difference differs significantly from 0. Each genetic correlation was compared in a pairwise fashion with all other genetic correlations in the set (i.e. for each PTSD phenotype). Differences between genetic correlations were considered statistically significant if they passed the Bonferroni adjusted threshold (p<0.008).

LDSC genetic correlation 1	LDSC genetic correlation 2	r_g difference	SE	*P* (diff 0)
UKB-PTSD and MDD with reported trauma	UKB-PTSD and MDD without reported trauma	-0.0039	0.1677	0.99
UKB-PTSD and MDD with reported trauma	UKB-PTSD and recurrent MDD	-0.1316	0.0825	0.08
UKB-PTSD and MDD with reported trauma	UKB-PTSD and single-episode MDD	-0.0659	0.1234	0.55
UKB-PTSD and MDD without reported trauma	UKB-PTSD and recurrent MDD	-0.1277	0.1238	0.25
UKB-PTSD and MDD without reported trauma	UKB-PTSD and single-episode MDD	-0.0620	0.1271	0.57
UKB-PTSD and recurrent MDD	UKB-PTSD and single-episode MDD	0.0657	0.1091	0.51
PGC1.5-PTSD and MDD with reported trauma	PGC1.5-PTSD and MDD without reported trauma	0.2462	0.2851	0.40
PGC1.5-PTSD and MDD with reported trauma	PGC1.5-PTSD and recurrent MDD	-0.0462	0.1724	0.76
PGC1.5-PTSD and MDD with reported trauma	PGC1.5-PTSD and single-episode MDD	0.1233	0.2004	0.58
PGC1.5-PTSD and MDD without reported trauma	PGC1.5-PTSD and recurrent MDD	-0.2924	0.2110	0.16
PGC1.5-PTSD and MDD without reported trauma	PGC1.5-PTSD and single-episode MDD	-0.1229	0.2305	0.57
PGC1.5-PTSD and recurrent MDD	PGC1.5-PTSD and single-episode MDD	0.1695	0.1774	0.36
PGC2-PTSD and MDD with reported trauma	PGC2-PTSD and MDD without reported trauma	0.0802	0.1959	0.65
PGC2-PTSD and MDD with reported trauma	PGC2-PTSD and recurrent MDD	-0.1329	0.1125	0.23
PGC2-PTSD and MDD with reported trauma	PGC2-PTSD and single-episode MDD	-0.0334	0.1348	0.87
PGC2-PTSD and MDD without reported trauma	PGC2-PTSD and recurrent MDD	-0.2131	0.1448	0.12
PGC2-PTSD and MDD without reported trauma	PGC2-PTSD and single-episode MDD	-0.1136	0.1626	0.49
PGC2-PTSD and recurrent MDD	PGC2-PTSD and single-episode MDD	0.0995	0.1250	0.36
MVP-PTSD and MDD with reported trauma	MVP-PTSD and MDD without reported trauma	0.0313	0.1261	0.82
MVP-PTSD and MDD with reported trauma	MVP-PTSD and recurrent MDD	-0.0906	0.0736	0.14
MVP-PTSD and MDD with reported trauma	MVP-PTSD and single-episode MDD	-0.0486	0.0992	0.60
MVP-PTSD and MDD without reported trauma	MVP-PTSD and recurrent MDD	-0.1219	0.0954	0.15
MVP-PTSD and MDD without reported trauma	MVP-PTSD and single-episode MDD	-0.0799	0.0953	0.40

MVP-PTSD and recurrent MDD	MVP-PTSD and single-episode MDD	0.0420	0.0904	0.52

Table S2.8. Results table from PRSice analysis. Polygenic Risk Scores (PRS) for postraumatic stress disorder (PTSD) were calculated based on the Million Veteran Program (MVP-PTSD) summary statistics. Table includes the best fitting *p*-value threshold used in the analyses (Threshold), the standardised beta coefficient from regression (Standardised beta) and standard error (SE), the odds ratio (OR) and 95% confidence intervals (CI), *p*-value from regression (*P*), the number of single nucleotide polymorphisms used in creating the risk scores (Number of SNPs) and empirical *p*-value accounting for testing at multiple thresholds (Empirical *P*). Regression coefficients were considered statistically significant if they surpassed the Bonferroni adjusted alpha (p<0.025). Significant *p*-values are shown in bold.

Regression	Threshold	Standardised beta	SE	OR (95% CI)	*P*	Number of SNPs	Empirical *P*
MDD with reported trauma vs. MDD without reported trauma	0.4	0.04	0.014	1.04 (1.01 – 1.07)	0.003	101111	**0.02**
Recurrent MDD vs. single-episode MDD	0.001	-0.03	0.013	0.97 (0.95 – 0.99)	0.01	646	0.08

Appendix 3. Supplementary material for chapter 3

Supplementary material

Supplementary methods

Major depressive disorder and anxiety disorder diagnostic criteria

COPING NBR participants who met symptom-based diagnostic criteria for major depressive disorder (MDD) and/or any anxiety disorder were combined with GLAD participants to create a cohort of "affected" participants. The category of "any anxiety disorder" included generalised anxiety disorder (GAD), specific phobia, social phobia, panic disorder, and agoraphobia. COPING NBR participants who did not meet criteria for MDD or any anxiety disorder were categorised as unaffected participants. These criteria were based upon the MDD, GAD, specific phobia, social anxiety disorder, panic disorder, and agoraphobia modules from an adapted version of the short form Composite International Diagnostic Interview (CIDI-SF) in the COPING baseline survey. Diagnostic algorithms were written in R to categorise COPING NBR participants as having a lifetime symptom-based diagnosis for these disorders if their responses on the CIDI-SF corresponded to Diagnostic Statistical Manual V (DSM-V) criteria. Further detail of these symptom-based diagnoses have been described elsewhere (Davies *et al.*, 2022).

Self-reported bipolar diagnosis

GLAD participants have two opportunities to self-report a diagnosis of bipolar disorder. The first is in the Mental Health Disorders (MHD) section and the second is in the MDQ section of the survey. Participants who self-reported that they had received a diagnosis of bipolar disorder by a professional in *one* of the two questions, but had missing data in the other, were categorised as a bipolar disorder case. Participants who answered "No" to one question but had missing data on the other question, had answers that did not match, or had missing data on both questions were excluded from analyses. NBR participants only have one opportunity to self-report a diagnosis of bipolar disorder in the MHD section of the COPING survey. Participants who had missing data for this question were excluded from analyses.

Mood Disorder Questionnaire screener

A positive screen in the MDQ requires that seven or more items are endorsed, that at least several of the items co-occurred, and that the symptoms caused at least moderate impairment (Hirschfeld *et al.*, 2000, 2003).

Statistical analyses

Exploratory and confirmatory factor analyses

Exploratory factor analysis (EFA) is a classic latent variable technique which finds latent variables based on the correlation structure of the manifest input variables (here, the MDQ items). EFA is distinguished from confirmatory factor analysis (CFA) in that in CFA the model is determined by the researcher based on an underlying theory or from a model identified in the EFA (Mair, 2018).

In our paper, each sample of participants was randomly split into two (without replacement) using the "rsample" R package; 70% for EFA and 30% for CFA. Before splitting the sample, the nearZeroVar R package was used to diagnose variables that have one unique value (i.e., are zero variance predictors) or are near zero variance predictors.

Polychoric correlation matrices were then computed for all items using the "polycor" R package; the answers to the MDQ were binary and tetrachoric correlations are a special case of polychoric correlations. To check for multicollinearity of the items, the determinant of the matrix was computed. If the determinant was greater than 0.00001 then the matrices were inspected for highly correlated items and were removed accordingly. To test for singularity, the matrices were inspected for values <0.3. Any items that correlated <0.3 with all other items were removed from the analysis. The ordinal alpha statistic, the Kaiser-Meyer-Olkin (KMO) statistic (Kaiser, 1974) and Bartlett's Test of Sphericity (Bartlett, 1950) were computed to assess whether the data were appropriate for exploratory factor analysis (EFA). Parallel analysis (Horn, 1965), Very Simple Structure (VSS) (Revelle and Rocklin, 1979), and Velicer's Minimum Average Partial (MAP) criterion (Velicer, 1976) were estimated to gain an initial idea of the number of factors suited to the data.

The EFA was performed in the 70% of the sample using the minimum residuals method in the "psych" R package. Factors were allowed to correlate using oblimin rotation. First, we performed EFA specifying a one factor solution to investigate whether all the items loaded onto one factor and therefore represent a single latent construct. We then continued performing EFA, by adding one more factor at a time, until the minimum number of items loading onto a factor reached one.

A number of fit criteria were used to assess which factor model fit the data best: the root mean square error of approximation (RMSEA) <0.5 (good fit) or 0.06-0.08 (fair fit), Tucker–Lewis Index (TLI) ≥0.90, root mean square residuals (RMSR) as small as possible (preferably <0.08) (Mair, 2018), and small Bayesian Information Criterion (BIC) relative to other factor solutions. Items were retained in a factor if the loading was ≥0.3 and greater than all loadings on other factors. Where multiple models showed comparable fit statistics, the model that encompassed the highest number of items was selected and models where the minimum number of items per factor was greater than three were preferred.

To validate the EFA-derived model, CFA was conducted on the remaining 30% of the sample using the "lavaan" R package. A number of fit statistics were interpreted: the TLI ≥0.9, Standardised Root Mean Square of the Residuals (RSMR) as small as possible, a smaller BIC compared to other factor solutions and Comparative Fit Index (CFI) ≥0.9. The CFA was then applied to the whole sample using the lavaan R package to provide overview fit statistics.

Factor scores for each factor in the best-fitting model were computed for every participant to gain individual values for participants which represent their "placement" on each factor. Factor scores were computed with the "lav_predict" function from the "lavaan" R package using the Empirical Bayes Modal (EBM) method. The factor scores were then transformed using a rank-based inverse normal transformation (INT) using the "RNOmni" R package and standardised using base R functions. Prior to transformation, ties in the data were broken using the "surveillance" R package.

Genetic analyses

Genotyping, imputation, and quality control

Genotyping

Quality assurance measures were calculated by ThermoFisher: samples with a dish QC value ≥0.82 (capturing the resolution of true signal from background noise on the genotyping array) and an initial call rate ≥0.97 were retained. Variants were recommended for inclusion if they were genotyped with high resolution (classified as "PolyHighResolution", "NoMinorHom", or "MonoHighResolution" by ThermoFisher). Data passing quality assurance was transferred to the Social, Genetic, and Developmental Psychiatry Centre at King's College London for further quality control, adapted from previous pipelines (Coleman *et al.*, 2016).

Data for GLAD and COPING NBR were processed separately following the same pipeline. An initial set of quality control was performed to determine sample ancestry. This consisted of excluding variants with a minor allele frequency (MAF) <0.01, variants and individuals with a call rate <95%, and variants with Hardy-Weinberg $p<10^{-10}$. Additional checks were performed on individuals to exclude outliers for sex discrepancies, heterozygosity, and relatedness. Samples were merged with data from Phase 3 of the 1000 Genomes project and principal component analyses were performed on genome-wide genotype data. Samples clustering with known individuals from European ancestries in the 1000 Genomes project formed the majority of the genotyped GLAD and COPING NBR cohorts (96% and 98% respectively; **figure S3.18**) and so further analyses were restricted only to these participants. Quality control was repeated, on raw data restricted to European ancestry participants. This comprised the same measures as above.

Imputation

For GLAD and COPING NBR separately, high quality genotype data was lifted to build 38 of the human genome and imputed to TopMed freeze 8, using version 1.5.7 of the dedicated imputation server provided by the University of Michigan (Taliun et al., 2021) with prior phasing using EAGLE2 (Loh et al., 2016). Following imputation, data (in variant call format [VCF] files) was restricted to variants with MAF ≥ 0.001 and imputation R^2 ≥ 0.3. Post-imputation VCFs were updated to include sex information and rsIDs, which were collected from the Single Nucleotide Polymorphism Database, build 153.

GLAD & NBR merge

Data from GLAD and COPING NBR were merged post-imputation using bcftools, and converted to PLINK2 pfile format, retaining genotype dosage information (Chang *et al.*, 2015). Only bi-allelic

SNPs were retained in the resulting merged pfiles. Post-merge, the data was filtered with a MAF threshold of 0.01 and a variant missingness of 0.02. Duplicate samples, related individuals with pihat > 0.1875, and samples with mismatched sex were also excluded.

Post-hoc phenotypic correlation between manic symptoms and PTSD symptoms

Based on the finding that the lifetime manic symptom sum score, which represents the total number of lifetime manic symptoms that a participant reported, was most genetically correlated with PTSD (**figure 3**), we calculated the phenotypic correlation between the lifetime manic symptom sum score and current PTSD symptoms in affected participants. PTSD symptoms were based on answers to the 6-item PTSD checklist (PCL-6), scored 6-30. As per the MDQ, GLAD participants answered the PCL-6 in the GLAD sign-up questionnaire and COPING participants answered the PCL-6 in the COPING baseline questionnaire. Participants were asked six questions relating to their experience of PTSD symptoms:

1. *Repeated, disturbing memories, thoughts, or images of a stressful experience?*
2. *Feeling very upset when something reminded you of a stressful experience?*
3. *Avoiding activities or situations because they reminded you of a stressful situation?*
4. *Feeling distant or cut off from other people?*
5. *Feeling irritable or having angry outbursts?*
6. *Difficulty concentrating?*

Participants could answer with *"Not at all"*, *"A little bit"*, *"Moderately"*, *"Quite a bit"*, and *"Extremely"* based on how they were feeling over the past month. These were coded numerically 1-6. Answers were then summed to create a sum score ranging 6-30 with lower scores representing lower levels of PTSD symptoms and higher scores representing higher levels of PTSD symptoms. Only participants with complete data on all six questions were included in the sum score. We used R to calculate Pearson's correlation between the lifetime manic symptom sum score and the current PTSD symptom sum score.

<u>Supplementary results</u>

Factor analyses

Concurrent manic symptoms in participants affected by MDD and/or anxiety

As mentioned in the main paper, the item "concurrent more active" was removed from the factor analysis due to having a correlation of 0.87 with "concurrent more energy" ("concurrent more active" was removed due to lower endorsement rate of the two items) (**table 3.1**). This left 12 concurrent MDQ items remaining. In order to avoid including participants who (now) only endorsed one out of the 12 concurrent MDQ items, we removed 82 participants. These were participants who initially had a concurrent manic symptom sum score of 2 and now only had a score of 1 (because "concurrent more active" had been removed from the score). This left a final N of 29,899. None of the remaining items were correlated <0.3 so no further items were removed. None of the items had near zero variance which showed that the data was suitable to be randomly split into an EFA and CFA sample. After the sample was split, the determinant of the matrix of the EFA sample was 0.003 which suggested that multicollinearity was not a problem and therefore no further items were removed (**table S3.2**). The ordinal alpha statistic, Kaiser-Meier-Olkin (KMO) statistic and Bartlett's test of sphericity p-value also demonstrated that the data were suitable for factor analysis (**table S3.2**).

After performing EFA in 70% of the sample (N=20,929), the one factor model showed that concurrent irritability, racing thoughts, and concentration problems did not load onto the factor. A decision was made to keep these items in the factor analysis because the correlation matrix suggested they would form their own factor and to keep comparability with the lifetime MDQ items factor analysis. The three factor solution was selected as the final model because it showed the best fit statistics while retaining at least three items per factor (**figure 3.2; table S3.3**). All fit statistics demonstrated that the three factor model was a good fit for the data (**table S3.4a**). Factor one, two and three included six, three, and three items respectively. We named the three factors according to their loaded items: *energy/activity, cognitive, and impulsivity* (**figure 3.2; figure S3.6**). The *energy/activity* and *impulsivity factors* correlated with each other at 0.54. As expected, concurrent "irritability", "racing thoughts", and "concentration problems" formed their own factor. This factor (*concurrent cognitive*) did not correlate with the other two factors (**figure 3.2; figure S3.7; table S3.4b**) which reflected the absence of correlations between "concentration problems", "racing thoughts", and "irritability" and the other MDQ items (**figure S3.1**). It is important to note that "irritability" had a fairly weak loading onto the cognitive factor (**figure S3.6**). The model was confirmed in CFA on the remaining 30% of the sample (N=8,970) and showed good fit statistics

(**table S3.5**). The three factor model was then applied to the full sample (N=29,899) to provide overall model fit statistics (**table S3.5**).

Lifetime manic symptoms in participants affected by MDD and/or anxiety

As mentioned in the main paper, the item "more active" was removed from the factor analysis due to having a correlation of 0.90 with "more energy". None of the remaining items were correlated <0.3 so no further items were removed. None of the items had near zero variance which showed that the data was suitable to be randomly split into an EFA and CFA sample. After splitting the sample, the determinant of the matrix in the EFA sample was 0.0002 which suggested that multicollinearity was not a problem and therefore no further items were removed (**table S1**). The ordinal alpha statistic, KMO statistic and Bartlett's test of sphericity p-value also demonstrated that the data were suitable for factor analysis (**table S3.1**).

After performing EFA in 70% of the sample (N=33,450), the three factor solution was selected as the final model because it showed the best fit statistics while retaining at least three items per factor (**figure 3.2; table S3.7**). Factor one, two and three included six, three, and three items respectively. We named the three factors according to the loaded items: *energy/activity, cognitive and impulsivity*. These factors perfectly mirrored those identified in the concurrent MDQ items factor analysis, with the exception that all factors correlated with each other ($r \geq 0.55$) (**table S3.7b**). CFA on the remaining 30% of the sample (N=14,337) confirmed that the three factor model fit the data very well (**table S3.8**). CFA was then applied to the full sample (N=47,787) to provide overall model fit statistics (**table S3.8**).

Lifetime manic symptoms in participants unaffected by MDD and/or anxiety

As mentioned in the main paper, the item "more active" was removed from the factor analysis due to having a correlation of 0.90 with "more energy". None of the remaining items were correlated <0.3 so no further items were removed. Five of the items had near zero variance (the items "hyperactivity", "more talkative", "more sociable", "risky behaviour", and "reckless spending") which showed that the data may not have been suitable to be randomly split into an EFA and CFA sample. Therefore, after splitting the sample, the frequencies of these five items were compared between the EFA and CFA sample. We were satisfied that the frequencies were comparable and

we continued with the factor analysis. The determinant of the matrix in the EFA sample was 0.0004 which suggested that multicollinearity was not a problem and therefore no further items were removed. The ordinal alpha statistic, KMO statistic and Bartlett's test of sphericity p-value also demonstrated that the data were suitable for factor analysis (**table S3.2**).

After performing EFA in 70% of the sample (N=4,283), none of the factor solutions showed adequate fit (**table S3.10**). Therefore, no solution was carried forward to CFA.

Post-hoc phenotypic correlation between manic symptoms and posttraumatic stress disorder symptoms

Based on the unexpected finding that the lifetime manic symptom sum score was most strongly genetically correlated with PTSD (**figure 3.3; table S3.15**), we calculated the phenotypic correlation between lifetime manic symptom sum score and current PTSD symptoms in affected participants. PTSD symptoms were based on answers to the 6-item PTSD checklist (PCL-6) and were scored 6-30. The phenotypic correlation was far lower than the genetic correlation (r_{ph}=0.41, $p<2x10^{-16}$). See supplementary methods for more details of the PCL-6 and **figure S3.27** for a scatter plot of the two measures.

Supplementary discussion

Genetics of manic symptom subgroups

We hypothesised that the genetics of the symptom subgroups identified in the factor analyses would differentially genetically correlate with other psychiatric and behavioural traits. We did not make any *a priori* assumptions about which subgroups would correlate with which traits. The concurrent symptom subgroups had no significant genetic correlations with other traits but *concurrent energy/activity* and *concurrent impulsivity* significantly genetically correlated with each other. Neither of these subgroups genetically correlated with *concurrent cognitive* (**table S3.12**). This mirrors the pattern of phenotypic correlations between the factors found in our study (**figure 3.2**). The *concurrent cognitive* symptom subgroup was also not genetically correlated with the overall concurrent manic symptom sum score. This, combined with the phenotypic results, suggest that the MDQ items within this subgroup do not capture the same trait as the other MDQ

items when measured as concurrent. This suggests that the concurrent MDQ lacks internal consistency and is therefore unreliable.

This lack of internal consistency may explain why the concurrent manic symptom phenotypes were not significantly genetically correlated with any of the external traits in our study (**table S3.14**). Compared to the lifetime MDQ items, the GWASs of the concurrent MDQ items already had reduced statistical power (smaller N [11,568 vs. 19,859] and smaller variance [range 2-12 items vs. 0-12 items]). This was likely exacerbated by the three concurrent *cognitive symptoms* being included in the composite score despite, as learned from our findings, operationalising a fundamentally different trait to the other nine items. This may also explain the non-significant heritability of the sum score.

The results with the lifetime manic symptom subgroups, where we *did* find significant genetic correlations with other traits, confirms that our hypothesis (that manic symptoms would show significant positive genetic correlations with bipolar disorder) was not supported. The symptom subgroups showed broadly similar r_g estimates with the traits to each other as well as the overall sum score and none differed from each other significantly. This is reflected by the finding that the genetic correlations between the lifetime manic symptom subgroups all hovered around 1 (**table S3.13**). These estimates far exceed the phenotypic correlations found in the factor analysis (**figure S3.6**).

Supplementary references

Bartlett, M.S. (1950) 'Tests of significance in factor analysis', *British journal of psychology* , 3, pp. 77–85.

Chang, C.C. *et al.* (2015) 'Second-generation PLINK: rising to the challenge of larger and richer datasets', *GigaScience*, 4, p. 7.

Coleman, J.R.I. *et al.* (2016) 'Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray', *Briefings in functional genomics*, 15(4), pp. 298–304.

Davies, M.R. *et al.* (2022) 'Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts', *Journal of anxiety disorders*, 85, p. 102491.

Hirschfeld, R.M. *et al.* (2000) ‘Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire’, *The American journal of psychiatry*, 157(11), pp. 1873–1875.

Hirschfeld, R.M.A. *et al.* (2003) ‘Validity of the mood disorder questionnaire: a general population study’, *The American journal of psychiatry*, 160(1), pp. 178–180.

Horn, J.L. (1965) ‘A RATIONALE AND TEST FOR THE NUMBER OF FACTORS IN FACTOR ANALYSIS’, *Psychometrika*, 30, pp. 179–185.

Kaiser, H.F. (1974) ‘An index of factorial simplicity’, *Psychometrika*, 39(1), pp. 31–36.

Loh, P.-R. *et al.* (2016) ‘Reference-based phasing using the Haplotype Reference Consortium panel’, *Nature genetics*, 48(11), pp. 1443–1448.

Mair, Patrick 2018. Modern Psychometrics with R, *Springer*

Revelle, W. and Rocklin, T. (1979) ‘Very Simple Structure: An Alternative Procedure For Estimating The Optimal Number Of Interpretable Factors’, *Multivariate behavioral research*, 14(4), pp. 403–414.

Taliun, D. *et al.* (2021) ‘Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program’, *Nature*, 590(7845), pp. 290–299.

Velicer, W.F. (1976) ‘Determining the number of components from the matrix of partial correlations’, *Psychometrika*, 41(3), pp. 321–327.

Supplementary figures

Figure S3.1. Correlations between concurrent Mood Disorder Questionnaire (MDQ) items in individuals affected by major depressive disorder (MDD) and/or an anxiety disorder.
Tetrachoric correlation matrix of concurrent manic symptoms from answers to the Mood Disorder Questionnaire (MDQ) in participants affected by major depressive disorder (MDD) and/or an anxiety disorder (N=31,427). Participants can answer with "Yes" or "No" which were coded as 1 or 0 respectively. Correlation matrix was computed using the hetcor R package. Participants can answer with "Yes" or "No" which were coded as 1 or 0 respectively. Correlations are therefore tetrachoric (which is a special case of polychoric). "More active" was removed from analysis following inspection of the correlation matrix due to its correlation of 0.87 with "more energy". Then, 83 participants were excluded from further analyses because they previously endorsed two concurrent items and now, after the removal of the item "more active", only endorsed one concurrent item. This left a final N of 29,889.

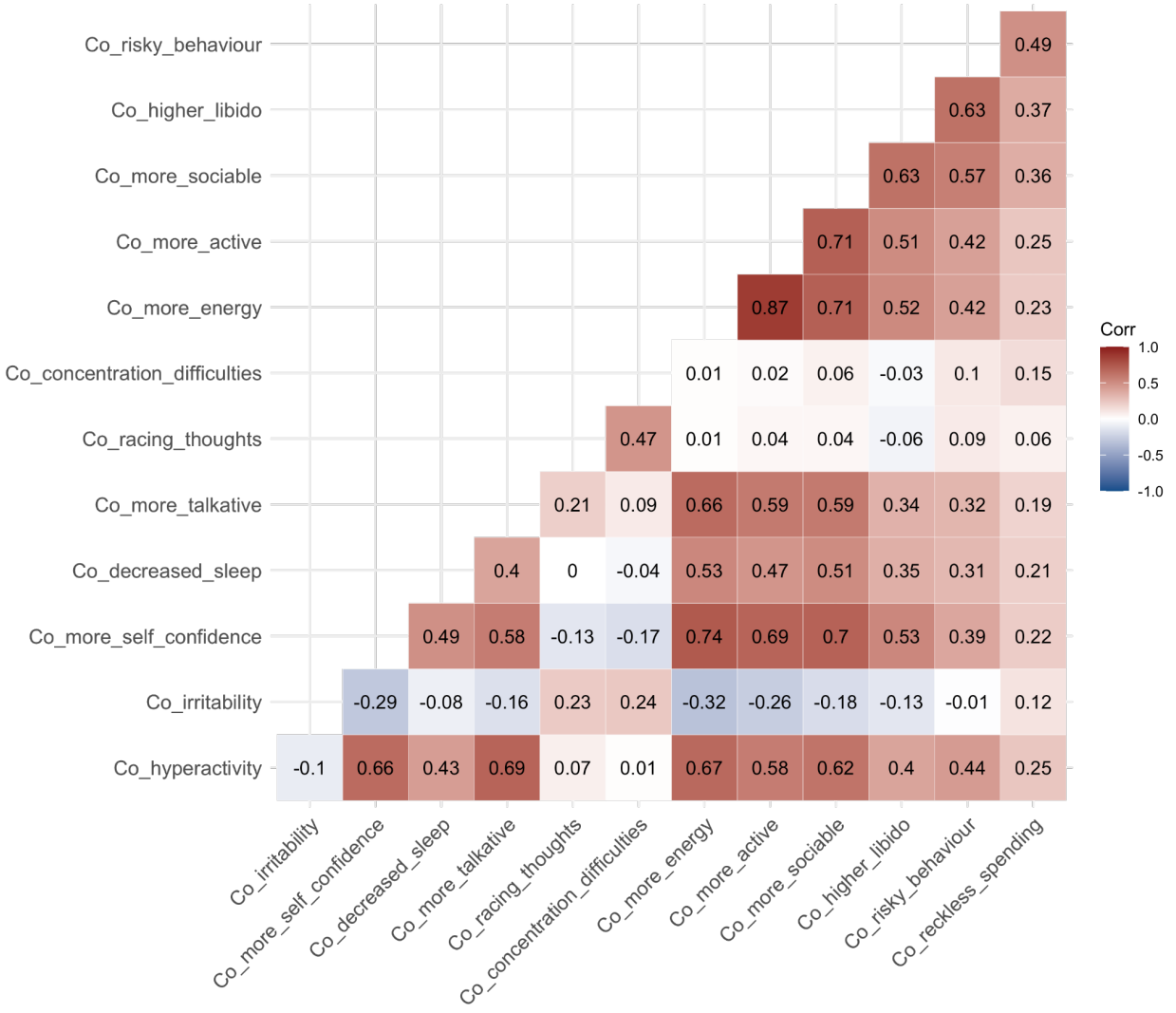


Figure S3.2. Correlations between lifetime Mood Disorder Questionnaire (MDQ) items in individuals affected by major depressive disorder (MDD) and/or an anxiety disorder.
Tetrachoric correlation matrix of lifetime manic symptoms from answers to the Mood Disorder Questionnaire (MDQ) in participants affected by major depressive disorder (MDD) and/or an anxiety disorder (N=47,787). Participants can answer with "Yes" or "No" which were coded as 1 or 0 respectively. Correlation matrix was computed using the hetcor R package. Participants can answer with "Yes" or "No" which were coded as 1 or 0 respectively. Correlations are therefore tetrachoric (which is a special case of polychoric). "More active" was removed from analysis following inspection of the correlation matrix due to its correlation of 0.90 with "more energy".

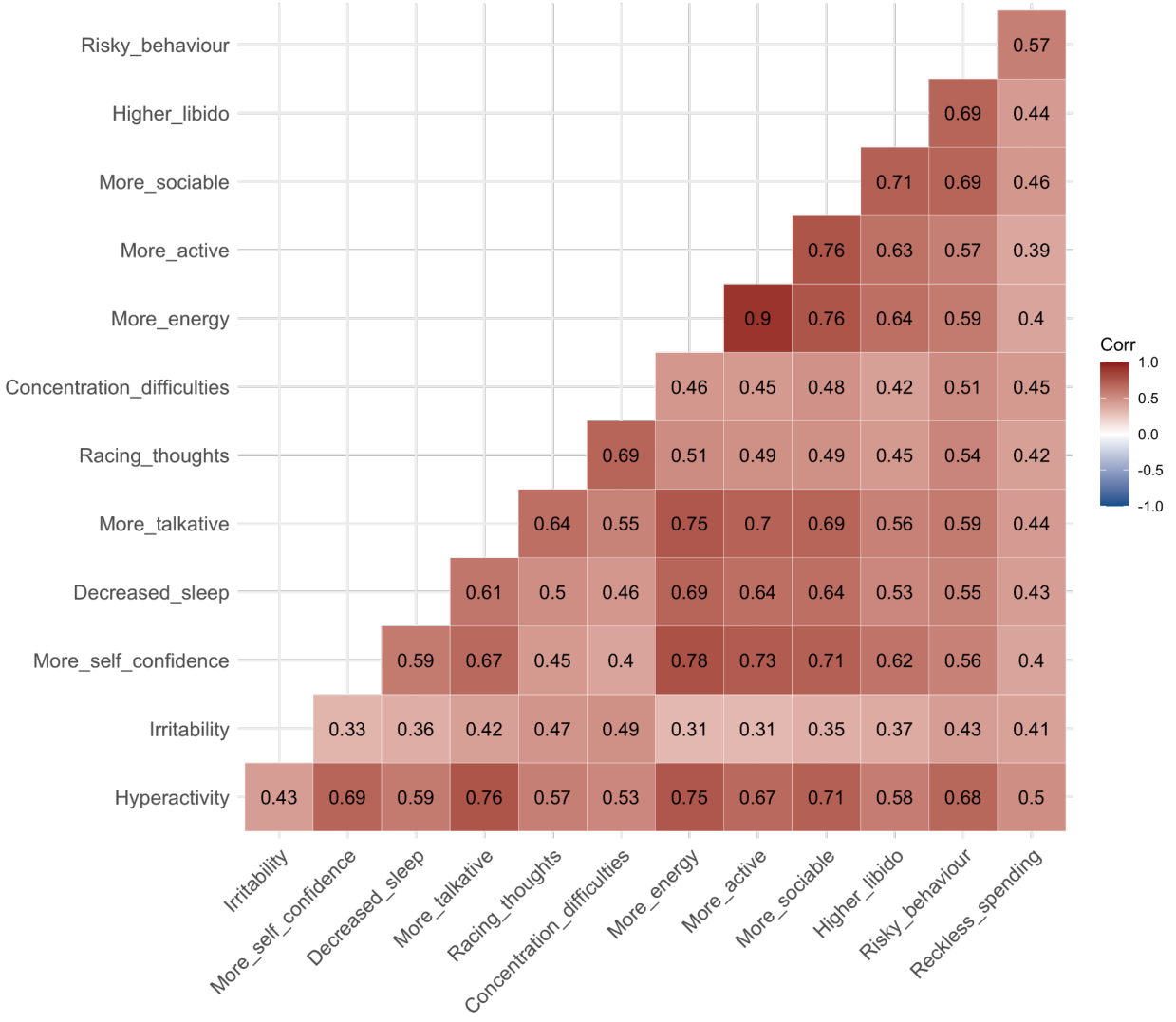


Figure S3.3. Correlations between lifetime Mood Disorder Questionnaire (MDQ) items in individuals unaffected by major depressive disorder (MDD) and/or an anxiety disorder.
Tetrachoric correlation matrix of lifetime manic symptoms from answers to the Mood Disorder Questionnaire (MDQ) in participants unaffected by major depressive disorder (MDD) and/or an anxiety disorder (N=6,119). Correlation matrix was computed using the hetcor R package. Participants can answer with "Yes" or "No" which were coded as 1 or 0 respectively. Correlations are therefore tetrachoric (which is a special case of polychoric). "More active" was removed from analysis following inspection of the correlation matrix due to its correlation of 0.90 with "more energy".

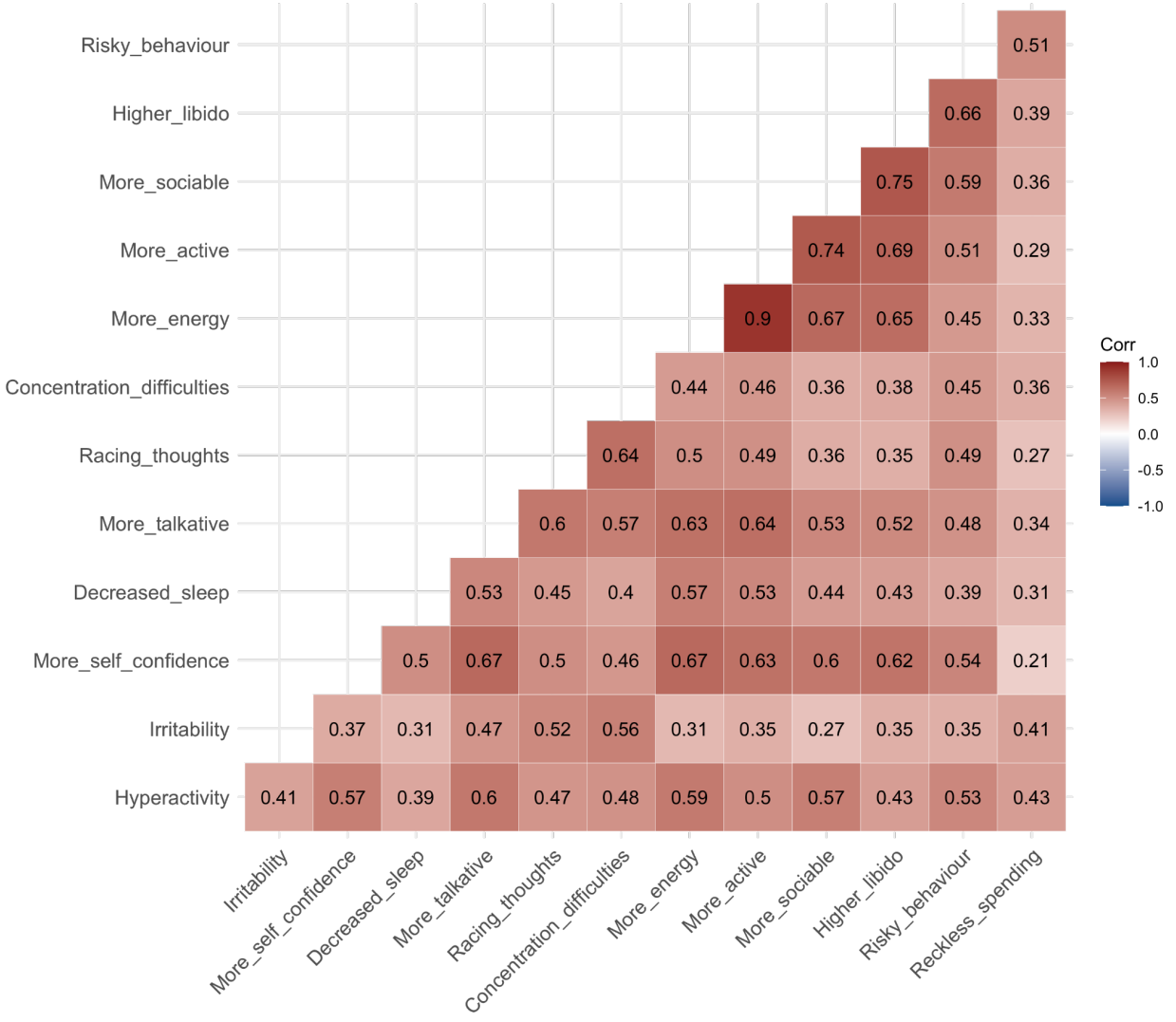


Figure S3.4. Exploratory factor analysis (EFA): one factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

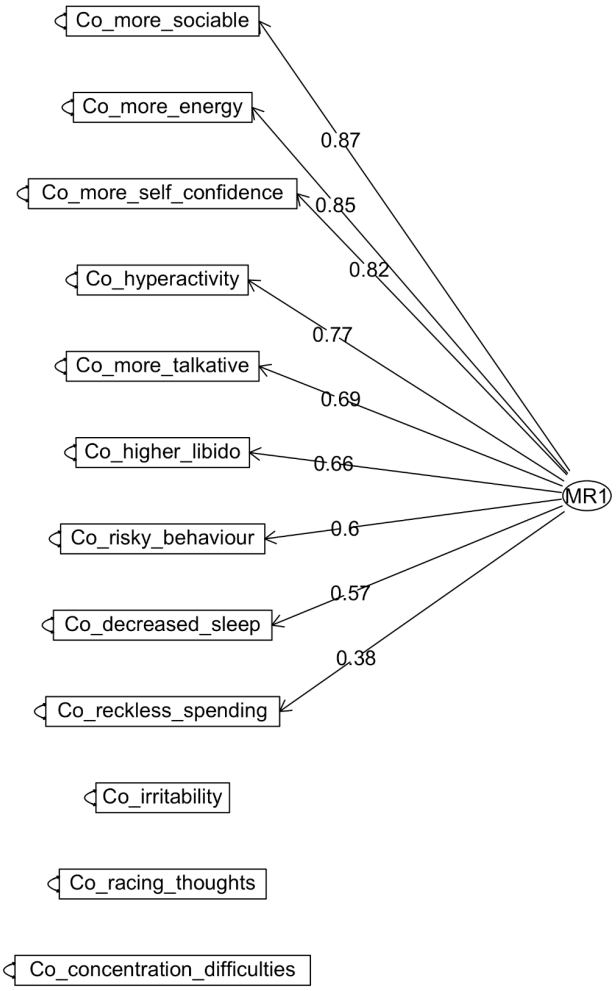


Figure S3.5. Exploratory factor analysis (EFA): two factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

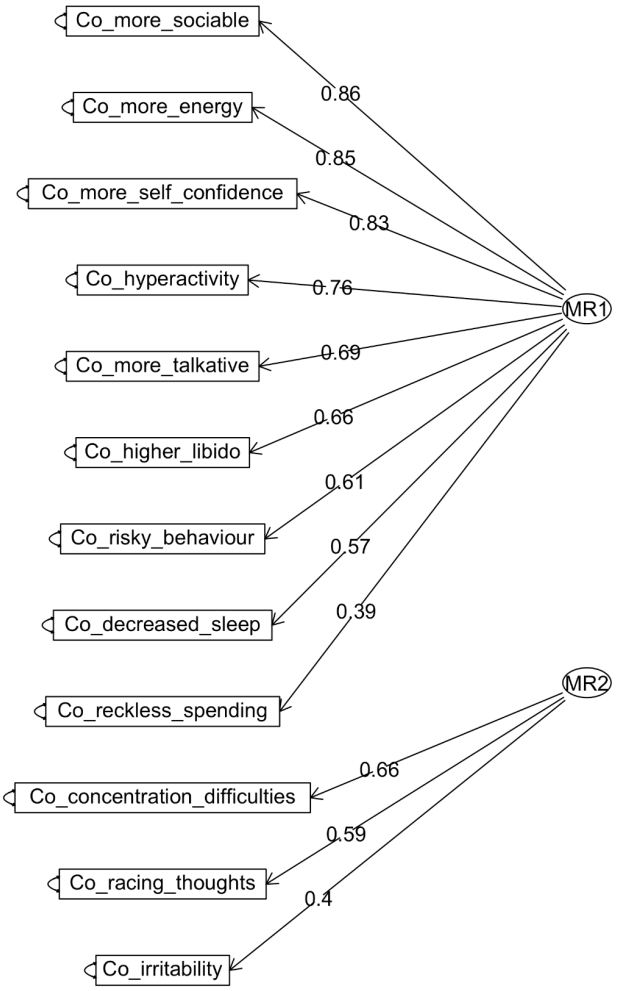


Figure S3.6. Exploratory factor analysis (EFA): three factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

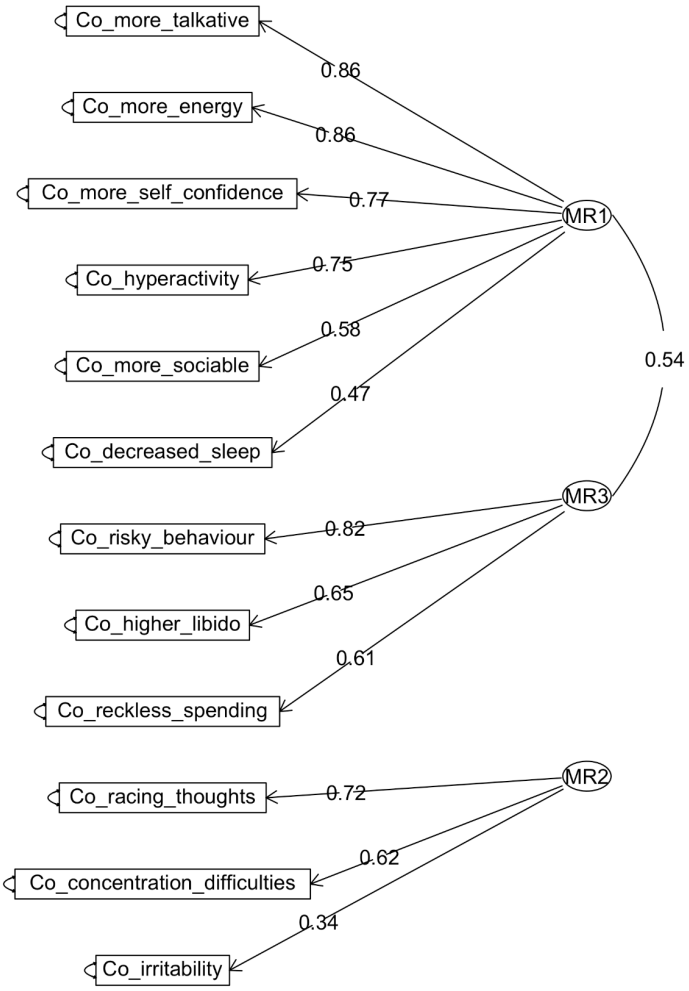


Figure S3.7. Exploratory factor analysis (EFA): four factor solution of 12 concurrent Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

