

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Leveraging genome-wide data to understand the risk and treatment of common mental disorders

ter Kuile, Abigail

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Leveraging genome-wide data to understand the risk and treatment of common mental disorders

Abigail Roos ter Kuile

Social, Genetic and Developmental Psychiatry Centre
Institute of Psychiatry, Psychology and Neuroscience
King's College London



This thesis is submitted for the degree of
Doctor of Philosophy in Statistical Genetics

Supervised by Professor Thalia Eley and Professor Gerome
Breen

March 2023

Abstract

Common mental disorders, including anxiety and depression, have a detrimental impact on global disease burden and quality of life. Both the risk and treatment of anxiety and depression are associated with the complex interplay between genetic, psychological, and socio-environmental factors. Integrating these factors has posed a challenge in research efforts aimed at improving the prevention and treatment of these disorders. The recent growth of genomic data and advancements in multi-trait methodology allows for the incorporation of genetics with psycho-social factors at an unprecedented level. This thesis applies statistical genetic approaches to further understand the genetic influences on the risk and resolution of common mental disorders. The first empirical study (**Chapter 2**) presents the largest genome-wide association study (GWAS) meta-analysis of lifetime fear-based anxiety disorders (N total =188,812; N cases =30,861) and examines the shared and distinct genetic relationship with generalised anxiety disorder (N total =172,248; N cases =54,928) and broad domains of other complex traits. **Chapter 3** builds on previous research finding a genetic component of reported trauma, a major socio-environmental risk factor for anxiety and depression. By leveraging pre-existing GWAS summary statistics, genetic correlations and genomic multiple regression analyses are used to identify heritable psychological and behavioural traits that capture the common genetic variant-based heritability of reported childhood maltreatment (N=185,414). **Chapter 4** represents the largest GWAS meta-analysis of outcomes following psychological treatment for anxiety and depression (N total =15,131; N reported positive outcomes =11,408). The utility of genetic factors is also assessed by incorporating polygenic scores of complex traits into multivariable prediction models of psychological treatment outcomes alongside known clinical and demographic predictors. The final chapter discusses the implications of the findings from this thesis within the context of challenges emerging in anxiety and depression genomics. With the ongoing expansion of GWAS data, consideration should be taken into how phenotyping approaches influence downstream analyses, including the genetic structure observed across psychopathologies and the psycho-social components involved in gene-environment interplay.

Table of contents

Abstract.....	2
Table of contents	3
List of tables and figures.....	5
Acknowledgments	7
Statement of authorship.....	8
Publications related to this thesis.....	10
Chapter 1. General introduction	12
<i>Anxiety and depression comorbidity is extensive and associated with adverse outcomes</i>	12
Research into understanding drivers of comorbidity whilst accounting for clinical heterogeneity.....	14
<i>Transdiagnostic and disorder-specific genetic influences</i>	16
Twin study findings	16
From twin studies to genome-wide methods to understand biological aetiology.....	19
Progress in depression and anxiety disorder genomics	26
Leveraging genomic data to understand pervasive comorbidity and clinical heterogeneity	28
<i>Social environmental influences on anxiety and depression.....</i>	31
The heritability of environmental measures	32
<i>Heritable psychological influences on anxiety and depression: from risk to treatment</i>	34
Heritable influences on subjective self-reports of adverse experiences	34
Heritable influences on positive experiences and application to psychological treatment	35
<i>Summary and aims.....</i>	38
<i>References</i>	39
Chapter 2. Exploring the genome-wide genetic overlap between anxiety and fear disorders	51
<i>Abstract</i>	52
<i>Introduction.....</i>	53
<i>Methods.....</i>	56
<i>Results.....</i>	62
<i>Discussion</i>	71
<i>References</i>	76
Chapter 3. Genetic decomposition of the heritable component of reported childhood maltreatment	81
<i>Abstract</i>	82
<i>Introduction.....</i>	83
<i>Methods and Materials</i>	86
<i>Results.....</i>	88
<i>Discussion</i>	93

<i>References</i>	99
Chapter 4. Genetic influences on self-reported outcomes following psychological treatment for depression and anxiety disorders	103
<i>Abstract</i>	104
<i>Introduction</i>	105
<i>Methods</i>	109
<i>Results</i>	117
<i>Discussion</i>	123
<i>References</i>	129
Chapter 5. General discussion	134
<i>General overview of findings</i>	134
<i>Theme 1: Phenotyping depth and the utility of self-reported measures</i>	137
<i>Theme 2: the genetic structure of internalising disorders and implications for disorder grouping</i>	139
<i>Theme 3: interpreting GWAS findings in the context of psychosocial influences</i>	143
<i>Personality, behaviour and cognitive biases</i>	144
<i>Socioeconomic status</i>	146
<i>General Limitations</i>	147
<i>Lack of replication and sources of bias</i>	147
<i>Lack of ancestral diversity</i>	149
<i>Future directions</i>	150
<i>Establishing causality, and disentangling individual from familial genomic effects</i>	150
<i>Expanding and integrating genome-wide data</i>	152
<i>Conclusions</i>	153
<i>References</i>	154
Appendix A. Supplementary materials for Chapter 2	160
Appendix B. Supplementary materials for Chapter 3	211
Appendix C. Supplementary materials for Chapter 4	241

List of tables and figures

Chapter 1

Figure 1: An overview of the three empirical chapters presented in this thesis and the overlap among genomic, socio-environmental and psychological factors.....	16
Box 1: Definitions and approaches to estimate trait heritability	18
Box 2: Definitions of types and consequences of human genetic variation.....	20
Box 3: Structural properties of the human genome and related population genetic concepts leveraged in genome-wide analyses	21
Figure 2: An illustration of the various stages of genome-wide analyses.....	23
Box 4: Definitions of pleiotropy, bivariate and multi-trait genetic analyses.....	24
Figure 3: An illustration of the genomic continuum of the mood disorder spectrum from mania genomics to internalising genomics.....	30
Figure 4: An illustration of the potential mechanisms underpinning the course and resolution of common mental disorders, which are influenced by the interplay between genetic, socio-environmental and psychological factors.....	31

Chapter 2

Table 1. Sample sizes in each study, GWAS dataset, and meta-analysis of fear-based disorders and GAD.....	59
Figure 1: Genome-wide association study Manhattan plots for anxiety disorder phenotypes meta-analysed across the GLAD+, QIMR and UKB datasets.....	63
Table 2: Independent genome-wide significant loci associated with anxiety disorder phenotypes	64
Table 3: Gene-level associations with anxiety disorder phenotypes.....	67
Table 4: Genetic correlations (r_g) between fear-based disorder and GAD phenotypes.....	69
Figure 2: Genetic correlations between anxiety disorder phenotypes and external traits estimated in LDSC regression.....	70

Chapter 3

Figure 1. Top bivariate genetic correlations (r_g) between reported childhood maltreatment and various heritable traits.	89
Figure 2. Path diagrams representing results from genomic multiple regression analyses. .	91

Chapter 4

Table 1. Number of participants from the GLAD+ and AGDS datasets with data on each measures of self-reported outcomes following therapy (Total N = 15,131).....	112
Table 2. Descriptive table for GLAD+ and AGDS datasets, total N = 15,131.....	117
Table 3. LDSC SNP-based heritability estimates (using summary-level data) in GLAD+ and AGDS datasets, total N = 15,131	118
Table 4. GCTA-GREML SNP-based heritability estimates (using individual-level data) in GLAD+ and AGDS datasets	118
Figure 1: Q-Q plot (left) and Manhattan plot (right) of associations from a meta-analysis of retrospectively self-reported treatment outcomes following psychological therapy (N = 15,131; 75% reported positive outcomes).....	119
Table 5. Univariable associations between predictors and retrospectively self-reported treatment outcomes in the GLAD+ dataset (N = 4,439)	121
Table 6. Models predicting self-reported outcomes following psychological therapy for depression/anxiety using genetic, sociodemographic and clinical predictors in the GLAD+ dataset (N = 4,439).....	122

Chapter 5

Figure 1: An illustration of the emergence of three key themes based on the findings of this thesis and the way they relate to various stages of GWAS.....	137
Figure 2: An illustration of the genetic continuum across the mood disorder spectrum to the internalising disorder spectrum, with the addition of fear and GAD.....	141

Acknowledgments

Thank you to the study participants who shared their life experiences, and all the individuals involved in designing and managing the studies and data, without whom the research presented in this thesis would not have been possible. Additionally, I am thankful to the NIHR Maudsley Biomedical Research Centre for providing me with the opportunity and funding to pursue this PhD.

I am deeply grateful to my supervisors, Thalia and Gerome, for their unwavering support and guidance. Your mentorship has provided me with numerous exciting opportunities to grow as a researcher and has been instrumental in shaping my research and enabling me to achieve my academic goals. To Thalia, thank you for your kindness, encouragement, and especially your support during the challenging period when I was unwell with long covid. You gave me invaluable insight to see the bigger picture of my research and communicate my findings more broadly, which I will take with me throughout my career. To Gerome, thank you for helping me find my passion and introducing me to the world of psychiatric genetics. You consistently had faith in my ability to tackle challenges and encouraged me to persevere. I am thankful for the opportunity to work with you both.

I feel fortunate to have had the opportunity to work at the SGDP Centre and meet so many wonderful people. Despite the pandemic, the support and sense of community that I experienced (even online) was truly special. I am grateful to my colleagues in the EDIT and TNG research teams for being a part of my journey. Working alongside such a friendly, sociable, and supportive group has been an enriching and rewarding experience.

A special thank you to Jess, Katie, Helena, Alicia, and Anna, whose friendship and collaboration have been a highlight of my PhD journey. You all have been an inspiration to me and have lifted me up during the most challenging times. Your support and encouragement have helped me grow as a scientist and as a person, and I have learned so much from each of you. Thank you for being there to celebrate every achievement and for encouraging me to be as proud of myself as I am of each of you.

To my dear friends, and especially, my housemates at Greig and Sandlings, thank you for taking care of me, reminding me to take breaks, and ensuring that I went outside to get some fresh air during periods of intense focus. Your kindness and positivity were crucial in helping me maintain my wellbeing.

I am also deeply grateful to my wonderful family - my Mum, Dad, and Naomi. Your unwavering love, support, and inspiration have brought me to where I am today. We spent a considerable portion of my PhD in lockdown together which I will always look back on fondly. Your care and encouragement when I was unwell with long covid helped me to pace myself until I was back to my normal self, enabling me to complete this hurdle.

Finally, to Oscar. I can't thank you enough for everything. You have been my rock and sounding board, providing me with love, encouragement, patience, and support every step of the way. You have kept me grounded and reminded me that there is life outside of the PhD when I needed it most. Thank you for being by my side throughout this journey.

Statement of authorship

All the work presented in this thesis is my own except where acknowledged in the text. Data collection for all study samples was completed by the respective research teams. Genetic data quality control (QC) was carried out by respective studies. In **Chapters 2 and 4**, I conducted additional QC in subsamples of the GLAD+ datasets (Genetic Link to Anxiety and Depression (GLAD) study and the COVID-19 Psychiatric and Neurological Genetics (COPING) study). During the duration of this PhD, I was involved in setting up and maintaining three pipelines in collaboration with colleagues to generate repositories of data used in this thesis and for future studies. First, I contributed to the phenotypic data cleaning pipeline of the GLAD+ dataset (used in **Chapters 2 and 4**). Second, I helped maintain and update a GWAS summary statistics QC pipeline and repository of hundreds of complex traits (used in **Chapters 2, 3 and 4**). Third, using the GWAS summary statistics from this repository, I contributed to setting up a polygenic score repository in the GLAD+ dataset (**Chapter 4**).

The studies presented in **Chapters 2 and 3** were conceived and conducted by me as first author (A.R.T.K), in collaboration with colleagues included in the author list presented at the start of each chapter. The work presented in **Chapter 4** was a joint first author study conducted with Alicia J. Peel, reflecting equal contributions from myself and A.J.P. I contributed to the work in **Chapter 4** by leading on genome-wide association analyses in the GLAD+ dataset, GWAS meta-analysis across datasets, post-GWAS analyses, genetic quality control procedures and additionally the creation of polygenic risk scores for prediction modelling, with input from A.J.P. I assisted A.J.P who led on preparation of phenotypic data, prediction modelling, preparing the pre-registration and an original draft of the manuscript. **Chapters 2 and 4** include GWAS meta-analyses that were conducted in collaboration with researchers involved in the Australian Genetics of Depression study (AGDS) based at Queensland Institute of Medical Research Berghofer (QIMR). Analyses in the AGDS dataset were conducted by researchers at the QIMR. I led on coordinating and generating scripts to harmonise data analysis with AGDS collaborators, including phenotypic definitions and genome-wide analysis scripts. A summary of chapter-specific author contributions is shown below.

Abigail ter Kuile

Chapter 2

A.R.T.K, T.C.E and G.B. were responsible for the study conception and design. T.C.E, G.B, G.K., M.R.D, M.H., N.R.W, S.E.M, E.M.B, N.G.M were responsible for the acquisition of the data. A.R.T.K, C.H., A.J.P, H.L.D, J.M., Y.L., J.Z, A.P. were responsible for phenotypic data preparation and cleaning. Preparation and quality control of genetic data was conducted by A.R.T.K, S.H.L, J.R.I.C and B.A. B.M carried out data analysis in the AGDS QIMR dataset. A.R.T.K conducted all other data analyses with support from G.M.V, S.H.L, H.L.D, J.M, M.S, J.R.I.C and A.E.F. A.R.T.K was responsible for the drafting and revising of the manuscript, under the close supervision of T.C.E. All co-authors reviewed the manuscript.

Chapter 3

A.R.T.K and G.B were responsible for the study conception and design, with input from all co-authors. A.R.T.K, G.B, C.H, J.R.I.C, D.F.L, J.G and M.B.S were responsible for the acquisition of the data. A.R.T.K conducted the analyses, with support from C.H. and C.R. A.R.T.K was responsible for the drafting and revising of the manuscript, under the close supervision of G.B and T.C.E. All co-authors contributed to the interpretation of the data, and reviewed and edited the manuscript.

Chapter 4

A.R.T.K, A.J.P, C.R and T.C.E conceived and designed the study. T.C.E, G.B, G.K., M.R.D, M.H., N.R.W, S.E.M, E.M.B, N.G.M were responsible for the acquisition of the data. Phenotypic data in the GLAD+ dataset was processed and cleaned by A.J.P, A.R.T.K, C.H., H.L.D, J.M. Preparation and quality control of genetic data was conducted by A.R.T.K, S.H.L, J.R.I.C and B.A. A.R.T.K led on genome-wide analyses, with support from J.M, M.S and S.H.L. A.R.T.K generated polygenic scores in the GLAD+ dataset, with support from H.L.D and Oliver Pain. A.J.P ran prediction modelling analysis, with support from A.R.T.K and Oliver Pain. B.M and K.H carried out data analysis in the AGDS dataset. All authors contributed to the interpretation of the results. A.J.P led on preparing an original draft of the manuscript submitted in A.J.P's PhD thesis, with input from A.R.T.K and supervisory input from T.C.E. A.R.T.K was responsible for preparing and revising the current version of the manuscript included in this thesis, with supervisory input from T.C.E and revised by G.B.

Publications related to this thesis

Publications resulting from the chapters in this thesis

Chapter 3 has been accepted for publication following peer-review in the journal *Biological Psychiatry Global Open Science*.

ter Kuile, A. R., Hübel, C., Cheesman, R., Coleman, J. R. I., Peel, A. J., Levey, D. F., Stein, M. B., Gelernter, J., Rayner, C., Eley, T. C., & Breen, G. (In press). Genetic decomposition of the heritable component of reported childhood maltreatment. *Biological Psychiatry Global Open Science*.
<https://doi.org/10.1016/j.bpsgos.2023.03.003>

Chapter 2 and **4** are currently in preparation for publication.

Chapter 2:

ter Kuile, A. R., Mitchell, B. L., Morneau-Vaillancourt, G., Lee, S. H., Davies, H. L., Mundy, J., Peel, A. J., Skelton, M., Hübel, C., Davies, M. R., Coleman, J. R. I., Fürtjes, A. E., Ahmad, Z., Lin, Y., Adey, B. N., McGregor, T., Purves, K., Palmos, A., Zvrskovec, J., Hotopf, M., Kalsi, G., Jones, I. R., Smith, D. J., Veale, D., Walters, J. T. R., Armour, C., Hirsch, C. R., McIntosh, A. M., Wray, N. R., Medland, S. E., Byrne, E. M., Martin, N. G., Breen, G., & Eley, T. C. (In prep). Exploring the genome-wide genetic overlap between anxiety and fear disorders.

Chapter 3:

Peel, A. J. *, **ter Kuile, A. R.***, Rayner, C., Hopkins, K., Mitchell, B. L., Medland, S. E., Lee, S. H., Adey, B. N., Armour, C., Buckman, J. E. J., Byrne, E. M., Coleman, J. R. I., Danese, A., Davies, M. R., Davies, H. L., Hickie, I. B., Hirsch, C., Hotopf, M., Hübel, C., Jones, I. R., Kalsi, G., Krebs, G., Martin, N. G., McIntosh, A. M., Mundy, J., Purves, K. L., Rogers, H., Skelton, M., Smith, D. J., Veale, D., Walters, J. T. R., Wray, N. R., Breen, G., & Eley, T. C. (In prep). Genome-wide analysis and prediction of self-reported outcomes following psychological treatment for depression and anxiety.

* *Joint first authors*

Additional publications

Co-authored publications in peer-reviewed journals arising from additional work conducted during this PhD

- Rayner, C., Coleman, J., Skelton, M., Armour, C., Bradley, J. R., Buckman, J. E. J., Davies, M., Hirsch, C., Hotopf, M., Hübel, C., Jones, I. R., Kingston, N., Krebs, G., Lin, Y., Monssen, D., McIntosh, A., Mundy, J., Peel, A., Rimes, K., Rogers, H., Smith, D. J., **ter Kuile, A. R.**, ... Breen, G. & Eley, T.C., (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, C., Parsons, C. E., Vassos, E., Adey, B. N., Bright, S., Hegemann, L., Lee, Y. T., Kalsi, G., Monssen, D., Mundy, J., Peel, A. J., Rayner, C., Rogers, H. C., **ter Kuile, A. R.**, ... , Eley, T. C. & Breen, G., (2022). Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK. *Psychological Medicine*.
- Peel, A. J., Purves, K. L., Baldwin, J. R., Breen, G., Coleman, J. R. I., Pingault, J. B., Skelton, M., **ter Kuile, A. R.**, Danese, A. & Eley, T. C., (2022). Genetic and early environmental predictors of adulthood self-reports of trauma. *British Journal of Psychiatry*.
- Davies, M. R., Buckman, J. E. J., Adey, B. N., Armour, C., Bradley, J. R., Curzons, S. C. B., Davies, H. L., Davis, K. A. S., Goldsmith, K. A., Hirsch, C. R., Hotopf, M., Hübel, C., Jones, I. R., Kalsi, G., Krebs, G., Lin, Y., Marsh, I., McAtarsney-Kovacs, M., McIntosh, A. M., Mundy, J., Monssen, D., Peel, A. J., Rogers, H. C., Skelton, M., Smith, D. J., **ter Kuile, A. R.**, ... , 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., Curzons, S. C. B., Davies, M. R., Hübel, C., Jones, I., Kalsi, G., McAtarsney-Kovacs, M., McIntosh, A. M., Monssen, D., Mundy, J., Rayner, C., Rogers, H. C., Skelton, M., **ter Kuile, A. R.**, Thompson, K. N., Breen, G., 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*.
- Davies, H. L., Hübel, C., Herle, M., Kakar, S., Mundy, J., Peel, A. J., **ter Kuile, A. R.**, ... 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. *The International Journal of Eating Disorders*.
- Davies, M. R., Glen, K., Mundy, J., **ter Kuile, A. R.**, Adey, B. N., Armour, C., Assary, E., Coleman, J. R. I., Goldsmith, K. A., Hirsch, C. R., Hotopf, M., Hübel, C., Jones, I. R., Kalsi, G., Krebs, G., McIntosh, A. M., Morneau-Vaillancourt, G., Peel, A. J., Purves, K. L., ... Eley, T. C. (2023). Factors associated with anxiety disorder comorbidity. *Journal of Affective Disorders*.

Chapter 1. General introduction

Anxiety and depression comorbidity is extensive and associated with adverse outcomes

Common mental disorders account for a significant proportion of the global disease burden¹. They comprise multiple anxiety and depressive disorders and are a leading cause of disability, with a global lifetime prevalence of approximately 16% and 11%, respectively²⁻⁴. Anxiety, fear and sadness are typical emotions experienced in everyday life, but when experienced at pathological levels can interfere with day-to-day functioning and become debilitating⁵. The impairing features of anxiety and depression lead to a widespread disability affecting broad aspects of quality of life, including problems maintaining relationships, poor educational attainment, unemployment and financial difficulties⁵⁻⁸.

In addition to the detrimental impact on quality of life, there is a strong economic case for improving the prevention and treatment of common mental disorders. In the UK, anxiety and depression contribute 17.5% (£20.5 billion) and 22.5% (£26.6 billion), respectively, to a minimum annual cost of £117.9 billion arising from mental health problems. This equates to 5% of the total gross domestic product. These costs include the effect of mental health problems on disability, occupational productivity and health services. Moreover, current UK mental health services are under extreme demand and struggle to meet the needs of all those who seek a diagnosis or treatment⁹. When left untreated, anxiety disorders and depression are often experienced for extensive periods, becoming either chronic or recurrent, contributing to a large proportion of suicide^{1,10-12}.

Anxiety disorders are characterised by excessive fear and anxiety, accompanied by impairing features such as avoidance behaviours, physical anxiety reactions and panic attacks. In the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), specific clusters of symptoms differentiate one anxiety disorder from the other and include generalised anxiety disorder (GAD), agoraphobia, social anxiety disorder, specific phobia and panic disorder¹³. The breadth and focus of the threat that causes anxiety or fear differs across the disorders. General anxiety is the anticipation and pathological preoccupation with a range of future potential threats. Essential criteria for a GAD diagnosis include experiencing general anxiety persistently for at least six months and difficulties controlling the worry. In contrast, fear is an emotional and physiological reaction to immediate threats. The particular focus of perceived threat distinguishes the phobias from each other and must consistently be feared or avoided for a minimum period of six months for a diagnosis. The focus of fear is the broadest in

agoraphobia; the fear that escape might not be possible or help might not be available if something were to go wrong in a range of situational exposures. In social anxiety disorder (also known as social phobia), a narrower focus is placed on situations involving social scrutiny. In specific phobias, the focus is limited to a certain situation or object that elicits fear. A panic disorder diagnosis requires experiencing recurrent, unexpected panic attacks for at least one month and consistent worry and avoidance of situations associated with having panic attacks⁵.

Major depressive disorder (MDD) is the most common depressive disorder and has considerable diagnostic heterogeneity. For a diagnosis of MDD, a low mood or the reduced capacity to experience pleasure must be present and persist for at least two weeks, along with the presence of other symptoms⁵. Clinical subtype specifiers can be symptom-based (with, e.g. anxious-distress, mixed hypomanic, atypical, psychotic, melancholic features), temporal or etiological-based (e.g. seasonal patterning, postpartum onset) and recurrence-based (single episode or recurrent MDD)^{5,14}.

Anxiety and depressive disorders are highly comorbid with one another, as well as other psychiatric disorders and physical health problems. Estimates of comorbidity with another mental health disorder are as high as 89% for anxiety disorders¹⁵ and approximately 75% for major depressive disorder^{16,17}. This includes comorbidity with a range of emotional and behavioural psychopathologies, namely alcohol and substance use disorders, post-traumatic stress disorder, obsessive-compulsive disorder, personality disorders and attention deficit hyperactivity disorder^{5,15,18,19}. However, the most common co-occurrence is among individual anxiety disorders and depression. Approximately half of those who meet the diagnostic criteria for one anxiety disorder have a history of another anxiety disorder¹⁰. Similarly, over half of those with depression have at least one anxiety disorder and vice versa^{16,20,21}.

In addition to anxiety and depression comorbidities being extensive, they are also associated with various adverse clinical outcomes. A broad range of comorbid physical health problems is common, including cardiovascular disease, diabetes, cancer, and irritable bowel syndrome^{5,22–24}. The presence of anxiety and depression is also linked with poor prognostic outcomes in both chronic somatic diseases²⁴ and, to some extent, treatment response in other psychiatric disorders²⁵. Compared to individuals with one psychiatric disorder, anxiety-depression comorbidity is associated with increased symptom severity, impairment, chronicity, the likelihood of suicide attempts, and is harder to treat^{20,21,26,27}.

Research into understanding drivers of comorbidity whilst accounting for clinical heterogeneity

Due to the magnitude and clinical severity of anxiety-depression comorbidity, examining factors driving shared liability has been a focus in decades of research²⁸. Comorbidity can be modelled for clinical and research utility by examining covariance patterns within a structural equation modelling (SEM) framework. Modelling patterns of comorbidity highlight the distinctive and shared features between disorders and indicate overlapping or unique liability factors²⁹. Such research has implications for improving diagnostic systems and understanding transdiagnostic or disorder-specific risk factors and treatment targets. Previous research examining shared liability underpinning comorbidity patterns has challenged diagnostic categories and continues to do so today^{28,30,31}. These studies have aimed to address the problem of pervasive comorbidity whilst taking into account clinical heterogeneity.

The extensiveness of anxiety-depression comorbidity is reflected in well-replicated disorder-based and symptom-based factor analyses. These have revealed a broad shared liability to both anxiety and depression along with other disorders characterised by emotional symptoms (e.g. PTSD, OCD and eating disorders). Anxiety and depressive disorders are often conceptualised as being core to this internalising disorder dimension, thought to be driven by shared negative affect³¹. Disorders outside of this dimension are differentiated by other core components of psychopathology. For example, disorders characterised primarily by behavioural difficulties comprise an externalising disorder spectrum (e.g. substance use disorders, borderline personality disorder). Disorders characterised by disturbances in thought content with delusions, hallucinations or dissociative symptoms form the thought disorder spectrum (i.e. schizophrenia spectrum disorders)^{32–34}. The placement of some aspects of disorders varies, for example, neurodevelopmental disorders (e.g. ASD, ADHD) and mania (i.e. bipolar disorders)^{35,36}. As discussed above, anxiety disorders and depression are also comorbid with disorders in other spectra. This is captured by an overarching psychopathology liability dimension, known as the ‘p-factor’, which accounts for shared features and high rates of comorbidity observed across psychiatric disorders³⁷.

Within these spectra lie subfactors that represent disorder-specific clinical heterogeneity. These subfactors reflect differences in core features and comorbidity patterns among the specific disorders. For example, GAD has higher rates of comorbidity with depression than other anxiety disorders¹⁷. A well-replicated structure in the internalising spectrum is splitting anxiety disorders and depression into two highly correlated subfactors: distress and fear. Distress-based disorders are mainly characterised by broad, pervasive negative emotionality

(i.e. GAD is grouped with major depression). In contrast, fear-based disorders have more context-specific features of distress, with fearful arousal and behavioural avoidance of a narrower range of stimuli (i.e. agoraphobia, social anxiety disorder, specific phobia and panic disorder)^{31,38}. The splitting of internalising into distress and fear subfactors suggests an overarching shared liability and some subfactor-specific risk factors that are not shared across distress-fear dimensions.

The complex aetiology of common mental disorders poses a challenge to understanding what drives comorbid psychopathologies and within-disorder heterogeneity. Both the risk and treatment of anxiety and depression are associated with a broad range of genetic, psychological and socio-environmental factors^{11,13}. Such factors on their own are not enough to cause the disorders or explain why some individuals respond better to certain treatments than others. Furthermore, they do not act independently of one another, with a complex, multi-layered interplay across different types of biological, psychological and social/environmental factors. The interplay between these influences is dynamic, varying across the lifespan, from person to person and across different psychiatric disorders³⁹. Thus, resolving heterogeneity and comorbidity problems in psychiatric research requires not only data on a broad range of psychopathologies, but also a multi-level approach that integrates genetic, psychological, and environmental factors.

In the next sections of this thesis introduction, progress in research of each facet of the bio-psycho-social approach in the risk and treatment of anxiety and depression will be described. By leveraging the complex interplay between genetic, environmental and psychological influences, a focus will be placed on how genomic data can be used to understand each facet and how this thesis aims to address key gaps in anxiety and depression genomics research, depicted in **Figure 1**. A glossary of key statistical genetic terms and concepts relevant to this thesis is provided in **Boxes 1-4**.

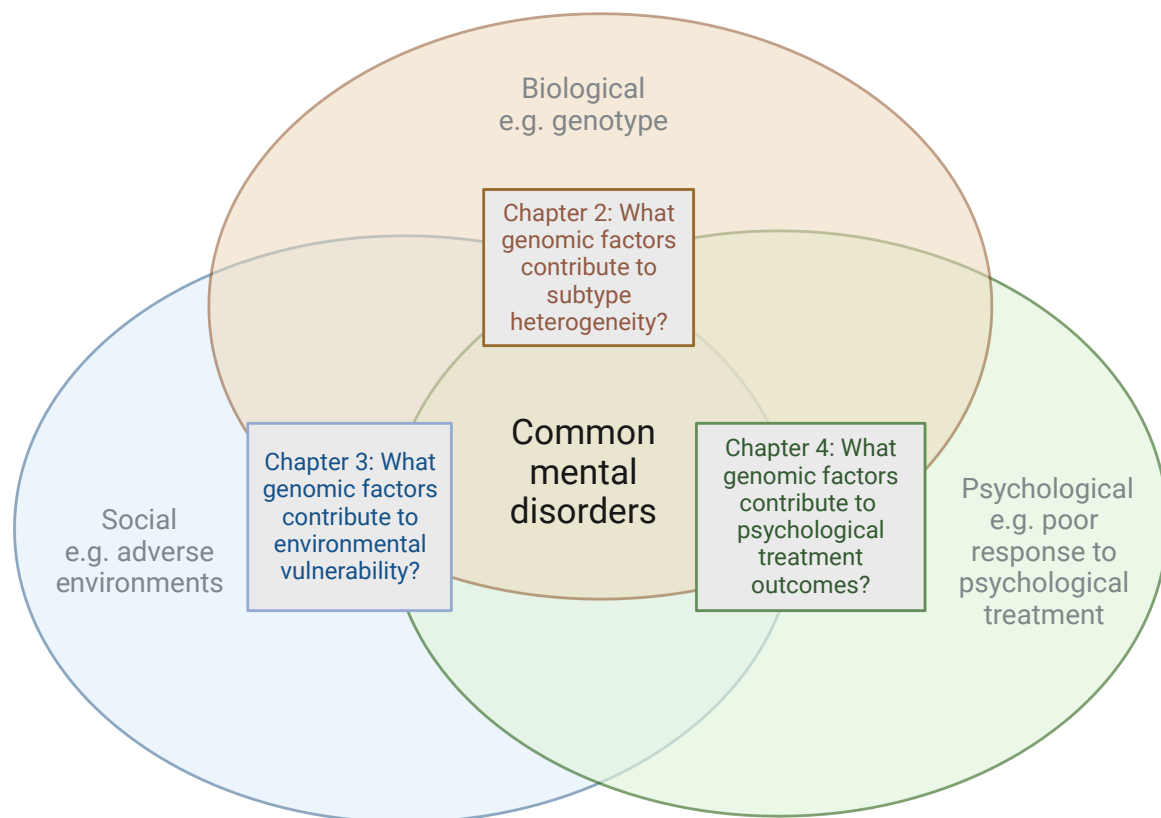


Figure 1: An overview of the three empirical chapters presented in this thesis and the overlap among genomic, socio-environmental and psychological factors.

Through the application of genome-wide data, the overlap between genomic and psychosocial factors can be harnessed to better understand risk factors and treatment outcomes for common mental disorders.

Transdiagnostic and disorder-specific genetic influences

Twin study findings

Pivotal findings from family and twin studies formed the basis for understanding genetic influences on anxiety and depression, which paved the way for future genomic studies. Research dating to the early-mid 20th century showed that anxiety and depressive disorders aggregate in families. A four to six-fold increased risk of anxiety disorders and a three-fold risk of depression were observed in individuals with an affected first-degree relative compared to unaffected controls^{40,41}. Co-aggregation across individual anxiety disorders and depression was also observed⁴².

Twin studies confirmed that the familial aggregation of anxiety and depressive disorders is largely due to heritable factors and, to a lesser extent, environmental factors^{40–42}. Definitions

and approaches to estimate heritability, including twin study estimates, are provided in **Box 1**. Twin study heritability estimates of anxiety and depressive disorders are moderate, ranging from 30-50%⁴⁰⁻⁴³. While substantial, these estimates are lower than rarer psychiatric disorders such as schizophrenia and bipolar disorder (65-80%)^{43,44}. Family and twin studies largely support that familial risk and heritability estimates are similar across anxiety and MDD subtypes; however, some find earlier-onset and recurrent MDD show greater familial risk and heritability^{40,42,43,45}. Twin studies have found that heritability estimates of anxiety and depression measures are broadly similar across age groups⁴³, although some report lower heritability for childhood than adult depressive symptoms⁴⁶. Stable genetic influences may largely explain the moderate, long-term stability of anxiety and depressive symptoms across the lifespan^{47,48}. Furthermore, both stable and time-specific genetic factors are shared among anxiety and depressive disorders⁴⁹. These findings highlight how shared genetic influences on anxiety and depression risk are important across the lifespan.

A major finding from twin studies is that high levels of shared genetic influences largely explain the high comorbidity across anxiety and depressive disorders and related symptom dimensions. Several twin studies found that GAD and MDD overlap almost entirely in genetic variance⁵⁰. A high genetic overlap across individual anxiety disorders is also consistently greater than disorder-specific genetic influences⁴². Twin models of genetic covariation have shown substantial shared genetic liability among anxiety and depression in childhood, adolescence and adulthood⁵⁰⁻⁵³. A broad, internalising genetic factor is well-replicated across twin datasets, supporting the overarching phenotypic structure of the internalising spectrum discussed earlier^{51,53-55}. The high genetic overlap between neuroticism, anxiety and depression indicates genetic influences on shared negative affectivity drive internalising psychopathology and anxiety-depression comorbidity⁵⁶. Twin studies have further explored the genetic relationship between anxiety and depression and disorders outside of the internalising spectrum, showing pervasive genetic overlap with practically all forms of psychopathology^{42,50}. This overlap indicates cross-disorder genetics and the genetic factors structure across disorders can be modelled as an overarching 'p-factor' liability that drives comorbidity^{57,58}. Studies also support partitioning psychiatric disorders into two distinct but correlated lower-order internalising and externalising clusters⁵⁹. As such, compared to their overlap with other disorders, the genetic overlap is highest between anxiety and depression.

Notably, some multivariate twin modelling also supports the distress-fear subfactors structure *within* the internalising dimension^{51,60,61}. Multiple studies find that the extent of genetic overlap between individual anxiety disorders and depression varies⁵⁰. GAD and MDD show the highest genetic overlap and are core to the internalising factor, indicating distress is a key

driver of general emotional problems⁵⁹. Panic disorder and phobias have robust genetic correlations with distress but also show some genetic distinction from GAD and MDD, forming a genetic liability fear factor^{51,60,61}. Together, twin study findings indicate that grouping disorders is appropriate to identify higher-order genetic influences that broadly influence internalising. However, genetic studies that group distress with fear disorders may also miss important sources of heterogeneity and insight into mechanistic differences.

Box 1: Definitions and approaches to estimate trait heritability

Heritability: The proportion of total phenotypic variance observed within a population (V_P) explained by genetic factors, with estimates ranging from zero to one. Heritability estimates of a trait can vary across time and populations due to differences in, for example, environmental contexts, the type of phenotypic instrument and age at measurement^{62,63}.

Broad-sense heritability: Defined as $H^2 = V_G / V_P$ with genetic values (V_G) attributed to both additive and non-additive genetic factors. Non-additive influences include interactions between alleles located within the same genetic locus (dominance) or located on different loci (epistasis)^{62–64}.

Narrow-sense heritability: Defined as $h^2 = V_A / V_P$ with genetic values (V_A) attributed only to additive genetic factors; the sum of the average effect of multiple alleles contributing to a trait^{62–64}.

Twin study heritability: Twin studies are a commonly used family-based approach to estimate trait heritability and environmental components. Phenotypic differences between genetically identical monozygotic twins (MZ) and ~50% genetically similar dizygotic twins (DZ) are compared. Traits that are more phenotypically similar in MZ than in DZ twins indicate the role of genetic variation. Variance components modelled in twin designs include additive genetics (A), non-additive genetics (D), the individual environment not shared among sibling pairs (E), and the common environment shared among siblings growing up in the same family, thus accounting for some within family resemblance (C). In the classical ACE twin model, narrow-sense heritability is estimated based on the assumption that the non-additive effects are zero. In the ADE model, the C component is replaced to estimate D, enabling broad-sense heritability to be calculated from the combined variance of A and D^{64,65}. Twin studies have found stronger evidence for the role of additive genetic influences on complex traits than non-additive^{43,66}.

SNP-based heritability (h^2_{SNP}): The proportion of phenotypic variance explained by a given set of genome-wide common genetic variants⁶⁷.

GREML h^2_{SNP} : Genomic relatedness restricted maximum likelihood (GREML) is a commonly used method to estimate narrow-sense h^2_{SNP} using individual-level genomic data in unrelated individuals⁶⁸. GREML is implemented in the software GCTA⁶⁹. Genome-wide common SNPs are leveraged to construct a genomic relatedness matrix (GRM) containing the relatedness coefficient between each pair of individuals across each SNP. The extent to which genomic similarity can predict pair-wise phenotypic similarity for a given trait is used to deconstruct trait variance into additive genetic variance and residual variance components (e.g. non-additive genetics, environment, error)^{64,67}.