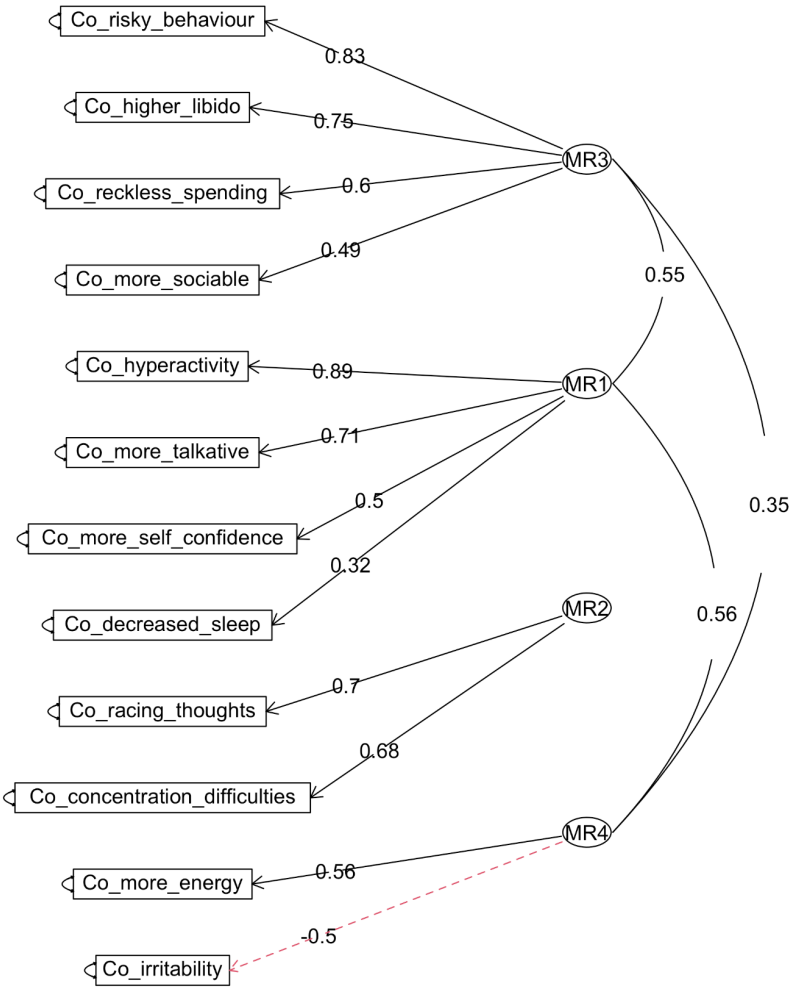


Figure S3.8. Exploratory factor analysis (EFA): one factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

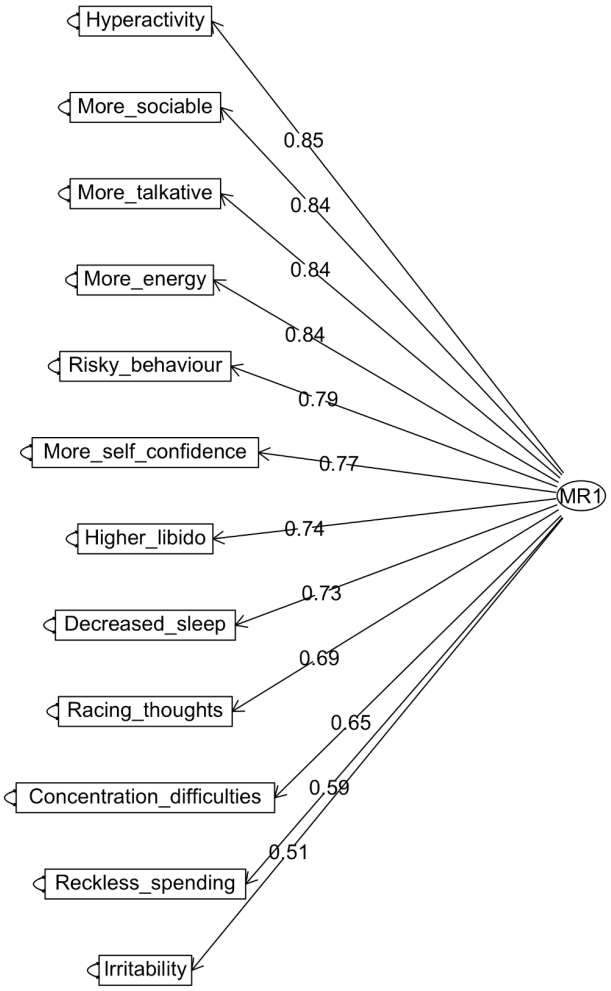


Figure S3.9. Exploratory factor analysis (EFA): two factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

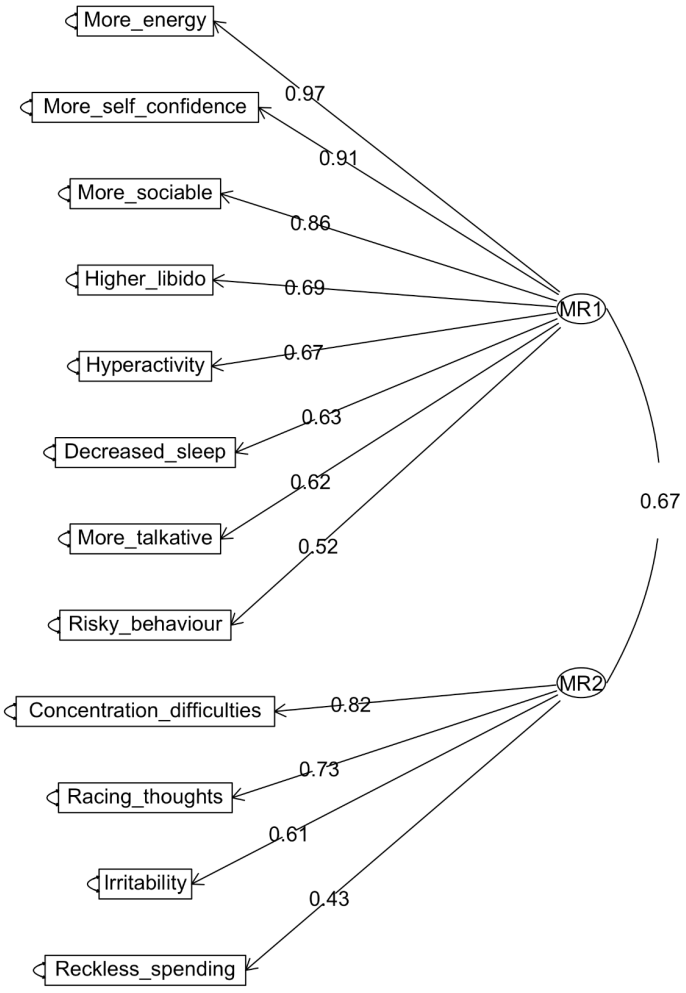


Figure S3.10. Exploratory factor analysis (EFA): three factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

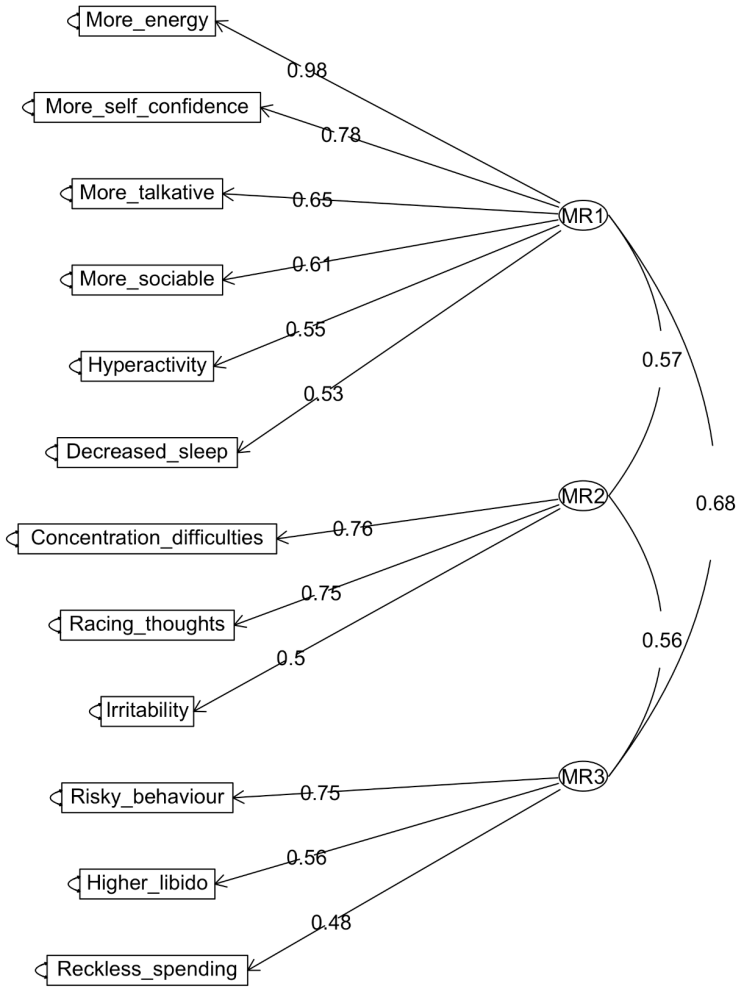


Figure S3.11. Exploratory factor analysis (EFA): four factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

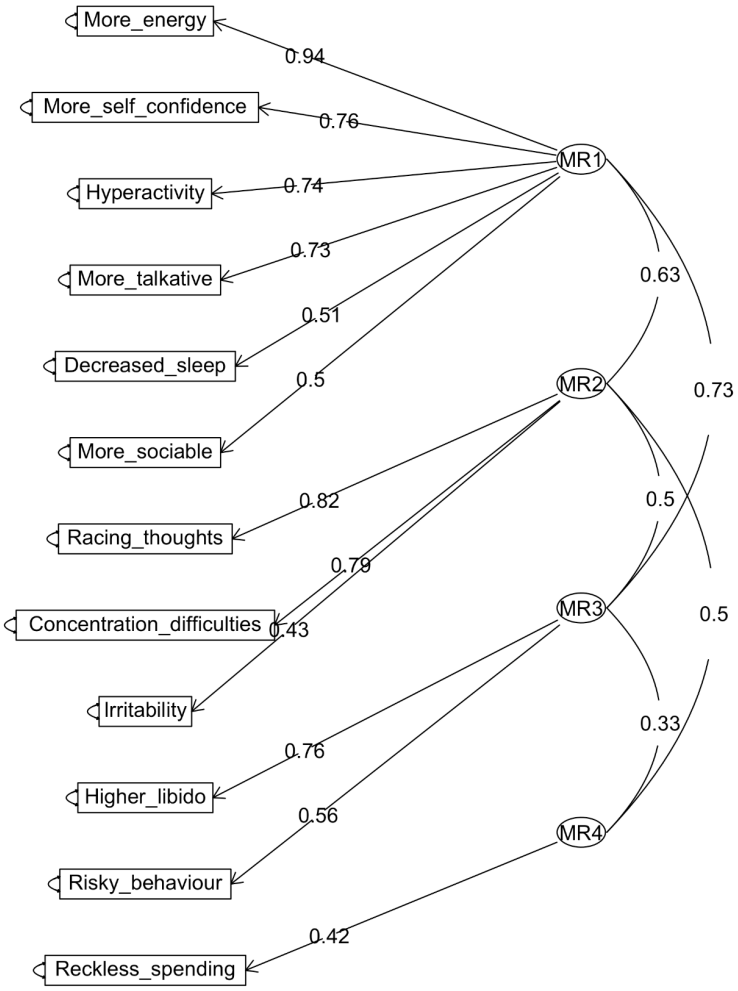


Figure S3.12. Exploratory factor analysis (EFA): one factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

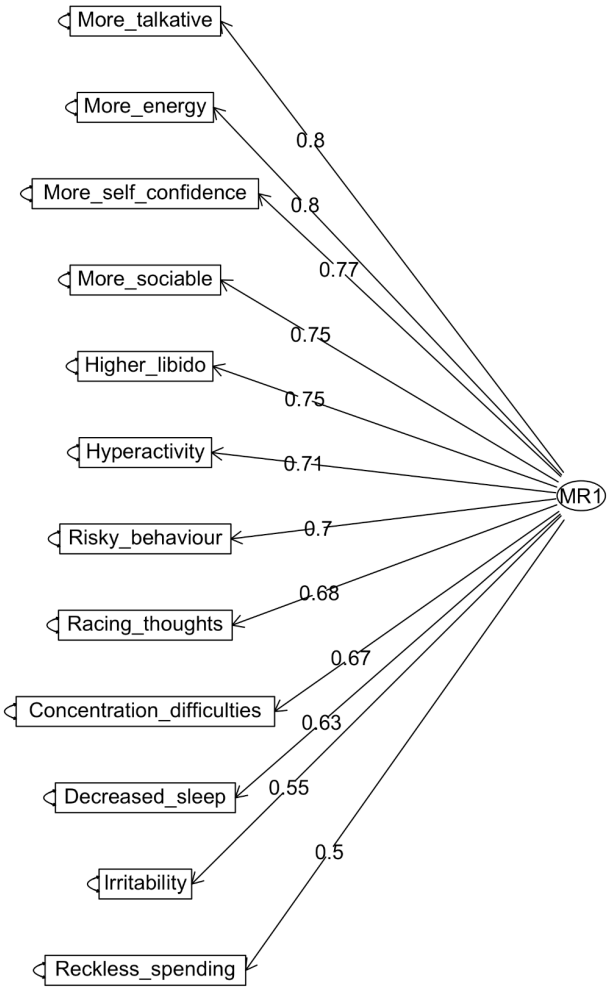


Figure S3.13. Exploratory factor analysis (EFA): two factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

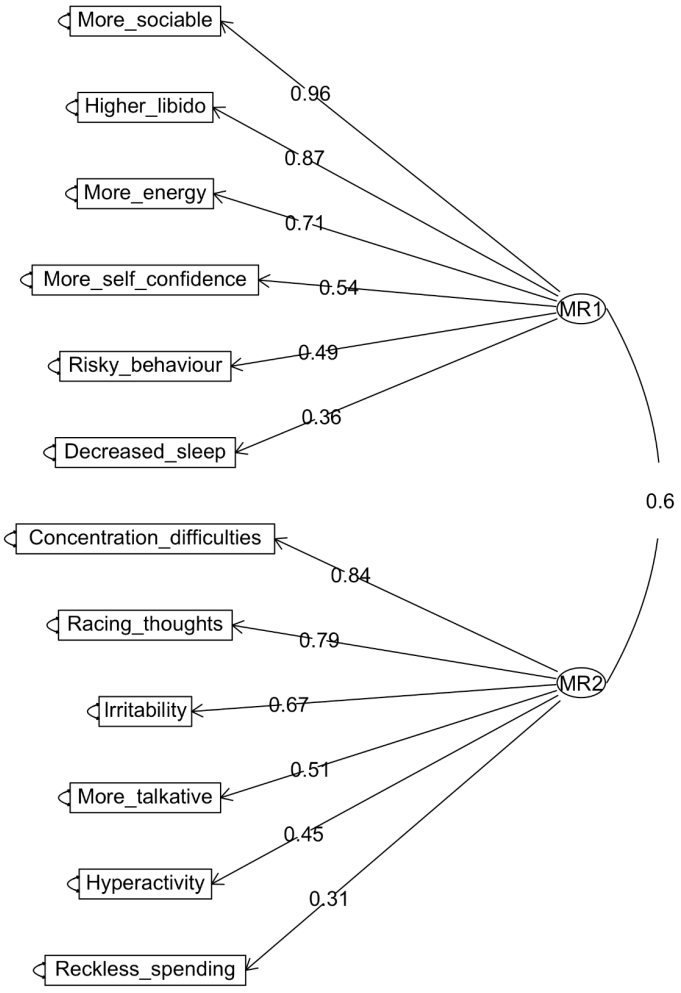


Figure S3.14. Exploratory factor analysis (EFA): three factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

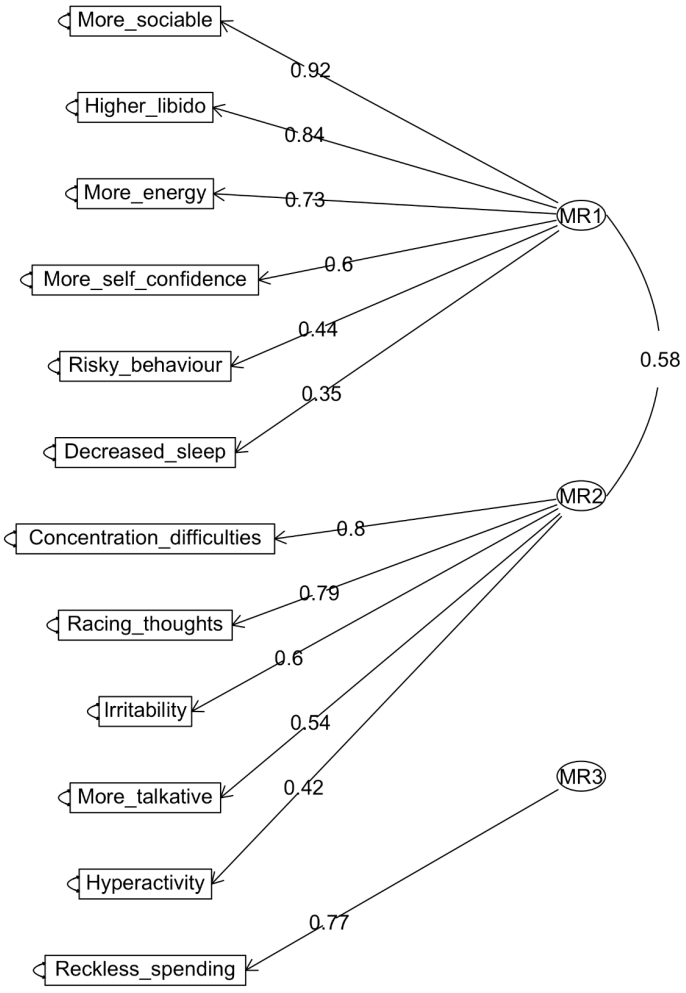


Figure S3.15. Exploratory factor analysis (EFA): four factor solution of 12 lifetime Mood Disorder Questionnaire (MDQ) items in unaffected participants.

EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was “minimum residuals”.

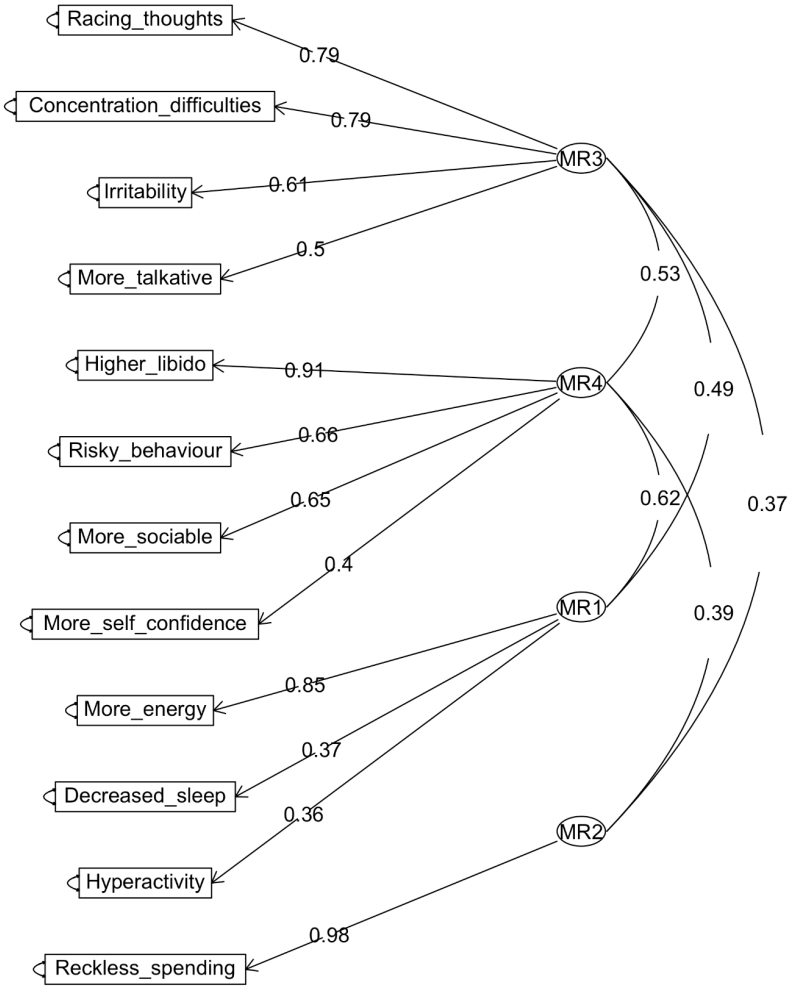


Figure S3.16. Raw factor scores from factor analysis of 12 concurrent manic symptoms measured by the Mood Disorder Questionnaire (MDQ) in affected participants.

The item “more active” was removed due to a correlation of 0.87 with “more energy”. Factor scores were computed, based on the best-fitting model identified in EFA, with the lavaan R package following confirmatory factor analysis (CFA).

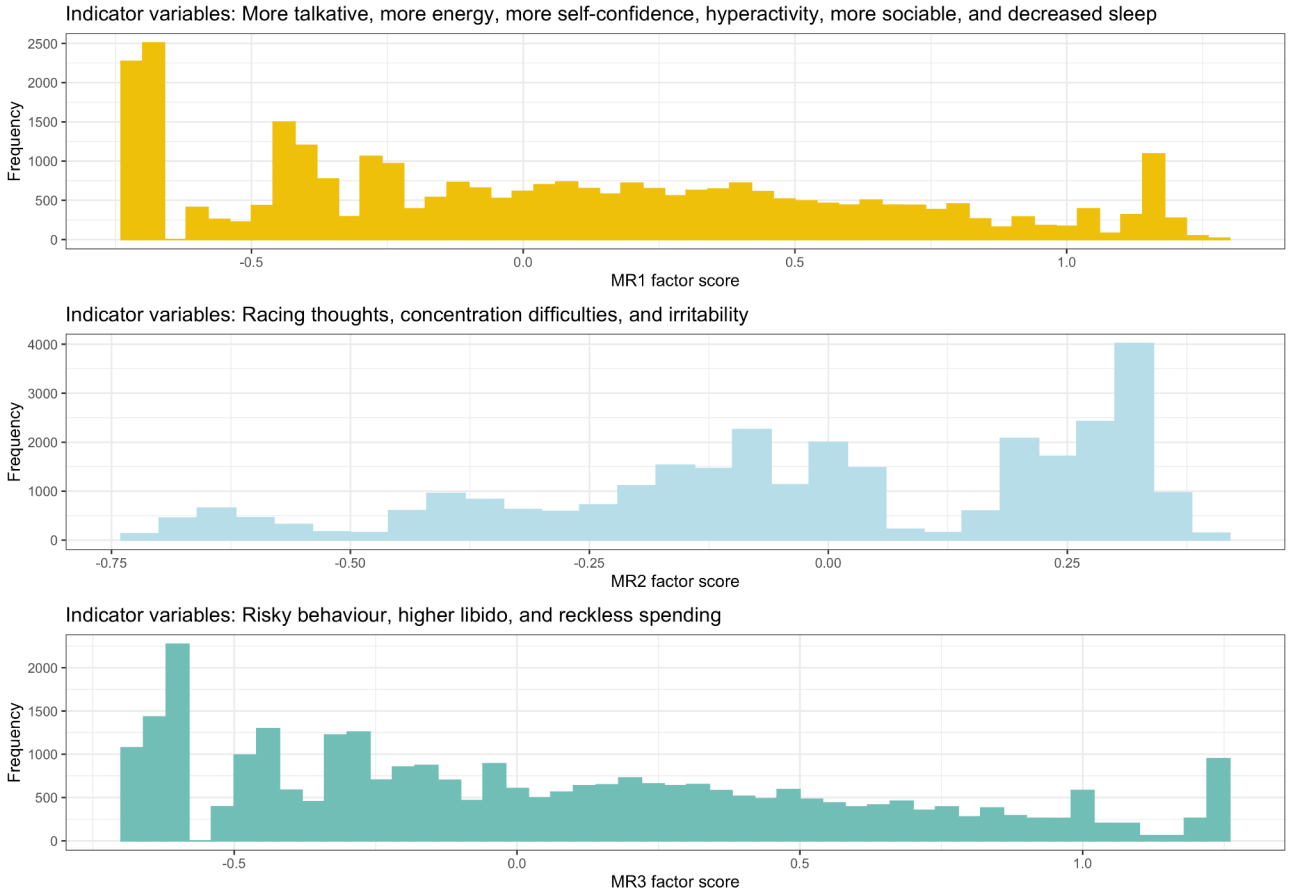


Figure S3.17. Raw factor scores from factor analysis of 12 lifetime manic symptoms measured by the Mood Disorder Questionnaire (MDQ) in affected participants.

The item “more active” was removed due to a correlation of 0.9 with “more energy”. Factor scores were computed, based on the best-fitting model identified in EFA, with the lavaan R package following confirmatory factor analysis (CFA).

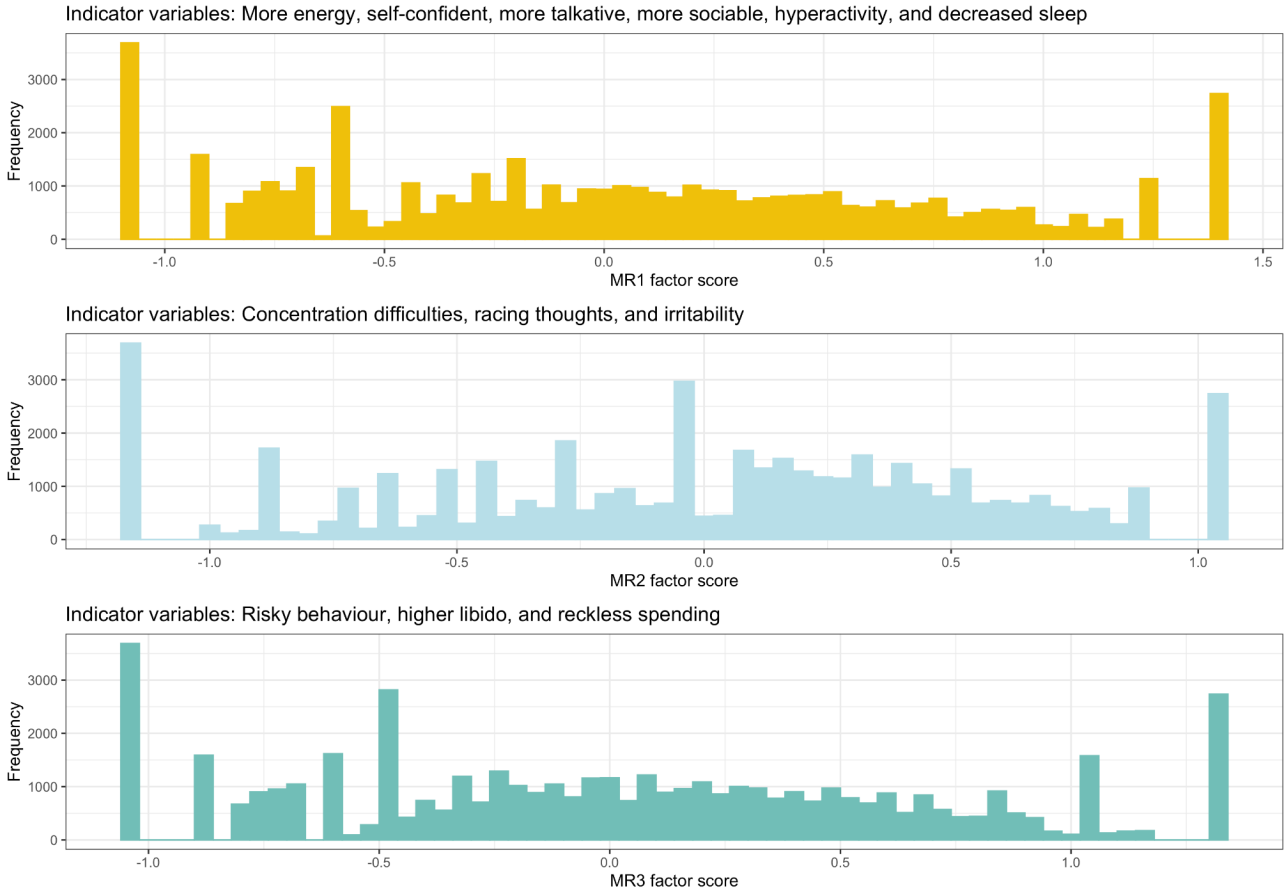
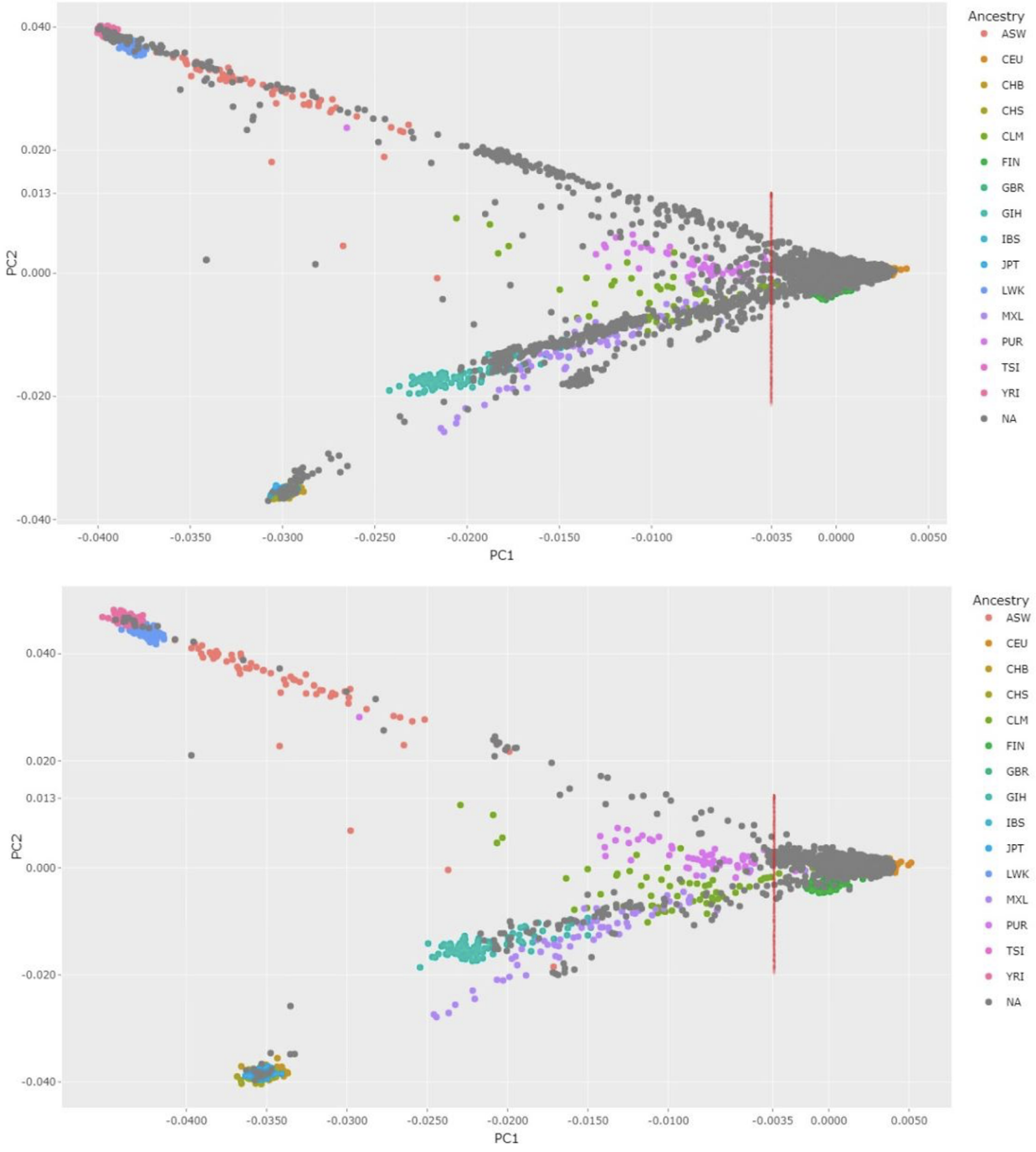


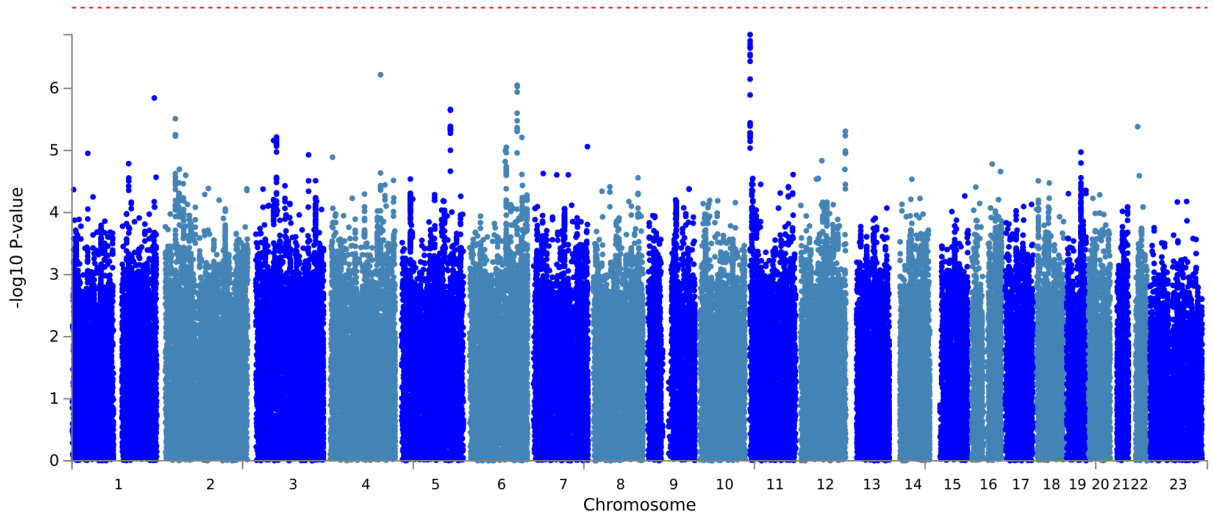
Figure S3.18. Principal component analysis (PCA) plots.
Top: Principal component analysis (PCA) plot of Genetic Links to Anxiety and Depression (GLAD) Study participants. Bottom: PCA plot of GLAD Study participants and COVID-19 Psychiatry and Neurological Genetics (COPING) Study participants.



Note: ASW (African Ancestry in SW USA), CEU (Europeans, from Utah), CHB (Northern Han Chinese from Beijing), CHS (Southern Han Chinese, from Shanghai), CLM (Colombian in Medellín, Colombia), FIN (Finnish in Finland), GBR (Western Europeans from Britain), IBS (Southern Europeans from Spain), JPT (Japanese in Tokyo, Japan), LWK (Luhya from Webuye,

Kenya), MXL (Mexican ancestry in Los Angeles, CA, USA), PUR (Puerto Rican in Puerto Rico), TSI (Southern Europeans from Tuscany in Italy), YRI (Yoruba in Ibadan, Nigeria), GIH (Gujarati Indians in Houston, Texas, USA), NA (GLADv2 or COPING NBRv1)

Figure S3.19. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the concurrent manic symptom sum score measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568).
GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



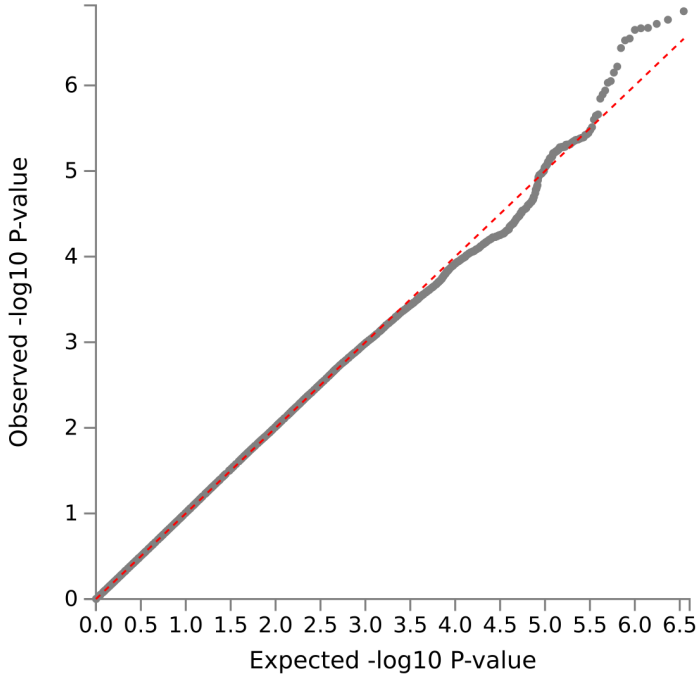
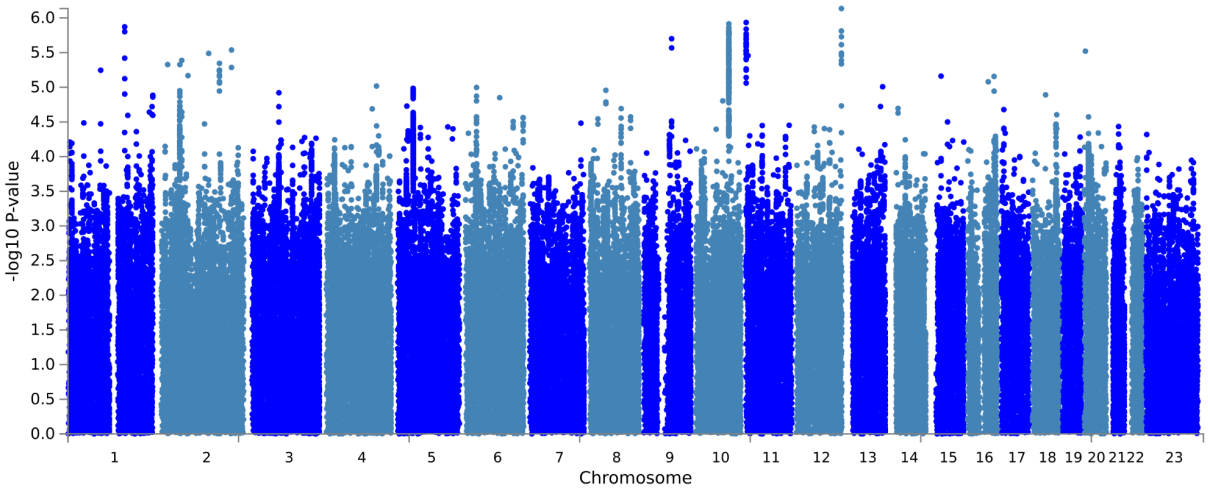


Figure S3.20. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of concurrent energy/activity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568).
GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



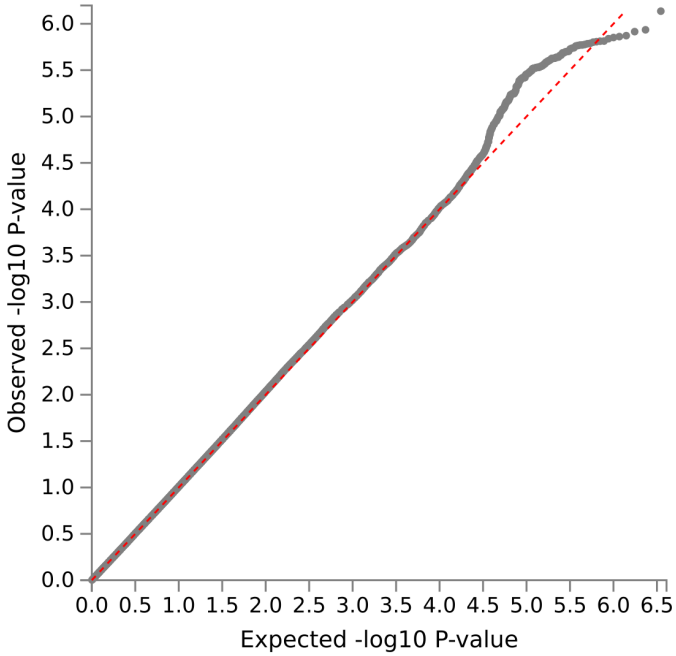
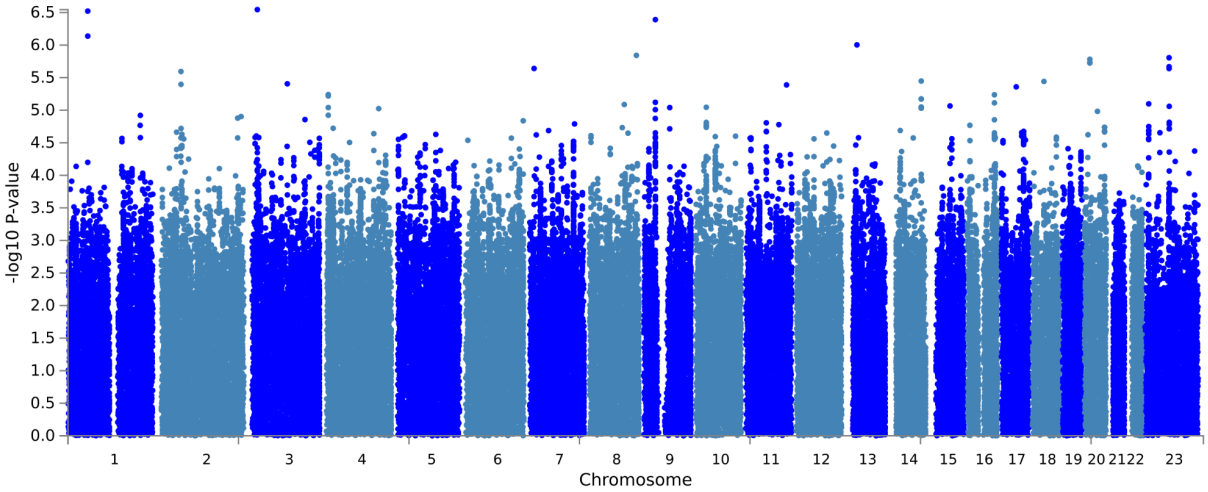


Figure S3.21. Quantile-quantile (QQ) plot and Manhattan plot of the genome-wide association study (GWAS) results of concurrent cognitive factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568). GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



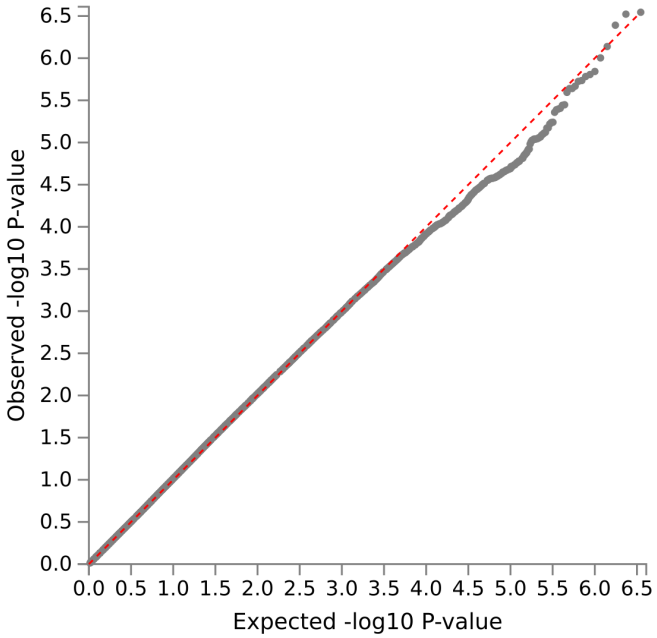
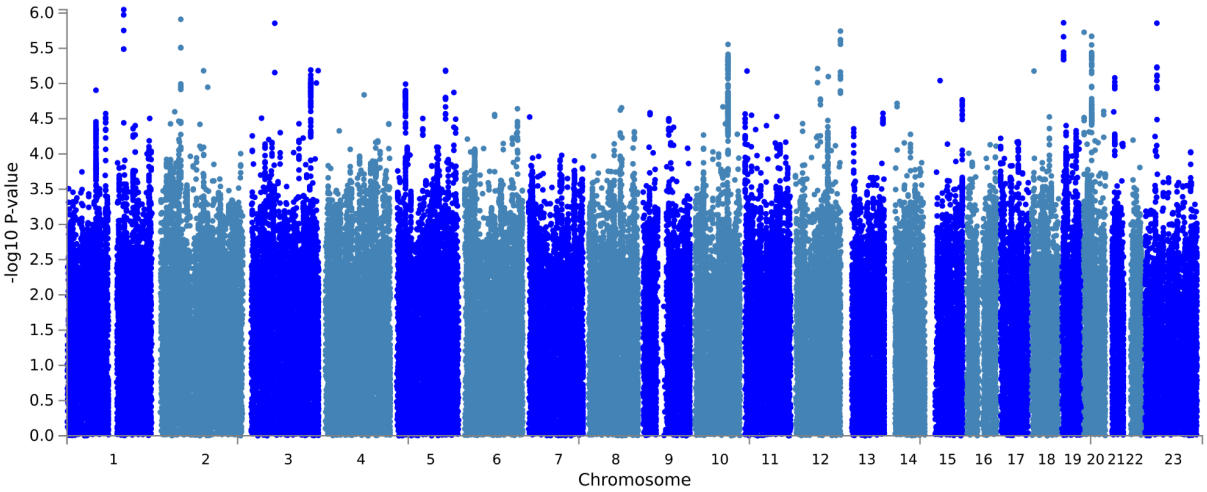