Box 1: Definitions and approaches to estimate trait heritability

LDSC regression h^2_{SNP} : Linkage disequilibrium score (LDSC) regression is a method to estimate narrow-sense h^2_{SNP} using GWAS summary statistics. LDSC regression assumes under a polygenic model that common variants in higher LD regions are, on average, more likely to be tagged by a causal variant in GWAS and therefore are more likely to be associated with a given trait. An LD score quantifies the level of tagging of variants. A higher score corresponds to a variant that tags more causal variants than those with a lower score and, therefore, on average, has a higher association test statistic. LD scores are created using a population reference panel representing the LD structure. For accurate estimates, the population must be well-matched to the ancestry of the target GWAS population. An estimate of h^2_{SNP} is quantified as the regression coefficient from regressing the observed mean chi-square association test statistic (X^2) against the LD score of each variant. The intercept can be used to quantify bias from population stratification but can arise from an increase in sample size and heritability of a trait. The attenuation ratio ($[\text{intercept} - 1] / [X^2 - 1]$) provides an estimate of inflation not due to polygenicity, with a higher estimate indicating confounding^{67,70,71}.

Liability scale h^2_{SNP} : Observed scale h^2_{SNP} estimates of categorical traits are converted to the liability scale by accounting for the assumed population prevalence. The liability threshold model presumes that genetic factors influence disorders under a normal distribution. An individual meets case disorder status once reaching above a certain threshold of liability (on the upper tail end of the normal distribution). The proportions of affected (cases) and unaffected individuals (controls) in a GWAS sample often do not reflect the incidence of a disorder observed in the population. The observed h^2_{SNP} scale is often lower than the liability h^2_{SNP} scale, as information is lost by dichotomising the trait^{62,72}.

From twin studies to genome-wide methods to understand biological aetiology

The heritable component and pervasive genetic overlap between anxiety and depressive disorders showed the potential for employing genetic techniques to unravel biological aetiology. Initial attempts to identify the genes associated with anxiety and depression involved candidate gene and linkage analysis studies. As with all complex traits, these failed to identify robust associations with anxiety and depression that survived replication⁴². As knowledge of the human genome expanded, it was soon realised that psychiatric and other complex traits are affected by many genetic variants, each of small effect size. In an attempt to uncover this polygenicity, the era of genome-wide association studies (GWAS) began⁷³.

A period of accelerated progress in genomics that led to the development of GWAS began with the completion of the first human genome sequence, revealing millions of genetic variants^{74,75}. **Box 2** provides a description of the various forms of human genetic variation. Single nucleotide polymorphisms (SNP) are the most extensively analysed in GWAS, followed by short insertion-deletion variants or 'indels'. The majority of variants analysed in GWAS are located within non-coding sequences of genes or in intergenic regions, which are potentially involved in gene regulation^{76–78}.

Box 2: Definitions of types and consequences of human genetic variation

Allele: Alternative versions of base pair changes of a variant⁷⁹.

Polygenic: A phenotype influenced by many genetic variants⁷⁹. Each genetic variant contributes to phenotypic variation with a small effect size. The sum of each small effect, along with environmental influences, contributes to the continuous distribution of complex trait variation observed in a population⁸⁰.

Polymorphism: Multiple definitions of genetic polymorphism exist in the literature⁸¹. Here, polymorphism is defined as a population-specific term related to the occurrence of two or more alleles at a position in a DNA sequence, with the rarer allele occurring in at least 1% of a given population^{82–84}.

Single nucleotide variant: Variation from a reference genome at a single base pair change. The most abundant form of genetic variation in the human genome^{76,77}.

Single nucleotide polymorphism (SNP): A single nucleotide variant with the minor allele present in at least 1% of the population^{77,85}.

Insertions and deletions (indels): A variant that corresponds to the insertion or deletion of one or more base pairs not found in the reference genome. The second most abundant form of genetic variation in the human genome^{76,77}.

Copy number variation: An intermediate-large scale form of structural variation, corresponding to additional copies (duplications) or deletions of segments of sequences > 1,000 base pairs^{85,86}.

Tandem repeats: A segment of sequences with a number of base pair repeats. Variants with repeat sequences of 2-6 bases are termed short tandem repeats, and those 7-10 bases are termed variable number tandem repeats⁷⁷.

Synonymous and non-synonymous variants: SNVs, and to a lesser extent SNPs, can occur in protein-coding sequences within genes (exons), and cause a change in the amino acid sequence, termed a non-synonymous variant. In contrast, variants within exons that do not change the amino acid sequence are termed synonymous variants⁷⁷.

Missense variant: A type of non-synonymous variant that alters one amino acid in the protein, potentially influencing protein function⁷⁷.

Nonsense variant: A type of non-synonymous variant that results in a premature STOP codon and an incomplete and potentially non-functional protein product⁷⁷.

Intergenic variant: A variant located between genes⁷⁷.

Intronic variant: A variant that falls within an intron - the non-coding sequences of genes⁷⁷.

Since the completion of the first human genome sequence, technological advancements have provided cost-effective tools to rapidly conduct GWAS⁷⁷. As the number of sequenced human genomes increased, the depth and detail of publicly available reference genomes improved^{87,88}. Such advancements have led to an improved understanding of the structural properties of the human genome and population genetic concepts that are leveraged for genome-wide analyses (**Box 3**). The cost of generating genotyping data for GWAS through microarrays has decreased substantially over the years. It is the most commonly used approach to gather genotype data as it costs approximately ~95% less than next-generation sequencing⁸⁹. Reference genomes are then used for imputation to expand the number of variants detected from microarrays, improving the overlap of variants across datasets using

different microarray technologies, thereby facilitating data harmonisation. The millions of variants generated by these approaches are used in GWAS to conduct a hypothesis-free study by testing across the entire genome if the frequency of alleles differs between cases and controls (e.g. an allele is more common in individuals with anxiety disorder than in those without), or along a trait dimension (e.g. an allele is associated with a higher GAD symptom score)⁷⁷. Variant associations may be causally or non-causally related to phenotype expression, but are indirectly detected in GWAS due to being correlated with nearby causal variants (in linkage disequilibrium [LD]). Haplotype blocks with regions of high LD are leveraged to identify causal variants without the need for genotyping all causal SNPs^{90–92}. However, interpreting GWAS results is complicated by LD, giving rise to multiple associated variants within a region, and subsequent analyses are required to determine causal associations^{92–94}.

Box 3: Structural properties of the human genome and related population genetic concepts leveraged in genome-wide analyses

Reference panel: A database of densely genotyped haplotypes measured in a given set of individuals from a specific population. Used to represent structural properties of the genome in a given population, including allele frequencies and LD patterning. Examples of commonly used reference panels include the Haplotype Reference Consortium, 1000 Genomes Project, and TOPMed⁹⁵.

Allele frequencies: Allele frequency is the incidence of a certain allele in a population. In a given population, minor allele frequency (MAF) is the frequency of the less common allele, while major allele frequency is the frequency of the more common allele of a variant. MAF is used to distinguish common from rare variants, with MAF > 1-5% considered common^{79,90}.

Linkage disequilibrium (LD): The non-random association between alleles located on two distinct loci in a given population. Linkage equilibrium (LE) implies independence, in which alleles at two loci are randomly associated by chance, with a theoretical disequilibrium coefficient of 0. When the frequency of association between alleles at two different loci deviates from expected when in LE, they are considered to be in LD. Levels of LD are dependent on recombination events during meiosis that occur across generations, natural selection, population bottlenecks, genetic drift, mutations, and genomic inversion. LD can occur at a short-range, whereby LD is higher among physically nearby genetic variants that are less likely to be separated by recombination events than more distant variants. LD can occur at a long range, whereby the correlation between genetic variants occurs at a larger distance, caused by various population processes^{91,96}.

Haplotype block: Across the human genome are chromosomal regions with block-like patterns of different LD levels. Low LD regions exist in recombination hot spots. In between these hotspots are haplotype blocks, regions with high LD (from a few kb-100kb) and low recombination, resulting in the inheritance of single-unit haplotype blocks across generations. Haplotype blocks, therefore, reflect population history and geographical subdivisions, with block sizes varying across ancestral populations⁹¹.

Genotype imputation: A statistical method that estimates genotypes of individuals at unmeasured variants based on data from a more densely-genotyped population-specific reference panel^{95,97}.

Box 3: Structural properties of the human genome and related population genetic concepts leveraged in genome-wide analyses

Genetic ancestry group: A label to denote groups of individuals who share more similar genetic ancestors. Genetic ancestry is estimated in GWAS participants by comparing genotypes with global reference panels. A continuum of genetic differences across ancestry populations arises as those with more common ancestors have more similar genomic structural patterns due to a number of influences. This includes ancestors mating within nearby geographic locations and the sharing of more recombination events than those of more distant ancestors. These differences confer variable frequencies of alleles and patterning of linkage disequilibrium across ancestry populations. It is, therefore, common practice to limit GWAS to more homogenous genetic ancestry groups. This reduces false-positive associations arising from allele-frequency differences across populations instead of reflecting phenotypic variation. It should be noted that ancestry groups are an oversimplification of the complex dimension of human genetic variation and demography^{95,98}.

Population structure and stratification: Population structure is the genetic structure (i.e. LD and allele frequencies) of an underlying population due to non-random mating and restricted geographical movement. When a phenotype is correlated with a population structure in a sample, this gives rise to population stratification that can confound genome-wide analyses⁹⁵.

Principal component analysis (PCA): A method used to estimate population structure within a GWAS dataset. PCA reduces the dimensions of large-scale data (i.e. millions of variants measured across thousands of individuals) into components that capture a large proportion of variance, reflecting the underlying structure of a dataset. PCs are used to estimate and assign participants to genetic ancestry groups by comparing the data with a population reference panel of known ancestries. Additionally, including PCs as covariates in genome-wide analyses is a commonly used approach to control for population stratification^{95,99,100}.

Genetic variants tested in GWAS are typically restricted to alleles more common in a given population, with a minor allele frequency of at least 1-5%^{76,90}. This approach has shown that many common variants of small effect size spread across the genome contribute to the genetic liability of complex traits and are, therefore, highly polygenic¹⁰¹. Consequently, large sample sizes are required to detect robust associations of small effect sizes across the millions of tested variants. To account for this multiple testing burden, a Bonferroni correction for one million independent statistical tests (5×10^{-8}) is applied to establish a genome-wide significance threshold and reduce the chance of false positives. An approximately linear relationship exists between the number of genome-wide significant associations and sample size^{73,102}.

Large, well-powered GWAS are pivotal to understanding the aetiology of complex traits. Summary results can be made publicly available and used for various downstream analyses (**Figure 2**), including elucidating the biological underpinnings of complex traits through systems biology approaches. GWAS summary results can be integrated with various functional genomic data types, including features from cells and tissues, such as gene expression data. The small effect of common genetic variant associations can be aggregated

to the level of the gene and then combined into sets and translated to biological functions and pathways enriched with a phenotype¹⁰³.

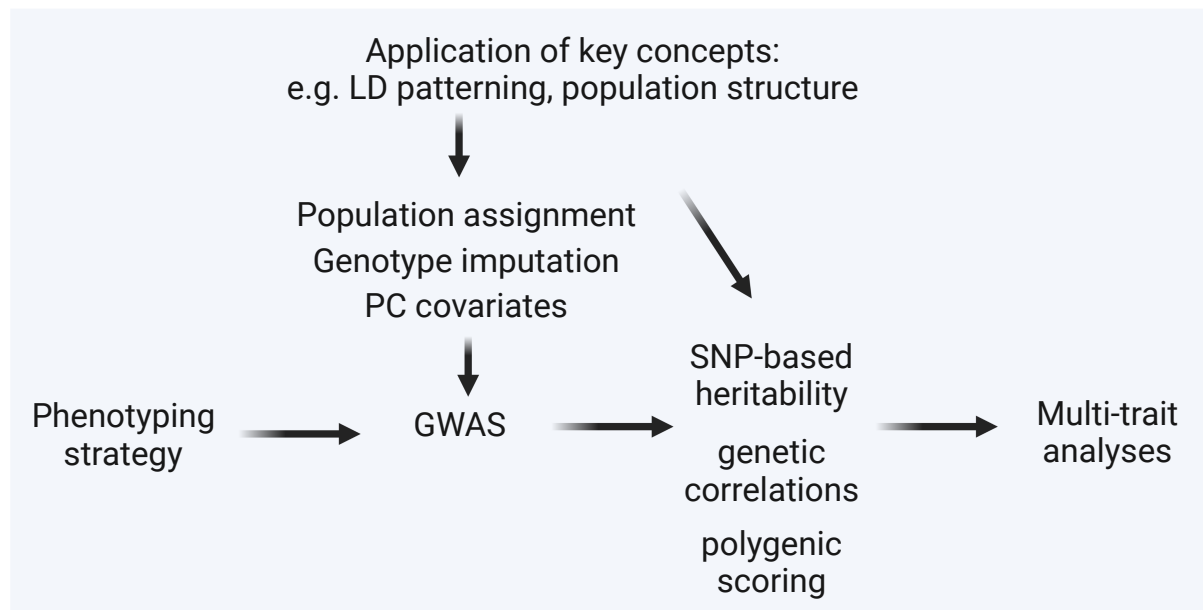


Figure 2: An illustration of the various stages of genome-wide analyses.

Key population genetic concepts are applied to genome-wide association analyses (GWAS), as described in Box 3. Summary results from GWAS can be used for several downstream analyses. GWAS summary statistics are often made publicly available, enabling data on various complex traits to be integrated into multiple-trait (e.g. multivariate or multivariable) analyses. PC = principal component, LD = Linkage disequilibrium.

In addition to uncovering the genomic mechanisms associated with complex traits, GWAS data can be used to estimate narrow-sense heritability captured by common genetic variants, most often termed SNP-based heritability (see **Box 1**)⁶⁷. Estimates of SNP-based heritability can guide researchers on the sample size and power required to detect genome-wide significant loci and the potential for future risk prediction based on GWAS data⁷⁰. More prevalent and highly polygenic psychiatric disorders with a lower SNP-based heritability, such as anxiety and depression, require larger sample sizes than more heritable, rarer disorders, such as schizophrenia, to achieve sufficient statistical power¹⁰⁴. Estimates of SNP-based heritability can be calculated directly from individual-level data used in GWAS, or through summary-level GWAS results. As individual-level-based methods use full SNP data instead of summary scores, they tend to yield higher estimates but are limited in computational speed, and load with large sample sizes^{64,67}. Using GWAS summary results is an alternative, computationally efficient approach to estimating lower bounds of SNP-based heritability. This has the additional benefit of not requiring access to individual-level genetic data for GWAS datasets, often not possible in large-scale, collaborative meta-analyses. A widely used summary-level method is linkage disequilibrium score (LDSC) regression (**Box 1**), which

exploits the fact that some SNPs are more correlated with each other than others (i.e. are in higher LD, as described in **Box 3**)^{67,70,71}.

A genetic correlation between two traits can be estimated using extensions of the same approaches used to estimate SNP-based heritability. To estimate genetic correlations across a wide range of complex traits, the computationally light and flexible approach of using GWAS summary statistics in LDSC regression is common. LDSC regression estimates global (**Box 4**) correlations by estimating genetic covariation across a pair of phenotypes (as captured by common genetic variants) and scaling by genetic variances (i.e. from SNP-based heritability)^{105,106}. A significant global genetic correlation between two phenotypes arises when the direction of the effect of common genetic variants is consistent across the genome¹⁰⁵. When genetic effects in both traits are, on average, in the same direction, this results in a positive correlation of up to 1, or when in the opposite direction, a negative correlation of up to -1¹⁰⁷. Of note, genetic correlations cannot be used to determine causality but can be used to generate novel hypotheses about the relationship between two traits to be then tested in follow-up analyses.

Box 4: Definitions of pleiotropy, bivariate and multi-trait genetic analyses

Pleiotropy: When a genetic variant (or locus, gene) is associated with more than one trait¹⁰⁶.

Horizontal pleiotropy: When a genetic variant influences two traits, either directly or indirectly, through an additional intermediate phenotype. Indicative of shared biological processes between two traits (also termed biological pleiotropy)¹⁰⁶.

Vertical pleiotropy: When a genetic variant influences one trait, which in turn is causally associated with a second trait (i.e. one trait mediates the effect of the genetic variant on another trait). Reverse causation of a causal cascade between two phenotypes is also possible. The presence of vertical pleiotropy can be useful for intervention purposes¹⁰⁶.

Environmentally mediated pleiotropy: A type of vertical pleiotropy when a genetic variant influences a phenotype that shapes the environment, which in turn influences a second phenotype¹⁰⁸.

Spurious pleiotropy: When a genetic variant is falsely associated with two traits. This can arise due to multiple different sources of biases. For example, design artefacts such as phenotype misclassification or ascertainment bias. High LD can also lead to spurious pleiotropy by a variant tagging multiple causal variants located in different genes with different functions, which are not causally related to both phenotypes^{106,109}. Variants in regions of extreme LD (e.g. the major histocompatibility complex) are sometimes removed for genetic correlation analyses¹¹⁰.

Global genetic correlations: A bivariate global genetic correlation between two traits is the average effect of pleiotropy across the whole genome¹⁰⁶.

Local genetic correlations: A bivariate local genetic correlation between two traits analyses region-specific pleiotropy (e.g. at a given locus). Different regions may show genetic correlations in opposite directions, contributing to a non-significant global genetic correlation¹¹¹.

Multi-trait analyses: Pervasive pleiotropy observed across multiple complex traits can be leveraged through multi-trait analyses^{112–114}. For example, multivariate genetic analyses (of more than one outcome phenotype) can be used to extend bivariate genetic correlations into multivariate Genomic Structural Equation Modelling¹¹⁵. Multivariable genetic analyses (of one outcome phenotype and multiple predictors) can be used to model multiple polygenic scores to predict a trait outcome jointly^{113,116}.

A significant global genetic correlation can reflect multiple forms of underlying mechanisms (**Box 4**). They may capture variants with a causal effect on multiple disorders, indicating shared biological pathways^{70,106}. Genetic correlations between psychiatric disorders and health-related lifestyle factors may reflect causal relationships⁷⁰. Genetic correlations observed between personality and psychiatric disorders may be environmentally mediated¹⁰⁸. As such, measuring genetic correlations from GWAS across broad domains of phenotypes is useful for understanding not only shared biological aetiology but also other sources of genetic signals¹⁰⁷.

To further strengthen hypotheses on shared and distinct mechanisms, a genetic overlap across phenotypes can be harnessed for several downstream multi-trait analyses (**Box 4**). Genetic correlations can be used to assess the genetic structure of traits related to psychiatric disorders through multivariate modelling of genetic covariances. One example is the software Genomic Structural Equation Modelling (Genomic SEM), which extends multivariate twin modelling to genomic data using GWAS summary statistics¹¹⁵. This method constructs estimates of SNP-based heritabilities and genetic correlations into matrices, often calculated in LDSC regression. The genetic covariance matrix contains SNP-based heritabilities of each trait and genetic correlations scaled relative to the heritabilities. A second sampling covariance matrix contains the precision of these estimates (i.e. the sampling variance of each estimate in the genetic covariance matrix), and the association between sampling errors to account for GWAS sample overlap. A number of different structural equation models can then be fit to these matrices. For example, this method can be used to evaluate hypotheses on why a group of phenotypes are correlated or the extent to which other traits explain the genetic variance of a trait in a model, or are unique to that phenotype¹¹⁵.

An alternative approach that leverages a genetic overlap between traits to build multi-trait models is through polygenic scoring¹¹³. Polygenic scores are an aggregate of the number of phenotype-associated alleles an individual carries. For example, for common mental disorders, both individuals affected and unaffected by a disorder will carry a range of alleles across the genome associated with an increased risk (coded as a 0, 1 or 2). To calculate a polygenic score, alleles are summed and weighted by the effect size estimate of their association with a given phenotype, derived from GWAS summary statistics in an independent

sample¹¹⁷. A polygenic score for one trait can be used to explain the phenotypic variance of another trait. This is particularly useful for examining the genetic overlap between two traits where one trait does not have a well-powered GWAS available for genetic correlation analysis. Even for polygenic scores derived from well-powered GWAS, the amount of phenotypic variance explained is relatively small^{73,102}. For example, one of the most powerful psychiatric polygenic scores is for schizophrenia, which explains 8.5% of the variance in schizophrenia liability¹¹⁸. As such, polygenic scores are not yet sufficiently powered for prediction at the individual level. Prediction at the group level can be improved by combining multiple well-powered polygenic scores and non-genetic predictors into multivariable models. As GWAS sample sizes grow, the hope is that the power of polygenic scores will improve to the extent that multivariable predictive models can be implemented for clinical application¹¹⁹.

Progress in depression and anxiety disorder genomics

Since the start of the GWAS era approximately 16 years ago, significant progress has been made in psychiatric genomics research¹²⁰. Collaborative efforts and large-scale data collection have enabled sample sizes to reach the level of power needed to start identifying the abundance of common genetic variants associated with psychiatric disorders. This has been facilitated by the Psychiatric Genomics Consortium (PGC), where large-scale meta-analyses of multiple GWAS have transformed psychiatric genomics research. Large national cohorts such as the UK Biobank and the Million's Veteran Programme (MVP) have also been crucial in these developments by contributing data to the largest GWAS of depression and anxiety to date^{45,121,122}.

Considerable progress has been made in depression genomics and led to a broad array of findings on the neurobiology of depression. The largest MDD GWAS identified 178 independent loci in a sample of approximately 370,000 cases¹²³. The translation of depression GWAS results have highlighted the role of several brain regions, nervous system development, synaptic processes, and immune-related function^{123–126}. Such findings have progressed our understanding of the aetiology of depression by corroborating previous hypotheses and generating new ones.

Anxiety disorder genomics, and thus an understanding of genomic-associated systems biology, is somewhat behind compared to depression. A lack of datasets with measures on the multiple types of anxiety disorders has hindered progress. This may partly be due to the complexity of phenotyping anxiety, a highly comorbid trait with somewhat vague boundaries between 'normal' to pathological states^{127,128}. The largest GWAS of anxiety identified five genome-wide significant loci associated with GAD symptoms in 175,000 individuals¹²⁹. Two

additional loci were associated with a case-control phenotype of anxiety and panic disorder in 34,000 cases¹²⁹. A follow-up study applied a novel phenotyping approach to increase statistical power by imputing GAD symptoms, identifying a further seven independent loci. Integrating these GWAS results with functional genomic data found associations with the dorsolateral prefrontal cortex and GABAergic neurons¹³⁰. Although these studies represent significant progress, larger datasets are still needed to further our understanding of the genomic mechanisms underpinning anxiety disorders. Efforts are underway to publish the first large-scale PGC anxiety disorder GWAS, representing a milestone in anxiety disorder genomics¹²⁷.

The SNP-based heritability estimates for anxiety and depression are broadly similar but are also notably smaller than twin study estimates. Individual-level methods report ~26% SNP-based heritability for any anxiety disorder¹³¹, and ~20% for MDD^{132,133}. The ‘missing’ heritability between SNP-based heritability and twin heritability estimates of complex traits are thought to be explained by the role of genetic factors contributing to heritability not measured in GWAS data, such as rare variants, non-additive genetics or gene-environment interactions⁷³. As noted earlier, summary-level SNP-based heritability estimates are even smaller than those calculated from individual-level SNP data and twin data. The LDSC SNP-based heritability estimate for MDD is 10.5%^{70,124} and 9% for broad depression¹²⁵. Similarly, the largest GWAS of anxiety reported an LDSC SNP-based heritability of 8.8%¹³⁴, which is in line with other anxiety disorder GWAS¹³⁵ and neuroticism GWAS¹³⁶. However, meta-analysis SNP-based heritability estimates of psychiatric disorders may also be biased downwards due to incorrect conversions to the liability scale, not taking into account cohort-specific ascertainment (e.g. an increase from 10.5% to 11.5% was observed for MDD)¹³⁷. Furthermore, smaller, single-study GWAS of anxiety disorders reported LDSC SNP-based heritability estimates as high as 28%¹³⁸. This highlights how SNP-based heritability can vary across populations and phenotypic measures⁷³. Such heterogeneity across samples may, in part, limit the ability to detect SNP-based heritability and contribute to missing heritability¹⁰².

As GWAS of anxiety and depression increase, a growing issue in the field is the balance between increasing sample size and consistent and *detailed* phenotyping^{121,139,140}. Many large-scale genetic datasets use *brief* phenotypic measures, often limited to a few self-report items to assess a broad psychiatric phenotype, as this approach is time and cost-effective. Brief phenotyping has been key for making progress in loci discovery for both anxiety and depression and will likely be vital for future risk prediction using GWAS data. The predictive power of depression polygenic scores derived from GWAS is more dependent on increased sample size than the detail of phenotyping. This indicates that sample size should be

prioritised over phenotyping depth for risk estimation¹⁴¹. However, when GWAS sample sizes are equal, polygenic scores for more detailed measures have captured more specific genetic influences on MDD^{141,142}. Loci identified by GWAS of briefly phenotyped broad depression are also less specific to MDD^{140,142}. A tradeoff thus exists between maximising sample size and the depth of phenotyping.

While research on the depth of GWAS phenotyping for anxiety disorders is lacking, lessons learnt from phenotyping strategies in depression genomics may apply to anxiety genomics. Larger samples ascertained from brief phenotyping may increase loci discovery, but some associations may not be disorder-specific. An improvement in power may lead to disorder-specific biology or non-specific treatment targets of broader phenotypes such as neuroticism. Genetic risk prediction may improve, but identify a broader group of individuals than the specific disorder under analysis¹⁴⁰. Therefore, both brief and detailed phenotyping are of research value. Genetic datasets with both brief and detailed phenotyping are needed to make significant progress in anxiety disorder genomics and ultimately translate GWAS findings for clinical utility¹³⁹.

Leveraging genomic data to understand pervasive comorbidity and clinical heterogeneity

As discussed earlier, with evidence from twin studies, the pervasive comorbidity across anxiety disorders and depression is likely partly driven by genetic overlap. Findings from GWAS data further support this notion. GWAS consistently report substantial polygenic overlap between anxiety and depression and other psychopathologies^{123,124,131,134}. The largest GWASs of anxiety and MDD show a genetic correlation of 0.72^{123,134}. Estimates from other GWAS are similar, with 0.78 reported between a clinically ascertained MDD phenotype and lifetime anxiety disorder and up to 0.88 with GAD symptoms^{124,131}. Neuroticism also shows high genetic correlations with anxiety and depression (0.69-0.73)^{123,125,131,134}, providing further support for genetic influences on shared negative affectivity driving comorbidity. Significant global genetic correlations have also been reported between anxiety and depression and a range of other psychiatric disorders, health, and behavioural traits^{123,125,131,134}. This broad array of significant genetic correlations calculated from GWAS data can be leveraged for various analyses to better understand shared and distinct aetiology¹⁰⁶.

Unlike twin studies, molecular genetic assessment of disorder subfactors, such as the distress-fear subdomain, is still in its infancy. A major limiting factor is the availability of detailed phenotyping to measure subtypes for GWAS sufficiently. Anxiety and depression

disorder subtype GWASs are needed to discern subtype-specific from transdiagnostic common genetic influences. Failure to consider clinical heterogeneity may result in overlooking specific biological pathways that can provide important insights into the genomic influences on psychopathology. Furthermore, variable clinical heterogeneity across studies, for example, inconsistent groupings of subtypes, may lead to inconsistent findings¹⁴⁰.

More progress has been made in examining the genomic differences between the *subtypes* of depression than for anxiety disorders. Subtyping of depression has found some evidence that genetic heterogeneity reflects clinical heterogeneity^{143–145}, although to a lesser extent in some studies^{133,146,147}. Differences in SNP-based heritability estimates have been reported. More severe subtypes, including those treated with ECT, atypical, recurrent, and depression with anxiety comorbidity, have shown higher heritability estimates than less clinically severe forms of depression^{143,145}. Genetic correlations across a range of subtypes from one study suggested that 30-70% of genetic influences are shared¹⁴³. Testing genetic differences between subtypes of differing severity with a range of other complex traits has revealed significant differences in genetic overlap with other psychiatric disorders, BMI, personality and cognitive-related traits^{143,145,148}. The more severe subtypes of depression are genetically less similar to neuroticism¹⁴⁵ and more similar to bipolar disorder and schizophrenia¹³². Such analyses shed light on the genetic spectrum from manic to distress clusters (**Figure 3**), showing both shared and unique genetic influences underlying depression and bipolar disorder subtypes¹⁴⁸. GWAS of depression subtypes have been largely underpowered for loci-discovery and thus comparison¹⁴⁹, but recent GWAS reveal evidence for some subtype-specific loci¹⁴³. Inconsistencies in subtyping have slowed progress in depression-subtype genomics¹⁴⁰.

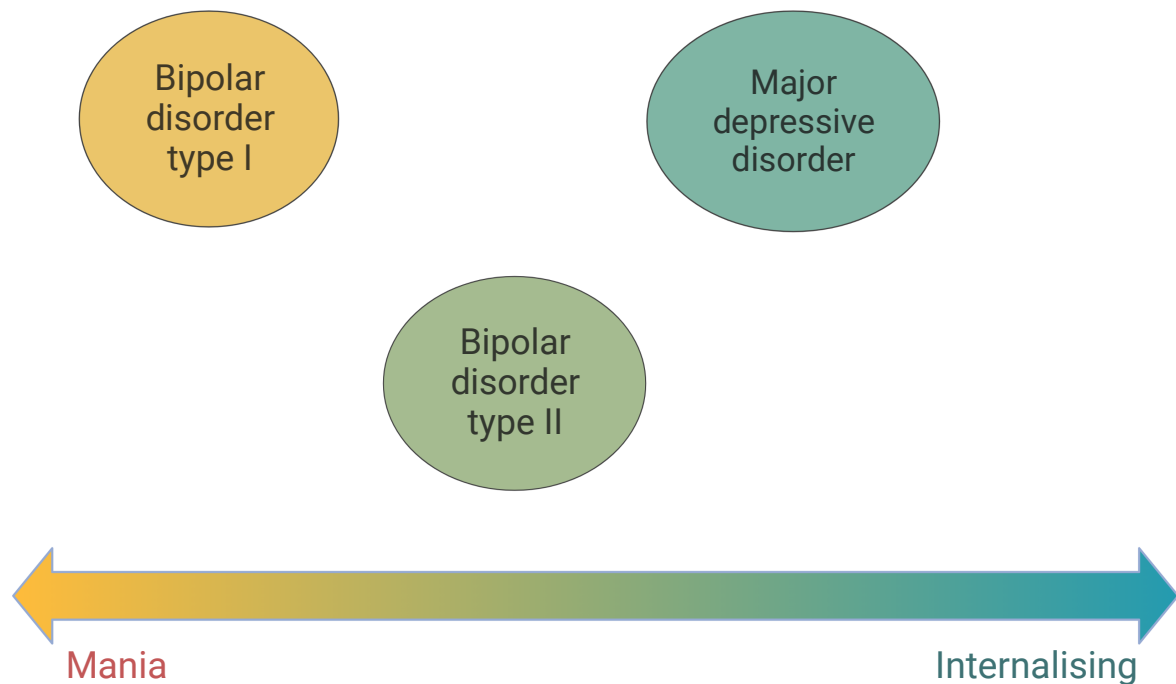


Figure 3: An illustration of the genomic continuum of the mood disorder spectrum from mania genomics to internalising genomics.

Adapted from Coleman et al.¹⁴⁸.

Examining anxiety disorder subtypes at the genome-wide level has been especially limited in psychiatric genomics. This is even though diagnostic subtypes of anxiety disorders are more well-established than depression subtypes⁵. Anxiety genomics has lagged behind due to the lack of genetic datasets with detailed phenotyping on anxiety. Datasets with detailed measures on all five anxiety disorders are even more scarce. Thus, a more prominent focus has been placed on ‘any’ anxiety disorder groupings to maximise power^{122,131,134,135}. It remains unclear what genomic influences are shared or specific to individual anxiety disorders without assessing comorbid conditions.

GWAS of individual anxiety disorders has been largely underpowered for loci-discovery and genetic correlation analyses to assess disorder-specific and transdiagnostic genetic liability. Recent progress has been made on GWAS of current GAD symptoms^{130,134}, but well-powered GWAS of lifetime GAD has been limited¹⁵⁰. GWAS of individual fear-based disorders is further lacking¹⁵¹. One genome-wide significant locus has been reported in a GWAS of panic disorder¹⁵², and one in a GWAS of agoraphobia symptoms¹⁵³. Therefore, a comprehensive assessment of the distress-fear model at the molecular genetic level is lacking. Preliminary GWAS findings in the UK Biobank indicate some genetic specificity among the fear-based disorders that is not shared with distress disorders¹⁵⁰. **Chapter 2** expands upon this study by conducting GWAS meta-analyses of fear-based disorders and GAD with newly established

datasets that have detailed phenotyping on all five anxiety disorders. An improvement in GWAS power enables a more comprehensive assessment of the common genetic variant architecture of GAD and fear-based disorders and their genetic overlap with depression and broad domains of other complex traits.

Social environmental influences on anxiety and depression

Research on the genetic basis of anxiety and depression is important for understanding disorder aetiology, particularly for advancing our understanding of biology. However, incorporating environmental information into genetic analyses is necessary to understand the nature of genetic signals and the interplay between biological and social factors¹⁴⁰. The following sections of this chapter will first discuss how *exposure* to social environmental factors is not independent of genetic influences. Psychological factors are discussed second, which are the lens through which we *experience* environmental exposures and act as a bridge between genetic and environmental influences on psychopathology. **Figure 4** illustrates the mechanisms involved in the interplay between genomics and different types of experiences in the risk and treatment of common mental disorders.

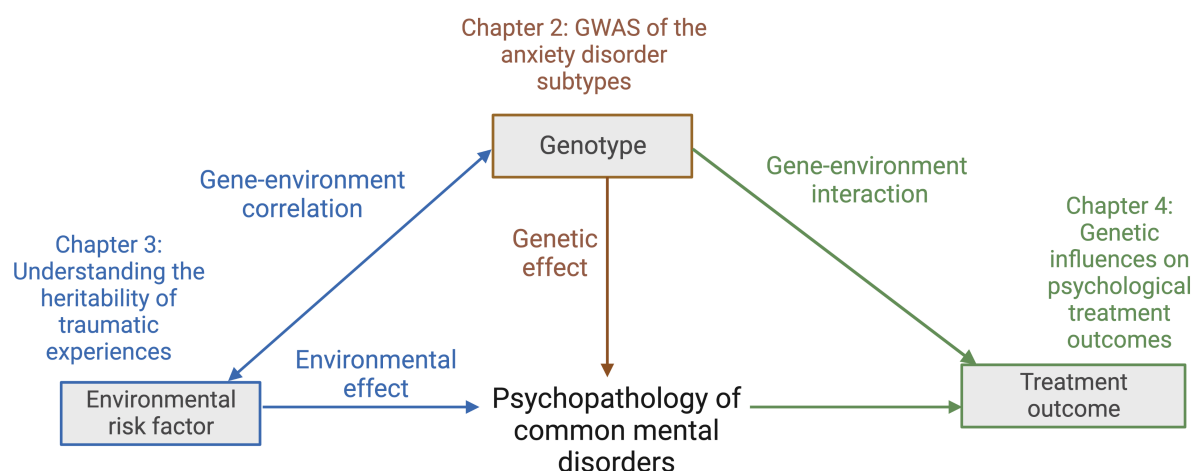


Figure 4: An illustration of the potential mechanisms underpinning the course and resolution of common mental disorders, which are influenced by the interplay between genetic, socio-environmental and psychological factors.

Brown arrows represent genetic influences on biological pathways associated with the risk of developing subtypes of common mental disorders, measured through genome-wide association studies (GWAS). Experiences of negative and positive exposures and their influence on the risk and treatment of psychopathology do not act independently of genetics. Genomics can therefore be leveraged to understand these experiences. Blue arrows represent the interplay between heritable characteristics associated with shaping adverse environmental exposures and the risk of psychopathology, known as gene-environment correlation. Green arrows represent the interaction between genetic liability to a disorder and a psychological intervention (depicted here as a positive environmental exposure), leading to variation in experiences and response, such as with psychological treatment.

Approximately half of the phenotypic variability of anxiety and depressive disorders is not explained by genetics. Twin studies (see **Box 1**) show that the remaining variance is primarily attributed to the individual environment (sometimes called “non-shared”). The common environment explains some variance during childhood but declines with age and is no longer present during adulthood^{47,154}. Individual-specific environmental liability is partially shared across anxiety disorders and depression, though some disorder-specific influences are also observed⁵¹. As such, unique individual-specific environmental factors are thought to explain why anxiety disorders and depression present as distinct conditions¹⁵⁵. Anxiety and depression are associated with both positive and negative social environmental contexts. Positive environments can mitigate the effects of adverse environmental risk factors. For example, social support is a particularly important transdiagnostic buffer for the effects of severe adverse environments, notably traumatic experiences¹⁵⁶.