Figure S3.22. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the concurrent impulsivity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=11,568).
GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



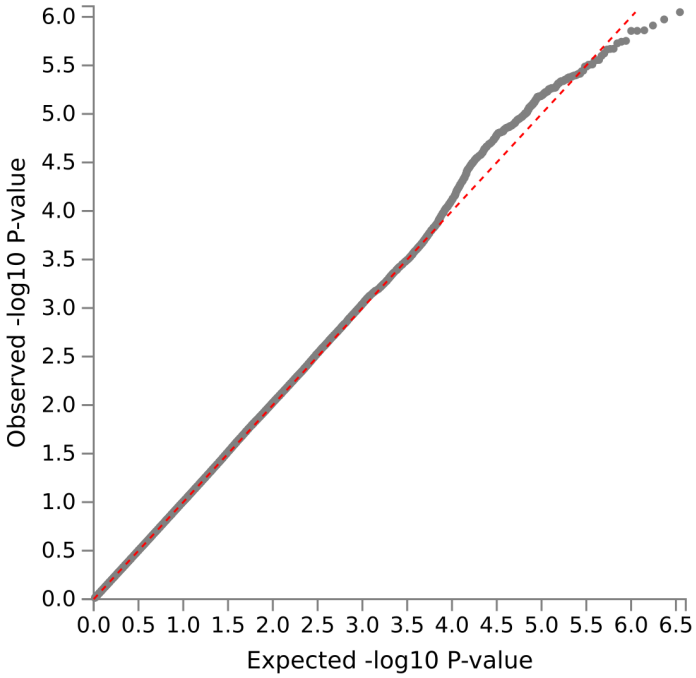
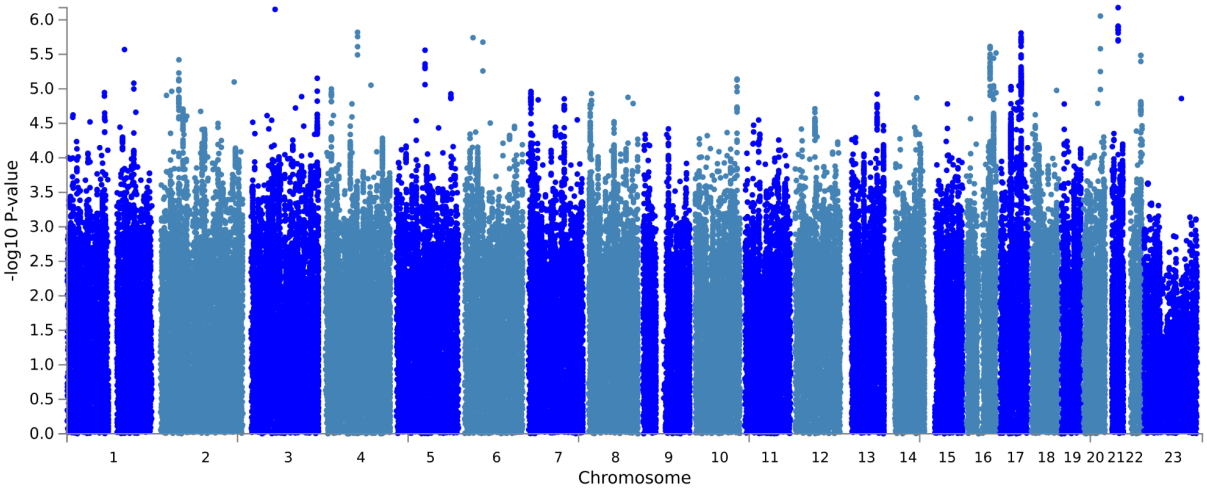


Figure S3.23. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime manic symptom sum score measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859). GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



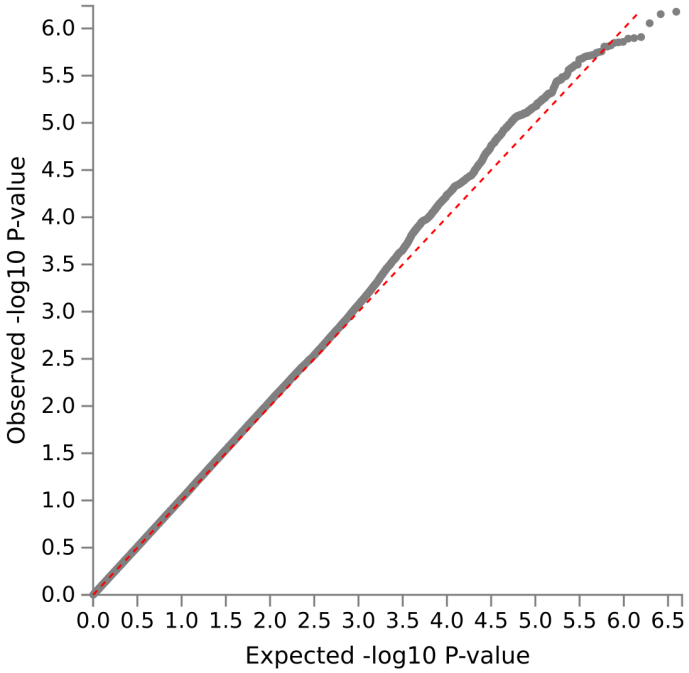
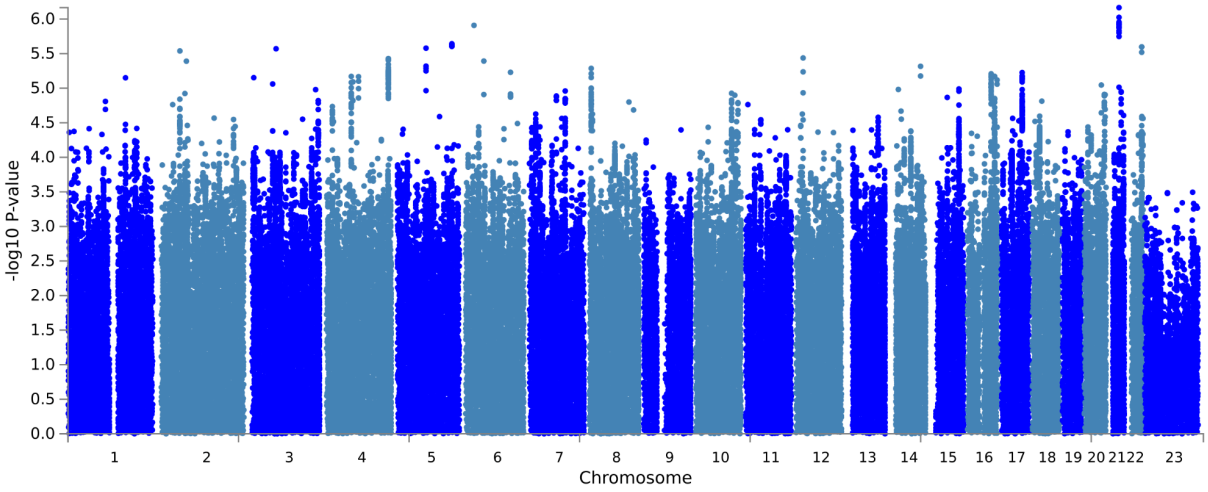


Figure S3.24. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime energy/activity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859).
GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



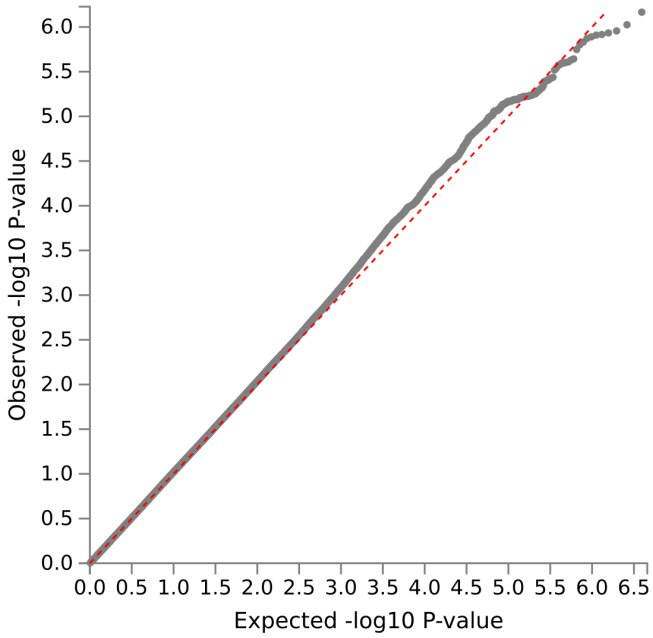
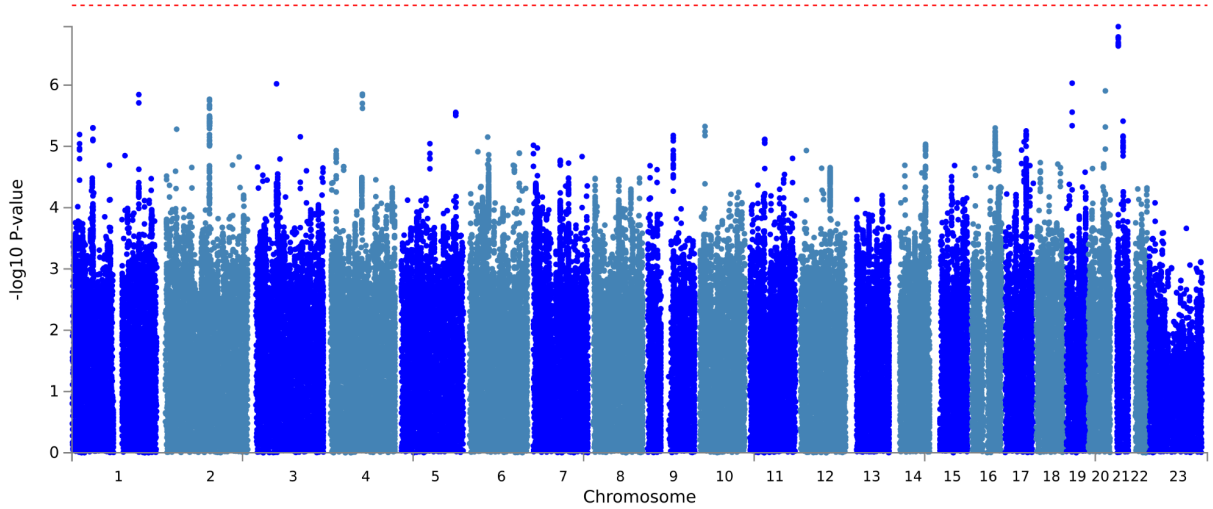


Figure S3.25. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime cognitive factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859).
GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



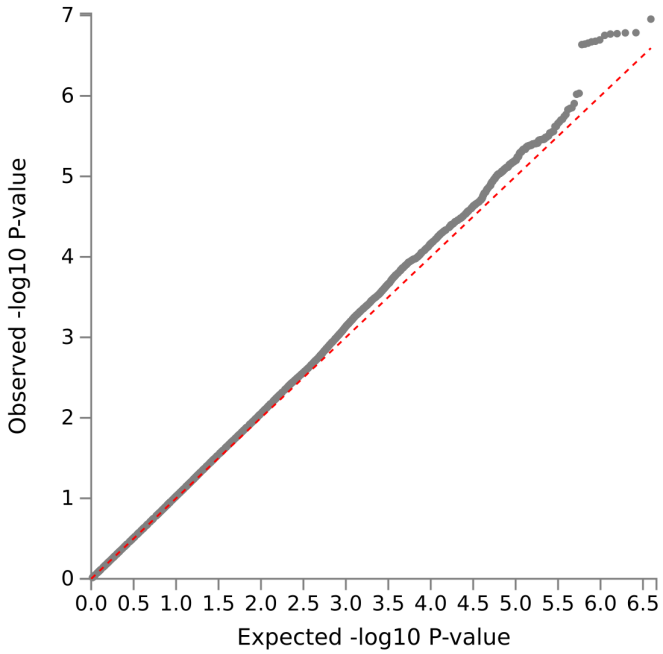
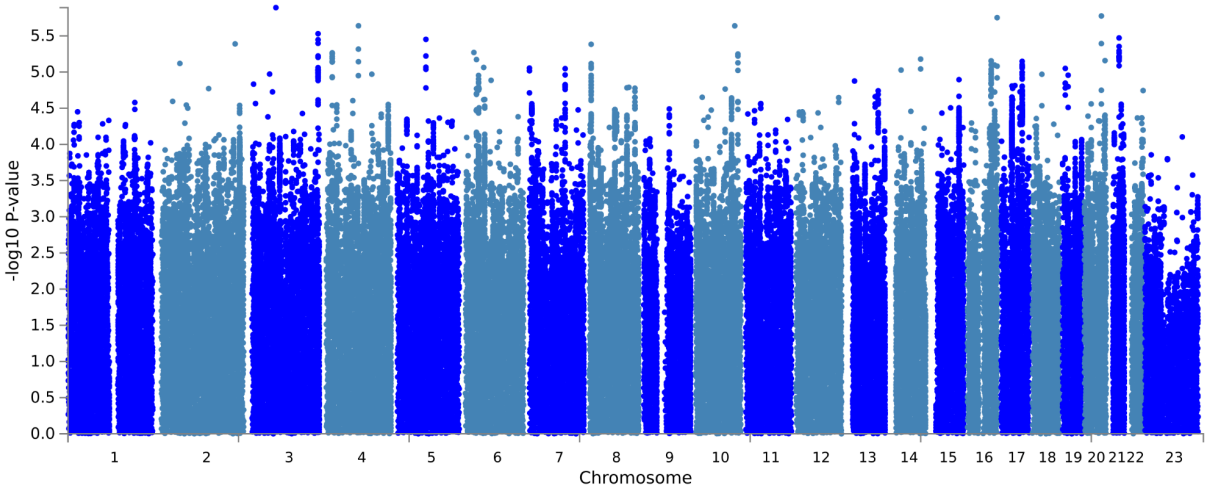


Figure S3.26. Quantile-quantile (QQ) plot and Manhattan plot of genome-wide association study (GWAS) results of the lifetime impulsivity factor measured by the Mood Disorder Questionnaire (MDQ) in affected participants of European ancestry (N=19,859).

GWAS was performed with REGENIE covarying for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA.



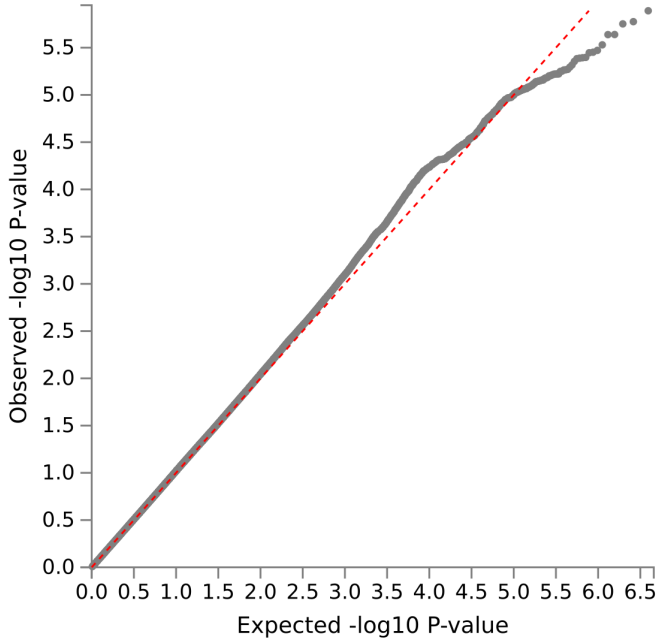


Figure S3.27. Scatter plot of affected participants' lifetime manic symptoms measured by the Mood Disorder Questionnaire (MDQ) and current posttraumatic stress disorder (PTSD) symptoms.

PTSD symptoms were measured by the six item PTSD Checklist (PCL-6). Lifetime manic symptoms were scored 0-12 and current PTSD symptoms were scored 6-30.

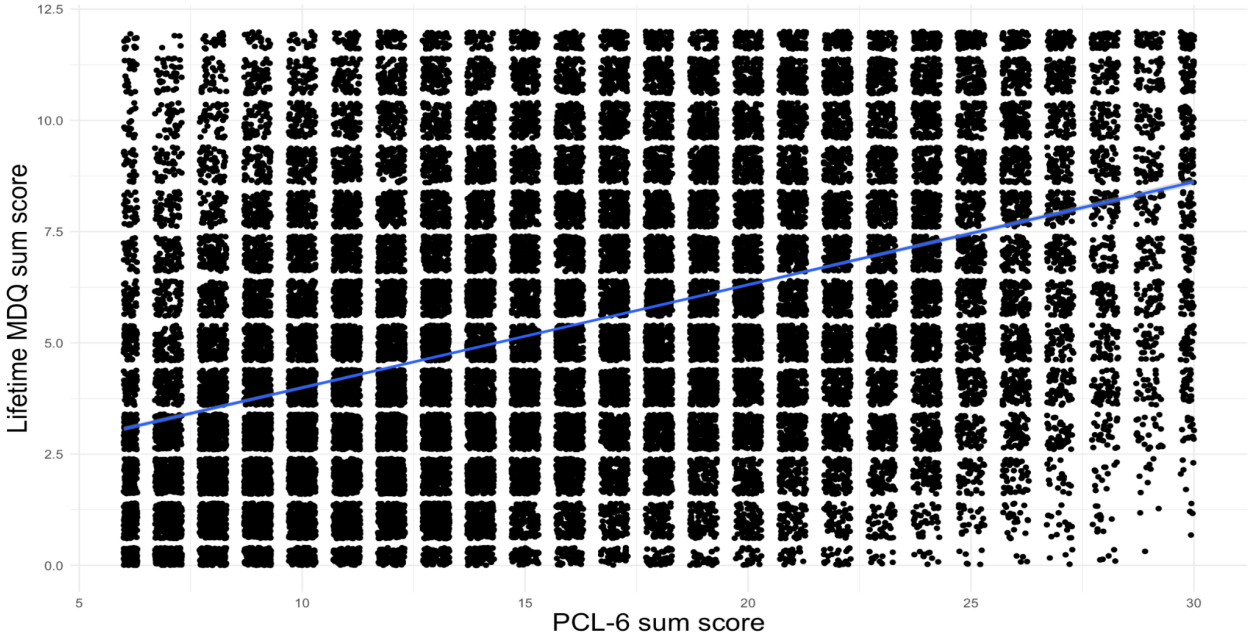
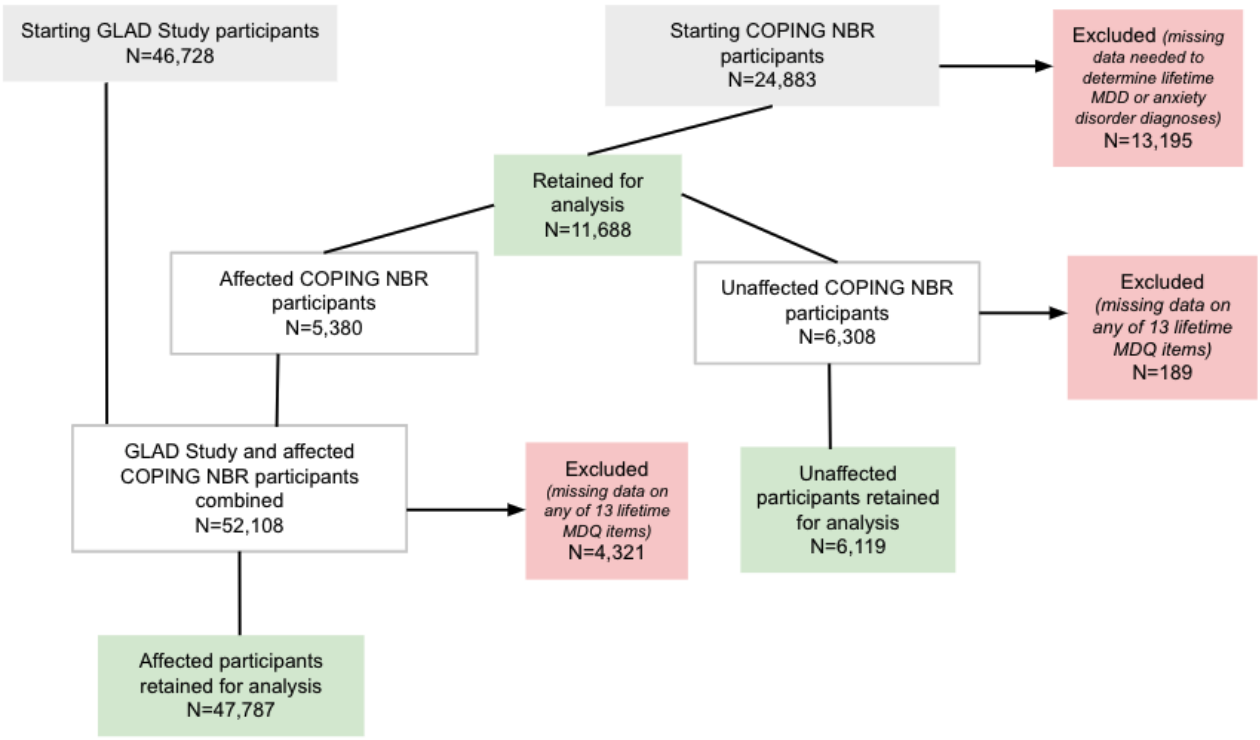


Figure S3.28. Flow-chart detailing how Genetic Links to Anxiety and Depression (GLAD) Study and COVID-19 Psychiatry and Neurological Genetics Study participants from the NIHR Bioresource (COPING NBR) were categorised as either "affected" or "unaffected' by major depressive disorder (MDD) and/or any anxiety disorder. MDQ=Mood Disorder Questionnaire.



Supplementary tables

Table S3.1. Descriptive statistics of the quantitative manic symptom phenotypes (sum scores) derived from the Mood Disorder Questionnaire (MDQ) in participants with and without a self-reported diagnosis of bipolar disorder (BD) by a professional.

Statistic	Concurrent manic symptoms (range 2-12)		Lifetime manic symptoms (range 0-12)	
	No BD diagnosis	***BD diagnosis***	***No BD diagnosis***	***BD diagnosis***
Minimum	2	2	0	0
Maximum	12	12	12	12
Mean	4.82113339	8.34659091	5.95056025	9.77390972
SD	2.43840507	2.85869768	2.96597055	2.59893268
Q1	3	6	3	9
Median	4	9	5	11
Q3	6	11	8	12
IQR	3	5	5	3
Skewness	1.0047145	-0.3971134	0.43410554	-1.421352
Kurtosis	0.37916083	-0.9299915	-0.9037836	1.67662905

SD=standard deviation
Q1=quartile 1
Q3=quartile 3
IQR=interquartile range

Table S3.2. Statistics from testing assumptions for factor analysis of the Mood Disorder Questionnaire (MDQ) items. "Affected" refers to participants affected by major depressive disorder (MDD) and/or an anxiety disorder.				
Statistic	**Alpha**	**KMO MSA**	**BTS p-value**	**EFA determinant**
Information	***>=0.7***	***>=0.8***	***<0.05***	***>=0.00001***
Concurrent manic symptoms in affected participants [12 MDQ items]	0.817082527	0.867492235	1.17E-36	0.002977618
Lifetime manic symptoms in affected participants [12 MDQ items]	0.934309472	0.939568417	5.85E-27	0.000184876
Lifetime manic symptoms in unaffected participants [12 MDQ items]	0.917241615	0.854195991	2.23E-27	0.000417324

Alpha = Ordinal alpha

KMO MSA = Kaiser-Meier-Olkin Measure of Sampling Adequacy

BTS p-value = Bartlett's Test of Sphericity p-value

EFA determinant = matrix determinant of MDQ items in EFA sample

Table S3.3. Item loadings for all factor solutions for the exploratory factor analysis (EFA) of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants. "concurrent more active" was previously removed due to having a correlation of 0.87 with "concurrent more energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

	Factor 1
Co_hyperactivity	0.76577364
Co_irritability	-0.215551363
Co_self_confident	0.817727099
Co_decreased_sleep	0.56864355
Co_more_talkative	0.689658203
Co_racing_thoughts	0.016184611
Co_concentration_difficulties	-0.007670273
Co_more_energy	0.854658142
Co_more_sociable	0.86526375
Co_interested_in_sex	0.658093279
Co_risky_behaviour	0.59878025
Co_spending_money	0.378508664

	Factor 1	Factor 2
Co_hyperactivity	0.763655784	0.059432698
Co_irritability	-0.220065062	0.397943677
Co_self_confident	0.833260526	-0.261960118
Co_decreased_sleep	0.566761575	-0.018941502
Co_more_talkative	0.689514939	0.112810134
Co_racing_thoughts	0.021112057	0.592485934
Co_concentration_difficulties	-0.004021691	0.656042199
Co_more_energy	0.853096615	-0.072788086
Co_more_sociable	0.863141366	0.077289698

Co_interested_in_sex	0.656927808	0.040564985	
Co_risky_behaviour	0.608306102	0.263580697	
Co_spending_money	0.386608791	0.291554874	

	Factor 1	Factor 2	Factor 3
Co_hyperactivity	0.74539424	0.079676456	0.079741139
Co_irritability	-0.337432605	0.341072103	0.17340554
Co_self_confident	0.768387426	-0.261071079	0.091585444
Co_decreased_sleep	0.472096269	-0.04067178	0.140755726
Co_more_talkative	0.860658721	0.221136578	-0.115428495
Co_racing_thoughts	0.1098626	0.723411141	-0.024548254
Co_concentration_difficulties	-0.061585168	0.620161875	0.141955722
Co_more_energy	0.859767267	-0.04804931	0.041459219
Co_more_sociable	0.579176971	-0.016889102	0.391642908
Co_interested_in_sex	0.171225298	-0.153974867	0.652236184
Co_risky_behaviour	0.016590215	0.071023084	0.820967538
Co_spending_money	-0.062837332	0.132421829	0.607023321

	Factor 1	Factor 2	Factor 3	Factor 4
Co_hyperactivity	0.89433639	-0.043419126	0.058109279	-0.093041183
Co_irritability	0.082789767	0.222133828	0.075735955	-0.502104922
Co_self_confident	0.503518681	-0.245414549	0.194118886	0.29073742
Co_decreased_sleep	0.319313336	-0.037129748	0.198085765	0.15655854
Co_more_talkative	0.709488779	0.175180576	-0.066929585	0.187347656
Co_racing_thoughts	0.101650013	0.700206038	-0.058139402	-0.00548232
Co_concentration_difficulties	-0.114440824	0.679283168	0.120545853	0.024847171
Co_more_energy	0.373425394	0.056451776	0.16677706	0.561300634
Co_more_sociable	0.288584558	0.028556664	0.488288393	0.277527372
Co_interested_in_sex	-0.062979382	-0.092117164	0.750828875	0.170354022
Co_risky_behaviour	0.02004378	0.061340508	0.832526291	-0.087602582

Co_spending_money	0.062537469	0.082565452	0.5983209	-0.212297168

Table S3.4a. Fit statistics for each factor solution in exploratory factor analysis (EFA) of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants. "concurrent more active" was previously removed due to having a correlation of 0.87 with "concurrent more energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

Fit statistics	TLI	RMSEA index (95% CIs)	BIC	RMSR	Min items
Information	*≥0.90*	*≤0.05 = good 0.06-0.08 = fair*	*As low as possible*	*As close to 0 as possible, preferably <0.08*	*Out of 12*
1 factor	0.694	0.164 (0.163 - 0.166)	29991.99	0.11	9
2 factors	0.773	0.141 (0.14 - 0.143)	17621.63	0.7	3
3 factors	0.917	0.085 (0.083 - 0.087)	4727.53	0.03	3
4 factors	0.944	0.07 (0.068 - 0.073)	2258.51	0.02	2

TLI = to Tucker Lewis Index

RMSEA index = Root Mean Squared Error of Approximation index

BIC = Bayes Information Criterion

RMSR = Root Mean Squared Residual

Min items = minimum number of concurrent MDQ items loading onto any factor.

Table S3.4b: Correlations between each factor in the best-fitting model (three factor solution) identified by exploratory factor analysis (EFA) of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants. "concurrent more active" was previously removed due to having a correlation of 0.87 with "concurrent more energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

	Energy/activity	Cognitive	Impulsivity
Energy/activity	1	-0.07	0.54
Cognitive		1	-0.02
Impulsivity			1

Table S3.5. Confirmatory Factor Analysis (CFA) fit statistics of the three factor model of 12 concurrent manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants. "concurrent more active" was previously removed due to having a correlation of 0.87 with "concurrent more energy". "CFA sample" refers to fit statistics for the remaining 30% of the sample, "whole sample" refers to fit statistics for all participants.

Fit statistic	CFI		TLI		RMSEA (90% CI)		SRMR	
	Standard	*Robust*	*Standard*	*Robust*	*Standard*	*Robust*	*Standard*	*Robust*
CFA sample	0.969	0.945	0.959	0.928	0.064 (0.061 - 0.066)	0.070 (0.068 - 0.073)	0.083	0.083
Whole sample	0.967	0.94	0.958	0.923	0.064 (0.063 - 0.065)	0.072 (0.071 - 0.073)	0.082	0.082

CFI = Comparative Fit Index

TLI = Tucker Lewis Index

RMSEA = Root Mean Squared Error of Approximation

SRMR = Standardised Root Mean Residual

Table S3.6. Item loadings for all factor solutions for the exploratory factor analysis (EFA) of the lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants. "More active" was previously removed due to having a correlation of 0.9 with "more energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

	Factor 1
Hyperactivity	0.85095301
Irritability	0.50738741
Self_confident	0.7741318
Decreased_sleep	0.73382358
More_talkative	0.83799407
Racing_thoughts	0.68908489
Concentration_difficulties	0.6489075
More_energy	0.83681652
More_sociable	0.84434475
Interested_in_sex	0.74098112
Risky_behaviour	0.78693801
Reckless_spending	0.58949702

	Factor 1	Factor 2
Hyperactivity	0.67304098	0.2345476
Irritability	0.00407216	0.60956683
Self_confident	0.90647271	-0.1115699
Decreased_sleep	0.6271817	0.14946448
More_talkative	0.62328079	0.27663306
Racing_thoughts	0.09286457	0.73281022
Concentration_difficulties	-0.0165301	0.81822703
More_energy	0.96766688	-0.1046592
More_sociable	0.86419293	0.01443488

Interested_in_sex	0.68525969	0.09210445	
Risky_behaviour	0.51554938	0.34024336	
Reckless_spending	0.24040797	0.4265651	

	Factor 1	Factor 2	Factor 3
Hyperactivity	0.55284399	0.23430669	0.17612479
Irritability	-0.0923223	0.49919478	0.24504357
Self_confident	0.78055843	-0.0568748	0.12077433
Decreased_sleep	0.53086456	0.16160149	0.13145418
More_talkative	0.65078596	0.35081601	-0.0476101
Racing_thoughts	0.13597407	0.75322784	-0.0121227
Concentration_difficulties	-0.001263	0.75944035	0.08643548
More_energy	0.98266049	-0.0102504	-0.0573819
More_sociable	0.61009724	-0.0234989	0.35466139
Interested_in_sex	0.33935754	-0.0488147	0.55776089
Risky_behaviour	0.06727452	0.13760035	0.74778725
Reckless_spending	-0.0347398	0.28029764	0.48051784

	Factor 1	Factor 2	Factor 3	Factor 4
Hyperactivity	0.74138335	0.05678441	-0.020297	0.25744074
Irritability	-0.0214545	0.42546073	0.09947175	0.22139872
Self_confident	0.76169326	-0.0832505	0.1564618	0.00164743
Decreased_sleep	0.51215719	0.1480174	0.14107328	0.02320955
More_talkative	0.7313219	0.25793414	-0.099981	0.05646338
Racing_thoughts	0.08203737	0.8188706	-0.0042084	-0.0453165
Concentration_difficulties	-0.0447389	0.79495729	0.06883547	0.02884334
More_energy	0.94110411	-0.0193559	0.04234262	-0.1041426
More_sociable	0.49521794	0.03862883	0.42918497	-0.0151276
Interested_in_sex	0.07340199	0.08289353	0.7630015	-0.0305878
Risky_behaviour	0.09353636	0.12536008	0.55786005	0.29286279

Reckless_spending	0.08914627	0.12806324	0.24594171	0.42247571

Table S3.7a. Fit statistics for each exploratory factor analysis (EFA) of the 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants. "More active" was previously removed due to having a correlation of 0.90 with "More energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

Fit statistics	TLI	RMSEA Index (95% CIs)	BIC	RMSR	Min items
Information	*>0.90*	*<0.05 = good, 0.06-0.08 = fair*	*As low as possible*	*As close to 0 as possible, preferably <0.08*	*Out of 12*
1 factor	0.835	0.147 (0.145 - 0.148)	38267.35	0.07	12
2 factors	0.909	0.109 (0.108 - 0.111)	16723.03	0.03	4
3 factors	0.96	0.072 (0.07 - 0.074)	5419.4	0.02	3
4 factors	0.97	0.062 (0.06 - 0.064)	2860.05	0.01	1

TLI = to Tucker Lewis Index

RMSEA index = Root Mean Squared Error of Approximation index

BIC = Bayes Information Criterion

RMSR = Root Mean Squared Residual

Min items = minimum number of lifetime MDQ items loading onto any factor

Table S3.7b: Correlations between each factor in the best-fitting model (three factor solution) identfied by exploratory factor analysis (EFA) of 12 lifetime Mood Disorder Questionnaire (MDQ) items in affected participants. "More active" was previously removed due to having a correlation of 0.9 with "more energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

	Energy/activity	Cognitive	Impulsivity
Energy/activity	1	0.57	0.68
Cognitive		1	0.56

Impulsivity			1

Table S3.8. Confirmatory Factor Analysis (CFA) fit statistics of the three factor model of 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in affected participants. "CFA sample" refers to fit statistics for the remaining 30% of the sample, "whole sample" refers to fit statistics for all participants.

Fit statistic	CFI		TLI		RMSEA (90% CI)		SRMR	
	Standard	*Robust*	*Standard*	*Robust*	*Standard*	*Robust*	*Standard*	*Robust*
CFA sample	0.995	0.986	0.994	0.982	0.038 (0.036 - 0.040)	0.049 (0.047 - 0.051)	0.042	0.042
Whole sample	0.995	0.986	0.993	0.982	0.039 (0.037 - 0.040)	0.049 (0.048 - 0.050)	0.042	0.042

CFI = Comparative Fit Index

TLI = Tucker Lewis Index

RMSEA = Root Mean Squared Error of Approximation

SRMR = Standardised Root Mean Residual

Table S3.9. Item loadings for all factor solutions for the exploratory factor analysis (EFA) of 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in unaffected participants. "More active" was previously removed due to having a correlation of 0.9 with "more energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

	Factor 1
Hyperactivity	0.70991762
Irritability	0.55315936
Self_confident	0.76928341
Decreased_sleep	0.62847847
More_talkative	0.80231045
Racing_thoughts	0.68064597
Concentration_difficulties	0.66700987
More_energy	0.80045915
More_sociable	0.75441632
Interested_in_sex	0.7512914
Risky_behaviour	0.69606546
Spending_money	0.50393777

	Factor 1	Factor 2
Hyperactivity	0.34152709	0.44824618
Irritability	-0.0260664	0.66677763
Self_confident	0.53985393	0.3095114
Decreased_sleep	0.35608866	0.34084029
More_talkative	0.38348565	0.50953345
Racing_thoughts	0.0021399	0.79028051
Concentration_difficulties	-0.0519544	0.83609632
More_energy	0.70617724	0.17883806
More_sociable	0.96004838	-0.1085601

Interested_in_sex	0.8700588	-0.0296635	
Risky_behaviour	0.49097034	0.27825387	
Spending_money	0.24669291	0.31434293	

	Factor 1	**Factor 2**	**Factor 3**
Hyperactivity	0.31976898	0.41513339	0.13171471
Irritability	-0.0634236	0.60464691	0.22833719
Self_confident	0.60271355	0.38088239	-0.2237536
Decreased_sleep	0.35351473	0.34660733	0.01141159
More_talkative	0.39764076	0.54477744	-0.0668074
Racing_thoughts	0.00593106	0.7894727	0.02282195
Concentration_difficulties	-0.0516867	0.79760892	0.10077476
More_energy	0.72965175	0.21657269	-0.0978341
More_sociable	0.9227459	-0.1163067	0.09842847
Interested_in_sex	0.8398451	-0.0571405	0.12868482
Risky_behaviour	0.43682489	0.17373529	0.36089031
Spending_money	0.09923871	0.13833721	0.77255623

	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**
Hyperactivity	0.36116894	0.20497774	0.35218241	0.02202473
Irritability	-0.0901625	0.149094	0.60940212	0.04146895
Self_confident	0.27048517	-0.2208162	0.39431537	0.40125727
Decreased_sleep	0.36631139	0.09078157	0.28837473	0.05501479
More_talkative	0.35955599	-0.0049177	0.50135652	0.11100108
Racing_thoughts	0.04701742	-0.0210933	0.79497387	-0.0075381
Concentration_difficulties	-0.0012909	0.04520876	0.79442782	-0.017709
More_energy	0.8532458	0.07489328	0.06172489	0.09010003
More_sociable	0.32050197	0.07688478	-0.0871199	0.65169728
Interested_in_sex	0.04274237	-0.0080708	-0.0291266	0.90917033
Risky_behaviour	-0.2065013	0.20213694	0.23893081	0.65892069

Spending_money	0.03905191	0.98222476	0.00825179	0.02010126

Table S3.10. Fit statistics for each exploratory factor analysis (EFA) of 12 lifetime manic symptoms derived from the Mood Disorder Questionnaire (MDQ) in unaffected participants. "More active" was removed due to correlation of 0.90 with "More energy". EFA was performed with the psych R package. Oblimin rotation method was used to allow the latent factors to correlate with each other and the factoring method was "minimum residuals".

Fit statistic	TLI	RMSEA index (95% CIs)	BIC	RMSR	Min items
Information	*≥0.90*	*≤0.05 = good 0.06-0.08 = fair*	*As low as possible*	*As close to 0 as possible, preferably <0.08*	*Out of 12*
1 factor	0.685	0.192 (0.189 - 0.196)	8150.9	0.09	12
2 factors	0.765	0.166 (0.163 - 0.17)	4777.31	0.06	5
3 factors	0.791	0.157 (0.152 - 0.161)	3225.49	0.04	1
4 factors	0.85	0.133 (0.128 - 0.138)	1638.01	0.02	1

TLI = to Tucker Lewis Index

RMSEA index = Root Mean Squared Error of Approximation index

BIC = Bayes Information Criterion

RMSR = Root Mean Squared
Residual

Min items = minimum number of lifetime MDQ items loading onto any factor.

Table S3.11. Information about each psychiatric and behavioural trait included in genetic correlations with the manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ). Genome-wide association summary statistics of each trait were used to calculate genetic correlations with the manic symptom phenotypes using Linkage Disequilibrum Score Regression (LDSC) (Bulik-Sullivan et al. 2015). "Nca" refers to number of cases, "Nco" refers to number of controls, and "N" refers to overall sample size.

Phenotype	Published paper	Nca	Nco	N
ADHD	Demontis et al. (2019)	19099	34194	53293
Alcohol dependence	Walters et al. (2018)	11569	11569	46568
Daily alcohol use	Schumann et al. (2016)			70460
Alzheimer's disease	Jansen et al. (2019)	71880	383378	455258
Anhedonia	Ward et al. (2019)			375,275
Anorexia nervosa	Watson et al. (2019)	16992	55525	73050
Anxiety (lifetime, probable)	Purves et al. (2020)	25453	58113	83566
Autism spectrum disorder	Grove et al. (2019)	18381	27969	46350
Bipolar disorder	Mullins et al. (2021)	41,917	371,549	413466
Bipolar disorder type I	Mullins et al. (2021)	25060	449978	475038
Bipolar disorder type II	Mullins et al. (2021)	6781	364075	370856
Body mass index	Hübel et al. (2019)			353972
Cannabis use (lifetime)	Stringer et al. (2016)	14374	17956	32330

Chronotype	Jones et al. (2016)			128266
Sleep duration	Jones et al. (2016)			128266
Oversleeper	Jones et al. (2016)			128266
Undersleeper	Jones et al. (2016)			128266
Major depressive disorder (PGC2 including 23andme)	Wray et al. (2018)	154649	394409	549058
Depressive symptoms	Okbay et al. (2016)			161460
Major depressive disorder (PGC2 excluding 23andme)	Wray et al. (2018)	59851	113154	173005
Years of education	Lee et al. (2018)			766345
Self-rated health	Harris et al. (2017)			111483
Household income	Hill et al. (2016)			112151
Insomnia	Hammerschlag et al. (2017)	32384	80622	113006
Cognitive ability	Savage (2018)			269867
Neuroticism	Hübel et al. (2019)			
Obsessive compulsive disorder	International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OCGAS) (2018)	2688	7037	9725
Physical activity	NA			66224

Posttraumatic stress disorder	Nievergelt et al. (2019)	32428	174227	206655
Posttraumatic stress disorder (military)	Stein et al. (2020)	36301	178107	214408
Posttraumatic stress disorder symptoms (military)	Stein et al. (2020)			186689
General risk tolerance (self-report)	Linner et al. (2019)			466571
Automobile speeding propensity	Linner et al. (2019)			404291
Number of sexual partners	Linner et al. (2019)			370711
Schizophrenia	Pardinas et al. (2018)	11260	24542	35802
Ever smoker	Linner et al. (2019)			518,663
Subjective well-being	Okbay et al. (2016)			298420

Table S3.12. Genetic correlations between the four concurrent manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected participants. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the 1000 Genomes Linkage Disequilibrium reference panel. *r*g=genetic correlation, SE=standard error, and p-value=p-value for rg difference from zero. Genetic correlations were significantly different to zero if the p-value surpassed the Bonferroni-adjusted alpha of 0.008 (0.05/6) to correct for the six tests. Significant p-values are shown in bold.

	Concurrent energy/activity factor			Concurrent cognitive factor			Concurrent impulsivity factor		
	*r*g	SE	p	*r*g	SE	p	*r*g	SE	p
Concurrent manic symptoms (sum score)	0.9307	0.08312	**4.20×10^{-29}**	-0.6045	0.4857	2.13×10^{-1}	0.8887	0.1019	**2.74×10^{-18}**
Concurrent energy/activity factor				-0.6607	0.3217	4.00×10^{-2}	0.8936	0.1018	**1.68×10^{-18}**
Concurrent cognitive factor							-0.395	0.488	0.42
Concurrent impulsivity factor									

Table S3.13. Genetic correlations between the four lifetime manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected participants. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the 1000 Genomes Linkage Disequilibrium reference panel. *r*g=genetic correlation, SE=standard error, and p-value=p-value for *r*g difference from zero. Genetic correlations were significantly different to zero if the p-value surpassed the Bonferroni-adjusted alpha of 0.008 (0.05/6) to correct for the six tests. Significant p-values are shown in bold.