Psychologically traumatic experiences, which can be defined as those perceived by the individual as devastating and overwhelming, are major risk factors for internalising disorders and psychopathology more broadly⁵. Traumas experienced during childhood are the most robust transdiagnostic environmental risk factors^{31,157}. Childhood trauma is reported to increase psychopathology risk by approximately two-fold^{158,159}. The mechanisms underlying the causal effects of trauma on psychopathology are an ongoing area of research. A recent meta-analysis found that after controlling for pre-existing risk factors such as environmental adversities and genetics, an attenuated but significant causal effect of childhood trauma on broad psychopathology remained¹⁶⁰. Transdiagnostic theories on the long-term biological effects of early-life trauma include altered brain structure and function, such as disruption to the development of the hypothalamic-pituitary-adrenal axis. This may lead to elevated glucocorticoid signalling in response to stress and abnormal brain development with long-lasting effects on emotion and behaviour^{161,162}. Difficulties with emotions include dysregulation and increased emotional reactivity, which, together with elevated informational processing of social cues as threatening, can influence maladaptive behavioural responses¹⁵⁶. Given its transdiagnostic effects, understanding the risk and consequences of trauma is an important area of research. Furthermore, targeting trauma exposure in prevention and intervention strategies may be more effective in reducing psychopathology risk than disorder-specific environmental risk factors¹⁵⁷.

The heritability of environmental measures

A complexity of disentangling the causal effects of environmental exposures on psychopathology is that the environment does not operate independently of genetics. Twin studies show that almost all measures of the environment are themselves genetically

influenced¹⁶³. This includes the reporting of positive, negative, and traumatic social exposures, which have a heritability ranging from 24-62%^{163–167}. Significant SNP-based heritability estimates have also been found for reported traumatic and stressful exposures (6-30%)^{132,168,169}. Twin studies find that life events dependent on one's behaviour are more heritable than independent events^{163,170,171}, likely due to heritable behavioural characteristics influencing the environments we are exposed to. These environments can be shaped passively by our parents during childhood, or we actively select them in later life and by how others engage with us. Correlations between genetic factors and environmental exposures can arise through these mechanisms, termed gene-environment correlation^{172,173}. As such, the association between environmental measures and psychopathology may partly be explained by shared genetic influences on both the likelihood of environmental exposure and the development of psychopathology.

Identifying the genetically influenced traits that contribute to the heritability of environmental measures would elucidate key traits involved in gene-environment correlations. Such findings help narrow down characteristics to target in follow-up analyses assessing causal pathways and ultimately inform intervention or prevention strategies. Multivariate twin modelling has been used to deconstruct the heritability of an environmental measure by jointly modelling genetically correlated personality and behavioural traits^{174,175}. For example, one study found that genetic propensity for non-cognitive traits is as important as genetic influences on cognitive ability in explaining the heritability of educational attainment. This study highlighted plausible alternative candidates for non-cognitive interventions to improve engagement with the schooling environment¹⁷⁴. As noted earlier, recent advancements in genomics research enable the systematic assessment and multi-trait modelling of genetic correlations between a broad range of traits, including environmental exposures and psychiatric and behavioural traits¹¹⁵. For example, genomic multiple regression models in Genomic SEM have been used to estimate genetic correlations between psychiatric disorders and smoking behaviours, independent of other heritable behavioural traits¹⁷⁶. This method could also be used to deconstruct the SNP-based heritability of adverse environments and identify uniquely genetically associated psychiatric and behavioural traits. Thus, such approaches provide the opportunity to better understand key environmental risk factors for psychopathology at the genomic level. **Chapter 3** explores the genomic influences of multiple traits on vulnerability to childhood trauma and aims to identify the traits that may be involved in gene-environment correlation mechanisms through genomic multiple regression modelling.

Heritable psychological influences on anxiety and depression: from risk to treatment

As outlined earlier, twin studies show that the individual environment explains more phenotypic variation in anxiety and depression than the common environment. This is partly due to the important role of individual differences in psychological influences that shape how people experience their environment, which is unique to each person^{177,178}. Genetics also plays a role in how sensitive we are and the context in which we perceive exposures, which in turn influences our response¹⁷⁹. Genetics can therefore be leveraged to understand characteristics associated with vulnerability to experiencing adverse exposures as traumatic and a greater risk of subsequent psychopathology¹⁸⁰.

Heritable influences on subjective self-reports of adverse experiences

Various genetically influenced psychological and cognitive factors impact sensitivity to adverse environments. The stress-diathesis model describes sensitivity to negative (i.e. stressful) environmental experiences and explains why not all those exposed to adverse environments develop a common mental health problem. Genetic pathogenic effects depend on environmental exposure and vice versa^{13,179,181,182}. It is plausible that psychological factors that influence *sensitivity* to adverse environments also influence *reporting* of events as stressful or traumatic. As such, using self-reports to capture environmental exposures is often conceptualised as a limitation to studies due to potential biases in reporting¹⁸³. Meta-analyses show that subjective and objective measures of adversity identify largely different groups of individuals¹⁸⁴. Thus, if research aims only to understand the influences on objective exposure, then subjective reporting may capture influences on reporting biases as opposed to mechanisms associated with exposure. Such research is important for prevention strategies. However, understanding the genetic components contributing to subjective reporting is also valuable. The subjective experience of trauma, as captured by retrospective self-reports, is key for developing psychopathology, more so than trauma exposure. Some individuals identified as exposed through objective measures, such as court documentation of childhood maltreatment, do not subjectively report trauma in later life. These individuals are less likely to develop post-traumatic psychopathology than those who retrospectively self-report trauma¹⁸⁴. This may reflect a group of individuals with low sensitivity to adverse environments, which acts as a protective factor and resilience to adversity. Thus, understanding the psychological and cognitive factors associated with retrospective and subjective reporting may be informative for psychopathology intervention.

Psychological factors, including personality traits and cognitive biases, may influence how individuals experience and engage with adverse environments and later risk of psychopathology. Neuroticism is associated with cognitive biases towards negative stimuli, elevated stress levels, and environmental sensitivity^{185,186}. Twin studies suggest that genetics primarily explain the correlation between neuroticism and adverse environmental sensitivity¹⁸⁶, highlighting the role of genetics in the subjective appraisal of adversity. Individuals with higher levels of neuroticism tend to self-report more adversity in the absence of prospective records. This is in contrast to agreeableness, which is associated with lower rates of retrospective reports than indicated by the level of prospective accounts¹⁸⁷. Higher neuroticism is also linked with reporting traumatic and stressful events as more central to one's identity, a risk factor for later psychopathology^{188,189}. A potential mechanism explaining the link between neuroticism and posttraumatic psychopathology is the increased emotional availability of trauma memories through more frequent retrieval and rehearsal and, thus, maintenance of the memory¹⁸⁹. A genetic overlap has been observed between neuroticism, depressive symptoms, and perceived stress¹⁹⁰. Thus, genetic influences on neuroticism may influence the reporting of events as traumatic¹⁶⁶. Indeed, genetic correlations between GWASs of neuroticism and retrospective self-reports of trauma have been observed¹³². **Chapter 3** explores the extent to which genomic influences on psychological traits such as neuroticism can explain the heritability of retrospectively self-reporting adverse environments, aiming to identify key traits involved in the *experience* of adverse environments.

Heritable influences on positive experiences and application to psychological treatment

Genetic influences on psychological factors impact not only vulnerability to respond negatively following trauma (as in the diathesis-stress model) but also sensitivity to positive environments. The differential susceptibility model extends environmental sensitivity to genetic influences on psychological factors affecting the experience and response to both negative and positive environments. Under this framework, we can leverage genetic information to understand who is more likely to benefit from support environments, providing a more positive outlook on intervention research of psychopathology. Indeed, a core focus on anxiety and depression genomics is understanding risk factors. However, genomics can also be harnessed to improve treatment selection. For example, individuals with high environmental sensitivity may translate to a better response to protective and supportive exposures, such as psychological intervention^{179,191,192}.

There is substantial individual variability observed in treatment response for anxiety and depression. Given the overlap in risk factors, it is not surprising that treatments for anxiety and

depression are highly similar, with comparable treatment response rates across disorders³¹. Medication and psychological therapy are both first-line treatments for anxiety and depression. However, not all individuals respond to them equally, with approximately only 50% responding to current treatments^{193,194}. Predicting who will respond best to which treatment is poor and is therefore often prescribed in a trial and error fashion^{13,194}. Research understanding the factors that influence treatment response is needed to aid in identifying suitable predictors to guide treatment selection.

Genetic influences on antidepressant response show the potential for using genetic information in guiding treatment. Psychiatric pharmacogenomics, which aims to harness genetic variants associated with variability in response to medication to guide treatment selection, has grown over recent years with some progress in antidepressant response research. Although findings have been mixed, a meta-analysis of randomised controlled trials of pharmacogenomic testing for antidepressant response showed a 71% increased likelihood of depressive symptom remission in individuals receiving pharmaco-guided antidepressant prescribing compared to standard practice prescribing¹⁹⁵. Based on individual-level estimates, the SNP-based heritability of antidepressant response is approximately 42%¹⁹⁶ and 13% for depressive symptom remission following antidepressant treatment¹⁹⁷. These findings represent considerable progress in the field of antidepressant response genetics. However, antidepressants are not always beneficial to patients, and an integrated approach with psychological therapy and medication outperforms pharmacotherapy alone. Thus, a combined understanding of factors associated with response to both forms of treatment would be of even greater clinical utility in guiding treatment selection.

Identifying predictors of improvement in anxiety and depressive symptoms following psychological therapy is an ongoing area of research. Cognitive behavioural therapy (CBT) is the gold standard psychological treatment for anxiety and depression¹⁹⁸, including disorder-specific and transdiagnostic approaches¹⁹⁹, which are equally as effective^{200,201}. Grouping internalising disorders may prove useful in improving the prediction of psychological treatment outcomes³¹. The ultimate goal is to incorporate a range of variables into predictive models to aid in treatment decision-making²⁰². Adding genetic factors to these models requires advancements in examining the genetics of psychological treatment outcomes.

Considering there are genetic influences on a range of behavioural, psychological and cognitive traits, which in turn influence response to positive environments¹⁶³, there is likely a genetic component to psychological treatment. No twin studies have found a significant heritability of outcomes following psychological therapy, also known as ‘therapygenetics’. However, evidence from examining related phenotypes suggests that a heritable component

exists. Twin studies report a significant heritability of fear extinction of 15-36%, a key mechanism involved in exposure-based CBT for anxiety disorders^{203,204}. Half of the genetic influences on fear acquisition also overlap with fear extinction. This suggests that some genetic influences on anxiety-related traits could be used to understand responses to psychological therapy²⁰³. For example, polygenic scores from existing well-powered GWAS of anxiety or depression may be useful in predicting psychological therapy outcomes alongside other predictors. As such, incorporating pre-existing polygenic scores from well-powered GWAS of related complex traits may be a promising avenue for therapygenetics research²⁰⁵.

Therapygenetics can be understood through the framework of gene-environment interaction, in which psychological intervention is the environmental exposure and genetics influence sensitivity to the exposure²⁰⁶. Preliminary findings using polygenic scores derived from a small GWAS of environmental sensitivity in twins indicates that a higher genetic propensity for environmental sensitivity is associated with a better response to more intensive CBT than less intensive²⁰⁷. As was the case for many polygenic traits in candidate gene studies, early attempts to identify specific genetic variants with robust gene-environment interaction effects on psychological treatment outcomes were unsuccessful^{206,208}. Response to psychological therapy is unlikely to be influenced by a few genes, but rather highly polygenic, as is the case for all heritable psychological traits.

A challenge for genome-wide therapygenetics is phenotyping at the scale required for GWAS. GWAS of prognostic outcomes following CBT for anxiety and depression have been limited in power due to the difficulty in ascertaining such phenotypes. The largest published GWAS meta-analysis thus far was small, with 2,724 participants, and did not detect a significant SNP-based heritability or genome-wide significant loci²⁰⁹. Alternative phenotyping strategies are required to achieve sufficient power to detect the thousands of loci underpinning complex traits. As with anxiety and depressive disorder genomics research, an improvement in power has largely been achieved through brief phenotyping. The utility of brief phenotyping in GWAS of psychological therapy outcomes has yet to be explored. Furthermore, subjective self-reports of benefitting from psychological therapy may be useful in capturing genetic influences that impact the perception of experiences, as with self-reports of environmental adversity. **Chapter 4** uses brief self-reports of psychological treatment outcomes to increase the sample size for GWAS and prediction models that incorporate polygenic scores of related traits.

Summary and aims

The overarching aim of this thesis was to further the understanding of the risk and treatment of common mental disorders by integrating bio-psycho-social components and leveraging large genomic datasets (**Figure 1**). **Chapters 2** and **3** focus on understanding genetic influences on risk factors, while **Chapter 4** explores genetic influences on treatment outcomes (see **Figure 4**). **Chapter 2** focuses more on the biological and genomic mechanisms associated with clinical heterogeneity, whilst **Chapters 3** and **4** harness genomic data to understand psychological influences on environmental experiences. To improve statistical power for genomic analyses of phenotypes that currently lack GWAS with sufficient sample sizes, **Chapters 2** and **4** leverage novel genotyped datasets with detailed measures of anxiety and depression outcomes. All three empirical chapters harness pre-existing, well-powered GWAS summary statistics across bio-psycho-social domains of complex traits related to common mental disorders. This includes psychopathologies and physical health problems commonly co-occurring with anxiety and depression, reported socio-environmental experiences, and cognitive-related and psychological traits. Examining a broad range of complex traits enables the assessment of shared and distinct genetic influences on common mental disorders and the interplay with psychosocial influences at a large scale.

References

1. Organization, W. H. & Others. *Depression and other common mental disorders: global health estimates*.
<https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf> (2017).
2. Kessler, R. C. *et al.* The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. *Epidemiol. Psychiatr. Soc.* **18**, 23–33 (2009).
3. Lim, G. Y. *et al.* Prevalence of Depression in the Community from 30 Countries between 1994 and 2014. *Sci. Rep.* **8**, 2861 (2018).
4. Somers, J. M., Goldner, E. M., Waraich, P. & Hsu, L. Prevalence and incidence studies of anxiety disorders: a systematic review of the literature. *Can. J. Psychiatry* **51**, 100–113 (2006).
5. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
6. Aquino, J. P. D., De Aquino, J. P., Londono, A. & Carvalho, A. F. An Update on the Epidemiology of Major Depressive Disorder Across Cultures. *Understanding Depression* 309–315 Preprint at https://doi.org/10.1007/978-981-10-6580-4_25 (2018).
7. Breslau, J., Lane, M., Sampson, N. & Kessler, R. C. Mental disorders and subsequent educational attainment in a US national sample. *J. Psychiatr. Res.* **42**, 708–716 (2008).
8. Fergusson, D. M. & Woodward, L. J. Mental health, educational, and social role outcomes of adolescents with depression. *Arch. Gen. Psychiatry* **59**, 225–231 (2002).
9. McDaid, D., Park, A. L., Davidson, G., John, A. & Knifton, L. The economic case for investing in the prevention of mental health conditions in the UK. (2022).
10. Craske, M. G. & Stein, M. B. Anxiety. *Focus* **15**, 199–210 (2017).
11. Malhi, G. S. & John Mann, J. Depression. *The Lancet* vol. 392 2299–2312 Preprint at [https://doi.org/10.1016/s0140-6736\(18\)31948-2](https://doi.org/10.1016/s0140-6736(18)31948-2) (2018).
12. Too, L. S. *et al.* The association between mental disorders and suicide: A systematic review and meta-analysis of record linkage studies. *J. Affect. Disord.* **259**, 302–313 (2019).
13. Craske, M. G. *et al.* Anxiety disorders. *Nat Rev Dis Primers* **3**, 17024 (2017).
14. Harald, B. & Gordon, P. Meta-review of depressive subtyping models. *J. Affect. Disord.* **139**, 126–140 (2012).
15. Toft, T. *et al.* Mental disorders in primary care: prevalence and co-morbidity among disorders. results from the functional illness in primary care (FIP) study. *Psychol. Med.* **35**, 1175–1184 (2005).
16. Kessler, R. C., Merikangas, K. R. & Wang, P. S. Prevalence, comorbidity, and service utilization for mood disorders in the United States at the beginning of the twenty-first century. *Annu. Rev. Clin. Psychol.* **3**, 137–158 (2007).
17. Kessler, R. C. *et al.* Comorbidity of DSM–III–R Major Depressive Disorder in the General Population: Results from the US National Comorbidity Survey. *Br. J. Psychiatry* **168**, 17–30 (1996).

18. Kessler, R. C. *et al.* The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* **289**, 3095–3105 (2003).
19. Kessler, R. C. *et al.* Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593–602 (2005).
20. Hirschfeld, R. M. A. The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care. *Prim. Care Companion J. Clin. Psychiatry* **3**, 244–254 (2001).
21. Lamers, F. *et al.* Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J. Clin. Psychiatry* **72**, 341–348 (2011).
22. Steffen, A., Nübel, J., Jacobi, F., Bätzing, J. & Holstiege, J. Mental and somatic comorbidity of depression: a comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data. *BMC Psychiatry* **20**, 142 (2020).
23. Penninx, B. W., Pine, D. S., Holmes, E. A. & Reif, A. Anxiety disorders. *Lancet* **397**, 914–927 (2021).
24. Prince, M. *et al.* No health without mental health. *Lancet* **370**, 859–877 (2007).
25. Kraus, C., Kadriu, B., Lanzenberger, R., Zarate, C. A., Jr & Kasper, S. Prognosis and Improved Outcomes in Major Depression: A Review. *Focus* **18**, 220–235 (2020).
26. Noyes, R., Jr. Comorbidity in generalized anxiety disorder. *Psychiatr. Clin. North Am.* **24**, 41–55 (2001).
27. Kalin, N. H. The Critical Relationship Between Anxiety and Depression. *Am. J. Psychiatry* **177**, 365–367 (2020).
28. Clark, L. A. & Watson, D. Distress and fear disorders: an alternative empirically based taxonomy of the ‘mood’ and ‘anxiety’ disorders. *Br. J. Psychiatry* **189**, 481–483 (2006).
29. Eysenck, M. W. & Fajkowska, M. Anxiety and depression: toward overlapping and distinctive features. *Cogn. Emot.* **32**, 1391–1400 (2018).
30. Crocq, M.-A. The history of generalized anxiety disorder as a diagnostic category. *Dialogues Clin. Neurosci.* **19**, 107–116 (2017).
31. Watson, D. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry* **21**, 26–54 (2022).
32. Conway, C. C., Forbes, M. K., South, S. C. & HiTOP Consortium. A Hierarchical Taxonomy of Psychopathology (HiTOP) Primer for Mental Health Researchers. *Clin. Psychol. Sci.* **10**, 236–258 (2022).
33. Krueger, R. F. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry* vol. 20 171–193 Preprint at <https://doi.org/10.1002/wps.20844> (2021).
34. Kotov, R. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. *World Psychiatry* **19**, 151–172 (2020).
35. Stanton, K. *et al.* An Initial Investigation of the Joint Classification of Hypomania- and Neurodevelopmental Disorder-Relevant Dimensions Within the Hierarchical Taxonomy of Psychopathology. *Assessment* **30**, 414–432 (2023).
36. Stanton, K., DeLucia, E. A., Brown, M. F. D. & McDonnell, C. G. Advancing understanding of the classification of broad autism phenotype and attention-

- deficit/hyperactivity disorder symptom dimensions within the Hierarchical Taxonomy of Psychopathology. *Personal. Ment. Health* **15**, 113–123 (2021).
37. Caspi, A. *et al.* The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? *Clin. Psychol. Sci.* **2**, 119–137 (2014).
 38. Watson, D. Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *J. Abnorm. Psychol.* **114**, 522–536 (2005).
 39. Kendler, K. S. The structure of psychiatric science. *Am. J. Psychiatry* **171**, 931–938 (2014).
 40. Hettema, J. M., Neale, M. C. & Kendler, K. S. A review and meta-analysis of the genetic epidemiology of anxiety disorders. *Am. J. Psychiatry* **158**, 1568–1578 (2001).
 41. Sullivan, P. F., Neale, M. C. & Kendler, K. S. Genetic epidemiology of major depression: review and meta-analysis. *Am. J. Psychiatry* **157**, 1552–1562 (2000).
 42. Smoller, J. W. The Genetics of Stress-Related Disorders: PTSD, Depression, and Anxiety Disorders. *Neuropsychopharmacology* **41**, 297–319 (2016).
 43. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
 44. Craddock, N., O'Donovan, M. C. & Owen, M. J. The genetics of schizophrenia and bipolar disorder: dissecting psychosis. *J. Med. Genet.* **42**, 193–204 (2005).
 45. Kendall, K. M. *et al.* The genetic basis of major depression. *Psychol. Med.* **51**, 2217–2230 (2021).
 46. Rice, F. Genetics of childhood and adolescent depression: insights into etiological heterogeneity and challenges for future genomic research. *Genome Med.* **2**, 68 (2010).
 47. Nivard, M. G. *et al.* Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. *Psychol. Med.* **45**, 1039–1049 (2015).
 48. Hannigan, L. J., Walaker, N., Waszczuk, M. A., McAdams, T. A. & Eley, T. C. Aetiological influences on stability and change in emotional and behavioural problems across development: a systematic review. *Psychopathol Rev* **4**, 52–108 (2017).
 49. Waszczuk, M. A., Zavos, H. M. S., Gregory, A. M. & Eley, T. C. The stability and change of etiological influences on depression, anxiety symptoms and their co-occurrence across adolescence and young adulthood. *Psychol. Med.* **46**, 161–175 (2016).
 50. Cerdá, M., Sagdeo, A., Johnson, J. & Galea, S. Genetic and environmental influences on psychiatric comorbidity: a systematic review. *J. Affect. Disord.* **126**, 14–38 (2010).
 51. Waszczuk, M. A., Zavos, H. M. S., Gregory, A. M. & Eley, T. C. The phenotypic and genetic structure of depression and anxiety disorder symptoms in childhood, adolescence, and young adulthood. *JAMA Psychiatry* **71**, 905–916 (2014).
 52. Middeldorp, C. M., Cath, D. C., Van Dyck, R. & Boomsma, D. I. The co-morbidity of anxiety and depression in the perspective of genetic epidemiology. A review of twin and family studies. *Psychol. Med.* **35**, 611–624 (2005).
 53. Mosing, M. A. *et al.* Genetic and environmental influences on the co-morbidity between depression, panic disorder, agoraphobia, and social phobia: a twin study. *Depress. Anxiety* **26**, 1004–1011 (2009).
 54. Hettema, J. M. *et al.* Genetic and environmental risk structure of internalizing

- psychopathology in youth. *Depress. Anxiety* **37**, 540–548 (2020).
55. Kendler, K. S. *et al.* The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *Am. J. Psychiatry* **168**, 29–39 (2011).
 56. Hettema, J. M., Neale, M. C., Myers, J. M., Prescott, C. A. & Kendler, K. S. A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am. J. Psychiatry* **163**, 857–864 (2006).
 57. Tackett, J. L. *et al.* Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *J. Abnorm. Psychol.* **122**, 1142–1153 (2013).
 58. Pettersson, E., Larsson, H. & Lichtenstein, P. Common psychiatric disorders share the same genetic origin: a multivariate sibling study of the Swedish population. *Mol. Psychiatry* **21**, 717–721 (2016).
 59. Kendler, K. S. & Myers, J. The boundaries of the internalizing and externalizing genetic spectra in men and women. *Psychol. Med.* **44**, 647–655 (2014).
 60. Hettema, J. M., Prescott, C. A., Myers, J. M., Neale, M. C. & Kendler, K. S. The structure of genetic and environmental risk factors for anxiety disorders in men and women. *Arch. Gen. Psychiatry* **62**, 182–189 (2005).
 61. Kendler, K. S., Prescott, C. A., Myers, J. & Neale, M. C. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Arch. Gen. Psychiatry* **60**, 929–937 (2003).
 62. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era—concepts and misconceptions. *Nat. Rev. Genet.* **9**, 255–266 (2008).
 63. Wray, N. & Visscher, P. Estimating trait heritability. *Nature education*.
 64. Barry, C.-J. S. *et al.* How to estimate heritability: a guide for genetic epidemiologists. *Int. J. Epidemiol.* (2022) doi:10.1093/ije/dyac224.
 65. Mayhew, A. J. & Meyre, D. Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. *Curr. Genomics* **18**, 332–340 (2017).
 66. Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* **4**, e1000008 (2008).
 67. Yang, J., Zeng, J., Goddard, M. E., Wray, N. R. & Visscher, P. M. Concepts, estimation and interpretation of SNP-based heritability. *Nat. Genet.* **49**, 1304–1310 (2017).
 68. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
 69. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
 70. Baselmans, B. M. L., Yengo, L., van Rheenen, W. & Wray, N. R. Risk in Relatives, Heritability, SNP-Based Heritability, and Genetic Correlations in Psychiatric Disorders: A Review. *Biol. Psychiatry* **89**, 11–19 (2021).
 71. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* vol. 47 291–295 Preprint at <https://doi.org/10.1038/ng.3211> (2015).
 72. Lee, S. H., Wray, N. R., Goddard, M. E. & Visscher, P. M. Estimating missing heritability for disease from genome-wide association studies. *Am. J. Hum. Genet.* **88**, 294–305 (2011).
 73. Andreassen, O. A., Hindley, G. F. L., Frei, O. & Smeland, O. B. New insights from the last decade of research in psychiatric genetics: discoveries, challenges and

- clinical implications. *World Psychiatry* **22**, 4–24 (2023).
74. Roses, A. D. The genome era begins. *Nat. Genet.* **33**, 217–217 (2003).
 75. Ikegawa, S. A short history of the genome-wide association study: where we were and where we are going. *Genomics Inform.* **10**, 220–225 (2012).
 76. Hofker, M. H., Fu, J. & Wijmenga, C. The genome revolution and its role in understanding complex diseases. *Biochim. Biophys. Acta* **1842**, 1889–1895 (2014).
 77. Michaelson, J. J. Genetic Approaches to Understanding Psychiatric Disease. *Neurotherapeutics* **14**, 564–581 (2017).
 78. Edwards, S. L., Beesley, J., French, J. D. & Dunning, A. M. Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet.* **93**, 779–797 (2013).
 79. Knopik, V. S., Neiderhiser, J. M., DeFries, J. C. & Plomin, R. *Behavioral Genetics*. (Macmillan Learning, 2016).
 80. Visscher, P. M., Yengo, L., Cox, N. J. & Wray, N. R. Discovery and implications of polygenicity of common diseases. *Science* **373**, 1468–1473 (2021).
 81. Karki, R., Pandya, D., Elston, R. C. & Ferlini, C. Defining ‘mutation’ and ‘polymorphism’ in the era of personal genomics. *BMC Med. Genomics* **8**, 37 (2015).
 82. Brookes, A. J. The essence of SNPs. *Gene* **234**, 177–186 (1999).
 83. Cavalli-Sforza, L. L. & Bodmer, W. F. *The Genetics of Human Populations*. (Courier Corporation, 1999).
 84. Arias, T. D., Jorge, L. F. & Barrantes, R. Uses and misuses of definitions of genetic polymorphism. A perspective from population pharmacogenetics. *Br. J. Clin. Pharmacol.* **31**, 117–119 (1991).
 85. Day, I. N. M. dbSNP in the detail and copy number complexities. *Hum. Mutat.* **31**, 2–4 (2010).
 86. Eichler, E. E. Copy number variation and human disease. *Nat Educ.*
 87. 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
 88. Consortium, †the International Hapmap & †The International HapMap Consortium. The International HapMap Project. *Nature* vol. 426 789–796 Preprint at <https://doi.org/10.1038/nature02168> (2003).
 89. Zhao, S., Jiang, L., Yu, H. & Guo, Y. GTQC: Automated Genotyping Array Quality Control and Report. *J Genomics* **10**, 39–44 (2022).
 90. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
 91. Slatkin, M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* **9**, 477–485 (2008).
 92. Schaid, D. J., Chen, W. & Larson, N. B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat. Rev. Genet.* **19**, 491–504 (2018).
 93. Cano-Gamez, E. & Trynka, G. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Frontiers in Genetics* vol. 11 Preprint at <https://doi.org/10.3389/fgene.2020.00424> (2020).
 94. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
 95. Peterson, R. E. *et al.* Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* **179**, 589–603 (2019).

96. Reich, D. E. *et al.* Linkage disequilibrium in the human genome. *Nature* **411**, 199–204 (2001).
97. Uffelmann, E. *et al.* Genome-wide association studies. *Nature Reviews Methods Primers* **1**, 1–21 (2021).
98. National Academies of Sciences, Engineering & Medicine. *Using Population Descriptors in Genetics and Genomics Research: A New Framework for an Evolving Field.* (2023).
99. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
100. Novembre, J. & Stephens, M. Interpreting principal component analyses of spatial population genetic variation. *Nat. Genet.* **40**, 646–649 (2008).
101. Visscher, P. M., Goddard, M. E., Derks, E. M. & Wray, N. R. Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Mol. Psychiatry* **17**, 474–485 (2012).
102. Fabbri, C. Genetics in psychiatry: Methods, clinical applications and future perspectives. *PCN Rep.* **1**, (2022).
103. Uffelmann, E. & Posthuma, D. Emerging Methods and Resources for Biological Interrogation of Neuropsychiatric Polygenic Signal. *Biol. Psychiatry* **89**, 41–53 (2021).
104. Wray, N. R. *et al.* Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol. Psychiatry* **17**, 36–48 (2012).
105. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
106. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* (2019) doi:10.1038/s41576-019-0137-z.
107. Abdellaoui, A. & Verweij, K. J. H. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nat Hum Behav* **5**, 686–694 (2021).
108. Avinun, R. The E Is in the G: Gene-Environment-Trait Correlations and Findings From Genome-Wide Association Studies. *Perspect. Psychol. Sci.* **15**, 81–89 (2020).
109. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495 (2013).
110. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* vol. 33 272–279 Preprint at <https://doi.org/10.1093/bioinformatics/btw613> (2017).
111. Ronald, A. & Pain, O. A Revolution Is Brewing in How We Understand the Shared Genetic Causes of Psychiatric Disorders. *The American journal of psychiatry* vol. 179 791–793 (2022).
112. Visscher, P. M. & Yang, J. A plethora of pleiotropy across complex traits. *Nature genetics* vol. 48 707–708 (2016).
113. Krapohl, E. *et al.* Multi-polygenic score approach to trait prediction. *Mol. Psychiatry* **23**, 1368–1374 (2018).
114. Grotzinger, A. D. Shared genetic architecture across psychiatric disorders. *Psychol. Med.* **51**, 2210–2216 (2021).
115. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513–525 (2019).

116. Allegrini, A. G. *et al.* Multivariable G-E interplay in the prediction of educational achievement. *PLoS Genet.* **16**, e1009153 (2020).
117. Choi, S. W., Mak, T. S.-H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).
118. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
119. Lewis, C. M. & Vassos, E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med.* **12**, 44 (2020).
120. Psychiatric GWAS Consortium Coordinating Committee *et al.* Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am. J. Psychiatry* **166**, 540–556 (2009).
121. Sullivan, P. F. & Kendler, K. S. The state of the science in psychiatric genomics. *Psychol. Med.* **51**, 2145–2147 (2021).
122. Ask, H. *et al.* Genetic contributions to anxiety disorders: where we are and where we are heading. *Psychol. Med.* **51**, 2231–2246 (2021).
123. Levey, D. F. *et al.* Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat. Neurosci.* **24**, 954–963 (2021).
124. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
125. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019).
126. Tubbs, J. D., Ding, J., Baum, L. & Sham, P. C. Immune dysregulation in depression: Evidence from genome-wide association. *Brain Behav Immun Health* **7**, 100108 (2020).
127. Smoller, J. W. Anxiety Genetics Goes Genomic. *The American journal of psychiatry* vol. 177 190–194 (2020).
128. Smoller, J. W. Who's afraid of anxiety genetics? *Biological psychiatry* vol. 69 506–507 (2011).
129. Levey, D. F., Gelernter, J. & Polimanti, R. Reproducible genetic risk loci for anxiety: results from ~ 200,000 participants in the Million Veteran Program. *American Journal of* (2020).
130. Wendt, F. R. *et al.* Using phenotype risk scores to enhance gene discovery for generalized anxiety disorder and posttraumatic stress disorder. *Mol. Psychiatry* (2022) doi:10.1038/s41380-022-01469-y.
131. Purves, K. L. *et al.* A major role for common genetic variation in anxiety disorders. *Mol. Psychiatry* **25**, 3292–3303 (2020).
132. Coleman, J. R. I. *et al.* Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Mol. Psychiatry* (2020) doi:10.1038/s41380-019-0546-6.
133. Hall, L. S. *et al.* Genome-wide meta-analyses of stratified depression in Generation Scotland and UK Biobank. *Transl. Psychiatry* **8**, 9 (2018).
134. Levey, D. F. *et al.* Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *Am. J. Psychiatry* **177**, 223–232 (2020).
135. Otowa, T. *et al.* Meta-analysis of genome-wide association studies of anxiety disorders. *Mol. Psychiatry* **21**, 1391–1399 (2016).

136. Nagel, M. *et al.* Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat. Genet.* **50**, 920–927 (2018).
137. Grotzinger, A. D., Fuente, J. de la, Privé, F., Nivard, M. G. & Tucker-Drob, E. M. Pervasive Downward Bias in Estimates of Liability-Scale Heritability in Genome-wide Association Study Meta-analysis: A Simple Solution. *Biol. Psychiatry* **93**, 29–36 (2023).
138. Meier, S. M. *et al.* Genetic Variants Associated With Anxiety and Stress-Related Disorders: A Genome-Wide Association Study and Mouse-Model Study. *JAMA Psychiatry* (2019) doi:10.1001/jamapsychiatry.2019.1119.
139. Derks, E. M., Thorp, J. G. & Gerring, Z. F. Ten challenges for clinical translation in psychiatric genetics. *Nat. Genet.* **54**, 1457–1465 (2022).
140. Flint, J. The genetic basis of major depressive disorder. *Mol. Psychiatry* (2023) doi:10.1038/s41380-023-01957-9.
141. Mitchell, B. L. *et al.* Polygenic Risk Scores Derived From Varying Definitions of Depression and Risk of Depression. *JAMA Psychiatry* **78**, 1152–1160 (2021).
142. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
143. Nguyen, T.-D. *et al.* Genetic heterogeneity and subtypes of major depression. *Mol. Psychiatry* **27**, 1667–1675 (2022).
144. Mitchell, B. L. *et al.* The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes. *Biol. Psychiatry* (2021) doi:10.1016/j.biopsych.2021.10.021.
145. Clements, C. C. *et al.* Genome-wide association study of patients with a severe major depressive episode treated with electroconvulsive therapy. *Mol. Psychiatry* **26**, 2429–2439 (2021).
146. Jermy, B. S., Glanville, K. P., Coleman, J. R. I., Lewis, C. M. & Vassos, E. Exploring the genetic heterogeneity in major depression across diagnostic criteria. *Mol. Psychiatry* **26**, 7337–7345 (2021).
147. Howard, D. M. *et al.* Genetic stratification of depression in UK Biobank. *Transl. Psychiatry* **10**, 163 (2020).
148. Coleman, J. R. I. *et al.* The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biol. Psychiatry* (2019) doi:10.1016/j.biopsych.2019.10.015.
149. Cai, N., Choi, K. W. & Fried, E. I. Reviewing the genetics of heterogeneity in depression: operationalizations, manifestations and etiologies. *Hum. Mol. Genet.* **29**, R10–R18 (2020).
150. Morneau-Vaillancourt, G. *et al.* The genetic and environmental hierarchical structure of anxiety and depression in the UK Biobank. *Depress. Anxiety* **37**, 512–520 (2020).
151. Forstner, A. J. *et al.* Genome-wide association study of panic disorder reveals genetic overlap with neuroticism and depression. *Mol. Psychiatry* **26**, 4179–4190 (2019).
152. Erhardt, A. *et al.* TMEM132D, a new candidate for anxiety phenotypes: evidence from human and mouse studies. *Mol. Psychiatry* **16**, 647–663 (2011).
153. Deckert, J. *et al.* GLRB allelic variation associated with agoraphobic cognitions, increased startle response and fear network activation: a potential neurogenetic pathway to panic disorder. *Mol. Psychiatry* **22**, 1431–1439 (2017).
154. Cheesman, R., Rayner, C. & Eley, T. The Genetic Basis of Child and Adolescent