	Lifetime energy/activity factor			Lifetime cognitive factor			Lifetime impulsivity factor		
	*r*g	SE	p	*r*g	SE	p	*r*g	SE	p
Lifetime manic symptoms (sum score)	1.0076	0.0134	**<0.008**	1.0135	0.0208	**<0.008**	1.0172	0.0204	**<0.008**
Lifetime energy/activity factor				0.967	0.0303	**7.64×10^{-223}**	1.0241	0.0242	**<0.008**
Lifetime cognitive factor							1.0298	0.0364	**1.30×10^{-175}**
Lifetime impulsivity factor									

Table S3.14. Genetic correlations between concurrent manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected paricipants and 36 psychiatric and behavioural phenotypes. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the extended 1000 Genomes Linkage Disequilibrium reference panel. *r*g=genetic correlation, SE=standard error, and p-value=p-value for rg difference from zero. Genetic correlations were significantly different to zero if the p-value surpassed the Bonferroni-adjusted alpha of 0.001 (0.05/36) to correct for the 36 tests for each phenotype. Significant genetic correlations are shown in bold.

Phenotype	Lifetime manic symptoms (sum score)			Lifetime energy/activity factor			Lifetime cognitive factor			Lifetime impulsivity factor		
	*r*g	SE	p-value	*r*g	SE	p-value	*r*g	SE	p-value	*r*g	SE	p-value
ADHD	0.4375	0.2183	4.51E-02	0.2665	0.1778	1.34E-01	0.3221	0.2365	1.73E-01	0.3988	0.1636	1.48E-02
Alcohol dependence	0.4731	0.4652	3.09E-01	0.3068	0.5369	5.68E-01	0.2194	0.451	6.27E-01	0.734	0.5455	1.78E-01
Daily alcohol use	-0.08212	0.4092	8.41E-01	NA	NA	NA	-0.2633	0.628	6.75E-01	NA	NA	NA
Alzheimer's disease	0.37	0.2205	9.34E-02	0.195	0.1935	3.14E-01	0.211	0.2264	3.51E-01	0.2012	0.198	3.09E-01
Anhedonia	0.2753	0.1093	1.18E-02	0.1634	0.0992	9.94E-02	0.2493	0.2627	3.43E-01	0.2821	0.113	1.25E-02
Anorexia Nervosa	0.4987	0.2769	7.16E-02	0.6014	0.6334	3.42E-01	-0.1071	0.2115	6.13E-01	0.3244	0.209	1.21E-01
Lifetime probable anxiety	0.2821	0.1568	7.21E-02	0.2512	0.1319	5.69E-02	-0.0159	0.1999	9.36E-01	0.3979	0.1689	1.85E-02
Autism spectrum disorder	0.3937	0.2019	5.12E-02	0.4199	0.1703	1.37E-02	0.06611	0.2341	7.78E-01	0.413	0.1907	3.03E-02
Bipolar disorder	0.1423	0.1243	2.52E-01	0.229	0.1086	3.49E-02	-0.279	0.1829	1.27E-01	0.1597	0.112	1.54E-01
Bipolar disorder type I	0.05769	0.1012	5.69E-01	0.103	0.1088	3.44E-01	-0.1856	0.159	2.43E-01	0.02146	0.1011	8.32E-01
Bipolar disorder type II	0.1289	0.2176	5.54E-01	0.3474	0.2283	1.28E-01	-0.5588	0.2774	4.40E-02	0.2897	0.253	2.52E-01
Body mass index	0.2629	0.09217	4.34E-03	0.04011	0.06467	5.35E-01	0.3848	0.2274	9.07E-02	0.2348	0.08844	7.92E-03

Lifetime cannabis use	1.115	3.104	7.19E-01	NA	NA	NA	-0.7794	0.7651	3.08E-01	0.6631	3.179	8.35E-01
Chronotype	0.1714	0.124	1.67E-01	0.1563	0.1175	1.83E-01	-0.025	0.1544	8.71E-01	0.2113	0.1268	9.56E-02
Sleep duration	-0.06384	0.1521	6.75E-01	0.0688	0.1348	6.10E-01	-0.495	0.2715	6.83E-02	-0.1783	0.148	2.28E-01
Oversleeper	0.04659	0.2274	8.38E-01	-0.07759	0.2219	7.27E-01	0.264	0.34	4.37E-01	-0.0027	0.2302	9.91E-01
Undersleeper	0.0673	0.1838	7.14E-01	-0.0325	0.1776	8.55E-01	0.5275	0.2955	7.42E-02	0.2404	0.1898	2.05E-01
Major depressive disorder (PGC2 including 23andme)	0.1694	0.2939	5.64E-01	0.04402	0.1955	8.22E-01	0.08546	0.1423	5.48E-01	0.3844	0.5109	4.52E-01
Depressive symptoms	0.216	0.3458	5.32E-01	0.1314	0.5267	8.03E-01	0.4604	0.6705	4.92E-01	0.2693	0.2265	2.35E-01
Major depressive disorder (PGC2 excluding 23andme)	0.4259	0.2847	1.35E-01	0.3778	0.25	1.31E-01	0.05304	0.1624	7.44E-01	0.5504	0.3767	1.44E-01
Years of education	-0.1568	0.08466	6.40E-02	0.06934	0.07296	3.42E-01	-0.4984	0.3153	1.14E-01	-0.1269	0.07995	1.12E-01
Self-rated health	-0.4559	0.1726	8.25E-03	-0.3278	0.1241	8.25E-03	-0.2338	0.2169	2.81E-01	-0.4929	0.1642	2.68E-03
Household income	-0.6602	0.2271	3.65E-03	-0.328	0.1734	5.86E-02	-0.4555	0.277	1.00E-01	-0.6741	0.2319	3.65E-03
Insomnia	0.2429	0.1871	1.94E-01	0.1808	0.1655	2.75E-01	0.2633	0.2543	3.01E-01	0.2999	0.1788	9.36E-02
Cognitive ability	-0.1278	0.07137	7.34E-02	0.01416	0.06967	8.39E-01	-0.3601	0.2699	1.82E-01	-0.0402	0.07258	5.80E-01
Neuroticism	0.05856	0.09642	5.44E-01	-0.02081	0.08586	8.09E-01	0.1346	0.1694	4.27E-01	-0.0052	0.08757	9.53E-01
Obsessive compulsive disorder	NA	NA	NA	NA	NA	NA	-0.3929	0.2982	1.88E-01	-0.5949	1.53	6.97E-01

Physical activity	0.02677	0.1381	8.46E-01	0.02188	0.1439	8.79E-01	-0.0222	0.2	9.12E-01	-0.0442	0.1482	7.65E-01
Posttraumatic stress disorder	1.125	0.527	3.27E-02	1.136	0.4635	1.42E-02	-0.3309	0.5086	5.15E-01	1.596	0.645	1.33E-02
Posttraumatic stress disorder (military)	0.3768	0.1742	3.06E-02	-0.00854	0.1613	9.58E-01	0.7971	0.3988	4.57E-02	0.2601	0.1444	7.17E-02
Posttraumatic stress disorder symptoms (military)	0.4403	0.1582	5.37E-03	0.206	0.1294	1.11E-01	0.2925	0.1761	9.66E-02	0.3261	0.1286	1.12E-02
General risk tolerance	0.4368	0.1486	3.28E-03	0.4354	0.1411	2.04E-03	-0.2096	0.174	2.28E-01	0.4614	0.155	2.92E-03
Automobile speeding propensity	-0.268	0.1204	2.61E-02	-0.1756	0.09515	6.50E-02	-0.2149	0.1804	2.33E-01	-0.2591	0.1171	2.70E-02
Number of sexual partners	0.2804	0.1119	1.22E-02	0.2753	0.1001	5.96E-03	-0.0807	0.1272	5.26E-01	0.346	0.1244	5.43E-03
Schizophrenia	0.6393	2.796	8.19E-01	0.673	1.48	6.49E-01	-0.1853	0.1283	1.49E-01	0.3452	0.3787	3.62E-01
Ever smoker	0.2492	0.08871	4.97E-03	0.1543	0.07924	5.15E-02	0.1455	0.1271	2.53E-01	0.2518	0.09508	8.08E-03
Subjective wellbeing	0.3161	0.5896	5.92E-01	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table S3.15. Genetic correlations between lifetime manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) in affected participants and 37 psychiatric and behavioural phenotypes. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the extended 1000 Genomes Linkage Disequilibrium reference panel. *rg*=genetic correlation, SE=standard error, and p-value=p-value for rg difference from zero. Genetic correlations were significantly different to zero if the p-value surpassed the Bonferroni-adjusted alpha of 0.001 (0.05/36) to correct for the 36 tests for each phenotype. Significant genetic correlations are shown in bold.

Phenotype	Lifetime manic symptoms (sum score)			Lifetime energy/activity factor			Lifetime cognitive factor			concurrent impulsivity factor		
	rg	SE	p-value	*rg*	SE	p-value	*rg*	SE	p-value	*rg*	SE	p-value
ADHD	0.6888	0.1251	**3.68E-08**	0.5975	0.1202	**6.73E-07**	0.6317	0.1057	**2.31E-09**	0.7742	0.1837	**2.51E-05**
Alcohol dependence	0.6162	0.2398	0.010186	0.5098	0.2321	0.028059	0.5979	0.2158	0.005597	0.8297	0.3161	0.008672
Daily alcohol use	0.1846	0.1668	0.26833	0.2277	0.1717	0.18466	0.0522	0.15	0.72789	0.2623	0.1844	0.1549
Alzheimer's disease	0.1053	0.1154	0.36137	0.072	0.1212	0.55245	0.087	0.1121	0.43757	0.0779	0.1365	0.56846
Anhedonia	0.4356	0.0943	**3.84E-06**	0.4963	0.102	**1.14E-06**	0.45	0.0877	**2.92E-07**	0.539	0.1236	**1.30E-05**
Anorexia Nervosa	0.043	0.1066	0.68654	-0.0004	0.1172	0.99755	0.0656	0.0994	0.50917	0.0811	0.1407	0.56433
Lifetime probable anxiety	0.0744	0.0962	0.43945	0.1336	0.101	0.1861	0.0909	0.0887	0.30576	0.1064	0.113	0.34621
Autism spectrum disorder	0.1664	0.1194	0.16313	0.1528	0.1212	0.20737	0.1705	0.1002	0.088946	0.2123	0.1579	0.17883
Bipolar disorder	0.0929	0.0748	0.21433	0.1279	0.0845	0.13015	-0.0177	0.0586	0.76217	0.1423	0.1009	0.15853
Bipolar disorder type I	0.04	0.0746	0.59167	0.0805	0.084	0.33764	-0.0335	0.0601	0.57713	0.0614	0.0968	0.52589
Bipolar disorder type II	0.158	0.1418	0.26521	0.236	0.1512	0.11868	-0.0196	0.1171	0.86706	0.2353	0.1778	0.18575
Body mass index	0.4013	0.0624	**1.24E-10**	0.3917	0.0658	**2.65E-09**	0.3951	0.0562	**2.03E-12**	0.467	0.0813	**9.14E-09**
Lifetime cannabis use	0.367	0.1928	0.057051	0.4291	0.2049	0.036253	0.2548	0.1587	0.10825	0.423	0.2393	0.077184

Chronotype	0.1347	0.0822	0.10109	0.0944	0.0864	0.27474	0.0779	0.0795	0.32743	0.1233	0.0952	0.19552
Sleep duration	-0.0363	0.0998	0.71608	-0.0405	0.1054	0.70103	-0.05	0.0968	0.60522	-0.0489	0.1162	0.67359
Oversleeper	0.3303	0.1683	0.049772	0.3482	0.1738	0.045128	0.4085	0.1619	0.011599	0.4771	0.2035	0.019063
Undersleeper	0.1994	0.1155	0.084179	0.2158	0.1225	0.078113	0.2754	0.1138	0.015524	0.3091	0.1383	0.025398
Major depressive disorder (PGC2 including 23andme)	0.2063	0.0789	0.008924	0.2007	0.0755	0.007837	0.1933	0.0679	0.004433	0.2322	0.0952	0.014723
Depressive symptoms	0.506	0.1438	**0.000434**	0.562	0.1586	**0.000394**	0.4367	0.1259	**0.000524**	0.6496	0.2034	0.001401
Major depressive disorder (PGC2 excluding 23andme)	0.4185	0.1033	**5.14E-05**	0.3998	0.0982	**4.69E-05**	0.3591	0.0868	**3.56E-05**	0.4505	0.1262	**0.000356**
Years of education	-0.4519	0.0633	**9.79E-13**	-0.4134	0.0679	**1.11E-09**	-0.4927	0.0658	**7.10E-14**	-0.5357	0.0948	**1.60E-08**
Self-rated health	-0.6004	0.101	**2.76E-09**	-0.6148	0.1162	**1.21E-07**	-0.5553	0.0918	**1.48E-09**	-0.742	0.1391	**9.53E-08**
Household income	-0.5947	0.1256	**2.19E-06**	-0.5475	0.1349	**4.92E-05**	-0.5968	0.1198	**6.28E-07**	-0.6832	0.1623	**2.56E-05**
Insomnia	0.5501	0.1495	**0.000232**	0.6235	0.1593	**9.04E-05**	0.5491	0.136	**5.38E-05**	0.722	0.197	**2.48E-04**
Cognitive ability	-0.2541	0.0577	**1.07E-05**	-0.2505	0.0593	**2.42E-05**	-0.2856	0.0552	**2.29E-07**	-0.3073	0.0748	**3.95E-05**
Neuroticism	0.1248	0.0627	0.046398	0.134	0.0638	0.035588	0.1675	0.0582	0.003966	0.1566	0.0746	0.035885
Obsessive compulsive disorder	-0.2513	0.1484	0.090377	-0.0783	0.1458	0.59139	-0.2725	0.1388	0.049513	-0.1471	0.1845	0.42525
Physical activity	0.0325	0.0875	0.71003	0.062	0.0936	0.50739	0.0505	0.0794	0.52522	0.0174	0.1034	0.8662

Posttraumatic stress disorder	1.037	0.306	**0.000702**	1.0367	0.3218	0.001276	0.9086	0.2819	0.001268	1.2328	0.3869	0.00144
Posttraumatic stress disorder (military)	0.5691	0.1167	**1.07E-06**	0.5898	0.1379	**1.90E-05**	0.5843	0.1053	**2.86E-08**	0.6997	0.1705	**4.05E-05**
Posttraumatic stress disorder symptoms (military)	0.6096	0.0976	**4.16E-10**	0.6151	0.1154	**9.84E-08**	0.5657	0.0833	**1.10E-11**	0.6943	0.1336	**2.04E-07**
General risk tolerance	0.3592	0.0761	**2.33E-06**	0.3781	0.0837	**6.19E-06**	0.2979	0.0687	**1.44E-05**	0.4244	0.0997	**2.08E-05**
Automobile speeding propensity	-0.1772	0.0692	0.010473	-0.1631	0.0722	0.023933	-0.1801	0.0681	0.008174	-0.1891	0.084	0.024265
Number of sexual partners	0.2304	0.062	**0.000203**	0.249	0.0648	**0.000122**	0.1671	0.056	0.002844	0.2786	0.0733	**1.45E-04**
Schizophrenia	0.1541	0.079	0.051233	0.2158	0.092	0.018919	0.0783	0.0661	0.23636	0.1927	0.0963	0.04533
Ever smoker	0.3138	0.064	**9.28E-07**	0.3414	0.0711	**1.58E-06**	0.2948	0.0623	**2.20E-06**	0.3955	0.0828	**1.78E-06**
Subjective wellbeing	0.0423	0.1266	0.7381	0.0066	0.1333	0.96033	-0.0006	0.1071	0.99525	-0.0434	0.1476	0.76856

Table S3.16. Single Nucleotide Polymorphism (SNP) based heritability estimates (h2SNP), standard errors (SE), lamdba GC, and mean chi-square statistic of the manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ) calculated by Linkage Disequilibrium Score Regression (LDSC). "N items" refers to the number of MDQ items included in the phenotype and "N" refers to the number of individuals included in the genome-wide association study (GWAS). Genetic correlations were significantly different to zero if the p-value surpassed the Bonferroni-adjusted alpha of 0.006 (0.05/8) to correct for the four tests within both groups (concurrent/lifetime) and are shown in bold. P-value is calculated in R using pchisq((h2/se)^2,1,lower.tail = FALSE) as recommended by LDSC developers.

MDQ items phenotype	N items	N	h^2_{SNP}	SE	z-score	p-value	Lambda GC	Mean X^2
Concurrent manic symptoms (sum score)	12	11,568	0.056	0.0293	1.911262799	0.05597082	1.0046	1.0087
Concurrent energy/activity factor	6	11,568	0.0676	0.0318	2.125786164	0.03352107	1.0046	1.0114
Concurrent cognitive factor	3	11,568	0.038	0.0315	1.206349206	0.2276829	1.0075	1.0115
Concurrent impulsivity factor	3	11,568	0.056	0.0279	2.007168459	0.04473173	1.0016	1.0079
Lifetime manic symptoms (sum score)	12	19,958	0.0716	0.0166	4.313253012	**1.61E-05**	1.0255	1.0306
Lifetime energy/activity factor	6	19,958	0.0648	0.0161	4.02484472	**5.70E-05**	1.0315	1.0286
Lifetime cognitive factor	3	19,958	0.0764	0.0175	4.365714286	**1.27E-05**	1.0315	1.0349
Lifetime impulsivity factor	3	19,958	0.0506	0.0158	3.202531646	**0.00136225**	1.0195	1.0258

Table S3.17. Block-jackknife results comparing genetic correlations between 16 traits and the lifetime manic symptom phenotypes derived from the Mood Disorder Questionnaire (MDQ). The block-jackknife tested two genetic correlations for a significant difference from 0. Genetic correlations were significantly different to each other if the block-jackknife p-value reached or surpassed the Bonferroni-adjusted alpha of 0.003 (0.05/16) to correct for the 16 sets of tests.

Genetic correlation 1	Genetic correlation 2	*r*g 1	*r*g 2	Difference	p-value
1. Attention deficit hyperactivity disorder (ADHD)					
ADHD and lifetime manic symptoms (sum score)	ADHD and lifetime energy/activity factor	0.6888	0.5975	0.0913	0.102483911
ADHD and lifetime manic symptoms (sum score)	ADHD and lifetime cognitive factor	0.6888	0.6317	0.0571	0.458902729
ADHD and lifetime manic symptoms (sum score)	ADHD and lifetime impulsivity factor	0.6888	0.7742	-0.0854	0.483056341
ADHD and lifetime energy/activity factor	ADHD and lifetime cognitive factor	0.5975	0.6317	-0.0342	0.527049603
ADHD and lifetime energy/activity factor	ADHD and lifetime impulsivity factor	0.5975	0.7742	-0.1767	0.150396951
ADHD and lifetime impulsivity factor	ADHD and lifetime cognitive factor	0.7742	0.6317	0.1425	0.344596659
2. Anhedonia					
Anhedonia and lifetime manic symptoms (sum score)	Anhedonia and lifetime energy/activity factor	0.4356	0.4963	-0.0607	0.057001895
Anhedonia and lifetime manic symptoms (sum score)	Anhedonia and lifetime cognitive factor	0.4356	0.45	-0.0144	0.58978234
Anhedonia and lifetime manic symptoms (sum score)	Anhedonia and lifetime impulsivity factor	0.4356	0.539	-0.1034	0.068062319
Anhedonia and lifetime energy/activity factor	Anhedonia and lifetime cognitive factor	0.4963	0.45	0.0463	0.386171571
Anhedonia and lifetime energy/activity factor	Anhedonia and lifetime impulsivity factor	0.4963	0.539	-0.0427	0.571070113

Anhedonia and lifetime impulsivity factor	Anhedonia and lifetime cognitive factor	0.539	0.45	0.089	0.245263636
3. Body mass index (BMI)					
BMI and lifetime manic symptoms (sum score)	BMI and lifetime energy/activity factor	0.4013	0.3917	0.0096	0.676458185
BMI and lifetime manic symptoms (sum score)	BMI and lifetime cognitive factor	0.4013	0.3951	0.0062	0.735055699
BMI and lifetime manic symptoms (sum score)	BMI and lifetime impulsivity factor	0.4013	0.467	-0.0657	0.13987387
BMI and lifetime energy/activity factor	BMI and lifetime cognitive factor	0.3917	0.3951	-0.0034	0.969783097
BMI and lifetime energy/activity factor	BMI and lifetime impulsivity factor	0.3917	0.467	-0.0753	0.092389254
BMI and lifetime impulsivity factor	BMI and lifetime cognitive factor	0.467	0.3951	0.0719	0.134187833
4. Cognitive ability					
Cognitive ability and lifetime manic symptoms (sum score)	Cognitive ability and lifetime energy/activity factor	-0.2541	-0.2505	-0.0036	0.864753715
Cognitive ability and lifetime manic symptoms (sum score)	Cognitive ability and lifetime cognitive factor	-0.2541	-0.2856	0.0315	0.302367996
Cognitive ability and lifetime manic symptoms (sum score)	Cognitive ability and lifetime impulsivity factor	-0.2541	-0.3073	0.0532	0.23673494
Cognitive ability and lifetime energy/activity factor	Cognitive ability and lifetime cognitive factor	-0.2505	-0.2856	0.0351	0.378957478
Cognitive ability and lifetime energy/activity factor	Cognitive ability and lifetime impulsivity factor	-0.2505	-0.3073	0.0568	0.24004897
Cognitive ability and lifetime impulsivity factor	Cognitive ability and lifetime cognitive factor	-0.3073	-0.2856	-0.0217	0.810876489
5. Depressive factor					
Depressive factor and lifetime manic symptoms (sum score)	Depressive factor and lifetime energy/activity factor	0.506	0.562	-0.056	0.429928182
Depressive factor and lifetime manic symptoms (sum score)	Depressive factor and lifetime cognitive factor	0.506	0.4367	0.0693	0.320219865

Depressive factor and lifetime manic symptoms (sum score)	Depressive factor and lifetime impulsivity factor	0.506	0.6496	-0.1436	0.282544126
Depressive factor and lifetime energy/activity factor	Depressive factor and lifetime cognitive factor	0.562	0.4367	0.1253	0.214041279
Depressive factor and lifetime energy/activity factor	Depressive factor and lifetime impulsivity factor	0.562	0.6496	-0.0876	0.503598285
Depressive factor and lifetime impulsivity factor	Depressive factor and lifetime cognitive factor	0.6496	0.4367	0.2129	0.165246308
6. Ever smoker					
Ever smoker and lifetime manic symptoms (sum score)	Ever smoker and lifetime energy/activity factor	0.3138	0.3414	-0.0276	0.394329412
Ever smoker and lifetime manic symptoms (sum score)	Ever smoker and lifetime cognitive factor	0.3138	0.2948	0.019	0.49127302
Ever smoker and lifetime manic symptoms (sum score)	Ever smoker and lifetime impulsivity factor	0.3138	0.3955	-0.0817	0.017493235
Ever smoker and lifetime energy/activity factor	Ever smoker and lifetime cognitive factor	0.3414	0.2948	0.0466	0.282138887
Ever smoker and lifetime energy/activity factor	Ever smoker and lifetime impulsivity factor	0.3414	0.3955	-0.0541	0.069369301
Ever smoker and lifetime impulsivity factor	Ever smoker and lifetime cognitive factor	0.3955	0.2948	0.1007	0.013945361
7. General risk tolerance					
General risk tolerance and lifetime manic symptoms (sum score)	General risk tolerance and lifetime energy/activity factor	0.3592	0.3781	-0.0189	0.571766873
General risk tolerance and lifetime manic symptoms (sum score)	General risk tolerance and lifetime cognitive factor	0.3592	0.2979	0.0613	0.077684687
General risk tolerance and lifetime manic symptoms (sum score)	General risk tolerance and lifetime impulsivity factor	0.3592	0.4244	-0.0652	0.260104317
General risk tolerance and lifetime energy/activity factor	General risk tolerance and lifetime cognitive factor	0.3781	0.2979	0.0802	0.086965984

General risk tolerance and lifetime energy/activity factor	General risk tolerance and lifetime impulsivity factor	0.3781	0.4244	-0.0463	0.447987548
General risk tolerance and lifetime impulsivity factor	General risk tolerance and lifetime cognitive factor	0.4244	0.2979	0.1265	0.056737171
8. Household income					
Household income and lifetime manic symptoms (sum score)	Household income and lifetime energy/activity factor	-0.5947	-0.5475	-0.0472	0.488628999
Household income and lifetime manic symptoms (sum score)	Household income and lifetime cognitive factor	-0.5947	-0.5968	0.0021	0.974594491
Household income and lifetime manic symptoms (sum score)	Household income and lifetime impulsivity factor	-0.5947	-0.6832	0.0885	0.230892979
Household income and lifetime energy/activity factor	Household income and lifetime cognitive factor	-0.5475	-0.5968	0.0493	0.63016136
Household income and lifetime energy/activity factor	Household income and lifetime impulsivity factor	-0.5475	-0.6832	0.1357	0.089055452
Household income and lifetime impulsivity factor	Household income and lifetime cognitive factor	-0.6832	-0.5968	-0.0864	0.350726386
9. Insomnia					
Insomnia and lifetime manic symptoms (sum score)	Insomnia and lifetime energy/activity factor	0.5501	0.6235	-0.0734	0.237582548
Insomnia and lifetime manic symptoms (sum score)	Insomnia and lifetime cognitive factor	0.5501	0.5491	0.001	0.952544561
Insomnia and lifetime manic symptoms (sum score)	Insomnia and lifetime impulsivity factor	0.5501	0.722	-0.1719	0.090351987
Insomnia and lifetime energy/activity factor	Insomnia and lifetime cognitive factor	0.6235	0.5491	0.0744	0.442609033
Insomnia and lifetime energy/activity factor	Insomnia and lifetime impulsivity factor	0.6235	0.722	-0.0985	0.32903187
Insomnia and lifetime impulsivity factor	Insomnia and lifetime cognitive factor	0.722	0.5491	0.1729	0.183515918
10. Major depressive disorder (MDD) excluding 23andMe					

MDD and lifetime manic symptoms (sum score)	MDD and lifetime energy/activity factor	0.4185	0.3998	0.0187	0.674353904
MDD and lifetime manic symptoms (sum score)	MDD and lifetime cognitive factor	0.4185	0.3591	0.0594	0.242903227
MDD and lifetime manic symptoms (sum score)	MDD and lifetime impulsivity factor	0.4185	0.4505	-0.032	0.803226335
MDD and lifetime energy/activity factor	MDD and lifetime cognitive factor	0.3998	0.3591	0.0407	0.48099871
MDD and lifetime energy/activity factor	MDD and lifetime impulsivity factor	0.3998	0.4505	-0.0507	0.602501987
MDD and lifetime impulsivity factor	MDD and lifetime cognitive factor	0.4505	0.3591	0.0914	0.332026549
11. Number of sexual partners					
Number of sexual partners and lifetime manic symptoms (sum score)	Number of sexual partners and lifetime energy/activity factor	0.2304	0.249	-0.0186	0.282683577
Number of sexual partners and lifetime manic symptoms (sum score)	Number of sexual partners and lifetime cognitive factor	0.2304	0.1671	0.0633	0.063800305
Number of sexual partners and lifetime manic symptoms (sum score)	Number of sexual partners and lifetime impulsivity factor	0.2304	0.2786	-0.0482	0.071245261
Number of sexual partners and lifetime energy/activity factor	Number of sexual partners and lifetime cognitive factor	0.249	0.1671	0.0819	0.034008295
Number of sexual partners and lifetime energy/activity factor	Number of sexual partners and lifetime impulsivity factor	0.249	0.2786	-0.0296	0.282297942
Number of sexual partners and lifetime impulsivity factor	Number of sexual partners and lifetime cognitive factor	0.2786	0.1671	0.1115	0.00879042
12. Posttraumatic stress disorder (PTSD) (military)					
PTSD (military) and lifetime manic symptoms (sum score)	PTSD (military) and lifetime energy/activity factor	0.5691	0.5898	-0.0207	0.799148448
PTSD (military) and lifetime manic symptoms (sum score)	PTSD (military) and lifetime cognitive factor	0.5691	0.5843	-0.0152	0.79276324

PTSD (military) and lifetime manic symptoms (sum score)	PTSD (military) and lifetime impulsivity factor	0.5691	0.6997	-0.1306	0.205064469
PTSD (military) and lifetime energy/activity factor	PTSD (military) and lifetime cognitive factor	0.5898	0.5843	0.0055	0.976363885
PTSD (military) and lifetime energy/activity factor	PTSD (military) and lifetime impulsivity factor	0.5898	0.6997	-0.1099	0.216628654
PTSD (military) and lifetime impulsivity factor	PTSD (military) and lifetime cognitive factor	0.6997	0.5843	0.1154	0.377923423
13. Posttraumatic stress disorder (PTSD) factor (military)					
PTSD factor (military) and lifetime manic symptoms (sum score)	PTSD factor (military) and lifetime energy/activity factor	0.6096	0.6151	-0.0055	0.852050555
PTSD factor (military) and lifetime manic symptoms (sum score)	PTSD factor (military) and lifetime cognitive factor	0.6096	0.5657	0.0439	0.250812548
PTSD factor (military) and lifetime manic symptoms (sum score)	PTSD factor (military) and lifetime impulsivity factor	0.6096	0.6943	-0.0847	0.458032758
PTSD factor (military) and lifetime energy/activity factor	PTSD factor(military) and lifetime cognitive factor	0.6151	0.5657	0.0494	0.474645755
PTSD factor (military) and lifetime energy/activity factor	PTSD factor (military) and lifetime impulsivity factor	0.6151	0.6943	-0.0792	0.373466893
PTSD factor (military) and lifetime impulsivity factor	PTSD factor (military) and lifetime cognitive factor	0.6943	0.5657	0.1286	0.202490917
14. Self-rated health					
Self-rated health and lifetime manic symptoms (sum score)	Self-rated health and lifetime energy/activity factor	-0.6004	-0.6148	0.0144	0.797000991
Self-rated health and lifetime manic symptoms (sum score)	Self-rated health and lifetime cognitive factor	-0.6004	-0.5553	-0.0451	0.450735929
Self-rated health and lifetime manic symptoms (sum score)	Self-rated health and lifetime impulsivity factor	-0.6004	-0.742	0.1416	0.063275416
Self-rated health and lifetime energy/activity factor	Self-rated health and lifetime cognitive factor	-0.6148	-0.5553	-0.0595	0.449166013

Self-rated health and lifetime energy/activity factor	Self-rated health and lifetime impulsivity factor	-0.6148	-0.742	0.1272	0.062671358
Self-rated health and lifetime impulsivity factor	Self-rated health and lifetime cognitive factor	-0.742	-0.5553	-0.1867	0.037495987
15. Years of education					
Years of education and lifetime manic symptoms (sum score)	Years of education and lifetime energy/activity factor	-0.4519	-0.4134	-0.0385	0.169497147
Years of education and lifetime manic symptoms (sum score)	Years of education and lifetime cognitive factor	-0.4519	-0.4927	0.0408	0.39069002
Years of education and lifetime manic symptoms (sum score)	Years of education and lifetime impulsivity factor	-0.4519	-0.5357	0.0838	0.163343592
Years of education and lifetime energy/activity factor	Years of education and lifetime cognitive factor	-0.4134	-0.4927	0.0793	0.135028893
Years of education and lifetime energy/activity factor	Years of education and lifetime impulsivity factor	-0.4134	-0.5357	0.1223	0.034293922
Years of education and lifetime impulsivity factor	Years of education and lifetime cognitive factor	-0.5357	-0.4927	-0.043	0.504565758

Highlighted cells indicates that that particular genetic correlation was not significant

Table S3.18a. Sub-cohorts of the National Institute for Health and Care Research (NIHR) BioResource who were included in the "affected" study sample (N=5,380).				
	N	**Recruitment methods**	**Eligibility criteria**	**Recruitment area**
Inflammatory Bowel Disease (IBD) cohort	1,283	IBD clinics in participating hospitals across the United Kingdom	16+, have a diagnosis of Crohn's disease, ulcerative colitis, indeterminate colitis, IBD type unspecified, or suspected IBD	England, Wales, Scotland, Northern Ireland
NHS Blood and Transplant studies (COMPARE, STRIDES, INTERVAL)	2,503	Blood donation centres	16+, live in England	England
Research Tissue Bank - Generic	1,594	Biomedical Research Centres, Clinical Research Facilities, hospital clinics, community recruitment, online	16+, live in England	England
Total	5,380			

Table S3.18b. Sub-cohorts of the National Institute for Health and Care Research (NIHR) BioResource who were included in the "unaffected" study sample (N=6,308).				
	N	**Recruitment methods**	**Eligibility criteria**	**Recruitment area**
Inflammatory Bowel Disease (IBD) cohort	749	IBD clinics in participating hospitals across the United Kingdom	16+, have a diagnosis of Crohn's disease, ulcerative colitis, indeterminate colitis, IBD type unspecified, or suspected IBD	England, Wales, Scotland, Northern Ireland
NHS Blood and Transplant studies (COMPARE, STRIDES, INTERVAL)	4,081	Blood donation centres	16+, live in England	England
Research Tissue Bank - Generic	1,478	Biomedical Research Centres, Clinical Research Facilities, hospital clinics, community recruitment, online	16+, live in England	England
Total	6,308			

Appendix 4. Supplementary material for chapter 4

Supplementary material

Supplementary methods

Phenotype definition of treatment-resistant depression

Treatment-resistant depression was measured continuously in GLAD participants with the Maudsley Staging Method (MSM) (Fekadu *et al.*, 2009) and the nine item Patient Health Questionnaire (PHQ9) (Kroenke, Spitzer and Williams, 2001). This is referred to as the "staged treatment-resistant depression" phenotype in the main text. The MSM is a tool to measure varying levels of treatment resistance in individuals currently affected by MDD. The MSM uses a points system that is based on three domains: 1) **severity of presenting illness**, 2) **duration of presenting illness**, and 3) **antidepressant treatment-response**. The MSM can be interpreted as a continuous measure of treatment resistance in relation to an individual's current or most recent depressive episode. Information regarding how the three domains are scored and how this was implemented into the GLAD study survey is shown below.

1) Severity of presenting illness (scored 1-4):

In the MSM, the scoring of depressive symptom severity ranges from subsyndromal depression (scored as 1) to severe syndromal depression with psychosis (scored as 5). We used the PHQ9 to measure this in the GLAD study. The PHQ9 asked nine questions relating to mood and feelings that individuals may have experienced over the past two weeks. Individuals could answer with: "*Not at all*"; "*Several days*"; "*More than half the days*"; and "*Nearly every day*" xfwhich are coded as 0, 1, 2 and 3 respectively. Numeric answers were summed to create a continuous measure of current depression symptoms ranging 0-27 with higher values reflecting more severe symptoms. The PHQ9 does not ask about the presence of psychotic symptoms (as is required in the MSM). So, for the purpose of scoring GLAD Study participants' current depression severity for use in the MSM, the PHQ9 was scored as follows:

- PHQ9 score of 5-9 = mild (scored as 1)
- PHQ9 score of 10-14 = moderate (scored as 2)
- PHQ9 score of 15-19 = moderately severe (scored as 3)
- PHQ9 score of 20-27 = severe (scored as 4)

Therefore, instead of ranging 1-5 (as is required for the MSM), we used an adapted version where current depression severity ranged 1-4. Note that individuals who scored 0-4 on the PHQ9 (i.e., a severity score of 0/"None") were not eligible to complete the MSM and were therefore excluded from analyses.

2) Duration of presenting illness (scored 1-3):

In the MSM, the duration of presenting illness is measured with the question *"How long ago did your current or most recent episode of depression or low mood begin?"*. Individuals can answer with the following:

- *"Less than 1 year ago"* (scored as 1)
- *"1-2 years ago"* (scored as 2)
- *"More than 2 years ago"* (scored as 3)

This question was implemented word-for-word in the GLAD study survey. Thus, participants who were experiencing "mild depression" or more (a severity score of at least 1 in the PHQ9) answered the MSM and could score between 1-3 to reflect the duration of their presenting illness.

3) Antidepressant treatment-response (scored 0-7)

In the MSM, antidepressant treatment-response is measured with three questions. The first, relating to how many antidepressant medications an individual has tried, is *"During the current or most recent episode of depression or low mood, how many antidepressant medications have you taken for 6 weeks or longer?"*. Individuals can answer with the following:

- *"None"* (scored as 0)
- *"One to two"* (scored as 1)
- *"Three to four"* (scored as 2)
- *"Five to six"* (scored as 3)
- *"Seven to ten"* (scored as 4)
- *"More than ten"* (scored as 5)

The second question, relating to additional/augmentation medications, is *"If individuals don't respond fully to antidepressants, doctors sometimes prescribe "add-on" or "augmentation" medications in addition to the antidepressants. During the current or most recent episode of depression or low mood, have you taken an add-on medication for 6 weeks or longer?"*. Individuals can answer with the following:

- *"No"* (scored as 0)
- *"Yes"* (scored as 1)

The final question was *"Have you ever received electroconvulsive therapy?"*. Individuals can answer with the following:

- *"No"* (scored as 0)
- *"Yes"* (scored as 1)

These three questions were implemented word-for-word in the GLAD study survey. Thus, participants could score anywhere between 0-7 in relation to their antidepressant treatment response. Note that individuals who answered "None" to the question about the number of antidepressant medications skipped the next question about add-on/augmentation medications.

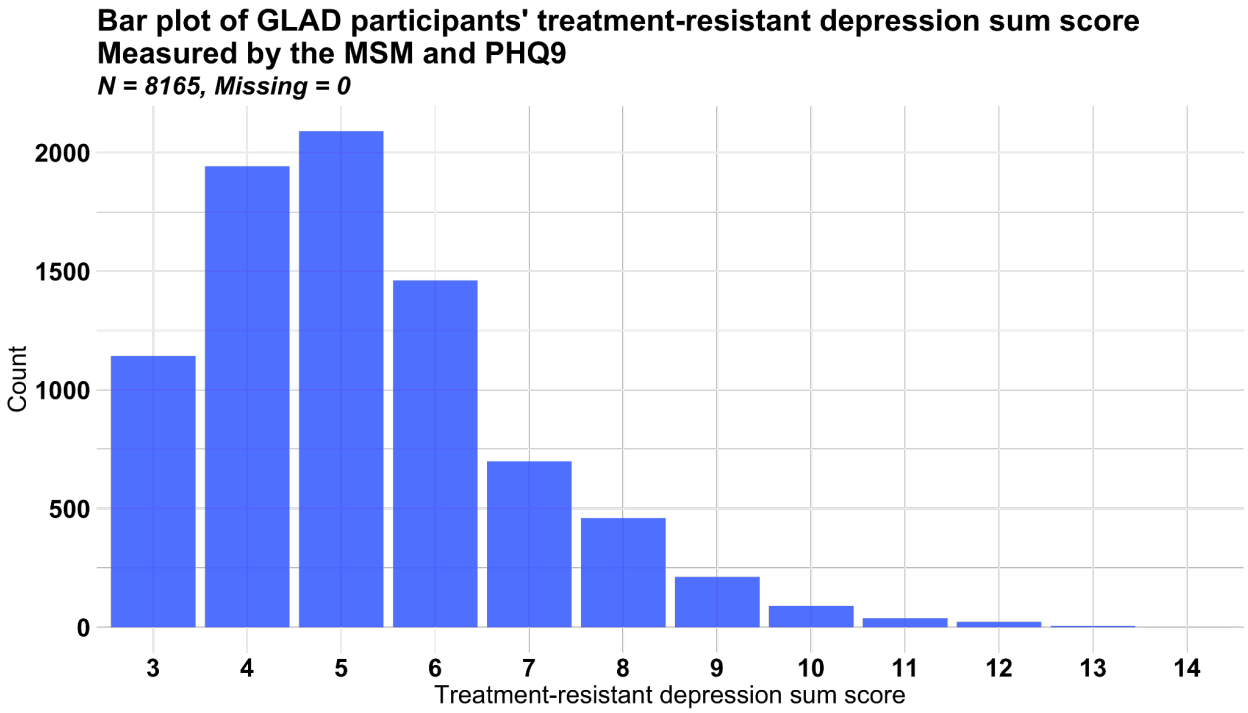
Clarification of scoring system (ranging 3-14)

The MSM ranges 3-14 (with 3 representing "mild resistance" and 14 representing "severe resistance") (Fekadu *et al.*, 2009). GLAD participants who did not meet the required threshold on the PHQ9 (i.e., their severity score was 0/"None") were not eligible to complete the MSM during

the GLAD study survey. These individuals were deemed to *not* be experiencing a current or most recent depressive episode and were therefore excluded. In those who were experiencing “mild depression” or more and *did* complete the MSM, the scores/points were attributed to each participant in the three domains described above. Their PHQ9 severity score (ranging 1-3) was added to their score from the MSM questions (ranging 2-10). Accordingly, GLAD participants could score anywhere between 3-14* for the treatment-resistant depression sum score.

Please also be reminded that the MSM is usually scored 3-15, however we used an adapted version which did not measure the presence of psychotic symptoms (explained above). Therefore, MSM was scored 3-14 in GLAD. This adapted version was confirmed via personal correspondence with Professor Anthony Cleare (one of the authors of the MSM).

*It is impossible to gain a score of two in the MSM because we excluded participants who reported that they have *not* taken any antidepressants for six weeks or more. We excluded these participants for two reasons. First, we cannot be sure of their treatment status as we do not know whether they have not taken antidepressants for six weeks or longer because a) they weren't prescribed any or b) they were prescribed them but they stopped taking them before six weeks. Second, participants who reported that they had not taken any antidepressants for six weeks or longer skipped the next question about add-on/augmentation medications. This means they have missing data on the MSM.



In order to gain a score of two, a participant would need to have been scored as having “mild depression” on the PHQ9 (+1 to score) and then, in the MSM, they would need to have reported that their current or most recent depressive episode had lasted less than one year (+1 to score). Additionally, they would have needed to score zero on the further two questions about their current or most recent depressive episode, by reporting that they had taken zero antidepressants for six weeks or longer and had not been prescribed any augmentation/add-on medications (+0 to score). They also would have needed to report that they had never received electroconvulsive therapy (+0 to score). This combination of answers would gain a participant a score of two. Since we excluded participants who scored zero on the question about taking antidepressants for six weeks or longer, no participant was able to gain a score of two. Therefore, the MSM ranged 3-14.

Exclusions

Anhedonic symptoms in COPING NBR: We first excluded participants who had complete data on the AD-MASQ-D30 at fewer than three COPING survey time points. We excluded these participants because, as part of our analyses, we calculated the highest, lowest, and mean anhedonic symptoms across all available COPING surveys for each participant. Therefore, we wanted to ensure that the participants had sufficient data points in order to robustly calculate

these values. Following this, we excluded anyone who's genetic data did not pass the standard genotype quality control (QC) (more detail below).

Treatment-resistant depression in GLAD: We excluded anyone who had missing data on either the PHQ9 or the MSM. This included participants who were not currently depressed at the time of completing the GLAD survey (i.e., a PHQ9 severity score of 0/"None") and therefore did not complete the MSM. It also included participants who reported that they had not taken any antidepressants for six weeks or longer for their current or most recent depressive episode (these participants were not shown the next question about add-on/augmentation medications). Additionally, we wanted to maximise the chance that all individuals were affected by major depressive disorder and not bipolar disorder. *(Individuals were eligible to complete the MSM if they were currently depressed at the time of completing the GLAD study survey. This means that individuals with bipolar disorder, who happened to be in a depressive episode, answered the MSM).* In the GLAD study survey, there were two questions about whether the participant has received a diagnosis of bipolar disorder by a professional. We retained anyone who self-reported that they had NOT received a diagnosis in BOTH questions (i.e., we excluded anyone who answered "Yes", "Don't know", "Prefer not to answer", or had missing data for either of these two questions). Following this, we excluded anyone who's genetic data did not pass the standard genotype QC (more detail below).

Genetic analyses

Genotyping, imputation, and quality control (QC)

Genotyping

Quality assurance measures were calculated by ThermoFisher: samples with a dish quality control (QC) value ≥0.82 (capturing the resolution of true signal from background noise on the genotyping array) and an initial call rate ≥0.97 were retained. Variants were recommended for inclusion if they were genotyped with high resolution (classified as "PolyHighResolution", "NoMinorHom", or "MonoHighResolution" by ThermoFisher). Data passing quality assurance was transferred to the Social, Genetic, and Developmental Psychiatry Centre at King's College London for further QC, adapted from previous pipelines (Coleman *et al.*, 2016).

Initial QC

Data for GLAD and COPING NBR were processed separately following the same pipeline. An initial set of quality control was performed to determine sample ancestry. This consisted of excluding variants with a minor allele frequency (MAF) <0.01, variants and individuals with a call rate <95%, and variants with Hardy-Weinberg $p<10^{-10}$. Additional checks were performed on individuals to exclude outliers for sex discrepancies, heterozygosity, and relatedness. Samples were merged with data from Phase 3 of the 1000 Genomes project and principal component analyses were performed on genome-wide genotype data. Samples clustering with known individuals from European ancestries in the 1000 Genomes project formed the majority of the genotyped GLAD and COPING NBR cohorts (96% and 98% respectively; figure S4.3) and so further analyses were restricted only to these participants. QC was repeated, on raw data restricted to European ancestry participants. This comprised the same measures as above.

Imputation

For GLAD and NBR separately, high quality genotype data was lifted to build 38 of the human genome and imputed to TopMed freeze 8, using version 1.5.7 of the dedicated imputation server provided by the University of Michigan (Taliun *et al.*, 2021) with prior phasing using EAGLE2 (Loh *et al.*, 2016). Following imputation, data (in variant call format [VCF]) were restricted to variants with MAF≥0.001 and imputation R^2≥0.3. Post-imputation VCFs were updated to include sex information and rsIDs, which were collected from the Single Nucleotide Polymorphism Database, build 153.

GLAD & NBR merge

Data from GLAD and COPING NBR were merged post-imputation using *bcftools*, and converted to PLINK2 pfile format, retaining genotype dosage information (Chang *et al.*, 2015). Only bi-allelic SNPs were retained in the resulting merged pfiles. Post-merge, the data was filtered with a MAF threshold of 0.01 and a variant missingness of 0.02. Duplicate samples, related individuals with pi-hat>0.1875, and samples with mismatched sex were also excluded.

Analyses-specific QC

Both study samples were subjected to further QC before GWAS using REGENIE. This included removing genotyped SNPs if missingness >5%, MAF<0.01, or Hardy Weinberg $p<10^{-10}$. SNPs imputed with low confidence (INFO<0.3) were also excluded. Individuals with missingness >5%, a mismatch between their self-reported assigned sex at birth and genetic sex, or whose genetic sex could not be determined were excluded. Individuals who were more than three SDs away from the mean pairwise relatedness (identical-by-descent [IBD]) of the data set (henceforth "individual IBD outliers") were also excluded. Lastly, for the analysis of anhedonic symptoms, one participant of each pair of duplicated participants between the GLAD and NBR cohorts was excluded.

Supplementary results

Exclusions

Anhedonic symptoms in COPING: The initial sample included 26,411 individuals who had participated in the COPING survey. A total of 4,103 participants were excluded for having complete data on the AD-MASQ-D30 at fewer than three COPING survey timepoints. A further 8,343 were excluded for having no available genetic data. This left a total of 13,965 COPING participants. Of which, 289 were removed due to missing information about PCs, 16 were removed due to being duplicates between the GLAD and NBR cohorts, 85 were removed due to having a mismatch between their biological sex and self-reported sex or because their genetic sex could not be determined, and 158 individual IBD outliers were removed. This left a final sample size of 13,433.

Treatment-resistant depression in GLAD: The initial sample included 46,725 participants who participated in the GLAD study. A total of 4,455 and 13,363 participants were excluded for having missing data on the PHQ9 or the MSM respectively. A further 6,648 participants were excluded for not being currently depressed at the time of completing the MSM (i.e., a PHQ severity score of zero). Following this, 1,469 participants were excluded for self-reporting a diagnosis of bipolar disorder by a professional. A total of 12,201 were then excluded for having no available genetic data. This left a total of 8,652 GLAD participants. Of which, 330 were removed due to missing information about PCs, 27 were removed due to having a mismatch between their biological sex and self-reported sex or because their genetic sex could not be determined, and 130 individual IBD outliers were removed. This left a final sample size of 8,165.

Phenotypic analyses

Correlations between highest, lowest, and mean anhedonic symptoms

There were medium-to-high phenotypic correlations between the three measures of anhedonic symptoms and they all were significantly different to zero. The correlation between the participants' highest and lowest anhedonic symptoms was 0.69 (SE=0.004, p=0.0). The correlation between highest and mean anhedonic symptoms was larger at 0.89 (SE=0.002, $p=1.67x10^{-301}$). The correlation between lowest and mean anhedonic symptoms was larger again at 0.90 (SE=0.002, $p=1.60x10^{-214}$). A plot of these correlation coefficients is presented in **figure S4.2**.

Correlations between anhedonic symptoms, concurrent depression and anxiety symptoms, and the single anhedonia item from the PHQ9

We assessed whether there was an association between the participants' highest anhedonic symptoms and the depressive and anxiety symptoms they were experiencing at that time. These concurrent symptoms were assessed via the PHQ9 and GAD7 respectively. There were medium-to-high phenotypic correlations between the participants' highest anhedonic symptoms and concurrent depression and anxiety symptoms and both were significantly different to zero. Anhedonic symptoms were correlated with concurrent depression symptoms at 0.59 (SE=0.006, $p<2.22x10^{-16}$) and with concurrent anxiety symptoms at 0.49 (SE=0.008, $p<2.22x10^{-16}$). Depressive and anxiety symptoms were much more strongly correlated with each other at 0.82 (SE=0.003, $p=3.74x10^{-258}$).

We also assessed whether there was an association between two different measures of anhedonia: anhedonic symptoms measured via the AD-MASQ-D30 and the single anhedonia item from the PHQ9. This single item is based on the question "*Over the past two weeks, how often have you had little interest or pleasure in doing things?*". Participants could answer with "*Not at all*", "*Several days*", "*More than half the days*", and "*Nearly every day*" which were coded as 0, 1, 2, and 3 respectively. The correlation between the participants' highest anhedonic symptoms and their anhedonia score from the single PHQ9 item was 0.59 (SE=0.006, $p<2.22x10^{-16}$). Depression and anxiety symptoms were much more strongly correlated with the single

anhedonia item from the PHQ9: depressive symptoms were correlated at 0.86 (SE=0.002, $p<2.22x10^{-16}$) and anxious symptoms were correlated at 0.68 (SE=0.006, $p<2.22x10^{-16}$). A plot of these correlation coefficients is presented in **figure S4.3**.

Correlation between anhedonic symptoms and treatment-resistant depression

A total of 2,669 GLAD participants who were included in the analysis of treatment-resistant depression also took part in COPING and were included in the analysis of anhedonic symptoms. In these individuals, we calculated the correlation between their highest and mean anhedonic symptoms and their treatment-resistant depression sum score (to reflect the genetic correlations performed in our study). We found low but highly significant correlations. Highest anhedonic symptoms were correlated at 0.23 (SE=0.02, $p=1.70x10^{-32}$). Mean anhedonic symptoms were correlated at 0.29 (SE=0.02, $p=2.12x10^{-54}$).

Supplementary references

Chang, C.C. *et al.* (2015) 'Second-generation PLINK: rising to the challenge of larger and richer datasets', *GigaScience*, 4, p. 7.

Coleman, J.R.I. *et al.* (2016) 'Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray', *Briefings in functional genomics*, 15(4), pp. 298–304.

Fekadu, A. *et al.* (2009) 'A multidimensional tool to quantify treatment resistance in depression: the Maudsley staging method', *The Journal of clinical psychiatry*, 70(2), pp. 177–184.

Kroenke, K., Spitzer, R.L. and Williams, J.B. (2001) 'The PHQ-9: validity of a brief depression severity measure', *Journal of general internal medicine*, 16(9), pp. 606–613.

Loh, P.-R. *et al.* (2016) 'Reference-based phasing using the Haplotype Reference Consortium panel', *Nature genetics*, 48(11), pp. 1443–1448.

Taliun, D. *et al.* (2021) 'Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program', *Nature*, 590(7845), pp. 290–299.

Supplementary figures

Figure S4.1. Phenotypic correlation matrix between the participants' highest, lowest, and mean anhedonic symptoms.
Anhedonic symptoms were measured with the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30). The correlation matrix of Pearson's product-moment correlations was computed in R.

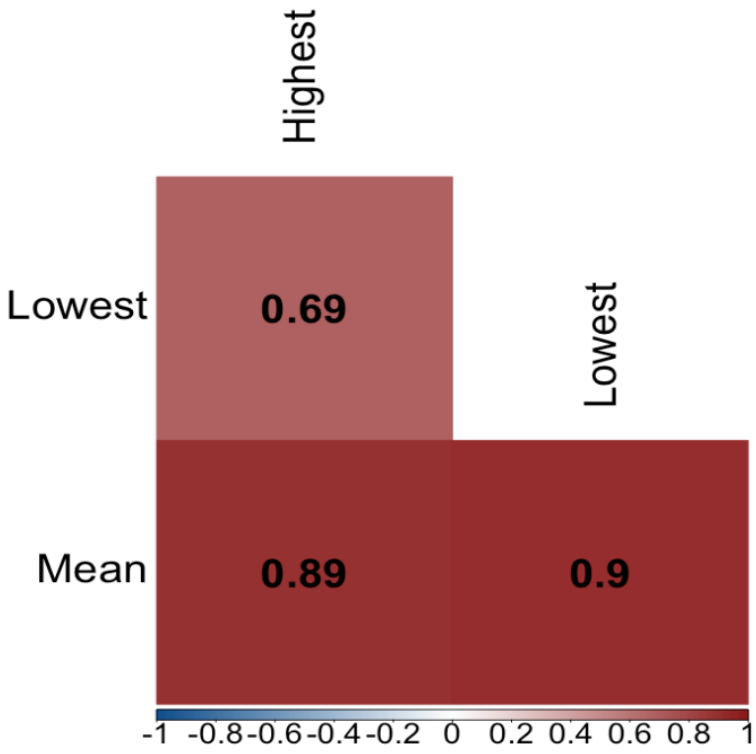
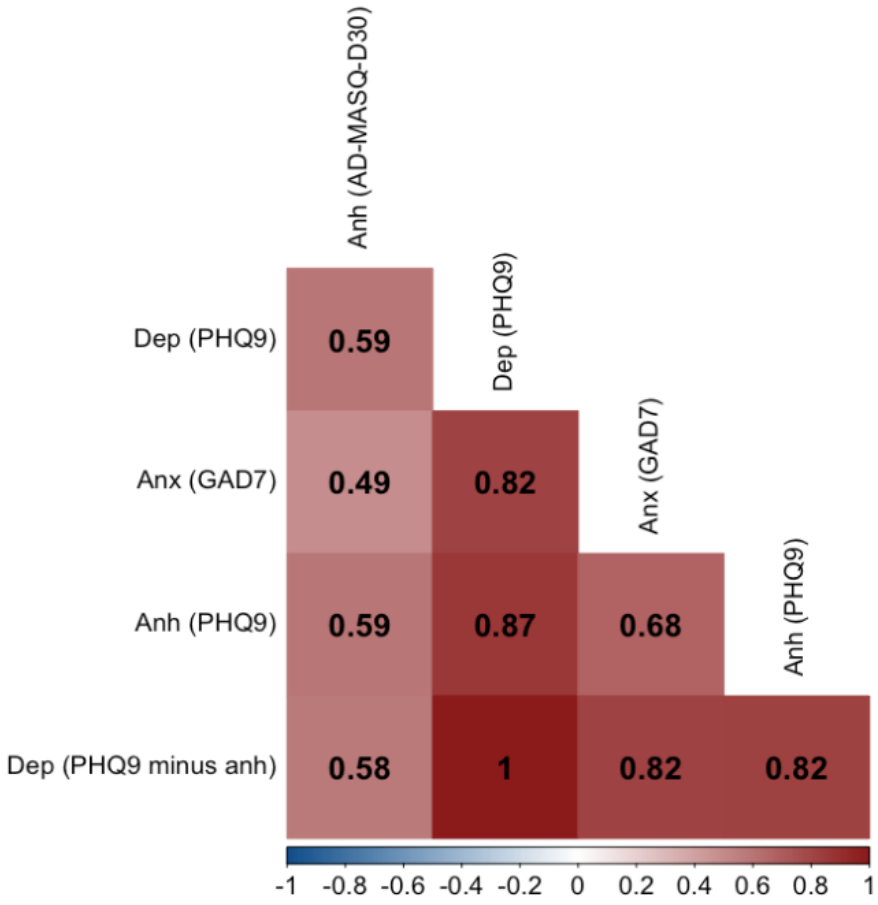


Figure S4.2. Phenotypic correlation matrix between the participants' highest anhedonic symptoms, their depressive and anxiety symptoms at that time (i.e., concurrent symptoms), and their concurrent score on the single anhedonia item from the nine item Patient Health Questionnaire (PHQ9).

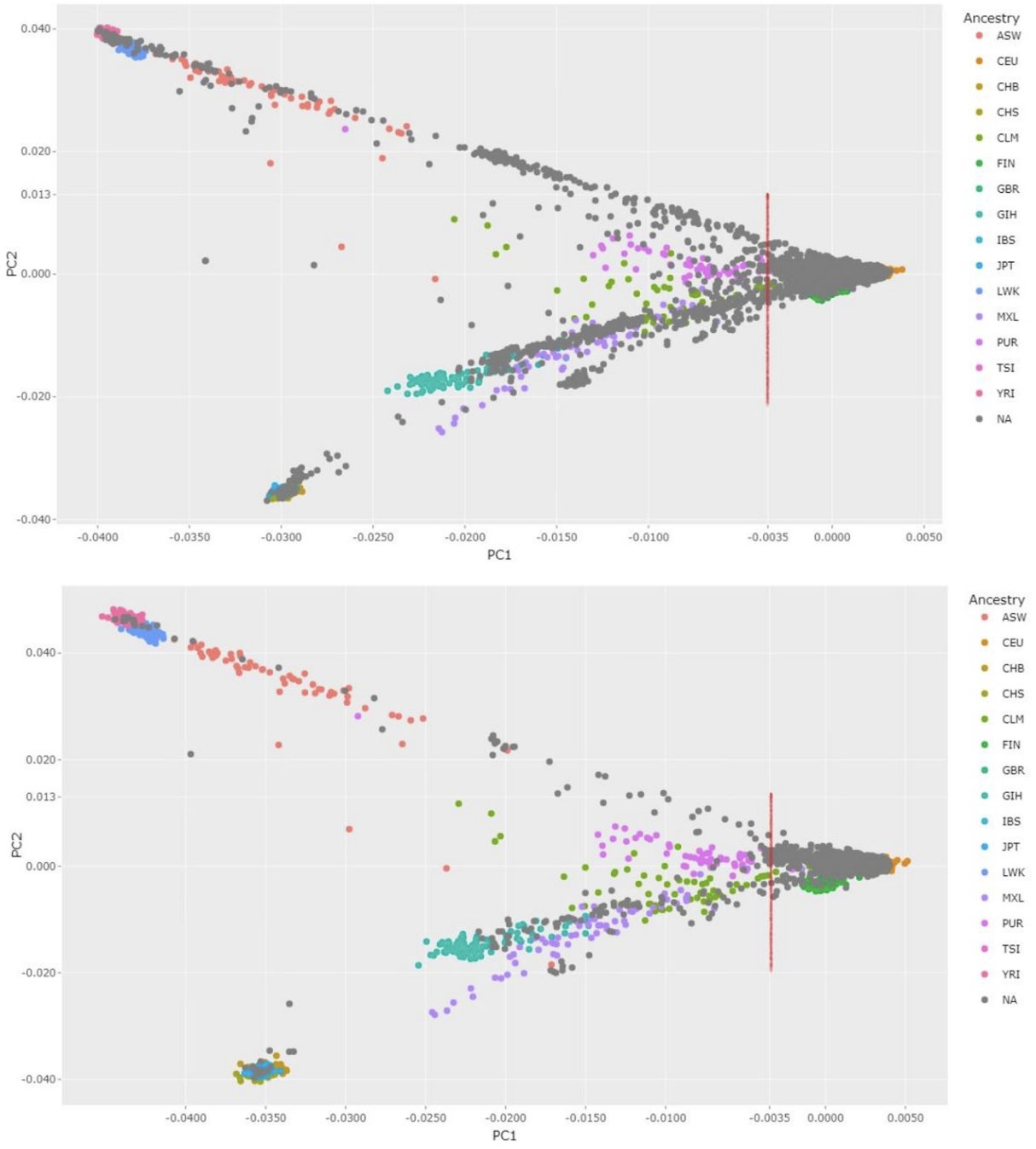
Current anhedonic symptoms were measured with the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30). Concurrent depressive symptoms were measured with the PHQ9 and concurrent anxiety symptoms were measured with the seven item Generalised Anxiety Disorder Questionnaire (GAD7). We also calculated the correlations with current depressive symptoms (measured via the PHQ9) with the anhedonia PHQ9 item removed. The correlation matrix of Pearson's product-moment correlations was computed in R.



Anh (AD-MASQ-D30) = Anhedonic symptoms assessed via AD-MASQ-D30
Dep (PHQ9) = Concurrent depressive symptoms assessed via PHQ9
Anx (GAD7) = Concurrent anxiety symptoms assessed via GAD7
Anh (PHQ9) = Concurrent anhedonia assessed via single item in PHQ9
Dep (PHQ9 minus anh) = Concurrent depression symptoms assessed via PHQ9 with single item anhedonia removed

Figure S4.3. Principal component analysis (PCA) plots.

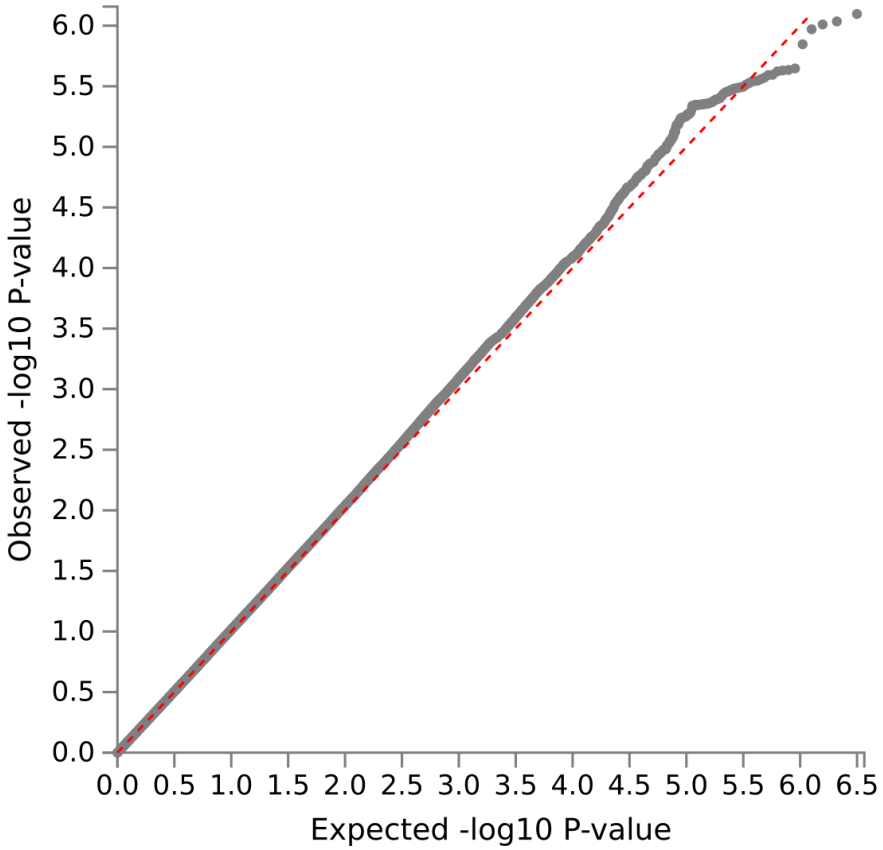
Above: principal component analysis (PCA) plot of genotype data of Genetic Links to Anxiety and Depression (GLAD) study participants. Below: PCA plot of genotype data of GLAD study participants and NIHR BioResource (NBR) study participants (merged data set). In both plots, the first principal component is plotted against the second principal component.



Note: ASW (African Ancestry in SW USA), CEU (Europeans, from Utah), CHB (Northern Han Chinese from Beijing), CHS (Southern Han Chinese, from Shanghai), CLM (Colombian in Medellín, Colombia), FIN (Finnish in Finland), GBR (Western Europeans from Britain), IBS

(Southern Europeans from Spain), JPT (Japanese in Tokyo, Japan), LWK (Luhya from Webuye, Kenya), MXL (Mexican ancestry in Los Angeles, CA, USA), PUR (Puerto Rican in Puerto Rico), TSI (Southern Europeans from Tuscany in Italy), YRI (Yoruba in Ibadan, Nigeria), GIH (Gujarati Indians in Houston, Texas, USA), NA (GLADv2 or COPING NBRv1)

Figure S4.4. Quantile-quantile (QQ) plot and Manhattan plot of Single Nucleotide Polymorphism (SNP) associations with the COVID-19 Psychiatry and Neurological Genetics (COPING) participants' maximum anhedonic symptoms (sum score) (N=13,433).
Anhedonic symptoms were measured by the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30) in COPING study participants with available genetic data that passed the standard genotype quality control (QC). The genome-wide association study (GWAS) was performed with REGENIE. We covaried for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA. The genome-wide significance threshold was set at $5x10^{-8}$.



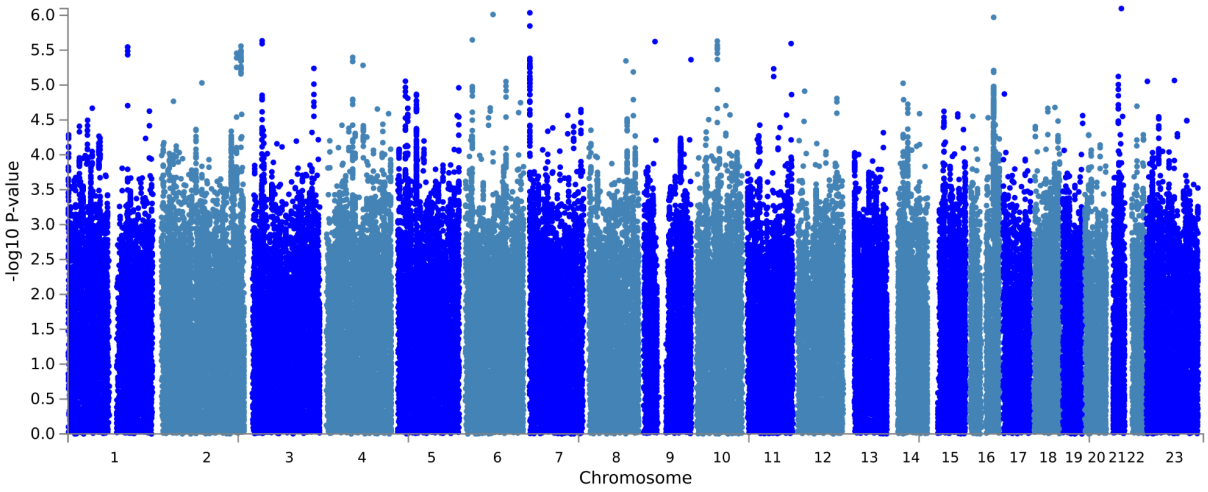
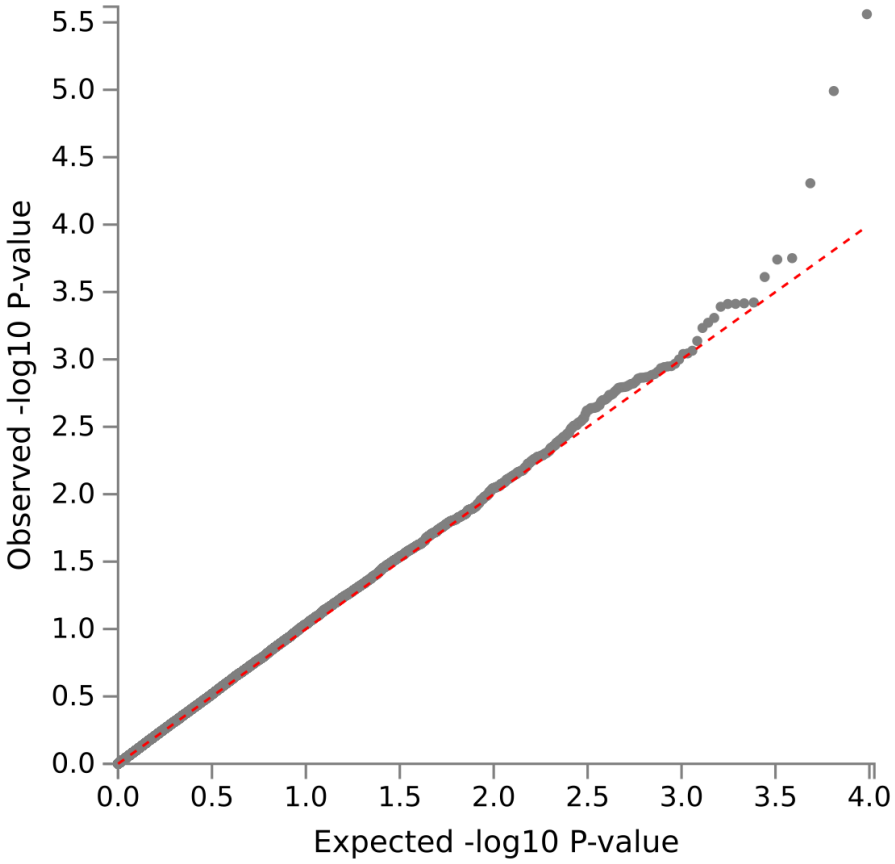

-log10 P-value
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Chromosome

Figure S4.5. Quantile-quantile (QQ) plot and Manhattan plot of gene associations with the COVID-19 Psychiatry and Neurological Genetics (COPING) participants' maximum anhedonic symptoms (sum score) (N=13,433).

Anhedonic symptoms were measured by the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30) in COPING study participants with available genetic data that passed the standard genotype quality control (QC). The genome-wide association study (GWAS) was performed with REGENIE. We covaried the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using MAGMA (implemented in FUMA). The input SNPs from the GWAS summary statistics were mapped to 19,397 protein coding genes. Genome-wide significance (red dashed line in the plot) was defined at P=0.05/19397= 2.578×10^{-6} to account for the 19,397 protein coding genes tested.



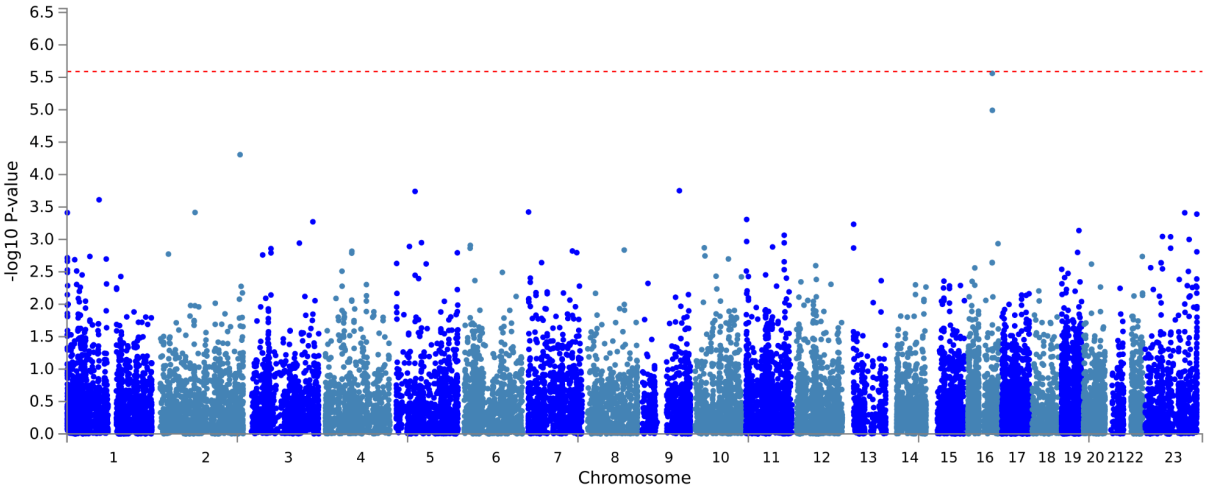
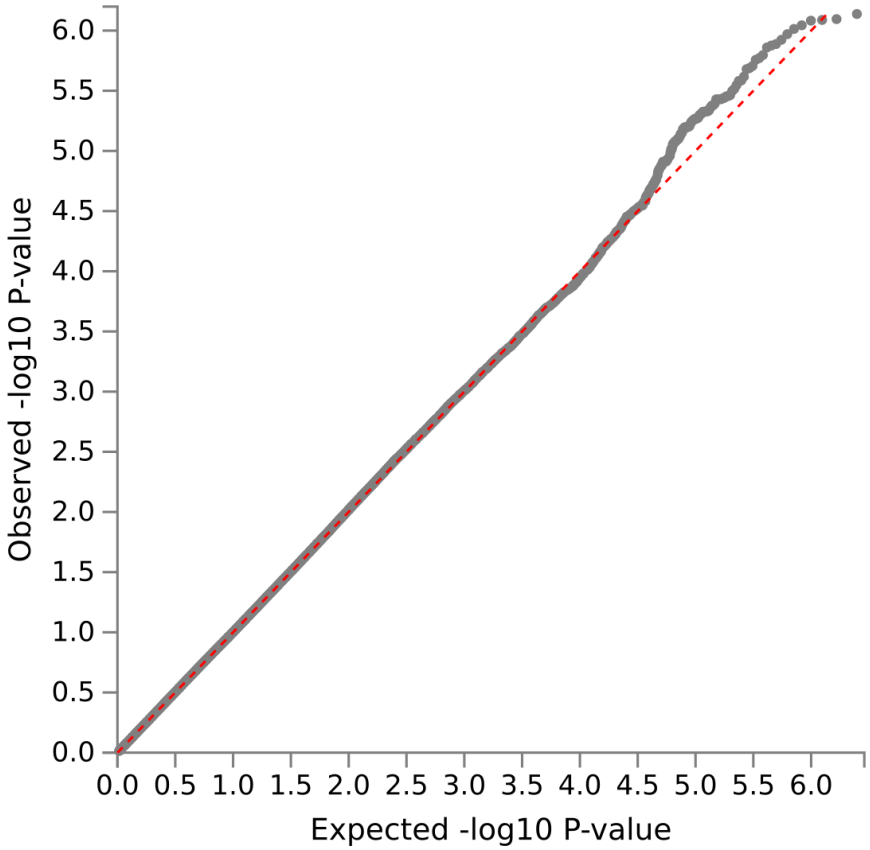

-log10 P-value
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Chromosome

Figure S4.6. Quantile-quantile (QQ) plot and Manhattan plot of Single Nucleotide Polymorphism (SNP) associations with the Genetic Links to Anxiety and Depression (GLAD) study participants' treatment-resistant depression sum score (N=11,234).

Treatment-resistant depression was measured with the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in GLAD participants with available genetic data that passed the standard genotype quality control (QC). The genome-wide association study (GWAS) was performed with REGENIE. We covaried for the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using FUMA. The genome-wide significance threshold was set at $5x10^{-8}$.



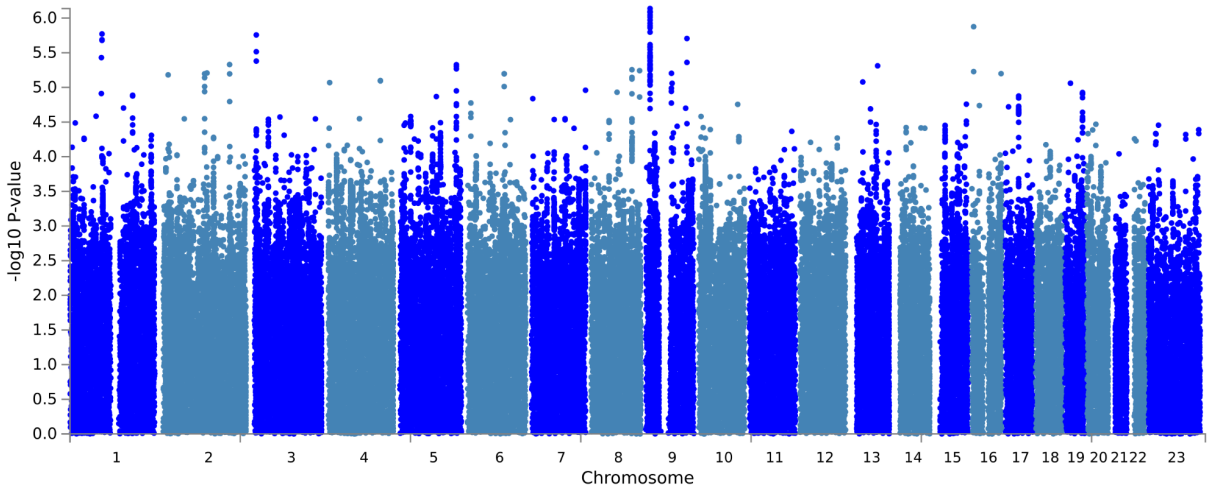
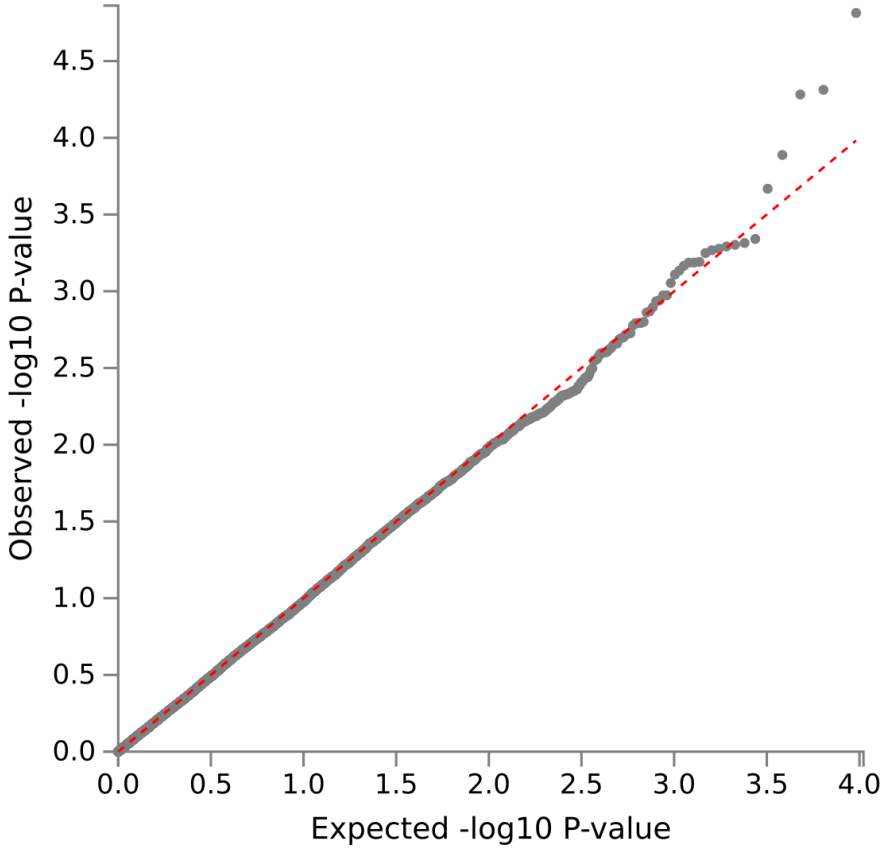

-log10 P-value
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Chromosome

Figure S4.7. Quantile-quantile (QQ) plot and Manhattan plot of gene associations with the Genetic Links to Anxiety and Depression (GLAD) study participants' treatment-resistant depression sum score (N=8,165).

Treatment-resistant depression was measured with the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in GLAD participants who had available genetic data that passed the standard genotype quality control (QC). The genome-wide association study (GWAS) was performed with REGENIE. We covaried the first ten ancestry principal components and genotyping batch. Manhattan and QQ plots were produced using MAGMA (implemented in FUMA). The input SNPs from the GWAS summary statistics were mapped to 19,188 protein coding genes. Genome-wide significance (red dashed line in the plot) was defined at P=0.05/19188= 2.606×10^{-6} to account for the 19,188 protein coding genes tested.



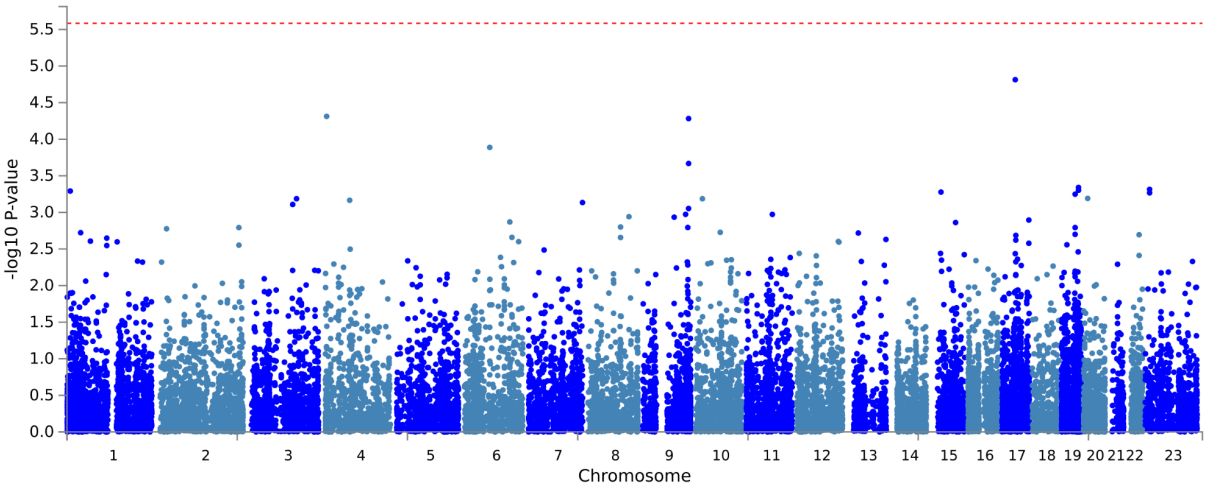
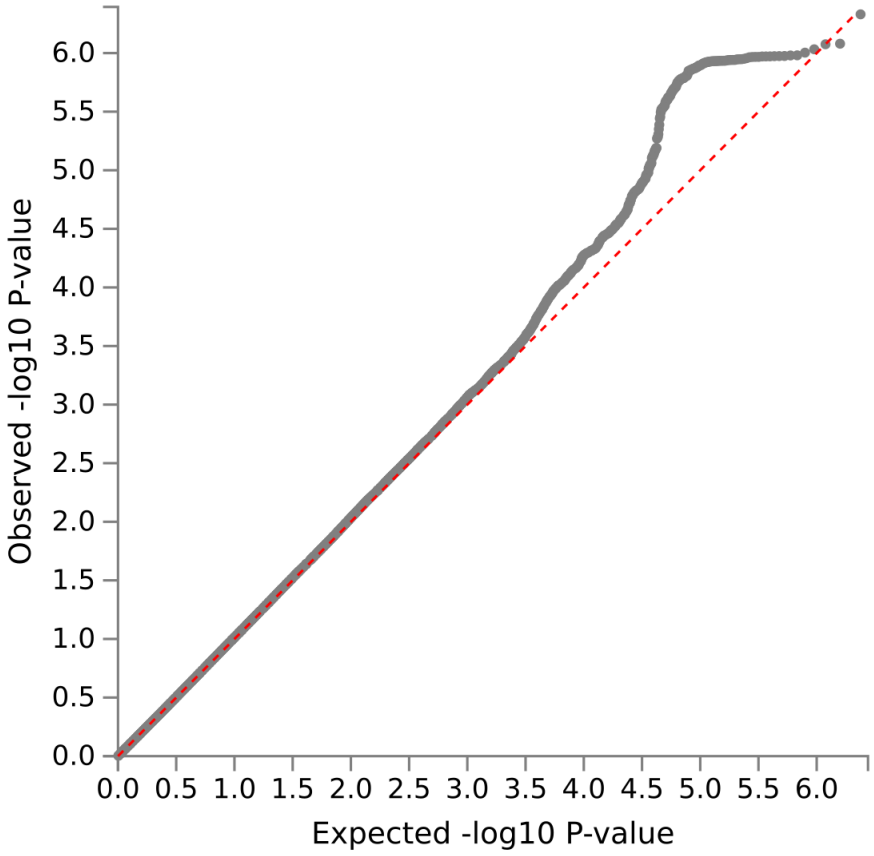

-log10 P-value
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Chromosome

Figure S4.8. Quantile-quantile (QQ) plot and Manhattan plot of Single Nucleotide Polymorphism (SNP) associations with the meta-analysed treatment-resistant depression phenotype (N=24,537).

Treatment-resistant depression was measured with the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in participants of the Genetic Links to Anxiety and Depression (GLAD) study (N=8,165). This was meta-analysed with the results of a genome-wide association study (GWAS) of treatment-resistant depression by Fabbri et al. (2021) (N=16,372). The total N for the meta-analysed phenotypes was 24,537. The meta-analysis was performed with the METAL software. Manhattan and QQ plots were produced using FUMA. The genome-wide significance threshold was set at $5x10^{-8}$.