- Anxiety. *Pediatric Anxiety Disorders* 17–46 Preprint at <https://doi.org/10.1016/b978-0-12-813004-9.00003-7> (2019).
155. van Dongen, J., Slagboom, P. E., Draisma, H. H. M., Martin, N. G. & Boomsma, D. I. The continuing value of twin studies in the omics era. *Nat. Rev. Genet.* **13**, 640–653 (2012).
 156. McLaughlin, K. A., Colich, N. L., Rodman, A. M. & Weissman, D. G. Mechanisms linking childhood trauma exposure and psychopathology: a transdiagnostic model of risk and resilience. *BMC Med.* **18**, 96 (2020).
 157. Waszczuk, M. A., Kotov, R., Ruggero, C., Gamez, W. & Watson, D. Hierarchical structure of emotional disorders: From individual symptoms to the spectrum. *J. Abnorm. Psychol.* **126**, 613–634 (2017).
 158. Lewis, S. J. *et al.* The epidemiology of trauma and post-traumatic stress disorder in a representative cohort of young people in England and Wales. *Lancet Psychiatry* **6**, 247–256 (2019).
 159. McLaughlin, K. A. *et al.* Childhood adversities and first onset of psychiatric disorders in a national sample of US adolescents. *Arch. Gen. Psychiatry* **69**, 1151–1160 (2012).
 160. Baldwin, J. R. *et al.* Childhood Maltreatment and Mental Health Problems: A Systematic Review and Meta-Analysis of Quasi-Experimental Studies. *Am. J. Psychiatry* appiajp20220174 (2023).
 161. Koss, K. J. & Gunnar, M. R. Annual Research Review: Early adversity, the hypothalamic–pituitary–adrenocortical axis, and child psychopathology. *Journal of Child Psychology and Psychiatry* vol. 59 327–346 Preprint at <https://doi.org/10.1111/jcpp.12784> (2018).
 162. Danese, A. & Baldwin, J. R. Hidden Wounds? Inflammatory Links Between Childhood Trauma and Psychopathology. *Annu. Rev. Psychol.* **68**, 517–544 (2017).
 163. Kendler, K. S. & Baker, J. H. Genetic influences on measures of the environment: a systematic review. *Psychol. Med.* **37**, 615–626 (2007).
 164. Wootton, R. E., Davis, O. S. P., Mottershaw, A. L., Wang, R. A. H. & Haworth, C. M. A. Genetic and environmental correlations between subjective wellbeing and experience of life events in adolescence. *Eur. Child Adolesc. Psychiatry* **26**, 1119–1127 (2017).
 165. Peel, A. J. *et al.* A multivariate genetic analysis of anxiety sensitivity, environmental sensitivity and reported life events in adolescents. *J. Child Psychol. Psychiatry* **64**, 289–298 (2023).
 166. Stein, M. B., Jang, K. L., Taylor, S., Vernon, P. A. & Livesley, W. J. Genetic and environmental influences on trauma exposure and posttraumatic stress disorder symptoms: a twin study. *Am. J. Psychiatry* **159**, 1675–1681 (2002).
 167. Pezzoli, P., Antfolk, J., Hatoum, A. S. & Santtila, P. Genetic Vulnerability to Experiencing Child Maltreatment. *Front. Genet.* **10**, 852 (2019).
 168. Power, R. A. *et al.* Estimating the heritability of reporting stressful life events captured by common genetic variants. *Psychol. Med.* **43**, 1965–1971 (2013).
 169. Dalvie, S. *et al.* Genomic influences on self-reported childhood maltreatment. *Transl. Psychiatry* **10**, 38 (2020).
 170. Saudino, K. J., Pedersen, N. L., Lichtenstein, P., McClearn, G. E. & Plomin, R. Can personality explain genetic influences on life events? *J. Pers. Soc. Psychol.* **72**, 196–206 (1997).
 171. Kendler, K. S., Karkowski, L. M. & Prescott, C. A. The assessment of dependence in

- the study of stressful life events: validation using a twin design. *Psychol. Med.* **29**, 1455–1460 (1999).
172. Knafo, A. & Jaffee, S. R. Gene–environment correlation in developmental psychopathology. *Dev. Psychopathol.* **25**, 1–6 (2013).
 173. Jaffee, S. R. & Price, T. S. Gene–environment correlations: a review of the evidence and implications for prevention of mental illness. *Mol. Psychiatry* **12**, 432–442 (2007).
 174. Krapohl, E. *et al.* The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 15273–15278 (2014).
 175. McAdams, T. A., Gregory, A. M. & Eley, T. C. Genes of experience: explaining the heritability of putative environmental variables through their association with behavioural and emotional traits. *Behav. Genet.* **43**, 314–328 (2013).
 176. Barkhuizen, W., Dudbridge, F. & Ronald, A. Genetic overlap and causal associations between smoking behaviours and mental health. *Sci. Rep.* **11**, 14871 (2021).
 177. Werme, J., van der Sluis, S., Posthuma, D. & de Leeuw, C. A. Genome-wide gene–environment interactions in neuroticism: an exploratory study across 25 environments. *Transl. Psychiatry* **11**, 180 (2021).
 178. Turkheimer, E. Three Laws of Behavior Genetics and What They Mean. *Curr. Dir. Psychol. Sci.* **9**, 160–164 (2000).
 179. Assary, E., Vincent, J. P., Keers, R. & Pluess, M. Gene–environment interaction and psychiatric disorders: Review and future directions. *Semin. Cell Dev. Biol.* **77**, 133–143 (2018).
 180. Peel, A. J. *et al.* Genetic and early environmental predictors of adulthood self-reports of trauma. *Br. J. Psychiatry* 1–8 (2022).
 181. Kendler, K. S. *et al.* Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am. J. Psychiatry* **152**, 833–842 (1995).
 182. Dick, D. M. Gene–environment interaction in psychological traits and disorders. *Annu. Rev. Clin. Psychol.* **7**, 383–409 (2011).
 183. Pingault, J.-B. *et al.* Research Review: How to interpret associations between polygenic scores, environmental risks, and phenotypes. *J. Child Psychol. Psychiatry* **63**, 1125–1139 (2022).
 184. Danese, A. & Widom, C. S. Objective and subjective experiences of child maltreatment and their relationships with psychopathology. *Nat Hum Behav* **4**, 811–818 (2020).
 185. Uliaszek, A. A. *et al.* The role of neuroticism and extraversion in the stress–anxiety and stress–depression relationships. *Anxiety Stress Coping* **23**, 363–381 (2010).
 186. Assary, E., Zavos, H. M. S., Krapohl, E., Keers, R. & Pluess, M. Genetic architecture of Environmental Sensitivity reflects multiple heritable components: a twin study with adolescents. *Mol. Psychiatry* **26**, 4896–4904 (2021).
 187. Reuben, A. *et al.* Lest we forget: comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *J. Child Psychol. Psychiatry* **57**, 1103–1112 (2016).
 188. Gehrt, T. B., Berntsen, D., Hoyle, R. H. & Rubin, D. C. Psychological and clinical correlates of the Centrality of Event Scale: A systematic review. *Clin. Psychol. Rev.* **65**, 57–80 (2018).
 189. Ogle, C. M., Siegler, I. C., Beckham, J. C. & Rubin, D. C. Neuroticism Increases

- PTSD Symptom Severity by Amplifying the Emotionality, Rehearsal, and Centrality of Trauma Memories. *J. Pers.* **85**, 702–715 (2017).
190. Rietschel, L. *et al.* Perceived stress has genetic influences distinct from neuroticism and depression. *Behav. Genet.* **44**, 639–645 (2014).
 191. Belsky, J. & Pluess, M. Beyond diathesis stress: differential susceptibility to environmental influences. *Psychol. Bull.* **135**, 885–908 (2009).
 192. Jolicoeur-Martineau, A. *et al.* Distinguishing differential susceptibility, diathesis-stress, and vantage sensitivity: Beyond the single gene and environment model. *Dev. Psychopathol.* **32**, 73–83 (2020).
 193. Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M. & Huibers, M. J. H. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry* vol. 15 245–258 Preprint at <https://doi.org/10.1002/wps.20346> (2016).
 194. Rush, A. J. *et al.* Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am. J. Psychiatry* **163**, 1905–1917 (2006).
 195. Bousman, C. A., Arandjelovic, K., Mancuso, S. G., Eyre, H. A. & Dunlop, B. W. Pharmacogenetic tests and depressive symptom remission: a meta-analysis of randomized controlled trials. *Pharmacogenomics* **20**, 37–47 (2019).
 196. Tansey, K. E. *et al.* Contribution of common genetic variants to antidepressant response. *Biol. Psychiatry* **73**, 679–682 (2013).
 197. Pain, O. *et al.* Identifying the Common Genetic Basis of Antidepressant Response. *Biol Psychiatry Glob Open Sci* **2**, 115–126 (2022).
 198. David, D., Cristea, I. & Hofmann, S. G. Why Cognitive Behavioral Therapy Is the Current Gold Standard of Psychotherapy. *Front. Psychiatry* **9**, 4 (2018).
 199. Powers, M. B., de Kleine, R. A. & Smits, J. A. J. Core Mechanisms of Cognitive Behavioral Therapy for Anxiety and Depression: A Review. *Psychiatr. Clin. North Am.* **40**, 611–623 (2017).
 200. Newby, J. M., McKinnon, A., Kuyken, W., Gilbody, S. & Dalgleish, T. Systematic review and meta-analysis of transdiagnostic psychological treatments for anxiety and depressive disorders in adulthood. *Clin. Psychol. Rev.* **40**, 91–110 (2015).
 201. García-Escalera, J., Chorot, P., Valiente, R. M., Reales, J. M. & Sandín, B. Efficacy of transdiagnostic cognitive-behavioral therapy for anxiety and depression in adults, children and adolescents: A meta-analysis. *Revista de Psicopatología y Psicología Clínica* **21**, 147–175 (2016).
 202. Tanguay-Sela, M. *et al.* A systematic meta-review of patient-level predictors of psychological therapy outcome in major depressive disorder. *J. Affect. Disord.* **317**, 307–318 (2022).
 203. Purves, K. L. *et al.* Evidence for distinct genetic and environmental influences on fear acquisition and extinction. *Psychol. Med.* 1–9 (2021).
 204. Hettema, J. M., Annas, P., Neale, M. C., Kendler, K. S. & Fredrikson, M. A twin study of the genetics of fear conditioning. *Arch. Gen. Psychiatry* **60**, 702–708 (2003).
 205. Zanardi, R., Carminati, M., Attanasio, F., Fabbri, C. & Serretti, A. Genetics of nonpharmacological treatments of depression. *Psychiatr. Genet.* **33**, 1–7 (2023).
 206. Lester, K. J. & Eley, T. C. Therapygenetics: Using genetic markers to predict response to psychological treatment for mood and anxiety disorders. *Biol. Mood Anxiety Disord.* **3**, 4 (2013).

207. Keers, R. *et al.* A Genome-Wide Test of the Differential Susceptibility Hypothesis Reveals a Genetic Predictor of Differential Response to Psychological Treatments for Child Anxiety Disorders. *Psychother. Psychosom.* **85**, 146–158 (2016).
208. Lester, K. J. *et al.* Non-replication of the association between 5HTTLPR and response to psychological therapy for child anxiety disorders. *Br. J. Psychiatry* **208**, 182–188 (2016).
209. Rayner, C. *et al.* A genome-wide association meta-analysis of prognostic outcomes following cognitive behavioural therapy in individuals with anxiety and depressive disorders. *Transl. Psychiatry* **9**, 150 (2019).

Chapter 2. Exploring the genome-wide genetic overlap between anxiety and fear disorders

This chapter is a manuscript that is in preparation for peer-review. Supplementary materials for this chapter, as detailed in the text, are included in **Appendix A**.

Authors: **Abigail R. ter Kuile**^{1,2}, Brittany L. Mitchell³, Geneviève Morneau-Vaillancourt¹, Sang Hyuck Lee¹, Helena L. Davies¹, Jessica Mundy^{1,2}, Alicia J. Peel¹, Megan Skelton^{1,2}, Christopher Hübel^{1,2,9}, Molly R. Davies^{1,2}, Jonathan R. I. Coleman^{1,2}, Anna E. Fürtjes¹, Zain Ahmad¹, Yuhao Lin¹, Brett N. Adey^{1,2}, Thomas McGregor¹, Kirstin Purves^{1,2}, Alish Palmos¹, Johan Zvrskovec^{1,2}, Matthew Hotopf^{1,2,8}, Gursharan Kalsi^{1,2}, Ian R. Jones¹⁰, Daniel J. Smith¹¹, David Veale^{1,2,8}, James T. R. Walters¹⁰, Chérie Armour⁵, Colette R. Hirsch^{1,2,8}, Andrew M. McIntosh¹¹, Naomi R. Wray^{12,13}, Sarah E. Medland³, Enda M. Byrne⁶, Nicholas G. Martin³, Gerome Breen^{1,2}, Thalia C. Eley^{1,2}

1. Institute of Psychiatry, Psychology and Neuroscience, King's College London, Denmark Hill, Camberwell, London, UK
2. National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre, South London and Maudsley NHS Foundation Trust, London, UK
3. QIMR Berghofer Medical Research Institute, Brisbane, Australia
4. School of Biomedical Sciences, Faculty of Medicine, The University of Queensland, Brisbane, Australia.
5. Research Centre for Stress, Trauma, and Related Conditions (STARC), School of Psychology, Queen's University Belfast (QUB), Belfast, Northern Ireland, UK
6. Child Health Research Centre, The University of Queensland, Brisbane, Australia
7. Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia
8. South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital, Monks Orchard Road, Beckenham, Kent, UK
9. National Centre for Register-based Research, Aarhus Business and Social Sciences, Aarhus University, Aarhus, Denmark
10. MRC Centre for Neuropsychiatric Genetics and Genomics, National Centre for Mental Health, Cardiff University, Cardiff, UK
11. Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK
12. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Australia
13. Queensland Brain Institute, The University of Queensland, Brisbane, Australia

Abstract

Background: Twin studies have consistently shown a high genetic overlap amongst anxiety disorders and depression that contributes to the internalising spectrum. Research has also identified modest genetic specificity to fear-based anxiety disorders, not shared with general anxiety and depression (often grouped as distress disorders). Due to the lack of datasets with detailed phenotyping of all five anxiety disorders, genome-wide analysis of anxiety has typically been limited to “any anxiety diagnosis”. Further genome-wide evidence is needed to establish if fear-based disorders are genetically distinct from distress disorders.

Methods: We undertook a genome-wide association study (GWAS) meta-analysis of fear-based disorders (panic, social anxiety disorder, specific phobia, and agoraphobia) and generalised anxiety disorder (GAD). Cases and controls were defined using a combination of brief single-item and detailed symptom-based diagnoses from three large datasets. We explored two approaches to define controls. First, we screened for any anxiety disorder and depression. Second, we screened specifically for fear or GAD. We estimated genetic correlations between fear and GAD and compared their correlations with broad domains of other complex traits.

Results: Our GWAS meta-analyses identified a total of three independent loci associated with fear (up to 30,861 N_{cases} ; 157,951 N_{controls}) and four with GAD (up to 54,928 N_{cases} ; 117,320 N_{controls}). The genetic correlation between fear and GAD was not significantly different from one, except for when excluding a depression-enriched dataset and screening controls specifically ($r_g = 0.85$; $P = 4.57 \times 10^{-3}$). Most complex traits did not have a significantly different genetic correlation with fear versus GAD, including depression. The exceptions to this included general cognitive ability, educational attainment, and coronary artery disease with stronger negative genetic correlations with fear than GAD. Bipolar disorder type I, anorexia nervosa, and neuroticism had stronger positive genetic correlations with GAD than fear.

Conclusions: Our findings partially support a distress-fear genetic distinction. However, we found stronger evidence for an overarching genetic liability to internalising psychopathology that drives comorbidity across anxiety disorders and depression.

Introduction

Anxiety disorders, including panic disorder, agoraphobia, specific phobia, social anxiety disorder, and generalised anxiety disorder (GAD), are among the most common mental health conditions, with an international lifetime prevalence of 16%^{1,2}. The core symptoms of anxiety disorders are persistent and excessive fear and worry, accompanied by impairing features such as avoidance behaviours, physiological arousal, and panic attacks³. Individual disorder diagnoses across the anxiety spectrum are based upon their distinct clinical features, although high rates of comorbidity exist between them^{4,5}. GAD is characterised by excessive and chronic worry about a broad range of future potential threats. Panic disorder is an extreme version of a flight or fight response with repeated experiences of a sudden unexpected feeling of fear in the absence of a threat (i.e. panic attack). Recurrent, unexpected panic attacks are necessary for a diagnosis of panic disorder. Panic attacks, however, are also transdiagnostic and can occur in certain situations, as in phobic disorders. The phobias, including agoraphobia, social anxiety disorder and specific phobia, are characterised by excessive fear and avoidance of particular signals of perceived threat^{4,6}. As they are all characterised by fear, panic disorder and the phobias are often grouped together as 'fear-based' disorders.

Shared risk factors for anxiety disorders include being female, exposure to stressful experiences, and a family history of anxiety and depression³. Familial aggregation of these disorders reflects a partial genetic basis, with twin study estimates of heritability ranging from 20-60%^{3,7,8}. Twin studies consistently show that most genetic influences on anxiety disorders overlap, contributing to a broad anxiety-related genetic factor, which has considerable overlap with genetic risk for depression⁹⁻¹². However, evidence suggests this higher-order internalising genetic liability can be differentiated into lower-order genetic components. Fear-based disorders are reported to be partly genetically distinct from those characterised by distress (GAD and depression), forming two distinct but correlated genetic liability fear-distress subfactors^{9,13}. Overall, twin studies show that anxiety disorders mostly share common genetic factors but also have some genetic specificity.

Although twin studies have been crucial in revealing the overall genetic structure of the anxiety disorder spectrum, they do not identify the specific genetic variants that contribute to the heritability and high genetic overlap between the disorders. Genome-wide association studies (GWAS) have begun to identify common genetic variants underpinning anxiety disorders. Fewer GWAS have been performed in anxiety disorders than in other psychiatric disorders¹⁴. However, recent developments have enabled several large-scale GWAS of broadly defined anxiety and GAD symptoms¹⁵⁻¹⁸. Such studies have identified significant genetic correlations with all other psychiatric disorders, including depression, schizophrenia, bipolar disorder, and

attention deficit hyperactivity disorder (ADHD)^{15–17}. Consistent with evidence from twin studies⁹, preliminary genome-wide findings indicate that a lifetime diagnosis of any fear-based disorder (as a grouping) and depression are less genetically correlated with one another than they each are with lifetime GAD¹⁹. However, the common genetic variants associated with fear-based disorders remain unclear, which limits our understanding of their genetic overlap at the molecular level.

Previous GWAS of fear-based disorders have been underpowered for loci discovery and to assess genetic correlations with a broad range of traits^{19–21}. Such analyses would provide more molecular genetic evidence and fine-grained detail to the current hierarchical structures of the psychiatric spectrum, which could have implications for improving diagnostic systems²². A formal comparison of fear and GAD genetic correlations with other psychiatric, behavioural, cognitive, or health traits could suggest transdiagnostic or subtype-specific mechanisms. Further well-powered genome-wide evidence of the genetic distinction between distress and fear is needed in the broader context of their relationships with other complex traits.

The way in which anxiety disorders are measured is an important consideration when identifying genetic similarities and differences. Dealing with comorbidity is a particular challenge in identifying anxiety-specific genetic influences, as anxiety disorders are among the most highly comorbid disorders across psychopathology²³. Screening controls for co-occurring traits in GWAS could inflate genetic correlations²⁴. A comparison of different control screening approaches is needed. Furthermore, genetic influences captured by using brief phenotyping measures may be less specific to the focal trait than those captured via detailed phenotyping, reflecting more general psychopathology^{25,26}. The previous GWAS reporting distress-fear genetic correlations was limited to brief, single-item reports of receipt of a diagnosis by a health professional for each fear disorder¹⁹. A recent study found that rates of fear-based disorder diagnoses were lower when using brief self-report diagnoses than detailed symptom-based measures, and observed the opposite pattern for GAD²⁷. Since identifying genetic variants requires large sample sizes, a trade-off exists between maximising sample size and the level of detail in phenotyping. One way to achieve larger sample sizes while retaining phenotypic specificity is by combining detailed symptom-based diagnoses with brief single-item self-report diagnoses²⁸.