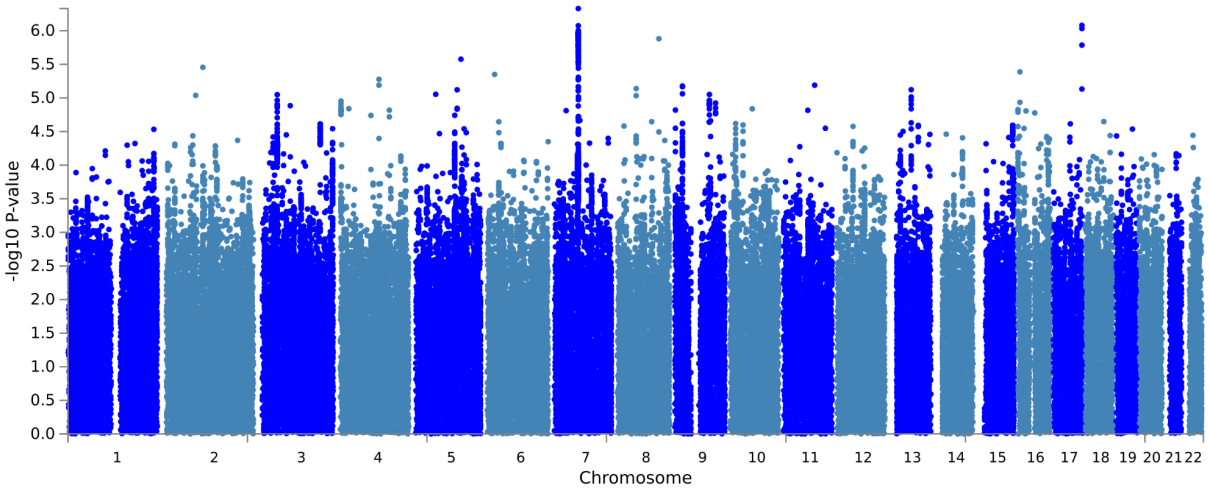
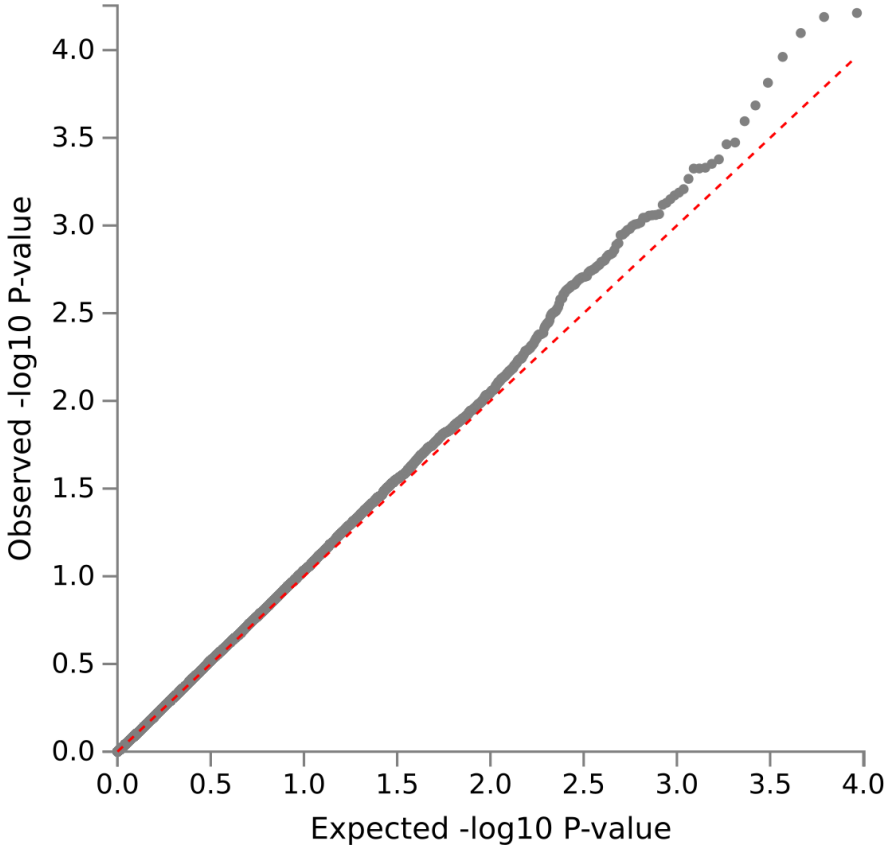

-log10 P-value
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome

Figure S4.9. Quantile-quantile (QQ) plot and Manhattan plot of gene associations with the meta-analysed treatment-resistant depression phenotype (N=24,537).

Treatment-resistant depression was measured with the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in participants of the Genetic Links to Anxiety and Depression (GLAD) study (N=8,165). This was meta-analysed with the results of a genome-wide association study (GWAS) of treatment-resistant depression by Fabbri et al. (2021) (N=16,372). The total N for the meta-analysed phenotype was 24,537. The meta-analysis was performed with the METAL software. Manhattan and QQ plots were produced using MAGMA (implemented in FUMA). The input SNPs from the GWAS summary statistics were mapped to 18,465 protein coding genes. Genome-wide significance (red dashed line in the plot) was defined at P=0.05/18465=2.708e-6 to account for the 18,465 protein coding genes tested.



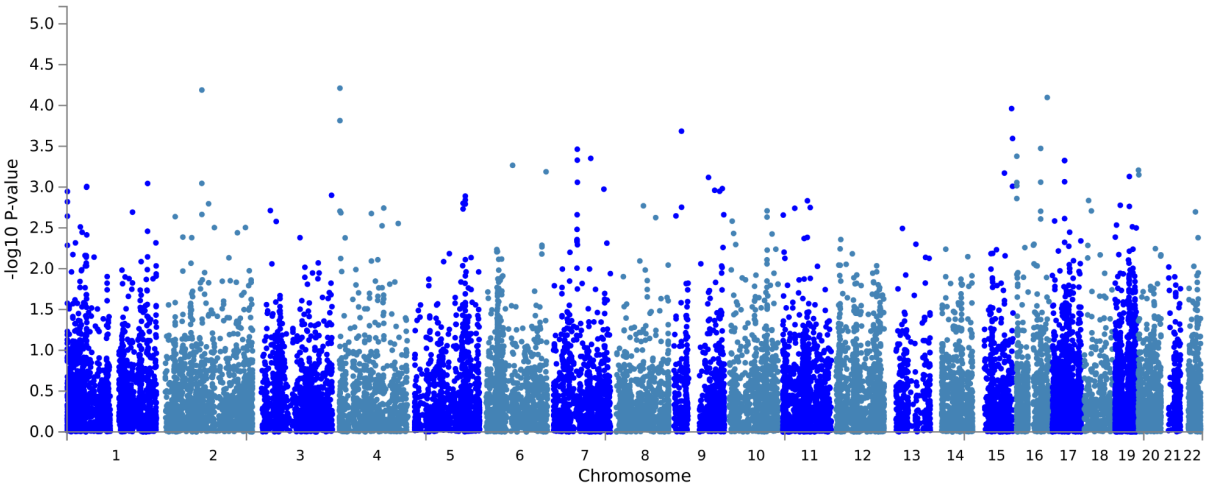

-log10 P-value
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome

Figure S4.10. Flow-chart of participant exclusions in the COVID-19 Psychiatry and Neurological Genetics (COPING) study sample.

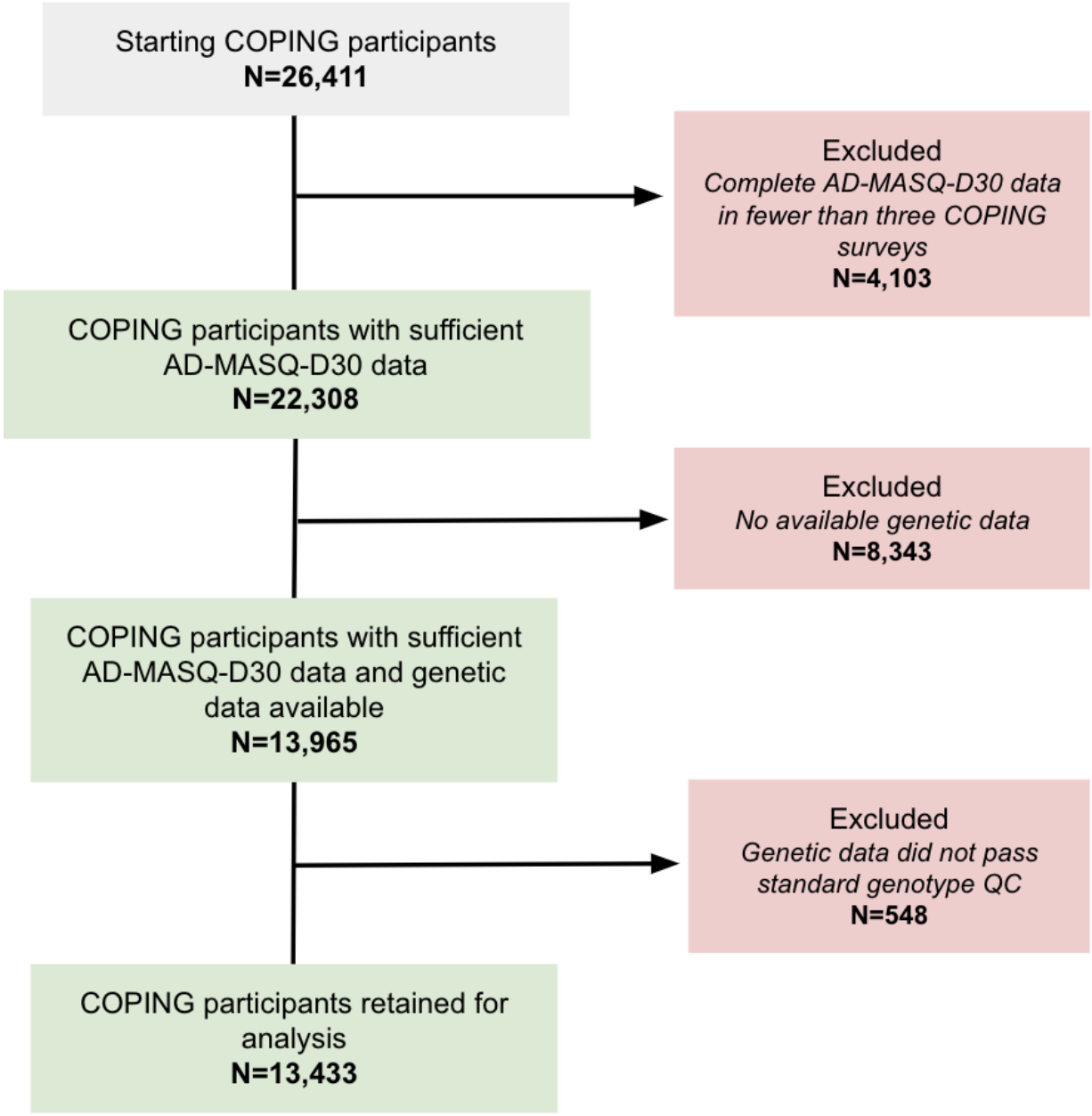
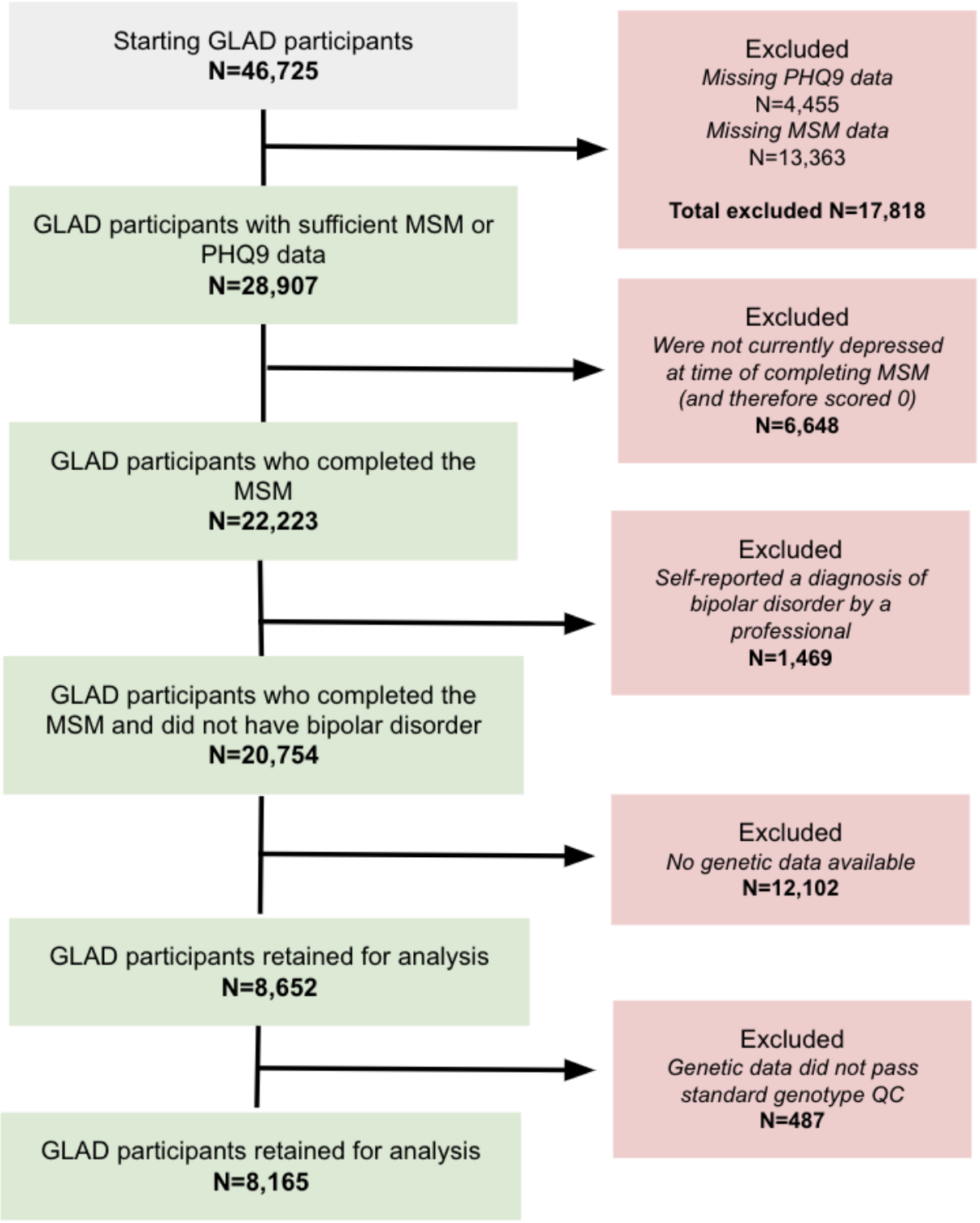


Figure S4.11. Flow-chart of participant exclusions in the Genetic Links to Anxiety and Depression (GLAD) study sample.



Supplementary tables

Table S4.1. Anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30) and questions included in the Maudsley Staging Method (MSM) to stage treatment-resistant depression (plus the measure of depression severity from the nine-item Patient Health Questionnaire [PHQ9]). In the AD-MASQ-D30, each item is preceded with the question *"Below is a list of feelings, sensations, problems, and experiences that people sometimes have. Read each item and then select the appropriate choice that best describes how much you have felt or experienced things this way during the past two weeks, including today."* The numeric values (reverse coded for AD-MASQ-D30) used to calculate the sum scores are shown in brackets next to the possible answers for each item in the questionnaire.

Items measuring positive affect in the AD-MASQ-D30	Answer values for each of the items included in the AD-MASQ-D30	Items measuring treatment-resistant depression in the MSM	Answer values for each of the items included in the MSM
Felt successful	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)	Severity of current depressive episode measured via the PHQ9	Mild (1) Moderate (2) Moderately severe (3) Severe (4)
Felt really happy	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)	How long ago did your current or most recent episode of depression or low mood begin?	Less than 1 year ago (1) 1-2 years ago (2) More than 2 years ago (3)
Felt optimistic	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)	During the current or most recent episode of depression or low mood, how many antidepressant medications have you taken for 6 weeks or longer?	None (0) One to two (1) Three to four (2) Five to six (3) Seven to ten (4) More than ten (5)
Felt like I was having a lot of fun	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)		

Felt like I accomplished a lot	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)	If individuals don't respond fully to antidepressants, doctors sometimes prescribe "add-on" or "augmentation" medications in addition to the antidepressants. During the current or most recent episode of depression or low mood, have you taken an add on medication for 6 weeks or longer?	No (0) Yes (1)
Felt like I had a lot to look forward to	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)		
Felt really 'up' or lively	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)	Have you ever received electroconvulsive therapy?	No (0) Yes (1)
Felt like I had a lot of energy	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)		
Felt really good about myself	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)		
Felt really talkative	Not at all (4) A little bit (3) Moderately (2) Quite a bit (1) Extremely (0)		
Total	0-40	**Total**	2-14

Table S4.2. Descriptive statistics of the three anhedonic symptoms (N=13,433). Sum scores were created by summing participant answers to the anhedonic depression subscale of the 30-item short adaptation of the Mood and Anxiety Symptoms Questionnaire (AD-MASQ-D30). Only participants with non-missing data at at least three COVID-19 Psychiatry and Neurological Genetics (COPING) survey time points were included. The mean sum score was computed by summing all non-missing sum scores for each participant and then dividing this total by the number of non-missing sum scores.

Statistic	Highest anhedonic symptoms	Lowest anhedonic symptoms	Mean anhedonic symptoms
Minimum	0	0	0
Maximum	40	40	40
Mean	31.7764461	18.6270379	18.6270379
Standard deviation	7.21378464	8.338824446	8.33882445
Quartile 1	28	13	20.1666667
Median	33	18	25.9375
Quartile 3	38	24	31.33333333
Interquartile range	10	11	11.1666667
Skewness	-0.9312074401	0.1787904	-0.3228312
Kurtosis	0.38878431	-0.4169021	-0.3969597

Table S4.3. Descriptive statistics of staged treatment-resistant depression (N=8,165). A sum score was created with a combination of answers to the Patient Health Questionnaire (PHQ9) which measures current depression severity and the Maudsley Staging Method (MSM) (both in the Genetic Links to Anxiety and Depression sign-up questionnaire). A detailed explanation of how this sum score was created is presented in the supplementary methods.

Statistic	Staged treatment-resistant depression
Minimum	3
Maximum	14
Mean	5.21518677
Standard deviation	1.68352672
Quartile 1	4
Median	5
Quartile 3	6
Interquartile range	2
Skewness	0.97772878
Kurtosis	1.22393656

Table S4.4. Genome-wide Complex Trait Analysis (GCTA) Single Nucleotide Polymorphism (SNP)-based heritability estimates (*h*2SNP), standard errors (SE), and z-scores of anhedonic symptoms and staged treatment-resistant depression. Anhedonic symptoms were measured in participants of the COVID-19 Psychiatry and Neurological Genetics (COPING) study and staged treatment-resistant depression was measured using the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in participants of the Genetic Links to Anxiety and Depression (GLAD) study. "N" refers to the number of individuals included in the genetic-relatedness matrix (GRM) used to calculate the *h*2SNP estimate. The *h*2SNP estimates were significantly different to zero if the p-value surpassed the Bonferroni-adjusted alpha of 0.025 (0.05/2) to correct for the two independent tests (and are shown in bold). P-values were calculated by GCTA.

Phenotype	Study sample	N	*h*2SNP	SE	Z-score	P-value
Highest anhedonic symptoms	COPING participants	13,271	0.108084	0.043656	2.475810885	**3.83E-03**
Lowest anhedonic symptoms	COPING participants	13,271	0.09435	0.045385	2.078880687	**2.19E-02**
Mean anhedonic symptoms	COPING participants	13,271	0.126469	0.044946	2.813798781	**1.96E-03**
Staged treatment-resistant depression	GLAD participants	8,062	0.025604	0.062608	0.4089573217	3.47E-01

Table S4.5. Linkage Disequilibrium Score Regression (LDSC) Single Nucleotide Polymorphism (SNP)-based heritability estimates (*h*2SNP), standard errors (SE), lambda GC, and mean chi-square statistic of anhedonic symptoms and staged treatment-resistant depression. Anhedonic symptoms were measured in participants of the COVID-19 Psychiatry and Neurological Genetics (COPING) study and staged treatment-resistant depression was measured using the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in participants of the Genetic Links to Anxiety and Depression (GLAD) study. "N" refers to the number of individuals included in the genome-wide association study (GWAS). The *h*2SNP estimates were significantly different to zero if the p-value surpassed the Bonferroni-adjusted alpha of 0.017 (0.05/3) to correct for the three independent tests (and are shown in bold). P-values were calculated in R using pchisq((h2/se)^2,1,lower.tail = FALSE) as recommended by LDSC developers.

Phenotype	Study sample	N	*h*2SNP	SE	Z-score	P-value	Lambda GC	Mean X2
Highest anhedonic symptoms	COPING participants	13,433	0.0763	0.0259	2.945945946	**0.003219687**	1.0135	1.017
Lowest anhedonic symptoms	COPING participants	13,433	0.0195	0.0242	0.805785124	0.4203668	1.0016	1.0096
Mean anhedonic symptoms	COPING participants	13,433	0.045	0.025	1.8	0.07186064	1.0105	1.0153
Staged treatment-resistant depression	GLAD participants	8,165	0.0499	0.0389	1.28277635	0.1995704	1.0075	1.0068
Meta-analysed GLAD-UK Biobank treatment-resistant depression	GLAD participants and participants from two UK cohorts	24,537	0.0242	0.014	1.728571429	0.08388583	1.0075	1.0158

Table S4.6. Genetic correlation between participants' mean anhedonic symptoms and staged treatment-resistant depression. The genetic correlation was estimated using the Genome-wide Complex Trait Analysis (GCTA) software (bivariate-REML). Anhedonic symptoms were measured in participants of the COVID-19 Psychiatry and Neurological Genetics (COPING) study and staged treatment-resistant depression was measured using the nine item Patient Health Questionnaire (PHQ9) and Maudsley Staging Method (MSM) in participants of the Genetic Links to Anxiety and Depression (GLAD) study. N=number of individuals included in esimation of GRM, *r*g=genetic correlation, SE=standard error, and p-value=p-value for *r*g difference from zero. Genetic correlations were significantly different to zero if the p-value surpassed p<0.05 (and are shown in bold). P-values were calculated by GCTA.

Phenotypes	N	*r*g	SE	P-value
Mean anhedonic symptoms and staged treatment-resistant depression	21,211	0.39733	0.61646	1.50E-01

Table S4.7. Information about each psychiatric and behavioural trait included in genetic correlations with maximum anhedonic symptoms (sum score). Genome-wide association study (GWAS) summary statistics of each trait were used to calculate genetic correlations using Linkage Disequilibrum Score Regression (LDSC) (Bulik-Sullivan et al. 2015). "Nca" refers to number of cases, "Nco" refers to number of controls, and "N" refers to overall sample size.

Phenotype	Published paper	Nca	Nco	N
Attention deficit hyperactivity disorder	Demontis et al. (2019)	19099	34194	53293
Alcohol dependence	Walters et al. (2018)	11569	11569	46568
Daily alcohol use	Schumann et al. (2016)			70460
Alzheimer's disease	Jansen et al. (2019)	71880	383378	455258
Anhedonia	Ward et al. (2019)			375,275
Anorexia nervosa	Watson et al. (2019)	16992	55525	73050
Antidepressant response (% improvement)	Pain et al. (2022)			5218
Antidepressant response (non-remission vs. remission)	Pain et al. (2022)	1852	3299	5151
Treatment-resistant depression	Fabbri et al. (2021)	2165	14207	16372
Anxiety	Purves et al. (2020)	25453	58113	83566
Autism spectrum disorder	Grove et al. (2019)	18381	27969	46350
Bipolar disorder	Mullins et al. (2021)	41,917	371,549	413466
Bipolar disorder type I	Mullins et al. (2021)	25060	449978	475038
Bipolar disorder type II	Mullins et al. (2021)	6781	364075	370856
Body mass index	Hübel et al. (2019)			353972
Cannabis use (lifetime)	Stringer et al. (2016)	14374	17956	32330
Chronotype	Jones et al. (2016)			128266
Sleep duration	Jones et al. (2016)			128266
Oversleeper	Jones et al. (2016)			128266
Undersleeper	Jones et al. (2016)			128266
Major depressive disorder (PGC2 including 23andme)	Wray et al. (2018)	154649	394409	549058

Depressive symptoms	Okbay et al. (2016)			161460
Major depressive disorder (PGC2 excluding 23andme)	Wray et al. (2018)	59851	113154	173005
Years of education	Lee et al. (2018)			766345
Self-rated health	Harris et al. (2017)			111483
Household income	Hill et al. (2016)			112151
Insomnia	Hammerschlag et al. (2017)	32384	80622	113006
Cognitive ability	Savage (2018)			269867
Neuroticism	Hübel et al. (2019)			
Obsessive compulsive disorder	International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OCGAS) (2018)	2688	7037	9725
Physical activity	NA			66224
Posttraumatic stress disorder	Nievergelt et al. (2019)	32428	174227	206655
Posttraumatic stress disorder (military)	Stein et al. (2020)	36301	178107	214408
Posttraumatic stress disorder symptoms (military)	Stein et al. (2020)			186689
General risk tolerance (self-report)	Linner et al. (2019)			466571
Automobile speeding propensity	Linner et al. (2019)			404291
Number of sexual partners	Linner et al. (2019)			370711
Schizophrenia	Pardinas et al. (2018)	11260	24542	35802
Ever smoker	Linner et al. (2019)			518,663
Subjective well-being	Okbay et al. (2016)			298420
Staged treatment-resistant depression	Internal (GLAD participants in this study)			11,234
Meta-analysed GLAD-UK Biobank treatment-resistant depression	GLAD meta-analysed with Fabbri et al. (2021)			27,606

Table S4.8. Genetic correlations between participants' highest anhedonic symptoms and 42 psychiatric and behavioural traits. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the extended 1000 Genomes Linkage Disequilibrium reference panel. *r*g=genetic correlation, SE=standard error, and p-value=p-value for *r*g difference from zero (diff 0) or one (diff 1). Genetic correlations were significantly different to zero or one if the p-value surpassed the Bonferroni-adjusted alpha of 0.001 (0.05/42) to correct for the 42 tests (traits with genetic correlations that were significantly non-zero are shown in bold).

Phenotype	*r*g	SE	P-value (diff 0)	P-value (diff 1)
Attention deficit hyperactivity disorder	0.04403	0.1114	6.93E-01	**0.00E+00**
Alcohol dependence	0.3421	0.2624	1.92E-01	**1.22E-02**
Daily alcohol use	0.1812	0.2355	4.42E-01	**5.07E-04**
Alzheimer's disease	0.07512	0.135	5.78E-01	**7.34E-12**
Anhedonia (UK Biobank)	**0.4532**	**0.1171**	**1.08E-04**	**3.02E-06**
Anorexia nervosa	0.08617	0.1373	5.30E-01	**2.82E-11**
Antidepressant response (% improvement)	NA	NA	NA	NA
Antidepressant response (non-remission vs. remission)	-0.08444	0.32	7.92E-01	**7.02E-04**
Treatment-resistant depression	0.3152	0.4393	4.73E-01	1.19E-01
Anxiety	**0.5023**	**0.1355**	**2.10E-04**	**2.40E-04**
Autism spectrum disorder	0.09851	0.1433	4.92E-01	**3.16E-10**
Bipolar disorder	0.09103	0.08871	3.05E-01	**0.00E+00**
Bipolar disorder I	0.1055	0.09954	2.89E-01	**0.00E+00**
Bipolar disorder II	-0.02199	0.1865	9.06E-01	**4.26E-08**
Body Mass Index (BMI)	0.0799	0.05949	1.79E-01	**0.00E+00**
Lifetime cannabis use	0.04198	0.2075	8.40E-01	**3.89E-06**
Chronotype	-0.05809	0.08813	5.10E-01	**0.00E+00**
Sleep duration	-0.2204	0.1068	3.90E-02	**0.00E+00**
Oversleeper	0.1064	0.1677	5.26E-01	**9.90E-08**
Undersleeper	0.2122	0.1419	1.35E-01	**2.83E-08**
Major depressive disorder (including 23andme)	0.4204	0.1469	4.22E-03	**7.96E-05**
Depressive symptoms	**0.7094**	**0.2087**	**6.75E-04**	1.64E-01

Major depressive disorder (excluding 23andme)	0.2831	0.1095	9.74E-03	**5.87E-11**
Years of education	-0.1346	0.05122	8.57E-03	**0.00E+00**
Self-rated health	**-0.4214**	**0.121**	**4.98E-04**	**0.00E+00**
Household income	-0.325	0.1307	1.29E-02	**0.00E+00**
Insomnia	0.4452	0.1807	1.38E-02	**2.14E-03**
Cognitive ability	-0.1476	0.06715	2.79E-02	**0.00E+00**
Neuroticism	**0.459**	**0.1026**	**7.60E-06**	**1.34E-07**
Obsessive Compulsive Disorder	0.2734	0.1906	1.51E-01	**1.38E-04**
Physical activity	-0.2261	0.1048	3.10E-02	**0.00E+00**
Posttraumatic stress disorder	0.899	0.344	8.97E-03	**7.69E-01**
Posttraumatic stress disorder (military)	0.2392	0.1373	8.14E-02	**3.00E-08**
Posttraumatic stress disorder symptoms (military)	0.4259	0.1331	1.37E-03	**1.61E-05**
General risk tolerance	-0.09922	0.07649	1.95E-01	**0.00E+00**
Automobile speeding propensity	**-0.2833**	**0.0835**	**6.90E-04**	**0.00E+00**
Number of sexual partners	0.07351	0.06485	2.57E-01	**0.00E+00**
Schizophrenia	0.08304	0.08437	3.25E-01	**0.00E+00**
Ever smoker	0.1558	0.06789	2.17E-02	**0.00E+00**
Subjective wellbeing	-0.7129	0.2927	1.49E-02	**4.86E-09**
Staged treatment-resistant depression	0.1014	0.3373	7.64E-01	**7.72E-03**
Meta-analysed GLAD-UK Biobank treatment-resistant depression	0.2633	0.3237	4.16E-01	2.29E-02

NA = LDSC SNP-based heritability was out of bounds

Table S4.9. Genetic correlations between anhedonia in the UK Biobank (Ward et al. 2019) and five psychiatric and behavioural traits which were significantly genetically correlated with anhedonic symptoms in the COVID-19 Psychiatry and Neurological Genetics (COPING) study, three measures of treatment-resistant depression, and two measures of antidepressant response. Genetic correlations were estimated using Linkage Disequilibrium Score Regression (LDSC) and the 1000 Genomes Linkage Disequilibrium reference panel. *r*g=genetic correlation, SE=standard error, and p-value=p-value for *r*g difference from zero (diff 0) or one (diff 1). Genetic correlations were significantly different to zero or one if the p-value surpassed the Bonferroni-adjusted alpha of 0.005 (0.05/10) to correct for the 11 tests (traits with genetic correlations that were significantly non-zero are shown in bold).

Phenotype	*r*g	SE	P-value (diff 0)	P-value (diff 1)
Antidepressant response (% improvement)	NA	NA	NA	NA
Antidepressant response (non-remission vs. remission)	0.006463	0.1226	0.958	5.55E-16
Treatment-resistant depression	0.5744	0.5336	0.2817	4.25E-01
Anxiety	0.5962	0.03814	**4.37E-55**	0.00E+00
Depressive symptoms	0.9596	0.03077	**1.80E-213**	1.89E-01
Self-rated health	-0.6949	0.03793	**5.61E-75**	0.00E+00
Neuroticism	0.7044	0.02171	**7.20E-231**	0.00E+00
Automobile speeding propensity	-0.1851	0.03111	**2.70E-09**	0.00E+00
Staged treatment-resistant depression	0.2313	0.1737	0.1828	9.62E-06
Meta-analysed GLAD-UK Biobank treatment-resistant depression	0.3773	0.149	0.01131	2.93E-05

NA = LDSC SNP-based heritability was out of bounds

Table S4.10. Sub-cohorts of the National Institute for Health and Care Research (NIHR) BioResource who took part in the COVID-19 Psychiatry and Neurological Genetics (COPING) study who were included in the anhedonic symptoms study sample (N=7,843)

	N	Recruitment methods	Eligibility criteria	Recruitment area
Inflammatory Bowel Disease (IBD) cohort	573	IBD clinics in participating hospitals across the United Kingdom	16+, have a diagnosis of Crohn's disease, ulcerative colitis, indeterminate colitis, IBD type unspecified, or suspected IBD	England, Wales, Scotland, Northern Ireland
NHS Blood and Transplant studies (COMPARE, STRIDES, INTERVAL)	4,997	Blood donation centres	16+, live in England	England
Research Tissue Bank - Generic	2,273	Biomedical Research Centres, Clinical Research Facilities, hospital clinics, community recruitment, online	16+, live in England	England
Other	0	Recruitment from BioResource at King's College London	16+, live in England	England
Total	7,843			

Appendix 5. Perceptions of Psychiatric Risk (PerPsych) project

Open Science Framework pre-registration

GLAD lived experience. Jessica Mundy, Helena L. Davies, Molly R. Davies, Evangelos Vassos, Janet Treasure, Jehannine Austin, Danielle Dick, Morgan Driver, Thalia C. Eley, Gerome Breen

EDGI lived experience. Helena L. Davies, Jessica Mundy, Molly R. Davies, Evangelos Vassos, Janet Treasure, Jehannine Austin, Danielle Dick, Morgan Driver, Thalia C. Eley, Gerome Breen

Mental health professionals, trainees and students: Helena L. Davies*, Jessica Mundy*, Karla Mohoric, Molly R. Davies, Evangelos Vassos, Janet Treasure, Jehannine Austin, Danielle Dick, Morgan Driver, Thalia C. Eley, Gerome Breen
**Joint first author*

The PerPsych project encompasses two studies which will run in parallel to each other:

1) PerPsych: Lived experience (GLAD & EDGI)
2) Perpsych: Mental health professionals, students, and trainees

All data will be cleaned by Jessica Mundy, Helena Davies, and Karla Mohoric (all King's College London students). All data cleaning and analysis scripts will be uploaded to an online GitHub repository.

Ethical approval

The PerPsych project was approved by King's College London Research Ethics Management Application System committee on 15th July 2021 (reference number: HR/DP-20/21-22019).

PerPsych: Lived experience

Rationale

Thousands of genetic variants and environmental factors contribute to risk for psychiatric disorders (Tsuang *et al.*, 2004; Dick, 2011) therefore, unlike for Mendelian disorders, it is not possible for current genetic testing to confirm a psychiatric diagnosis (Palk *et al.*, 2019). Accordingly, communicating information about psychiatric genetic risk is challenging, especially to the general public. Genetic knowledge amongst the general public is lacking (Haga *et al.*, 2013); after almost a century of educational efforts, studies show that both school children and adults lack an accurate understanding of Mendelian genetics (Richards, 1996), as well as the role of additive genetics in complex traits (Condit and Shen, 2011).

Furthermore, the general public often erroneously evaluate the chances of future children developing a psychiatric illness as higher than what is suggested by their family history, which is indicative of a lack of understanding of complex genetics. This overestimation can influence the decision to have children (Meiser *et al.*, 2007; Austin, Hippman and Honer, 2012), with overestimation being correlated with the increasing likelihood of deciding against having children (Austin, Smith and Honer, 2006). In addition to this, individuals with psychiatric illness and their families often experience negative emotional responses, such as shame, which are heavily tied up with misconceptions about the cause of their illness (Inglis, Morris and Austin, 2017), as well as feelings of guilt or a heavy burden of personal responsibility (Corrigan *et al.*, 2002).

For these reasons, accurate and accessible information about the causes of psychiatric disorders needs to be communicated, particularly to individuals who are affected by mental illness. In a recent paper by Lewis and Vassos (2020) that evaluated the possibility of involving genetics in psychiatric clinical work, the authors outlined that "*Education for clinicians and the public will be necessary to increase understanding and genetic literacy" (Lewis and Vassos, 2020)*. The proposed work aims to take a step towards achieving this amongst individuals with depression, anxiety, and eating disorders.

Proposed work

The PerPsych (**Per**ceptions of **Psych**iatric Risk) project seeks to understand more about the way in which individuals with mental health disorders think about the contribution of genetic and environmental factors to their illness and how this impacts their emotions and behaviours, as well as the way information about psychiatric genetics could be communicated accurately and effectively to them. The PerPsych project will recruit participants from the Genetic Links to Anxiety and Depression (GLAD) Study and the Eating Disorders Genetics Initiative (EDGI). Both studies are part of the National Institute for Health Research (NIHR) Bioresource and have been developed to examine contributions of genetic and environmental factors to psychiatric disorders, specifically anxiety, depression, anorexia nervosa, bulimia nervosa, and binge-eating disorder. At sign-up, participants consent to be recontacted in the future about involvement in new research.

The project will be conducted entirely online. First, participants from the GLAD Study and EDGI will be invited to participate via email from the respective study teams. Individuals who consent to be contacted by the PerPsych team will be sent an email containing the full study information sheet and consent form.

Second, individuals who consent to participate in the PerPsych study will be taken to the **baseline** survey which consists of questions about current symptoms, medication and therapy, self-stigma and general mental health stigma, recent behaviours, belief in genetic determinism, genetic knowledge, experience of genetic counselling and perceptions of genetic and environmental contribution to their disorder.

Third, a randomly-selected half of the GLAD participants will view a short, animated video that explains the contribution of genes and environment to **anxiety** and **depression (the 'case'**

group) using the mental health jars' analogy developed by Professor Jehannine Austin. The video explains that having a mental health disorder is nobody's "fault" and that **feeling better** and managing one's mental health is possible. The other half of the GLAD participants will be invited to watch a control video **(the 'control' group)** developed by Mind (the mental health charity) that discusses causes of mental ill health more generally but does not mention the role of genetics. This animated video is of similar length and style to the case video.

Similarly, a randomly-selected half of the EDGI participants will view an almost identical video, with the major difference being that this video will explain the contribution of genes and environment to **eating disorders** and will explain that **recovery** is possible. Again, we will invite the other half of the participating EDGI participants to watch the control video.

Fourth, after participants have watched the video, they will then be presented with **immediate follow-up questions** which will measure whether participants' responses have changed (perceptions of genetic and environmental contributions to their disorder, self-stigma and general mental health stigma). They will also be asked whether they found the video helpful in any way.

Finally, after two weeks, participants will be sent a shorter survey **(follow up 2)** and again at a 1 month follow-up **(follow up 3)**.

<u>Questions included in the study:</u>

- 7-item version of Eating Disorder examination questionnaire (EDE-Q-7) (Grilo *et al.*, 2013) with the six bingeing and purging items from the EDE-Q 6.0 (Luce and Crowther, 1999) [to measure eating behaviour over the past 14 days] (EDGI only)
- Patient health questionnaire (PHQ-9) to measure current depressive symptoms (Kroenke and Spitzer, 2002) (GLAD only)
- Generalised anxiety disorder screener (GAD-7) to measure current anxiety symptoms (Spitzer *et al.*, 2006) (GLAD only)
- Public Understanding and Attitudes towards Genetics and Genomics (PUGGS) questionnaire (Carver *et al.*, 2017)
    - Selected questions to assess belief genetic in genetic determinism
- International Genetic Literacy and Attitudes Survey (iGLAS) (Chapman *et al.*, 2017)

    - Selected questions to assess genetic knowledge
- Perceptions of causes of mental illness
    - Question to assess perceived cause of mental illness, adapted from (Kalb *et al.*, 2017)) and (Michael *et al.*, 2020))
- Feelings related to disorder/impact on living adapted from (Kalb *et al.*, 2017)) and (Michael *et al.*, 2020))
- Pursuit of information about risk for mental disorders adapted from (Kalb *et al.*, 2017)) and (Michael *et al.*, 2020))
- Measures of self-care behaviours (adapted from the Repeated Assessment of Mental Health in Pandemics [RAMP] baseline questionnaire)
- Measures of concern for family and perceived level of support from the healthcare system (developed by a genetic counsellor task force to help people self identify if they might benefit from meeting with a genetic counsellor)
- Recent treatment/medication
    - Question about whether participants have ever been prescribed medication/enrolled in therapy for their disorder
    - Question about whether participants have recently started a new treatment for their disorder
    - Question about whether participants have spoken to a healthcare professional about starting a new treatment for their disorder
- Experience of genetic counselling adapted from (Kalb *et al.*, 2017)) and (Michael *et al.*, 2020))
    - Personal experience of genetic counselling
- Opinions on whether psychiatric genetic counselling would be helpful

Aims

We aim to:

- Understand how individuals with current or past depression, anxiety, and eating disorders evaluate the causes of their own disorder (i.e., perceptions of genetic and environmental contributions to their own mental health)
- Understand how this evaluation is associated with emotional wellbeing, stigma and behaviours

- Investigate whether exposure to a short, informative video about the way in which genes and environment combine to produce psychiatric disorder risk can improve emotional wellbeing and lead to re-evaluation of one's psychiatric disorder and the adoption of self-care behaviours in the short- and long-term

The case video has been created in collaboration with Professor Danielle Dick and the ALTLab at Virginia Commonwealth University (VCU) who previously created an educational, animated video about substance use disorders (which can be found here: https://rampages.us/coga/). The GLAD and EDGI videos have been adapted from this video by PhD students Jessica Mundy and Helena Davies.

Preliminary Work

No preliminary work has been carried out and there are no plans for pilot work.

Datasets

GLAD. The Genetic Links to Anxiety and Depression (GLAD; https://gladstudy.org.uk) Study was launched in September 2018 by the National Institute for Health Research. The GLAD Study is ongoing and aims to collect genetic and phenotypic data from 40,000 participants with lifetime anxiety and/or depression as defined in the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5). GLAD is a re-contactable data resource and participants are recruited via advertising on social and traditional media channels in addition to clinical recruitment through National Health Service (NHS) organisations. Additional details of the design and implementation of the GLAD Study are described elsewhere (Davies *et al.*, 2019).

EDGI. The Eating Disorders Genetics Initiative (EDGI; https://edgiuk.org/about/) Study was launched in February 2020 by the National Institute for Health Research. EDGI is ongoing and aims to collect genetic and phenotypic data from 10,000 participants with lifetime eating disorder as defined in the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5). EDGI is a re-contactable data resource and participants are recruited via advertising on social and traditional media channels in addition to clinical recruitment through National Health Service (NHS) organisations.

Descriptives

First, we will describe the attrition of the overall sample (number of participants at baseline versus each follow-up [immediate follow-up, follow up 2, and follow up 3]) as well as their demographic characteristics (i.e., age, sex, gender) at each phase. Next, we will create a heterogeneous correlation matrix to assess the relationship between all variables of interest, e.g., overall higher genetic attribution to risk than environmental (i.e., binary variable of any answer of "Only genetic factors", "Almost only genetic factors", "Mainly genetic factors" or "Slightly more genetic factors"), equal genetic and environmental attribution to risk (i.e., answers of "Equal genetic and environmental factors"), overall higher environmental attribution to risk than genetic (i.e., binary variable of any answer of "Only environmental factors", "Almost only environmental factors", "Mainly environmental factors" or "Slightly more environmental factors"), genetic knowledge score, and being female.

We will also generate plots for the following:

- Genetic knowledge measured by the 6-item iGLAS (Chapman et al., 2017)
    - Plot of percentage correct (baseline only)
        - NB: Any questions which have been skipped will be counted as incorrect.
- Belief in genetic determinism measured by a subset of the PUGGS questionnaire (Carver et al., 2017)
    - Histogram of continuous score (baseline only)
- Percentage of participants who endorsed each answer option to the question: "Please select an option according to how much you think your disorder was caused by inherited genetic factors and/or environmental factors" at baseline, immediate follow-up, follow up 2 (two week follow-up), and follow up 3 (one month follow-up).
- Percentage of participants who endorsed each answer option to each statement following the question: "This question is about your feelings towards mental health/mental ill health more generally. To what extent do you agree with the following statements:" at baseline, immediate follow-up, follow up 2, and follow up 3.
    - We will generate two plots for each time-point: participant in case versus the control video.
- Self-stigma (experienced and anticipatory)

    - Bar plot showing percentage of participants who endorsed each answer option category for each question (bars representing baseline, immediate follow up, follow up 2 and follow up 3) and case versus control
- Treatment adherence and self-care behaviours
    - Bar plot showing percentage of each category for every question (bars representing baseline, immediate follow up, follow up 2, and follow up 3) and case versus control

Research questions

In our papers, we will investigate **seven research questions** looking at the effectiveness of the case video in terms of 1) perceptions and beliefs, 2) stigma, 3) mental health, 4) treatment adherence and/or uptake, and 5) health behaviours, in comparison to the control video.

Impact on perceptions and beliefs

The **first** and broadest research question we will answer is whether participants found watching the case video helpful in *any way,* i.e., answered “Yes” to the question “Did you find this video helpful in any way?”. We will report the percentage of participants who said it was helpful and compare this percentage to that of the participants who found the control video helpful. We will also perform a chi-square statistical test to investigate whether this difference is significant. We will also report demographic information of those who reported that the case video was helpful and compare it to those who answered “No”, ”Not sure” or “Prefer Not To Answer”. We will conduct regression analyses to investigate *who* the video was helpful for, e.g., whether being a woman is associated with significantly higher odds of finding the video helpful compared to being a man. We will also repeat this analysis for those participants who found the control video helpful.

The **second** research question we will answer is whether watching the case video is associated with a change in participants’ perceived genetic and environmental contribution to their mental health disorder and their confidence in this perception. Plots will show the distribution of answers at baseline, immediate follow-up, follow up 2 and follow up 3. In addition to this, we will report on any shift in average (mean or median depending on score distribution) scores (i.e., baseline compared to follow-up surveys) and the percentage of participants who show a change in their perceptions or their confidence. We will also group participants into 3 categories based on

their baseline answers: “more genetic”, “more environmental”, “equal genetic and environmental”. Within each of these groups, we will report on the percentage who moved into another group (and which group), percentage who moved within a group (and in which direction), and percentage whose group did not change. We will do the same with the participants who watched the control video in order to draw conclusions about the influence of the case video.

We will also investigate whether participants would like to have a session with a psychiatric genetic counsellor and whether this is associated with genetic and/or environmental determinism.

Impact on stigma

The **third** research question we will answer is whether watching the case video is associated with changes in participants’ self-stigma and general mental health stigma. We will report on any shift in average (i.e., baseline compared to follow-up surveys), percentage of people whose self-stigma increased, percentage of people whose self-stigma decreased, and percentage of people whose self-stigma did not change. We will also comment on whose self-stigma was helped by watching the animated video (i.e., report demographics of those whose self-stigma level decreased). We will also compare this with the control video.

The **fourth** question we will answer is whether people perceive their mental health disorder as “more genetic” or “more environmental” (i.e., falling at the extremes of this scale) have higher levels of self-stigma and general mental health stigma compared to participants who perceive their mental health as due to equal genetic and environmental factors. We will perform regressions to investigate this. We will regress baseline self-stigma onto answers about perceptions of genetic and environmental contributions to their mental health disorder. We will then repeat this for follow-up 2 answers and follow-up 3 answers. Current symptoms will be included as covariates to control for possible reporting bias.

Impact on mental health

The **fifth** research question we will answer is whether watching the case video affected symptom levels. We will report on the percentage of people whose symptoms improved, got worse, or stayed the same (i.e., baseline compared to follow up 2 and follow up 3). We will also perform regressions to establish whether the type of video watched by the participants (i.e., case

vs. control video) influences change in symptoms at follow up 2 and follow up 3 (compared to baseline). In the regression of symptoms at follow up 2, we will use the baseline symptom scores as a covariate. In the regression of symptoms at follow up 3, we will use baseline and follow up 2 symptom scores as covariates.

Impact on treatment adherence and/or uptake

The **sixth** research question we will answer is whether watching the case video is associated with treatment adherence or new interest/uptake in psychological therapy and/or medication. First, we will report the percentage of participants who newly considered or started prescription medication and/or therapy at follow up 2 and follow up 3. We will use chi-squared tests to compare this to the percentages within the control group. Secondly, we will limit the analysis to anyone in the case group who, at baseline, reported being currently prescribed medication (but not taking it as prescribed) and/or enrolled in therapy (but not attending regularly). We will then comment on the percentage of people who started taking their medication as prescribed and/or started regularly attending therapy at follow up 2 and follow up 3. Finally, we will limit our analysis to those who, at baseline, self-reported that they were taking their medication and/or attending their therapy sessions. For each behaviour in this group, we will comment on the percentage of participants who stopped as well as the percentage of those who continued. We will compare this to the control video.

Impact on health behaviours

The **seventh** research question we will answer is whether watching the case video increased the adoption of health behaviours. We will report on the percentage of people whose health behaviours improved, got worse, or stayed the same (i.e., baseline compared to follow up 2 and follow up 3). We will then compare these results to those from the participants in the control group (i.e., watched the control video).

Sensitivity analysis

As a ***sensitivity*** analysis, we will identify to what extent people for whom watching the video was associated with improvements in health behaviours, medication and treatment

adherence/uptake, self-stigma, and current symptom levels are the same people who self-reported that the video was helpful (i.e., assess awareness of the helpfulness of video).

PerPsych: Mental health professionals, students, and trainees

Rationale

Recently, advances in genomic technologies have allowed researchers to identify genetic variants (e.g., single nucleotide polymorphisms) associated with many psychiatric disorders, including major depressive disorder (Wray *et al.*, 2018), anxiety (Purves *et al.*, 2019) and eating disorders (Watson *et al.*, 2019) via genome-wide association studies (GWAS). From GWAS results, polygenic risk scores (PRS) can be computed. A PRS is a weighted sum of the number of risk alleles an individual carries for a particular trait and disease and therefore represents their individual genetic risk (Lewis and Vassos, 2020), or individual genetic susceptibilities to diseases based on additive genetic variants (Wray *et al.*, 2021).

The results of PRS analyses arguably have the potential to become a key part of the clinical examinations and for prediction of prognosis. This has already been applied to breast cancer screening with some success (Mavaddat *et al.*, 2019). Research has shown that support for genetic testing in psychiatry is common. In a study of 352 psychiatrists, a large proportion indicated willingness to use genetic test results to confirm diagnoses in adult or child patients showing symptoms of schizophrenia (adults 75%, children 75%), bipolar disorder (74% and 72%), or panic disorder (66% and 63%) (Finn *et al.*, 2005). In addition, many of the psychiatrists surveyed showed keenness to offer genetic testing to asymptomatic patients (for preventative strategies) and 83% considered it their responsibility to discuss genetic information with patients and families. While this seems promising, that same study found that knowledge of psychiatric genetic principles was poor among psychiatrists, and fewer than 25% felt prepared or competent to discuss the topic with their patients(Finn *et al.*, 2005). This suggests that identifying the areas in which knowledge of psychiatric genetics needs improvement, as well as the best way of supporting clinicians in achieving this, is imperative.

As mentioned previously, people suffering from psychiatric illness may benefit from understanding their disorder in the context of their inherited genetic risk, especially for reducing self-stigmatisation(Laegsgaard *et al.*, 2010). Since clinicians are in a position to offer

conversations about psychiatric genetic risk to patients, this represents an important area in which clinicians are able to make a substantial difference to patients in terms of their perception of the genetic contribution to their illness, with the potential to reduce negative emotions including guilt, shame, fear and stigma.

Despite a clear need for such discussion, psychiatric disorders present distinct challenges due to their multifactorial aetiology and the absence of genetic tests to confirm diagnosis or predict risk(Palk *et al.*, 2019). For these reasons, accurate and accessible information about the causes of psychiatric disorders needs to be communicated by clinicians who work in this area to patients. In a recent paper by Lewis & (2020) evaluating the possibility of involving genetics in psychiatric clinical work, the authors emphasise that "*Education for clinicians and the public will be necessary to increase understanding and genetic literacy".* The proposed work aims to take a step towards achieving this amongst mental health professionals. We plan to conduct a project to understand more about the way in which a range of UK mental health clinicians think about the contribution of genetic and environmental factors to illness, and how information about psychiatric genetics could be communicated accurately and effectively.

Proposed Work

In the **PerPsych: Mental health professionals, students, and trainees** study, we plan to recruit UK-based mental health professionals (such as clinical psychologists, assistant psychologists, psychiatrists, psychological wellbeing practitioners [PWPs], social workers, General Practitioners [GPs] mental health nurses, therapists, and counsellors) as well as students and trainees who are entering into mental health careers.

We will use a survey delivered via Qualtrics to assess participants' knowledge of genetics and genomics and their belief in genetic determinism, and to identify whether they believe psychiatric genetic counselling would be helpful for their patients. In this survey, they will also be invited to watch a short animated video detailing how genes and environment combine to produce risk for psychiatric illnesses (i.e., one of the 'case' videos from the 'PerPsych: Lived experience' project). This will be followed by questions assessing whether the participants believe the content of the video would be beneficial for their patients/clients.

<u>Questions included in the study:</u>

- Demographic questions (e.g., gender, sex, age, and ethnic background)
- Questions about professional setting and year of qualification/starting role
- Questions about which disorders the participants work with and how often patients/clients and the relatives of patients/clients raise questions about the genetic and environmental risk of each disorder
- Public Understanding and Attitudes towards Genetics and Genomics (PUGGS) questionnaire (Carver *et al.*, 2017)
- International Genetic Literacy and Attitudes Survey (iGLAS) (Chapman *et al.*, 2017)
    - Selected questions to assess genetic knowledge
- Questions about knowledge of psychiatric genetics
- Questions about whether psychiatric genetic counselling would be beneficial for patients/clients
- Questions assessing whether the participants think the content of the animated video would be beneficial for their patients/clients

Aims

We aim to understand:

- How often patients/clients ask questions about the cause(s) of their disorder(s)
- How often the relatives of patients/clients ask questions about the causes(s) of their disorder(s)
- Whether mental health professionals and trainees have an adequate understanding of a) genetics and genomics and b) psychiatric genetics
- Whether mental health professionals and trainees feel confident discussing psychiatric genetics with their patients
- Whether exposure to the content of the animated video would be beneficial for their patients/clients
- Whether mental health professionals and trainees believe psychiatric genetic counselling would be beneficial for their patients/clients

Preliminary Work

No preliminary work has been carried out and there are no plans for pilot work.

Datasets

We will recruit mental health professionals, students, and trainees through a number of avenues. Participants must fulfil the criteria of working with or training to work with patients in the field of mental health in the UK. We will advertise the study on social media (e.g., Instagram and Twitter) and we will also send out emails to mailing lists at relevant institutions and organisations (such as SLaM and the Royal College of Psychiatrists). We aim to recruit a minimum of 100 participants for this study.

Descriptives

First, we will describe the overall sample's demographic characteristics (i.e., age, sex, gender, education level, ethnic background, and professional setting).

We will also generate plots for the following:

- Genetic knowledge measured by the 6-item International iGLAS questionnaire (Chapman et al., 2017)
    - Plot of percentage correct
        - NB: Any questions which have been skipped will be counted as incorrect.
- Belief in genetic determinism measured by a subset of the PUGGS questionnaire (Carver et al., 2017)
    - Histogram of continuous score
- Percentage of participants who endorsed each answer option to the question: "Please select an option according to how much you think [disorder] is caused by inherited genetic factors and/or environmental factors" per disorder.
- Percentage of each answer to the question "How often do patients/clients raise questions about their genetic predisposition for that disorder?" and the equivalent question about environmental risk
- Percentage of each answer to the question "How often do relatives of patients/clients raise questions about patient's/clients' genetic predisposition for that disorder" and the equivalent question about environmental risk per disorder.

- Percentage of each answer to the question “I think it would be beneficial for my patients/client to see a psychiatric genetic counsellor” and the equivalent question about environmental risk per disorder
- Percentage of how confident participants feel discussing genetic risk vs. environmental risk with their patients

Research questions

Generally, we are interested in understanding current mental health professionals’ and trainees’ belief in the value of delivering genetic risk information to patients/clients as well as their confidence in doing so. More specifically, we would like to assess whether there are any gaps in their belief or confidence, for instance, regarding a specific psychiatric disorder or within a specific clinical profession.

We will group participants by profession. Depending on our final sample size, this will either be a fine-grained grouping by specific profession or we will create overarching groups (e.g., “medical” [including GPs, psychiatrists, mental health nurses], “psychology” [including clinical psychologists, PWPs, therapists, and counsellors] and “social workers”).

Differences across profession

We will report the following:

1) Association between profession (covarying for age, sex, and time in job) and:
    a) How confident participants feel discussing genetic risk with patients
    b) How confident participants feel discussing environmental risk with patients
    c) Genetic knowledge scores
    d) Belief in genetic or environmental determinism

Differences across disorders

We will report the following per disorder:

- How often patients/clients (or relatives of patients/clients) raise questions about genetic/environmental risk for that disorder
- Percentage of participants who believe that psychiatric genetic counselling would be beneficial for their patients/clients with that disorder

References

Austin, J.C., Hippman, C. and Honer, W.G. (2012) 'Descriptive and numeric estimation of risk for psychotic disorders among affected individuals and relatives: implications for clinical practice', *Psychiatry research*, 196(1), pp. 52–56.

Austin, J.C., Smith, G.N. and Honer, W.G. (2006) 'The genomic era and perceptions of psychotic disorders: genetic risk estimation, associations with reproductive decisions and views about predictive testing', *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*, 141B(8), pp. 926–928.

Carver, R.B. *et al.* (2017) 'Young Adults' Belief in Genetic Determinism, and Knowledge and Attitudes towards Modern Genetics and Genomics: The PUGGS Questionnaire', *PloS one*, 12(1), p. e0169808.

Chapman, R. *et al.* (2017) 'Genetic Literacy and Attitudes Survey (iGLAS): International Population-Wide Assessment Instrument', *The European Proceedings of Social & Behavioural Sciences EpSBS*, 33(6), pp. 45–66.

Condit, C.M. and Shen, L. (2011) 'Public understanding of risks from gene-environment interaction in common diseases: implications for public communications', *Public health genomics*, 14(2), pp. 115–124.

Corrigan, P.W. *et al.* (2002) 'Challenging two mental illness stigmas: personal responsibility and dangerousness', *Schizophrenia bulletin*, 28(2), pp. 293–309.

Davies, M.R. *et al.* (2019) 'The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety', *Behaviour research and therapy*, 123, p. 103503.

Dick, D.M. (2011) 'Gene-environment interaction in psychological traits and disorders', *Annual review of clinical psychology*, 7, pp. 383–409.

Finn, C.T. *et al.* (2005) 'Psychiatric genetics: a survey of psychiatrists' knowledge, opinions, and practice patterns', *The Journal of clinical psychiatry*, 66(7), pp. 821–830.

Grilo, C.M. *et al.* (2013) 'Eating disorder examination-questionnaire factor structure and

construct validity in bariatric surgery candidates', *Obesity surgery*, 23(5), pp. 657–662.

Haga, S.B. *et al.* (2013) 'Public knowledge of and attitudes toward genetics and genetic testing', *Genetic testing and molecular biomarkers*, 17(4), pp. 327–335.

Inglis, A., Morris, E. and Austin, J. (2017) 'Prenatal genetic counselling for psychiatric disorders', *Prenatal diagnosis*, 37(1), pp. 6–13.

Kalb, F.M. *et al.* (2017) 'Genetic Counseling for Alcohol Addiction: Assessing Perceptions and Potential Utility in Individuals with Lived Experience and Their Family Members', *Journal of genetic counseling*, 26(5), pp. 963–970.

Kroenke, K. and Spitzer, R.L. (2002) 'The PHQ-9: a new depression diagnostic and severity measure', *Psychiatric annals*, 32(9), pp. 509–515.

Laegsgaard, M.M. *et al.* (2010) 'The perceived and predicted implications of psychiatric genetic knowledge among persons with multiple cases of depression in the family', *Acta psychiatrica Scandinavica*, 122(6), pp. 470–480.

Lewis, C.M. and Vassos, E. (2020) 'Polygenic risk scores: from research tools to clinical instruments', *Genome medicine*, 12(1), p. 44.

Luce, K.H. and Crowther, J.H. (1999) 'The reliability of the eating disorder examination—Self-report questionnaire version (EDE-Q)', *The International journal of eating disorders*, 25(3), pp. 349–351.

Mavaddat, N. *et al.* (2019) 'Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes', *American journal of human genetics*, 104(1), pp. 21–34.

Meiser, B. *et al.* (2007) 'Attitudes towards childbearing, causal attributions for bipolar disorder and psychological distress: a study of families with multiple cases of bipolar disorder', *Psychological medicine*, 37(11), pp. 1601–1611.

Michael, J.E. *et al.* (2020) 'Perceptions of genetic risk, testing, and counseling among individuals with eating disorders', *The International journal of eating disorders*, 53(9), pp. 1496–1505.

Palk, A.C. *et al.* (2019) 'Potential use of clinical polygenic risk scores in psychiatry - ethical implications and communicating high polygenic risk', *Philosophy, ethics, and humanities in*

medicine: PEHM, 14(1), p. 4.

Purves, K.L. *et al.* (2019) 'A major role for common genetic variation in anxiety disorders', *Molecular psychiatry* [Preprint]. Available at: https://doi.org/10.1038/s41380-019-0559-1.

Richards, M. (1996) 'Lay and professional knowledge of genetics and inheritance', *Public understanding of science* , 5(3), pp. 217–230.

Spitzer, R.L. *et al.* (2006) 'A brief measure for assessing generalized anxiety disorder: the GAD-7', *Archives of internal medicine*, 166(10), pp. 1092–1097.

Tsuang, M.T. *et al.* (2004) 'Gene-environment interactions in mental disorders', *World psychiatry: official journal of the World Psychiatric Association* , 3(2), pp. 73–83.

Watson, H.J. *et al.* (2019) 'Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa', *Nature genetics*, 51(8), pp. 1207–1214.

Wray, N.R. *et al.* (2018) 'Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression', *Nature genetics*, 50(5), pp. 668–681.

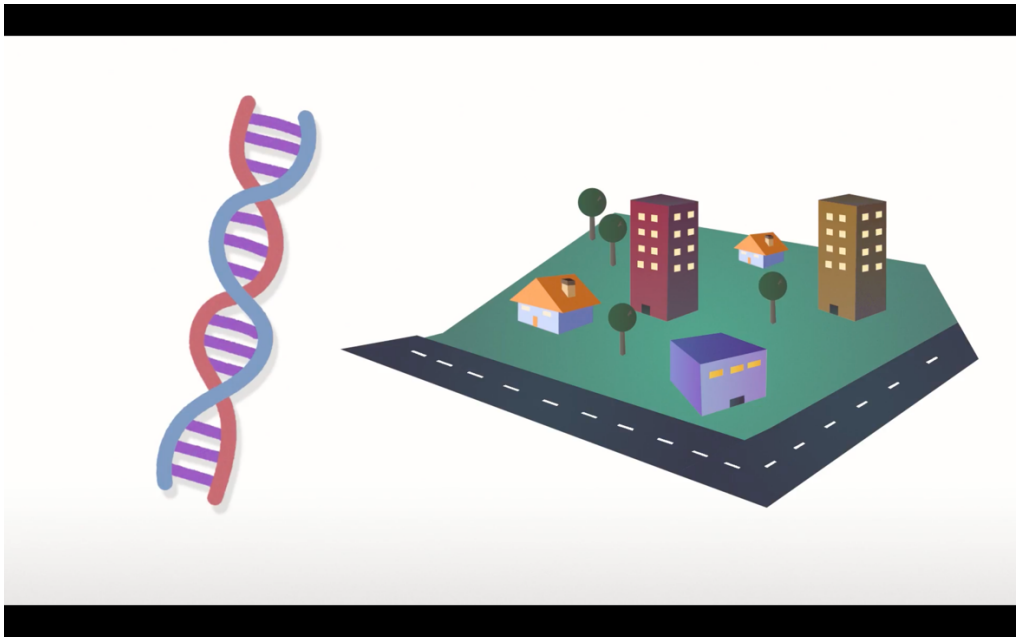
Wray, N.R. *et al.* (2021) 'From Basic Science to Clinical Application of Polygenic Risk Scores: A Primer', *JAMA psychiatry* , 78(1), pp. 101–109.

Script for animated video

Video script the GLAD study (depression and anxiety)

You may have heard people say…

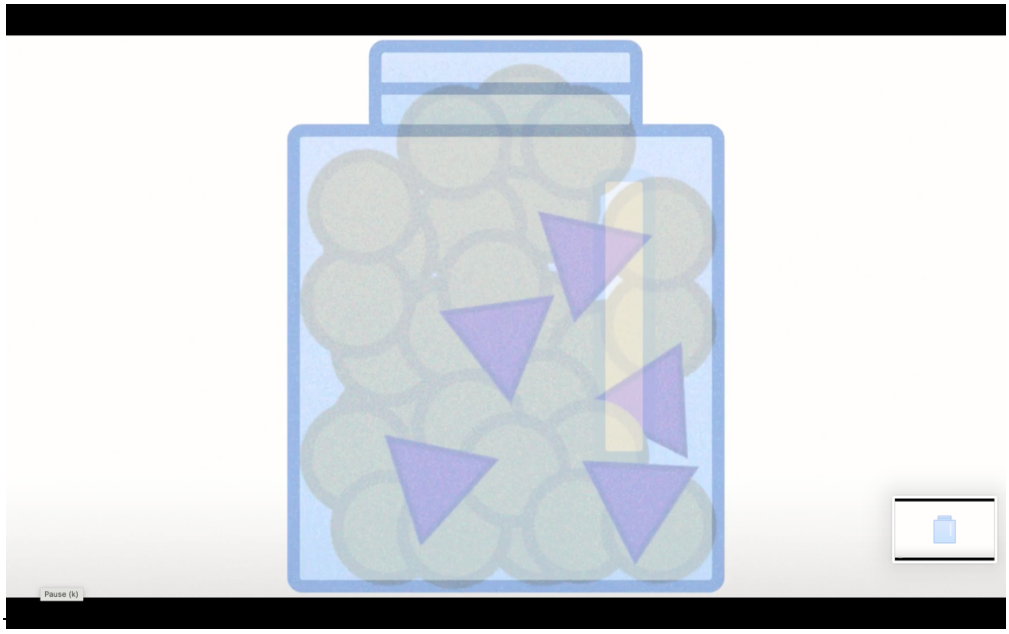
...that both genes and environment influence whether or not people develop mental health conditions such as depression or anxiety.



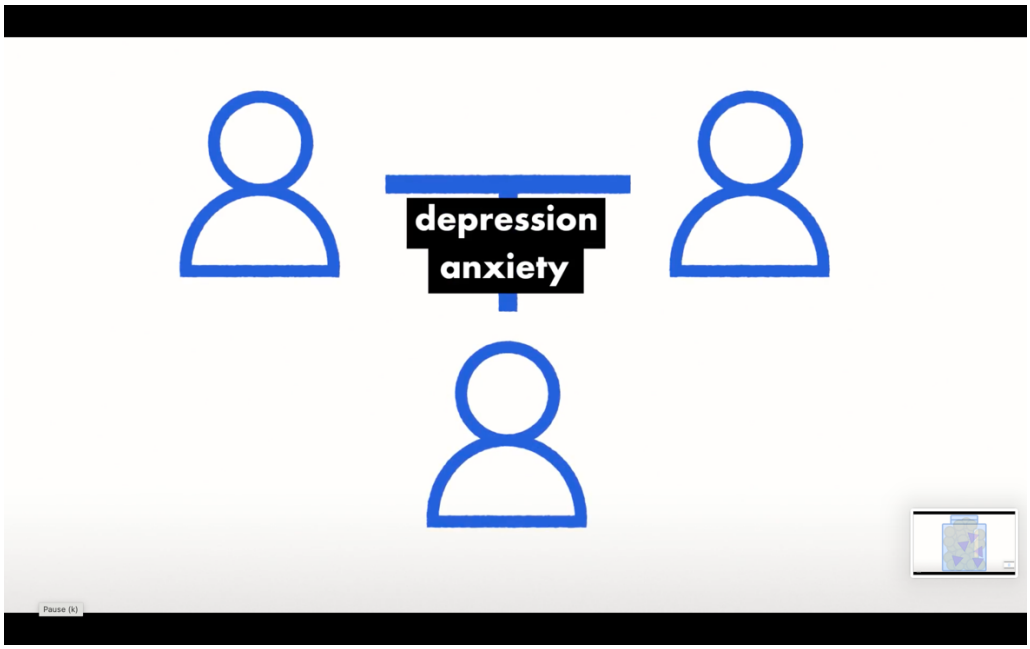
But what does that really mean?

For every one of us, our mental health can be thought of as a jar.

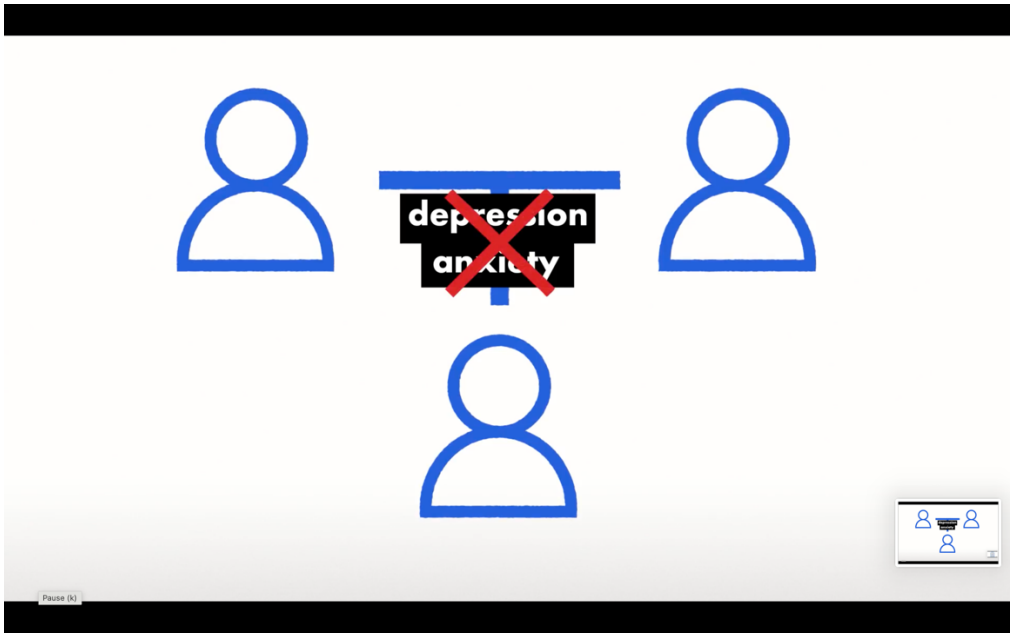
For a person to be experiencing an episode of depression or anxiety, their jar has to be full.



It’s important to understand that even though depression and anxiety are genetically influenced…

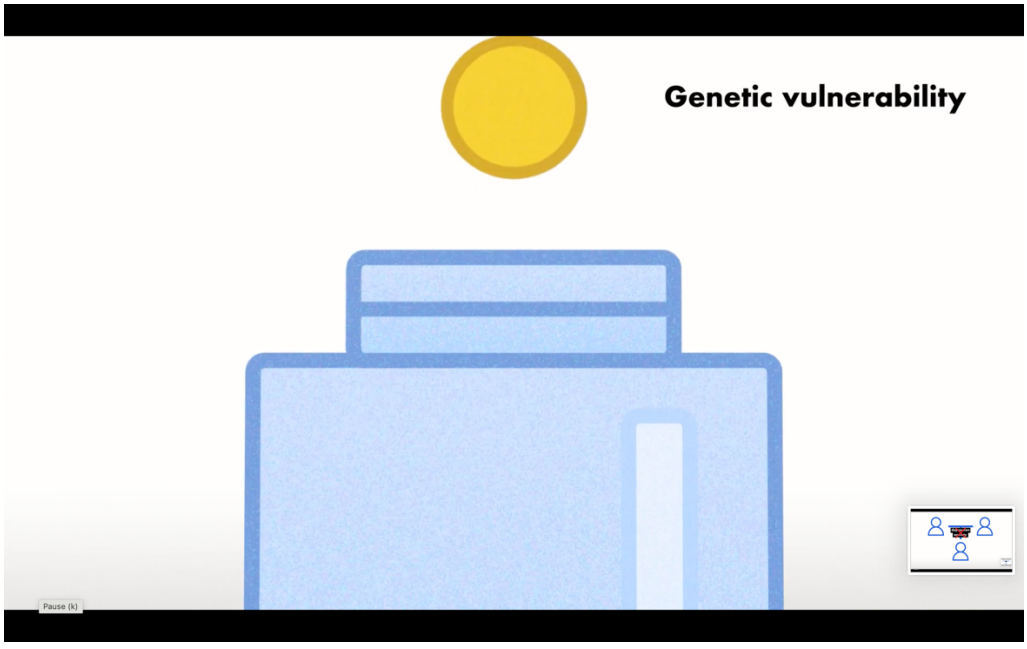


...you don’t inherit the disorders themselves.



What we inherit is a *vulnerability* to them. All of us are born with a certain amount of genetic vulnerability.

There is no single “gene for” depression or anxiety.



There are probably thousands of common genetic variants in hundreds of genes that each contribute a tiny amount to overall risk. Because they are so common in the population, all of us

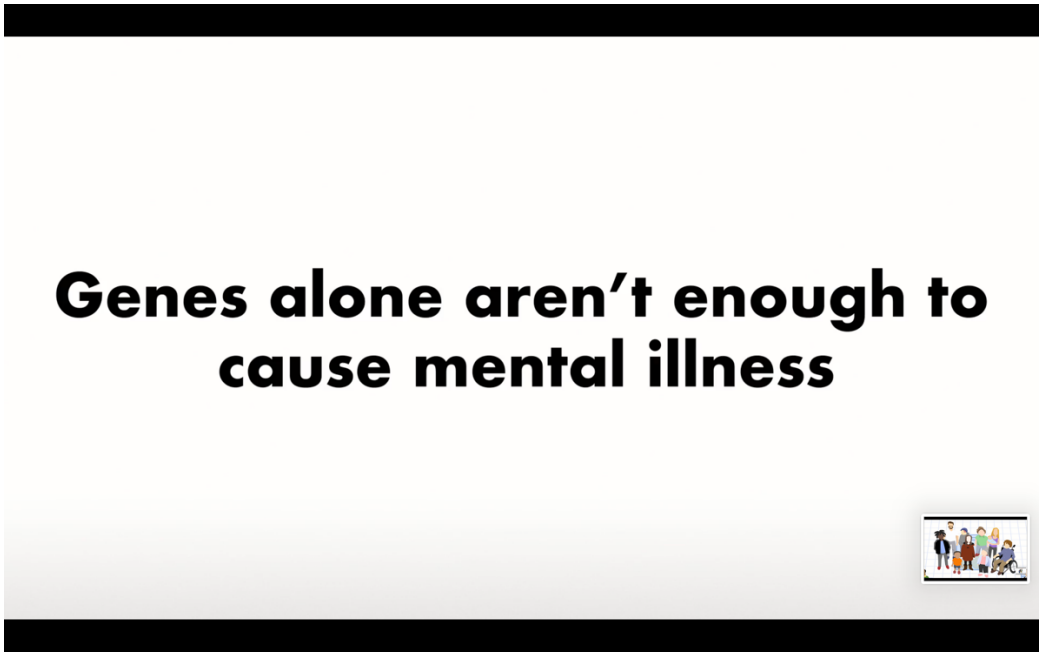
have some or even many of them.

A few people have very little genetic vulnerability…

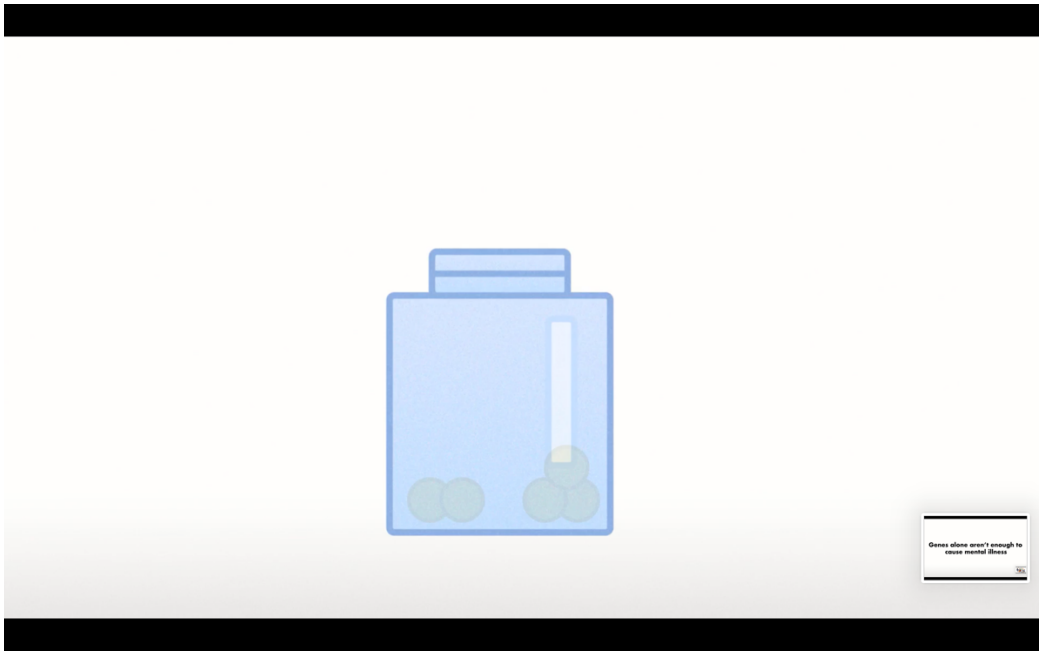
….a few people inherit a lot of genetic vulnerability…

...but most people fall somewhere in the middle.

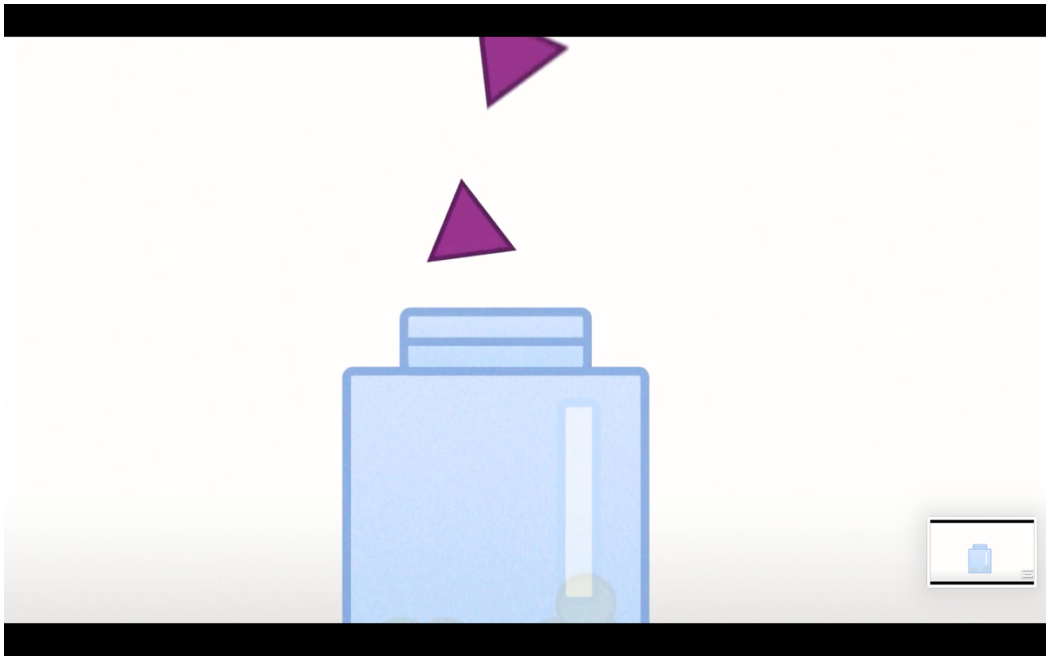
Genes alone aren’t enough to cause mental illness.



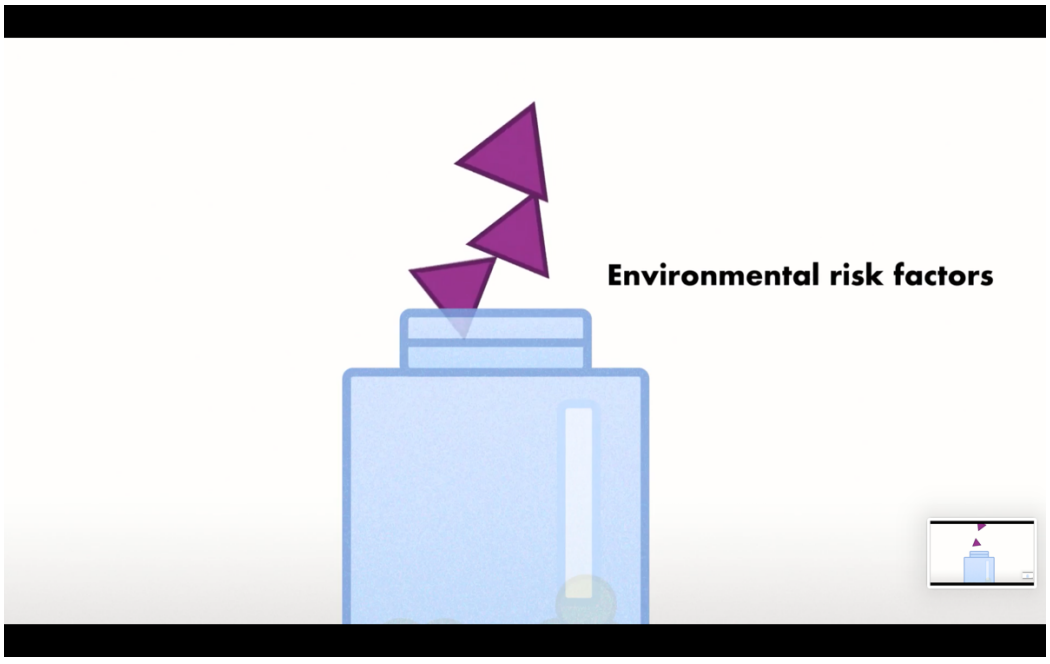
We all start with some level of genetic vulnerability for mental ill health and over time we may experience environmental risk factors.



What these environmental risk factors are will be different for each person.



They could include stressful life experiences or traumatic events.



Day-to-day events, such as financial worries or stress, are also risk factors.

Stressful life experiences

Traumatic events

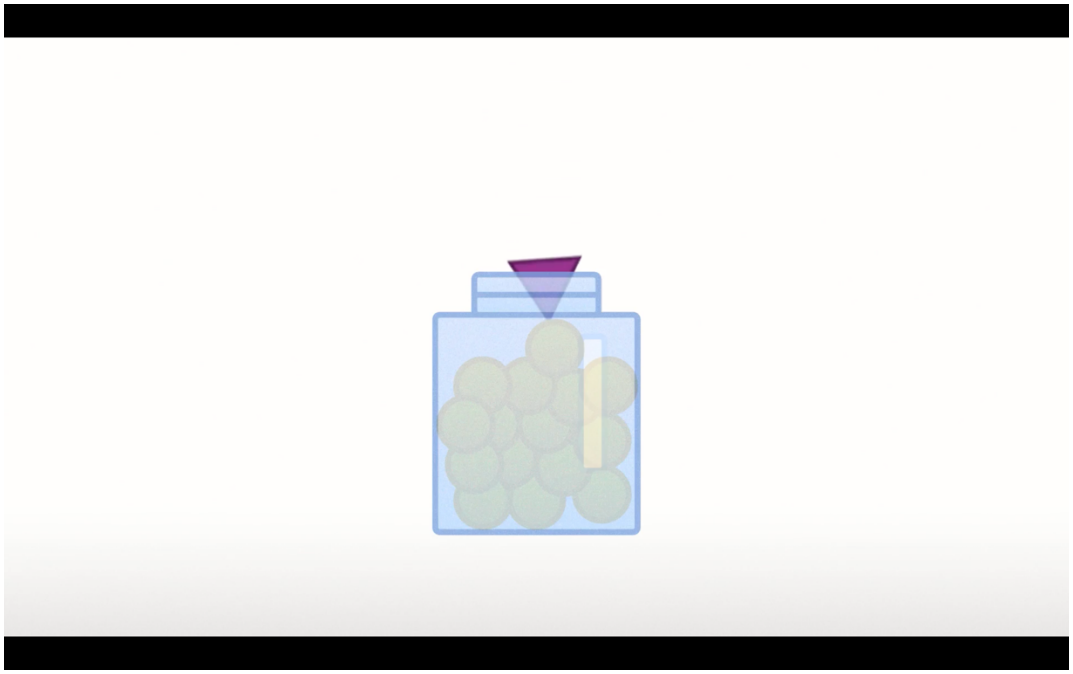
Day-to-day events

- **Financial worries**
- **Stress**

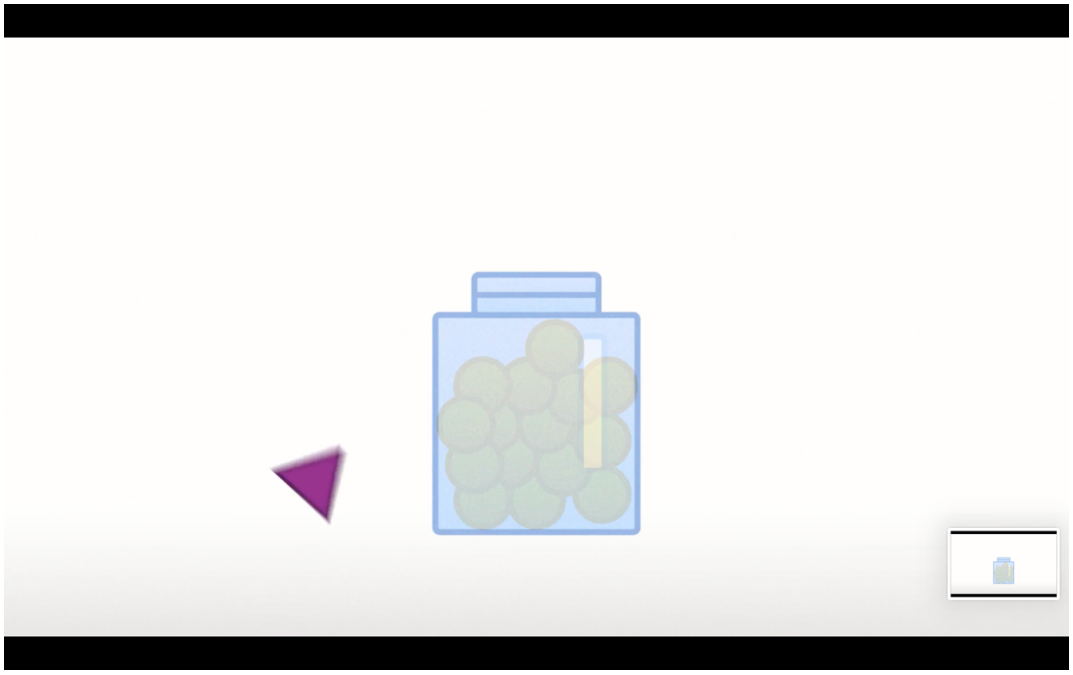
These stressful or adverse life events are unique to each of us.

You can start life with a lower genetic vulnerability, but still experience a lot of environmental risk, which could lead to your jar becoming full and experiencing mental ill health.

Or, you could start with a high genetic vulnerability. Then, it doesn't take as many stressful environmental events to fill up your jar.



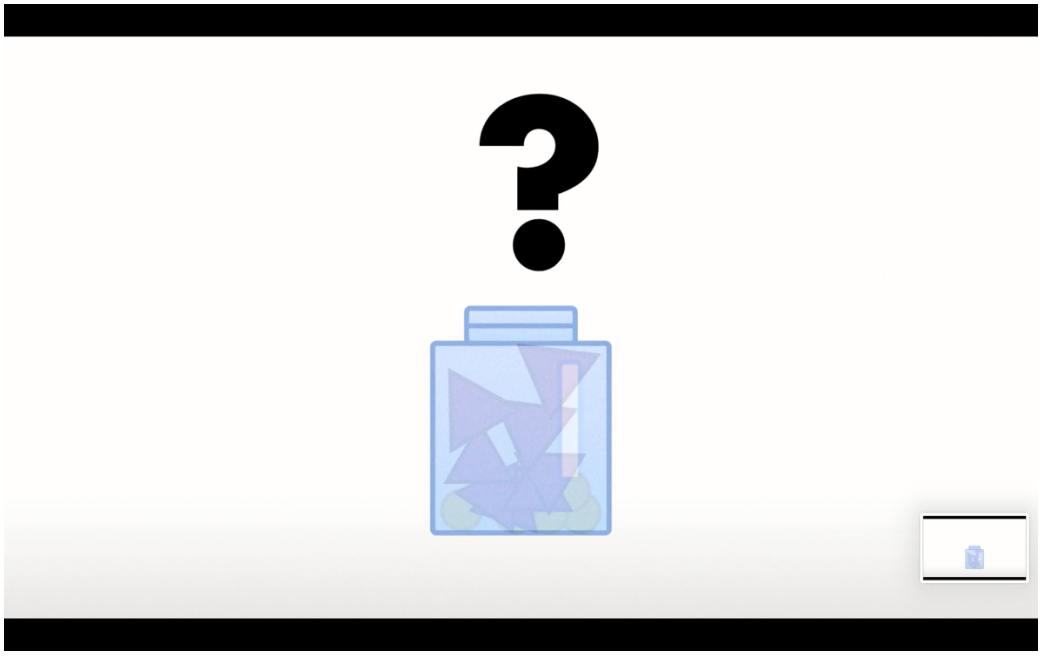
But you can also have a lot of genetic risk, and not encounter many stressful environments, so your jar may never fill up, and you may never experience mental ill health.



Understanding genetic and environmental risk can also help you understand how it's possible to reduce risk for developing depression or anxiety.

By adding protective factors, it makes it possible for a person to experience more environmental risk without their jar becoming full and without actively experiencing mental ill health.

What are these protective factors?



Many of them are things that are good for all of us…

Adopting a positive sleep pattern…

...nutrition…

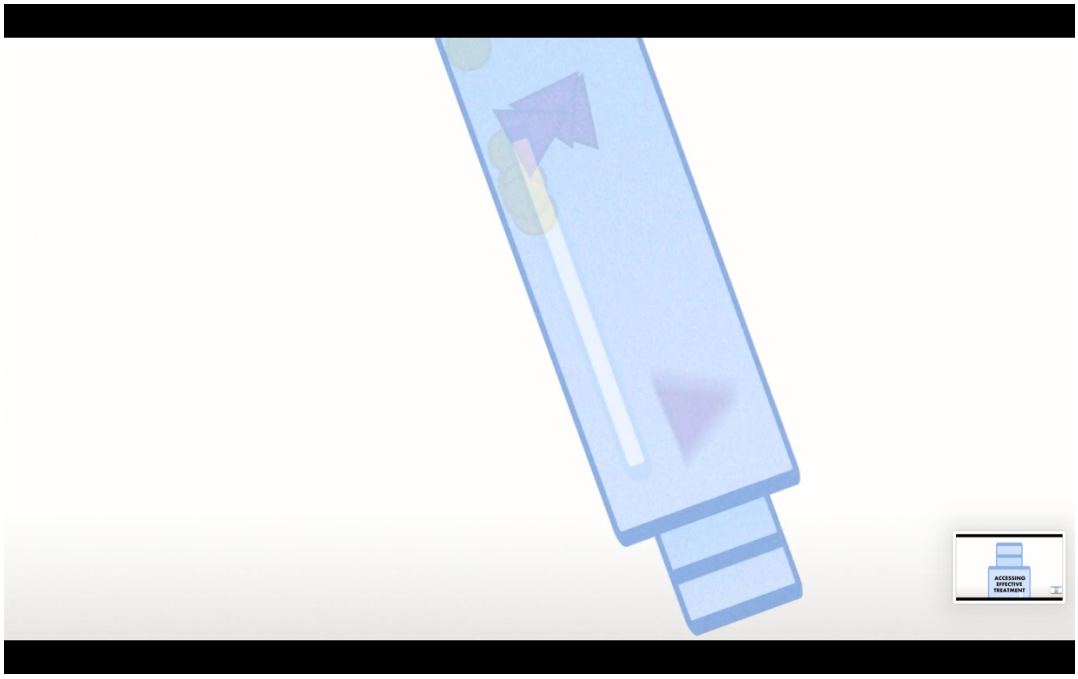
...being physically active in our own personal way…

…social support…

...finding more effective ways to manage stress…

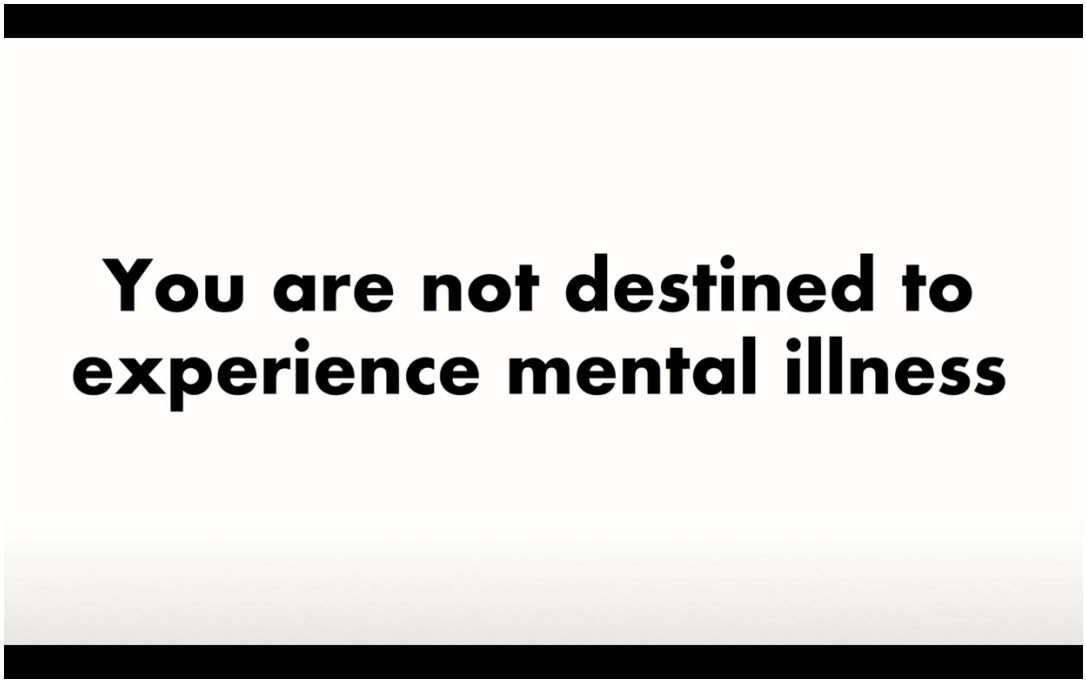
...or accessing effective treatment.

We may also be able to work to remove environmental risk factors.



In this way, regardless of your genetic vulnerability, you are not destined to experience mental illness. And even if you have experienced depression or anxiety, by adding protective factors

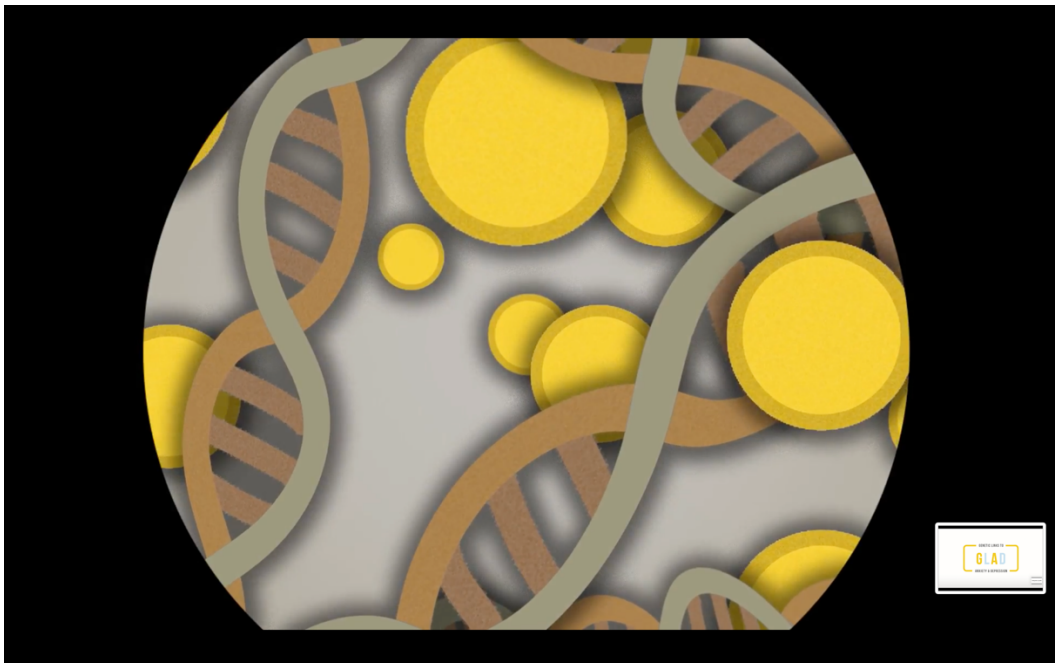
and removing environmental risk, managing your mental health…



...and feeling better is possible!

Scientists, such as those working on the Genetic Links to Anxiety and Depression study, are still in the process of finding the many genetic variants that influence risk for depression and anxiety

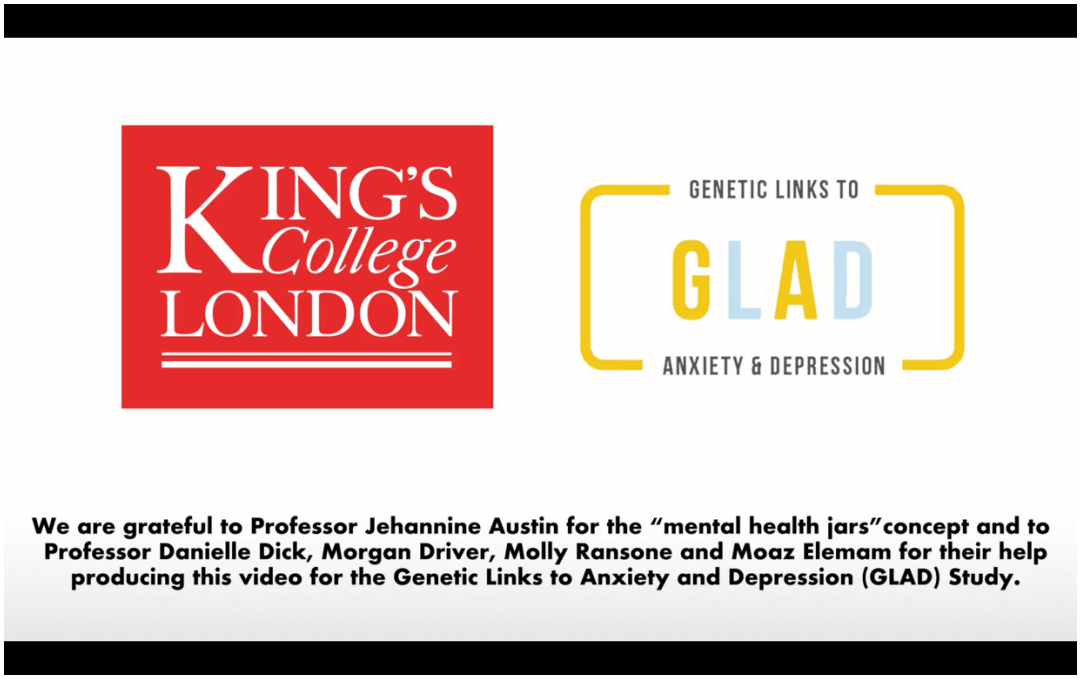
and other mental illnesses.



As the science advances we'll be able to do a better job of giving people an estimate of how much genetic vulnerability is in their jar.

But this will always be only a part of the story. Our experiences are the other part.

By participating in this study, you are making a huge difference to future treatment and understanding of depression and anxiety.


KING'S College LONDON
GENETIC LINKS TO
GLAD
ANXIETY & DEPRESSION
We are grateful to Professor Jehannine Austin for the "mental health jars"concept and to Professor Danielle Dick, Morgan Driver, Molly Ransone and Moaz Elemam for their help producing this video for the Genetic Links to Anxiety and Depression (GLAD) Study.

Appendix 6. Data cleaning example

Below is an example of an R script used to clean one of the surveys in the GLAD study and EDGI.

PHQ9 GLAD & EDGI data cleaning

Jessica Mundy

30/07/2021

Set up

```
rm(list=ls())

source(file = "scripts/functions/add_numeric.R")

source(file = "scripts/functions/sumscores.R")

source(file = "scripts/functions/package_check.R")

source(file = "scripts/functions/remove_duplicates.R")

source(file = "scripts/functions/imp_check.R")
```

Note: always load tidyverse last

```
packages = c(

  "summarytools",

  "sjlabelled",

  "Amelia",

  "knitr",

  "gtsummary",

  "tidyverse"

  )

package_check(packages)

date <- Sys.Date()

date

[1] "2022-12-06"

source("scripts/credentials/paths.R")
```

Read in data

GLAD data

```
glad_dat <- read_rds(
  file = paste0(ilovedata,"/data_raw/latest_freeze/glad/phq9_glad.rds")
  )


#Check
glad_dat %>%
  colnames()
 [1] "externalDataReference"
 [2] "startDate"
 [3] "endDate"
 [4] "phq9.little_interest_or_pleasure_in_doing_things"
 [5] "phq9.feeling_down_depressed_or_hopeless"
 [6] "phq9.staying_asleep_sleeping_trouble"
 [7] "phq9.feeling_tired_or_having_little_energy"
 [8] "phq9.poor_appetite_or_overeating"
 [9] "phq9.feeling_bad_failure_family"
[10] "phq9.trouble_concentrating_reading_newspaper"
[11] "phq9.moving_fidgety_noticed_opposite"
[12] "phq9.dead_hurting_thoughts"
glad_dat %>%
  dim()
[1] 46853        12
```

Specify columns to be excluded from add_numeric function

Continuous variables should be excluded, as they are already numeric

```
exclude_cols_numeric <- c(
  "ID",
```

```
  "sample",
  "startDate",
  "endDate"
  )
```

Select & rename relevant columns

```
glad_dat_id <- glad_dat %>% #new dataset with ID
  drop_na(externalDataReference) %>% # drop NAs in ID
  remove_duplicates("externalDataReference") %>%

  add_column(sample = "GLAD",
             .after = "externalDataReference") %>% #create new column with sample
  select(
             ID = externalDataReference, # ID
             startDate,
             endDate,
             sample,
             phq9.dead_hurting_thoughts,
             phq9.feeling_bad_failure_family,
             phq9.feeling_down_depressed_or_hopeless,
             phq9.feeling_tired_or_having_little_energy,
             phq9.little_interest_or_pleasure_in_doing_things,
             phq9.moving_fidgety_noticed_opposite,
             phq9.poor_appetite_or_overeating,
             phq9.staying_asleep_sleeping_trouble,
             phq9.trouble_concentrating_reading_newspaper
             ) %>%
  add_numeric(exclude = exclude_cols_numeric)
```

```
# Inspect colnames
colnames(glad_dat_id)
 [1] "ID"
 [2] "startDate"
 [3] "endDate"
 [4] "sample"
 [5] "phq9.dead_hurting_thoughts"
 [6] "phq9.feeling_bad_failure_family"
 [7] "phq9.feeling_down_depressed_or_hopeless"
 [8] "phq9.feeling_tired_or_having_little_energy"
 [9] "phq9.little_interest_or_pleasure_in_doing_things"
[10] "phq9.moving_fidgety_noticed_opposite"
[11] "phq9.poor_appetite_or_overeating"
[12] "phq9.staying_asleep_sleeping_trouble"
[13] "phq9.trouble_concentrating_reading_newspaper"
[14] "phq9.dead_hurting_thoughts_numeric"
[15] "phq9.feeling_bad_failure_family_numeric"
[16] "phq9.feeling_down_depressed_or_hopeless_numeric"
[17] "phq9.feeling_tired_or_having_little_energy_numeric"
[18] "phq9.little_interest_or_pleasure_in_doing_things_numeric"
[19] "phq9.moving_fidgety_noticed_opposite_numeric"
[20] "phq9.poor_appetite_or_overeating_numeric"
[21] "phq9.staying_asleep_sleeping_trouble_numeric"
[22] "phq9.trouble_concentrating_reading_newspaper_numeric"
# Inspect dimensions
glad_dat_id %>%
  dim()
[1] 46725    22
#Differences
```

```
glad_excluded <- dim(glad_dat_id)[1]-dim(glad_dat)[1]
glad_excluded
[1] -128
```

Inspect numeric variables

Add summary table with questions as columns, factor levels as rows, with absolute frequencies

```
glad_dat_id %>%
  select(all_of(ends_with("numeric"))) %>%
  tbl_summary(missing_text = "Missing")
Table printed with {flextable}, not {gt}. Learn why at
http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in the code chunk header.
```

Characteristic	N = 46,725[1]
Over the last 2 weeks, how often have you been bothered by any of the following problems? Feeling bad about yourself or that you are a failure or have let yourself or your family down	
-77	44 (<0.1%)
0	9,695 (22%)
1	14,069 (32%)
2	8,630 (19%)

3	11,838 (27%)
Missing	2,449

Over the last 2 weeks, how often have you been bothered by any of the following problems? Feeling down, depressed or hopeless

-77	50 (0.1%)
0	8,563 (19%)
1	18,765 (42%)
2	8,384 (19%)
3	8,514 (19%)
Missing	2,449

Over the last 2 weeks, how often have you been bothered by any of the following problems? Feeling tired or having little energy

-77	30 (<0.1%)
0	3,249 (7.3%)

1	13,023 (29%)
2	10,305 (23%)
3	17,669 (40%)
Missing	2,449

Over the last 2 weeks, how often have you been bothered by any of the following problems? Little interest or pleasure in doing things

-77	36 (<0.1%)
0	9,759 (22%)
1	18,028 (41%)
2	8,353 (19%)
3	8,100 (18%)
Missing	2,449

Over the last 2 weeks, how often have you been bothered by any of the following problems? Poor appetite or overeating

-77	49 (0.1%)
0	10,208 (23%)
1	12,345 (28%)
2	9,521 (22%)
3	12,153 (27%)
Missing	2,449

Over the last 2 weeks, how often have you been bothered by any of the following problems? Trouble falling or staying asleep, or sleeping too much

-77	35 (<0.1%)
0	6,318 (14%)
1	13,424 (30%)
2	9,327 (21%)
3	15,172 (34%)

Missing	2,449
Over the last 2 weeks, how often have you been bothered by any of the following problems? Trouble concentrating on things, such as reading the newspaper or watching television	
-77	42 (<0.1%)
0	11,472 (26%)
1	14,588 (33%)
2	9,154 (21%)
3	9,020 (20%)
Missing	2,449

[1]n (%)

Check missingness by missmap

```
glad_miss_map <- glad_dat_id %>%
  missmap()
```

```
Warning: Unknown or uninitialised column: `arguments`.
```

```
Unknown or uninitialised column: `arguments`.
```

```
Warning: Unknown or uninitialised column: `imputations`.
```

```
glad_miss_map
```

```
NULL
```

EDGI data

Note: this is an optional questionnaire in EDGI.

```
edgi_dat <- readRDS(
  file = paste0(ilovedata, "/data_raw/latest_freeze/edgi_opt/phq_edgi_opt.rds")
  )


#Check
edgi_dat %>%
  colnames()
  [1] "externalDataReference"
  [2] "startDate"
  [3] "endDate"
  [4] "phq9.little_interest_or_pleasure_in_doing_things"
  [5] "phq9.feeling_down_depressed_or_hopeless"
  [6] "phq9.staying_asleep_sleeping_trouble"
  [7] "phq9.feeling_tired_or_having_little_energy"
  [8] "phq9.poor_appetite_or_overeating"
  [9] "phq9.feeling_bad_failure_family"
 [10] "phq9.trouble_concentrating_reading_newspaper"
 [11] "phq9.moving_fidgety_noticed_opposite"
 [12] "phq9.dead_hurting_thoughts"
 [13] "tb.felt_sad_blue_row"
 [14] "tb.hobbies_work_lost_interest"
 [15] "tb.day_feelings"
 [16] "tb.did_you_feel_this_way"
 [17] "tb.tired_energy_usual_low"
 [18] "tb.didyour_weight_change"
```

[19] "tb.10lbs_weight_change"

[20] "tb.did_your_sleep_change"

[21] "tb.trouble_falling_asleep"

[22] "tb.waking_too_early"

[23] "tb.sleeping_too_much"

[24] "tb.depression_episode_sleeping"

[25] "tb.sleep_hours_average_day.txt"

[26] "tb.hours_average_sleep_depressed.txt"

[27] "tb.appetite_change_experience"

[28] "tb.mood_brighten_positive_events"

[29] "tb.arms_legs_experience_heavy"

[30] "tb.forhow_hours_day_heaviness.txt"

[31] "tb.were_you_overly_sensitive_to_interpersonal_rejection"

[32] "tb.was_your_mood_worse"

[33] "tb.trouble_concentrating_usual_lot"

[34] "tb.good_worthless_people_feel"

[35] "tb.death_general_lot_elses"

[36] "tb.long_altogether_feel"

[37] "tb.longest_episode_low_mood"

[38] "tb.longest_period_low_mood"

[39] "tb.roles_problems_interfere_activities"

[40] "tb.low_mood_life_lasting"

[41] "tb.low_mood_life_lasting.1"

[42] "tb.approximate_age_age_fine.txt"

[43] "tb.approximate_age_age_fine.txt.1"

[44] "tb.giving_birth_suggested_episodes"

[45] "tb.distressing_event_significant_close"

[46] "tb.aprofessional_problems"

[47] "tb.medication_prescribed_weeks"

[48] "tb.unprescribed_medication_more_than_once"

[49] "tb.drugs_or_alcohol_more_than_once"

[50] "tb.talking_therapy_psychotherapy"

[51] "tb.structured_wellbeing_activity"

[52] "tb.regular_physical_exercise"

[53] "tb.none_of_the_above"

[54] "tb.prefer_not_to_answer"

[55] "tb.enrolled_psychotherapy_nhs_funded"

[56] "tb.did_you_take_your_medication_as_advised"

[57] "tb.did_you_find_the_medication_helpful"

[58] "tb.counselling"

[59] "tb.mindfulness"

[60] "tb.relationship_therapy"

[61] "tb.group_therapy"

[62] "tb.guided_selfhelp"

[63] "tb.family_therapy"

[64] "tb.cbt"

[65] "tb.workshops"

[66] "tb.online_therapy"

[67] "tb.other"

[68] "tb.othertext.txt"

[69] "tb.never"

[70] "tb.dont_know"

[71] "tb.prefer_not_to_answer.1"

[72] "tb.talking_therapy_complete_psychotherapy"

[73] "tb.talking_therapy_helpful_find"

[74] "tb.low_mood_begin_long"

[75] "tb.antidepressantmedications_recent_episode_low"

[76] "tb.low_mood_haveyou_addition"

```
 [77] "tb.have_you_received_electroconvulsive_therapy"
 [78] "tb.hyper_normal_trouble_people"
 [79] "tb.shouted_started_fights_irritable"
 [80] "tb.usual_felt_selfconfident"
 [81] "tb.found_miss_usual_sleep"
 [82] "tb.faster_spoke_usual_talkative"
 [83] "tb.head_mind_couldnt_slow"
 [84] "tb.staying_track_easily_distracted"
 [85] "tb.you_had_much_more_energy_than_usual"
 [86] "tb.usual_active_things"
 [87] "tb.middle_outgoing_telephoned_friends"
 [88] "tb.sex_usual_interested"
 [89] "tb.excessive_foolish_risky_unusual"
 [90] "tb.spending_money_trouble_family"
 [91] "tb.ticked_happened_period_time"
 [92] "tb.prefer_not_to_answer.2"
 [93] "tb.dont_know.1"
 [94] "tb.hyper_normal_trouble_people.1"
 [95] "tb.shouted_started_fights_irritable.1"
 [96] "tb.usual_felt_selfconfident.1"
 [97] "tb.found_miss_usual_sleep.1"
 [98] "tb.faster_spoke_usual_talkative.1"
 [99] "tb.head_mind_couldnt_slow.1"
[100] "tb.staying_track_easily_distracted.1"
[101] "tb.you_had_much_more_energy_than_usual.1"
[102] "tb.usual_active_things.1"
[103] "tb.middle_outgoing_telephoned_friends.1"
[104] "tb.sex_usual_interested.1"
[105] "tb.excessive_foolish_risky_unusual.1"
```

```
[106] "tb.spending_money_trouble_family.1"
[107] "tb.longest_time_lasted_quothighquot"
[108] "tb.legal_troubles_unable_fights"
[109] "tb.bipolar_disorder_told_health"
[110] "tb.bipolar_disorder_blood_relatives"
[111] "tb.mania_irritable_periods_psychosis"
[112] "tb.diagnosis_psychosis_episode_primary"
edgi_dat %>%
  dim()
[1] 749 112
```

Select & rename relevant columns

```
edgi_dat_id <- edgi_dat %>% #new dataset with ID
  drop_na(externalDataReference) %>% # drop NAs in ID
  remove_duplicates("externalDataReference") %>%
  add_column(sample = "EDGI",
             .after = "externalDataReference") %>% #create new column
  select(
             ID = externalDataReference, # ID
             sample,
             startDate,
             endDate,
             phq9.dead_hurting_thoughts,
             phq9.feeling_bad_failure_family,
             phq9.feeling_down_depressed_or_hopeless,
             phq9.feeling_tired_or_having_little_energy,
             phq9.little_interest_or_pleasure_in_doing_things,
             phq9.moving_fidgety_noticed_opposite,
             phq9.poor_appetite_or_overeating,
```

```
          phq9.staying_asleep_sleeping_trouble,
            phq9.trouble_concentrating_reading_newspaper
          ) %>%
  add_numeric(exclude = exclude_cols_numeric)

# Inspect colnames
colnames(edgi_dat_id)
 [1] "ID"
 [2] "sample"
 [3] "startDate"
 [4] "endDate"
 [5] "phq9.dead_hurting_thoughts"
 [6] "phq9.feeling_bad_failure_family"
 [7] "phq9.feeling_down_depressed_or_hopeless"
 [8] "phq9.feeling_tired_or_having_little_energy"
 [9] "phq9.little_interest_or_pleasure_in_doing_things"
[10] "phq9.moving_fidgety_noticed_opposite"
[11] "phq9.poor_appetite_or_overeating"
[12] "phq9.staying_asleep_sleeping_trouble"
[13] "phq9.trouble_concentrating_reading_newspaper"
[14] "phq9.dead_hurting_thoughts_numeric"
[15] "phq9.feeling_bad_failure_family_numeric"
[16] "phq9.feeling_down_depressed_or_hopeless_numeric"
[17] "phq9.feeling_tired_or_having_little_energy_numeric"
[18] "phq9.little_interest_or_pleasure_in_doing_things_numeric"
[19] "phq9.moving_fidgety_noticed_opposite_numeric"
[20] "phq9.poor_appetite_or_overeating_numeric"
[21] "phq9.staying_asleep_sleeping_trouble_numeric"
[22] "phq9.trouble_concentrating_reading_newspaper_numeric"
```

```
# Inspect dimensions
dim(edgi_dat_id)
[1] 733  22
#Differences
edgi_excluded <- dim(edgi_dat_id)[1]-dim(edgi_dat)[1]
edgi_excluded
[1] -16
```

edgi.excluded EDGI participants excluded due to missing data

Inspect numeric variables

Add summary table with questions as columns, factor levels as rows, with absolute frequencies

```
edgi_dat_id %>%
  select(all_of(ends_with("numeric"))) %>%
  tbl_summary(missing_text = "Missing")
Table printed with {flextable}, not {gt}. Learn why at
http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in the code chunk header.
```

Characteristic	N = 733[1]
Over the last 2 weeks, how often have you been bothered by any of the following problems? Feeling bad about yourself or that you are a failure or have let yourself or your family down	
0	61 (17%)
1	89 (25%)

2	62 (18%)
3	138 (39%)
Missing	383

Over the last 2 weeks, how often have you been bothered by any of the following problems? Feeling down, depressed or hopeless

0	56 (16%)
1	143 (41%)
2	76 (22%)
3	75 (21%)
Missing	383

Over the last 2 weeks, how often have you been bothered by any of the following problems? Feeling tired or having little energy

0	40 (11%)
1	92 (26%)

2	66 (19%)
3	152 (43%)
Missing	383

Over the last 2 weeks, how often have you been bothered by any of the following problems? Little interest or pleasure in doing things

0	66 (19%)
1	131 (37%)
2	79 (23%)
3	74 (21%)
Missing	383

Over the last 2 weeks, how often have you been bothered by any of the following problems? Poor appetite or overeating

0	79 (23%)
1	82 (23%)

2	81 (23%)
3	108 (31%)
Missing	383

Over the last 2 weeks, how often have you been bothered by any of the following problems? Trouble falling or staying asleep, or sleeping too much

0	63 (18%)
1	91 (26%)
2	73 (21%)
3	123 (35%)
Missing	383

Over the last 2 weeks, how often have you been bothered by any of the following problems? Trouble concentrating on things, such as reading the newspaper or watching television

0	70 (20%)

1	114 (33%)
2	73 (21%)
3	93 (27%)
Missing	383

[1]n (%)

Check missingness by missmap

```
edgi_miss_map <- edgi_dat_id %>%
  missmap()
Warning: Unknown or uninitialised column: `arguments`.
Unknown or uninitialised column: `arguments`.
Warning: Unknown or uninitialised column: `imputations`.
edgi_miss_map
NULL
```

Bind rows of GLAD and EDGI data

```
dat <- bind_rows(
        glad_dat_id,
        edgi_dat_id
        )
```

There should be ID, sample and the 9 PHQ9 variables in categorical form and 9 PHQ9 variables in numeric form.

Data cleaning

Recode Non-answer values to 3 digits -555 'Not applicable' response from participant -777 Seen but not answered -888 Don't know -999 Prefer not to answer/Prefer not to say *NA* Were not shown the question (genuinely missing value)

```
dat <- dat %>%
  mutate(
        across(
        ends_with("numeric"),
        ~case_when(
        . == -55 ~ -555,
        . == -77 ~ -777,
        . == -88 ~ -888,
        . == -99 ~ -999,
        TRUE ~ .)
  )
  )
```

Numeric variables

Select numeric PHQ9 variables for cleaning

This will select all of the PHQ9 numeric variables

```
phq9_variables_numeric <- dat %>%
  select(
        contains("_numeric")
        ) %>%
  colnames()



# check
phq9_variables_numeric
[1] "phq9.dead_hurting_thoughts_numeric"
```

```
[2] "phq9.feeling_bad_failure_family_numeric"
[3] "phq9.feeling_down_depressed_or_hopeless_numeric"
[4] "phq9.feeling_tired_or_having_little_energy_numeric"
[5] "phq9.little_interest_or_pleasure_in_doing_things_numeric"
[6] "phq9.moving_fidgety_noticed_opposite_numeric"
[7] "phq9.poor_appetite_or_overeating_numeric"
[8] "phq9.staying_asleep_sleeping_trouble_numeric"
[9] "phq9.trouble_concentrating_reading_newspaper_numeric"
```

Vector of plausible numeric values for PHQ9 variables

```
phq9_vector_numeric <- c(
 0,
 1,
 2,
 3,
 -777,
 NA
 )
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
          variables = phq9_variables_numeric,
          values = phq9_vector_numeric)
[1] "There are no implausible values in the dataset. Can leave these variables as they are."
Table printed with {flextable}, not {gt}. Learn why at
http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in the code chunk header.
```

Characteristic	N = 47,458[1]
phq9.feeling_bad_failure_family_numeric	

-777	44 (<0.1%)
0	9,756 (22%)
1	14,158 (32%)
2	8,692 (19%)
3	11,976 (27%)
Missing	2,832
phq9.feeling_down_depressed_or_hopeless_numeric	
-777	50 (0.1%)
0	8,619 (19%)
1	18,908 (42%)
2	8,460 (19%)
3	8,589 (19%)
Missing	2,832
phq9.feeling_tired_or_having_little_energy_numeric	
-777	30 (<0.1%)

0	3,289 (7.4%)
1	13,115 (29%)
2	10,371 (23%)
3	17,821 (40%)
Missing	2,832
phq9.little_interest_or_pleasure_in_doing_things_numeric	
-777	36 (<0.1%)
0	9,825 (22%)
1	18,159 (41%)
2	8,432 (19%)
3	8,174 (18%)
Missing	2,832
phq9.poor_appetite_or_overeating_numeric	
-777	49 (0.1%)
0	10,287 (23%)

1	12,427 (28%)
2	9,602 (22%)
3	12,261 (27%)
Missing	2,832
phq9.staying_asleep_sleeping_trouble_numeric	
-777	35 (<0.1%)
0	6,381 (14%)
1	13,515 (30%)
2	9,400 (21%)
3	15,295 (34%)
Missing	2,832
phq9.trouble_concentrating_reading_newspaper_numeric	
-777	42 (<0.1%)
0	11,542 (26%)
1	14,702 (33%)

2	9,227 (21%)
3	9,113 (20%)
Missing	2,832

[1]n (%)

Categorical variables

Select categorical variables for cleaning

This will select all the non-numeric PHQ9 variables

```
phq9_variables <- dat %>%
  select(
         contains("phq9.") # select variables with phq9. at the start
         ) %>%
  select(
         !contains("_numeric") # deselect the numeric variables
         ) %>%
  colnames()


# check
phq9_variables
[1] "phq9.dead_hurting_thoughts"
[2] "phq9.feeling_bad_failure_family"
[3] "phq9.feeling_down_depressed_or_hopeless"
[4] "phq9.feeling_tired_or_having_little_energy"
[5] "phq9.little_interest_or_pleasure_in_doing_things"
[6] "phq9.moving_fidgety_noticed_opposite"
[7] "phq9.poor_appetite_or_overeating"
```

```
[8] "phq9.staying_asleep_sleeping_trouble"
[9] "phq9.trouble_concentrating_reading_newspaper"
```

Vector of plausible categorical values for PHQ9 variables

```
phq9_vector <- c(
 "Not at all",
 "Several days",
 "More than half the days",
 "Nearly every day",
 "Seen but not answered",
 NA
 )
```

Use imp_check function to find if any implausible values and obtain summary table of variables

```
imp_check(data = dat,
          variables = phq9_variables,
          values = phq9_vector)
[1] "There are no implausible values in the dataset. Can leave these variables as they are."
Table printed with {flextable}, not {gt}. Learn why at
http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in the code chunk header.
```

Characteristic	N = 47,458[1]
phq9.feeling_bad_failure_family	
Seen but not answered	44 (<0.1%)
Not at all	9,756 (22%)

Several days	14,158 (32%)
More than half the days	8,692 (19%)
Nearly every day	11,976 (27%)
Missing	2,832
phq9.feeling_down_depressed_or_hopeless	
Seen but not answered	50 (0.1%)
Not at all	8,619 (19%)
Several days	18,908 (42%)
More than half the days	8,460 (19%)
Nearly every day	8,589 (19%)
Missing	2,832
phq9.feeling_tired_or_having_little_energy	
Seen but not answered	30 (<0.1%)
Not at all	3,289 (7.4%)
Several days	13,115 (29%)

More than half the days	10,371 (23%)
Nearly every day	17,821 (40%)
Missing	2,832
phq9.little_interest_or_pleasure_in_doing_things	
Seen but not answered	36 (<0.1%)
Not at all	9,825 (22%)
Several days	18,159 (41%)
More than half the days	8,432 (19%)
Nearly every day	8,174 (18%)
Missing	2,832
phq9.poor_appetite_or_overeating	
Seen but not answered	49 (0.1%)
Not at all	10,287 (23%)
Several days	12,427 (28%)
More than half the days	9,602 (22%)

Nearly every day	12,261 (27%)
Missing	2,832
phq9.staying_asleep_sleeping_trouble	
Seen but not answered	35 (<0.1%)
Not at all	6,381 (14%)
Several days	13,515 (30%)
More than half the days	9,400 (21%)
Nearly every day	15,295 (34%)
Missing	2,832
phq9.trouble_concentrating_reading_newspaper	
Seen but not answered	42 (<0.1%)
Not at all	11,542 (26%)
Several days	14,702 (33%)
More than half the days	9,227 (21%)
Nearly every day	9,113 (20%)

Missing	2,832

[1]n (%)

Produce sumscores

Reference to scoring guidance here: Kroenke K, Spitzer RL, Williams JB; The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001 Sep 16(9):606-13.

Sumscore inputs

```
keys <- c(
  1, #1
  1, #2
  1, #3
  1, #4
  1, #5
  1, #6
```

```
 1, #7
 1, #8
 1 #9
 ) # should be 9 1s (none of the PHQ9 items are reverse coded)


sum_vars <- c(
 "phq9.dead_hurting_thoughts_numeric",
 "phq9.feeling_bad_failure_family_numeric",
 "phq9.feeling_down_depressed_or_hopeless_numeric",
 "phq9.feeling_tired_or_having_little_energy_numeric",
 "phq9.little_interest_or_pleasure_in_doing_things_numeric",
 "phq9.moving_fidgety_noticed_opposite_numeric",
 "phq9.poor_appetite_or_overeating_numeric",
 "phq9.staying_asleep_sleeping_trouble_numeric",
 "phq9.trouble_concentrating_reading_newspaper_numeric"
 )
```

Generate sumscores

Generate sumscores from questionnaire data and add to dat as new column

Sumscores assumes that all items in the questionnaire have the SAME minimum and maximum scores for ALL items, ensure that this is correct before proceeding

When adding the column name for your sumscore use “questionnaire.score_name”

```
dat <- dat %>%
 mutate(
        phq9.sum_score =
        sumscores(input = dat,
                sum_vars = sum_vars,
                coding_keys = keys,
                na_allowed = 0,
```

```
                min_item = 0,
                max_item = 3,
                min_score = 0,
                max_score = 27
                )$scores
        )
Warning in sumscores(input = dat, sum_vars = sum_vars, coding_keys = keys, :
Input contains non-answer values. These will be converted to NA_real_ for this calculation.
Warning in sumscores(input = dat, sum_vars = sum_vars, coding_keys = keys, :
Scores vector contains missing values.
# check
dat %>%
  descr(phq9.sum_score)
Descriptive Statistics
dat$phq9.sum_score
N: 47458


                phq9.sum_score
---------------- ----------------
        Mean            12.26
        Std.Dev          6.98
        Min              0.00
        Q1               7.00
        Median          12.00
           Q3           18.00
        Max             27.00
        MAD              7.41
        IQR             11.00
        CV               0.57
```

```
Skewness          0.23
SE.Skewness       0.01
Kurtosis         -0.86
N.Valid       44462.00
Pct.Valid        93.69
```

Clinical phenotyping of PHQ9

Current depression

Create binary phenotypes based on PHQ9 scoring criteria Validity has been assessed against an independent structured mental health professional (MHP) interview. PHQ-9 score ≥10 had a sensitivity of 88% and a specificity of 88% for major depression.

```
# numeric
dat <- dat %>%
  mutate(
        phq9.binary_depression_numeric =
        case_when(
        phq9.sum_score >= 10 ~ 1, # current depression
        phq9.sum_score < 10 ~ 0 # no current depression
        )
        )

# recode as categorical
dat <- dat %>%
  mutate(
        phq9.binary_depression =
        recode_factor(phq9.binary_depression_numeric,
        "1" = "Current depression", # current depression
        "0" = "No current depression" # no current depression
        )
```

```
        )

# check
dat %>%
  select(
        phq9.binary_depression_numeric,
        phq9.binary_depression
        ) %>%
  tbl_summary(missing_text = "Missing")
```

```
Table printed with {flextable}, not {gt}. Learn why at
http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in the code chunk header.
```

Characteristic	N = 47,458[1]
phq9.binary_depression_numeric	27,133 (61%)
Missing	2,996
phq9.binary_depression	
Current depression	27,133 (61%)
No current depression	17,329 (39%)
Missing	2,996

[1] n (%)

```
dat %>%
```

```
 select(
        phq9.binary_depression,
        phq9.binary_depression_numeric) %>% # variable is fine - just tbl_summary() giving
wrong result
 freq()
Frequencies
dat$phq9.binary_depression
Type: Factor

                          Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
------------------------- ------- --------- -------------- --------- --------------
        Current depression   27133    61.03          61.03   57.17          57.17
        No current depression   17329    38.97          100.00   36.51          93.69
                <NA>   2996                   6.31           100.00
                Total   47458   100.00         100.00 100.00          100.00

dat$phq9.binary_depression_numeric
Type: Numeric

        Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
---------- ------- --------- -------------- --------- --------------
        0   17329    38.97          38.97   36.51          36.51
        1   27133    61.03          100.00   57.17          93.69
        <NA>   2996                   6.31           100.00
        Total   47458   100.00         100.00 100.00          100.00
```

Severity of depression

```
# numeric
dat <- dat %>%
```

```
 mutate(
        phq9.severity_threshold_numeric =
        case_when(
        phq9.sum_score <= 4 ~ 0,
        phq9.sum_score > 4 & phq9.sum_score <= 9 ~ 1,
        phq9.sum_score > 9 & phq9.sum_score <= 14 ~ 2,
        phq9.sum_score > 15 & phq9.sum_score <= 19 ~ 3,
         phq9.sum_score > 19 & phq9.sum_score <= 27 ~ 4
        )
        )

# recode as categorical
dat <- dat %>%
 mutate(
        phq9.severity_threshold =
        recode_factor(phq9.severity_threshold_numeric,
        "0" = "None",
        "1" = "Mild",
        "2" = "Moderate",
        "3" = "Moderately severe",
        "4" = "Severe"
        )
        )

# check
dat %>%
 select(
        phq9.severity_threshold,
        phq9.severity_threshold_numeric
```

```
  ) %>%
  tbl_summary(missing_text = "Missing")
Table printed with {flextable}, not {gt}. Learn why at
http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
To suppress this message, include `message = FALSE` in the code chunk header.
```

Characteristic	N = 47,458[1]
phq9.severity_threshold	
None	6,731 (16%)
Mild	10,598 (25%)
Moderate	10,604 (25%)
Moderately severe	6,617 (16%)
Severe	8,050 (19%)
Missing	4,858
phq9.severity_threshold_numeric	
0	6,731 (16%)
1	10,598 (25%)
2	10,604 (25%)

3	6,617 (16%)
4	8,050 (19%)
Missing	4,858

[1]n (%)

Save cleaned data

Save variables for export

```
export_variables <- c("ID",
                      "startDate",
                      "endDate",
                      "sample",
                      "phq9.dead_hurting_thoughts",
                      "phq9.feeling_bad_failure_family",
                      "phq9.feeling_down_depressed_or_hopeless",
                      "phq9.feeling_tired_or_having_little_energy",
                      "phq9.little_interest_or_pleasure_in_doing_things",
                      "phq9.moving_fidgety_noticed_opposite",
                      "phq9.poor_appetite_or_overeating",
                      "phq9.staying_asleep_sleeping_trouble",
                      "phq9.trouble_concentrating_reading_newspaper",
                      "phq9.dead_hurting_thoughts_numeric",
                      "phq9.feeling_bad_failure_family_numeric",
                      "phq9.feeling_down_depressed_or_hopeless_numeric",
                      "phq9.feeling_tired_or_having_little_energy_numeric",
                      "phq9.little_interest_or_pleasure_in_doing_things_numeric",
```

```
                "phq9.moving_fidgety_noticed_opposite_numeric",
                "phq9.poor_appetite_or_overeating_numeric",
                "phq9.staying_asleep_sleeping_trouble_numeric",
                "phq9.trouble_concentrating_reading_newspaper_numeric",
                "phq9.sum_score",
                "phq9.binary_depression",
                "phq9.binary_depression_numeric",
                "phq9.severity_threshold",
                "phq9.severity_threshold_numeric"
                )
```

GLAD

```
dat %>%
  select(all_of(export_variables)) %>%
  filter(sample == "GLAD") %>%  # select only GLAD participants
  saveRDS(
          file = paste0(ilovedata, "/data/latest_freeze/glad/clinical/phq9_glad_clean.rds")
          )
```

EDGI

```
dat %>%
  select(all_of(export_variables)) %>%
  filter(sample == "EDGI") %>%  # select only EDGI participants
  saveRDS(
          file = paste0(ilovedata, "/data/latest_freeze/edgi/clinical/phq9_edgi_clean.rds")
          )
```

GLAD & EDGI

```
dat %>%
  select(all_of(export_variables)) %>%
```

```
 saveRDS(
          file = paste0(ilovedata,
"/data/latest_freeze/glad_edgi/clinical/phq9_glad_edgi_clean.rds")
          )
```