Here, we aimed to assess the shared and non-shared genetic influences on fear-based disorders and GAD at the level of common genetic variants. We increased sample size and thus power for loci-discovery and genetic correlation analyses by combining detailed symptom-based diagnoses with brief single-item diagnoses. We examined differences in genetic correlations between fear-based disorders and GAD with a range of other complex

traits, and also assessed broad, and more specific approaches to screening controls. Our study represents progress in the anxiety disorder genomics field and adds further detail on the distress-fear distinction at the genome-wide level.

Methods

Phenotypes

Anxiety disorder phenotypes

Analyses were part of a pre-registered plan on the Open Science Framework (accessible at <https://tinyurl.com/OSFfearGAD>). We conducted GWAS meta-analyses of two lifetime anxiety disorder phenotypes: i) fear-based disorders (referred to as ‘fear’ hereafter; panic, agoraphobia, specific phobia and social anxiety disorder; N cases = 30,861) and ii) GAD (N cases = 54,928). We used the term “panic” instead of “panic disorder” as brief diagnostic measures of self-reported panic attacks were used in addition to brief and detailed diagnostic measures of panic disorder. This increased power and data harmonisation, as the UK Biobank Mental Health Questionnaire only had brief measures on panic attacks²⁹ (**Supplementary Tables 1 & 2**). Although panic attacks are not specific to panic disorder, the brief diagnostic measure of panic attacks has shown reasonable agreement with detailed diagnostic measures of panic disorder²⁷. Five samples were included in the GWAS meta-analyses. The five samples resulted in three case and control datasets, as detailed below (see **Table 1** for a summary of sample sizes). Individuals were screened for an anxiety disorder using a combination of detailed symptom-based diagnoses and brief, self-report diagnoses. Fear cases were defined as participants who endorsed at least one fear disorder.

Detailed diagnostic measures

Detailed symptom-based measures of anxiety disorders were derived using previously described algorithms²⁷. Cases for each phenotype were defined as those who met DSM-5 criteria for a lifetime symptom-based disorder diagnosis. Each anxiety disorder was assessed using an online, self-report questionnaire version of the Composite International Diagnostic Interview short-form (CIDI-SF)^{30,31}. The availability of questionnaires in each study is summarised in **Supplementary Table 1**.

Brief diagnostic measures

Individuals meeting case status on brief measures endorsed a single-item self-report question for the corresponding disorder, e.g. “Have you ever been diagnosed with one or more of the following mental health problems by a professional, even if you do not have it currently?”. The exact phrasing of the question and responses in each study, as well as mapping to disorders, are provided in **Supplementary Table 2**.

Control groups

We explored two different ways of defining control groups. First, controls were screened for any anxiety disorder or depression ($\text{fear}_{\text{anx-dep}}$ and $\text{GAD}_{\text{anx-dep}}$). Due to the high genetic correlation between these disorders¹⁹, this was the most powerful approach for loci discovery. Detailed and brief measures were used to screen controls for any anxiety disorder (**Supplementary Tables 1 & 2**). Controls were screened for depression using brief, self-report diagnoses in all datasets. Second, to better distinguish genetic differences between GAD and fear, we defined specifically screened controls whereby participants were only screened for the specific disorder being analysed ($\text{fear}_{\text{specific}}$ and $\text{GAD}_{\text{specific}}$). This is because both fear and GAD have high genetic overlap with one another and with depression, thus screening controls for any anxiety disorder and depression could further increase the genetic correlation between them. $\text{Fear}_{\text{specific}}$ GWAS controls were only screened for any fear disorder, and $\text{GAD}_{\text{specific}}$ GWAS controls were only screened for GAD where possible. Screening controls specifically for only GAD or only fear-based disorders was not possible in the QIMR dataset as we were limited to using a broad, single-item self-report diagnosis of “anxiety”. Controls were not screened for the presence of other psychiatric disorders as this can bias genetic correlation estimates²⁴.

Samples

The GLAD Study and COPING Study (GLAD+ dataset)

Data from two studies within the National Institute for Health and Care Research (NIHR) BioResource were included. Cases were primarily ascertained from the Genetic Links to Anxiety and Depression (GLAD) study³². The GLAD study is an online research platform launched in September 2018 to recruit participants with lifetime anxiety and/or depressive disorder. Recruitment is ongoing and is open to individuals residing in the UK and aged at least 16 years. Initial recruitment involved a widespread media campaign, followed by ongoing social media advertising and participating NHS trusts and GP practices. Controls for analyses with GLAD cases were ascertained from a longitudinal survey run by the GLAD team, collecting data within GLAD and other NIHR BioResource cohorts called the COVID-19 Psychiatric and Neurological Genetics (COPING) study³³. The phenotyping in COPING for anxiety disorders was identical to GLAD. A proportion of COPING study participants met the criteria for an anxiety disorder and were defined as cases (**Table 1**). Cases and controls in the GLAD and COPING studies were defined using both brief and detailed measures of all five anxiety disorders. We refer here to the GLAD and COPING combined study samples as the GLAD+ dataset.

Queensland Institute of Medical Research (QIMR) Berghofer dataset

A second unrelated set of case-control analyses included data from two Australian QIMR Berghofer studies. Cases were obtained from the Australian Genetics of Depression Study (AGDS), launched in September 2016, to collect genetic and phenotypic data on individuals with a lifetime experience of depression³¹. Recruitment is ongoing and has been conducted via pharmaceutical prescription history provided by the Australian government and through media campaigns. Participation is open to individuals aged at least 18 years across Australia. AGDS cases were defined using detailed and brief self-report measures of all five anxiety disorders. Controls for analyses with AGDS cases were obtained from the QIMR Berghofer QSkin Sun & Health Study³⁴, defined using a brief self-report measure of any anxiety disorder.

The UK Biobank dataset

The UK Biobank is a population-based dataset of over 500,000 participants aged between 40 and 69 recruited from 2006 to 2010³⁵. In 2017, a subset of participants (N = 157,366) completed the online follow-up Mental Health Questionnaire (MHQ) assessing mental health and well-being through self-report measures²⁹. Participants were defined as cases or controls using brief measures of fear disorders and brief and detailed measures of GAD.

Table 1. Sample sizes in each study, GWAS dataset, and meta-analysis of fear-based disorders and GAD.

Anxiety disorder	GWAS Datasets	Study Sample	N _{cases}	Control criteria				N dataset total (N _{cases} + N _{controls})	
				Anx-dep controls		Specific controls		Anx-dep	Specific
				N _{controls}	N _{effective}	N _{controls}	N _{effective}		
Fear	GLAD+	GLAD COPING	12,243 1,186	- 4,419	13,300	- 6197	16,961	17,848	19,626
		UKB	10,789	83,236	38,204	137,209	40,010	94,025	147,998
	QIMR	AGDS Qskin	6,643 -	- 12,722	17,457	- 14,545	18,241	19,365	21,188
		Total meta-analysis N (GLAD+, UKB & QIMR)	30,861	100,377	68,960	157,951	75,212	131,238	188,812
		Total meta-analysis N (GLAD+ & UKB)	24,218	87,655	51,504	143,406	56,971	111,873	167,624
GAD	GLAD+	GLAD COPING	14,339 1,694	- 4,419	13,857	- 7060	19,606	20,452	23,093
		UKB	26,067	83,236	79,402	95,715	81,950	109,303	121,782
	QIMR	AGDS Qskin	12,828 -	- 12,722	25,550	- 14,545	27,265	25,550	27,373
		Total meta-analysis N (GLAD+, UKB & QIMR)	54,928	100,377	118,808	117,320	128,822	155,305	172,248
		Total meta-analysis N (GLAD+ & UKB)	42,100	87,655	93,259	102,775	101,556	129,755	144,875

Numbers shown in bold are total sample size used in GWAS meta-analyses. Effective sample size was calculated as: $N_{\text{effective}} = 4 / [(1/n_{\text{case}}) + (1/n_{\text{control}})]$, which converts the power of a study sample size to one with a balanced case-control ratio (i.e. equivalent to a study with a sample prevalence of 50%, with $N_{\text{cases}} = N_{\text{effective}}$ and $N_{\text{controls}} = N_{\text{effective}}$).

Genome-wide association analyses

Genotyping and quality control was performed separately in each study (**Supplementary Materials**). Common genetic variants in the GLAD+ and QIMR datasets were imputed to the Top Med imputation panel (Version R2 on GRC38) and the UK Biobank to the Haplotype Reference Consortium and the UK10K Consortium reference panels. GWAS analyses were restricted to common genetic variants (MAF > 1%) and high confidence-imputed variants (INFO score > 0.3 for TOPMed imputed variants; > 0.4 for UK Biobank).

We ran four anxiety disorder GWASs; $\text{fear}_{\text{specific}}$, $\text{GAD}_{\text{specific}}$, $\text{fear}_{\text{anx-dep}}$ and $\text{GAD}_{\text{anx-dep}}$. A separate GWAS was conducted for each of the three datasets (GLAD+, UK Biobank, QIMR). Sample sizes for each GWAS are shown in **Table 1**. The software REGENIE v3.1.3 was used for analyses in the GLAD+ and the UK Biobank datasets, whilst SAIGE v0.44 was used in the QIMR dataset^{36,37}. The first ten principal components and genotyping batch were included as covariates. Additionally, the assessment centre was included as a covariate for GWAS in the UK Biobank. GWAS were limited to participants of European-associated genetic ancestry clusters, defined using principal component analysis (PCA).

Meta-analyses and annotation

We used the software METAL to perform inverse-variance weighted meta-analyses of each anxiety disorder phenotype³⁸. In our primary analyses, we meta-analysed GWAS results from the GLAD+, QIMR Berghofer, and UK Biobank datasets (termed *full meta-analysis* throughout). As all cases from the AGDS also had comorbid depression, we conducted a second set of meta-analyses excluding the QIMR datasets (referred to as *GLAD+ and UKB meta-analysis*). A depression-enriched dataset might increase depression/distress genetics in anxiety disorder GWAS, elevating the genetic correlation between fear, GAD, and depression. Meta-analysis sample sizes are shown in **Table 1**. We restricted common genetic variants to those overlapping across datasets in each meta-analysis. Associated variants were mapped and annotated using FUMA v1.4.0 (with default parameters applied) and the UKB release2b 10K European reference panel³⁹. MAGMA v1.08 was used with default parameters to identify gene-level and biological pathway associations. Gene-level association analysis aggregates the combined effect of associated common genetic variants at the level of the gene. Results from gene-level analyses were then applied to gene-set analysis to conduct biological pathway analyses⁴⁰.

SNP-based heritability

We estimated the heritability captured by common genetic variants (h^2_{SNP}) of fear and GAD. Population prevalences for liability scale estimates were based on the assumption of accurate sampling from the COPING study, in which both detailed and brief measures were available to define cases and controls for all five anxiety disorders (**Supplementary Table 3**). GWAS meta-analysis summary results were used in LDSC regression to estimate h^2_{SNP} ⁴¹.

Genetic correlations

Using LDSC regression⁴², we calculated genetic correlations (r_g) between i) fear and GAD and ii) fear and GAD with 345 external complex traits respectively. Traits were considered sufficiently powered for genetic correlation analysis if they had a GWAS mean $X^2 > 1.02$ and a heritability Z score > 4 , calculated in LDSC regression. We tested if the genetic correlation between fear and GAD was significantly different from 0 (using default parameters in bivariate LDSC regression) and significantly different from 1 (calculated in R using the chi-squared distribution function and $[(|r_g| - 1)/\text{se}]^2$). To correct for multiple testing, we applied a Bonferroni corrected P -value threshold ($\alpha [0.05] / \text{the number of fear-GAD } r_g \text{ tested } [4] = P \leq 0.0125$). Fear and GAD GWAS with a genetic correlation significantly different from 1 were then tested for genetic correlations with external traits. We selected external traits in a hypothesis-free manner and applied a Bonferroni-corrected significance threshold ($0.05 / \text{number of external traits tested } [345] = P \leq 1.45 \times 10^{-4}$). We tested for differences between fear versus GAD in terms of their respective genetic correlations with external traits using the block-jackknife method implemented in LDSC regression^{42,43}. Fear and GAD were considered to have a significantly different genetic correlation with an external trait if the P -value derived from the block-jackknife Z statistic result exceeded the Bonferroni-corrected threshold ($P \leq 1.45 \times 10^{-4}$).

Results

Genome-wide association meta-analyses

Manhattan plots for the most well-powered GWAS are shown in **Figure 1** ($\text{fear}_{\text{anx-dep}}$ and $\text{GAD}_{\text{anx-dep}}$ in the full meta-analysis). Manhattan plots for all other GWAS with genome-wide significant loci are shown in **Supplementary Figures 2, 4, 14 and 16**. We found little evidence of confounding in all GWAS (LDSC intercept = 0.96-1.04). LDSC intercepts >1 indicate some inflation, which is expected as the GWAS sample size increases⁴⁴. The majority of inflation in our meta-analyses was due to polygenicity (90-97%), as indicated by LDSC attenuation ratio calculations. The genomic inflation factor, LDSC intercept, attenuation ratio, and Q-Q plots for each GWAS are shown in **Supplementary Figures 1, 3, 12, 13 and 15**. Regional plots for independent genome-wide significant loci are also in **Supplementary Figures 5-11**.

Fear-based disorders

Across the four GWASs, we found a total of three independent genome-wide significant loci associated with fear. In each fear GWAS, a different locus was identified, except for $\text{fear}_{\text{specific}}$ in the GLAD+ and UKB meta-analysis, where no loci reached $P < 5 \times 10^{-8}$. The locus with the most significant lead variant was rs10047892 on chromosome 14 in the $\text{fear}_{\text{anx-dep}}$ GWAS in the full meta-analysis ($P = 2.99 \times 10^{-8}$; **Figure 1, upper panel**). The $\text{fear}_{\text{specific}}$ GWAS in the full meta-analysis showed a locus on chromosome 5 as genome-wide significant (lead variant rs3996354; $P = 4.71 \times 10^{-8}$). A locus on chromosome 1 with lead variant rs11576254 was found in the $\text{fear}_{\text{anx-dep}}$ GWAS in the GLAD+ and UKB meta-analysis ($P = 4.70 \times 10^{-8}$; **Table 2**).

Generalised anxiety disorder

In all, we identified four independent genome-wide significant loci associated with GAD ($P < 5 \times 10^{-8}$), although not all four loci reached genome-wide significance in each GAD GWAS. The most significantly associated locus was on chromosome 9 (lead variant rs10120318; $P = 3.1 \times 10^{-13}$; **Figure 1, lower panel**) and was genome-wide significant in all four GAD GWASs (**Table 2**). The locus on chromosome 6 was the second most significant (lead variant rs3858; $P = 1.1 \times 10^{-9}$; **Figure 1, lower panel**) and reached genome-wide significance in all GWAS meta-analyses of GAD, except for $\text{GAD}_{\text{specific}}$ in the UKB and GLAD+ meta-analysis. The locus on chromosome 2 exceeded genome-wide significance in the $\text{GAD}_{\text{specific}}$ GWAS in the full meta-analysis (lead variant rs11688767; $P = 4.20 \times 10^{-8}$), and a locus on chromosome 5 in the $\text{GAD}_{\text{specific}}$ GWAS in the UKB and GLAD+ meta-analysis (lead variant rs3095951; $P = 4.33 \times 10^{-8}$; **Table 2**).

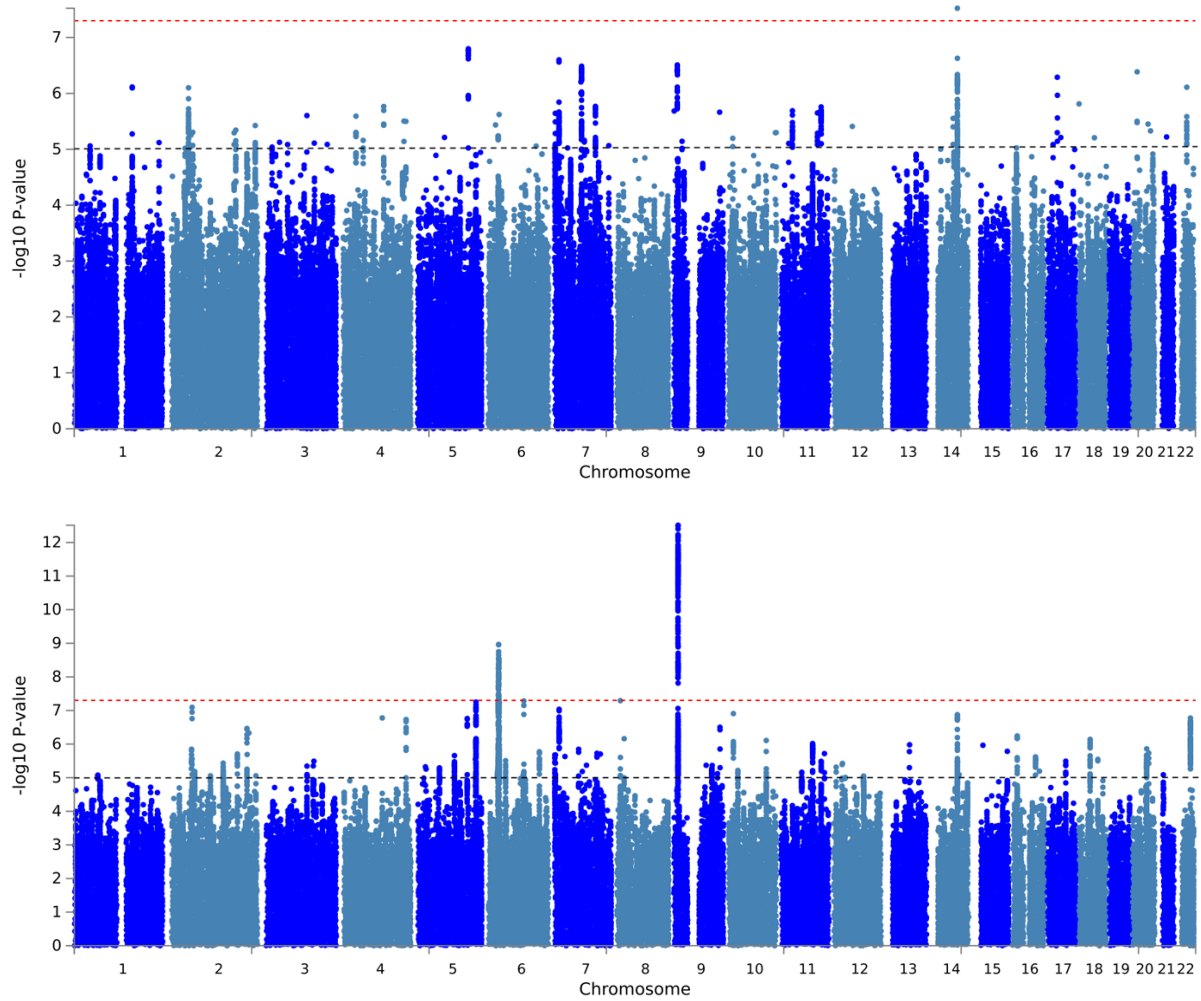


Figure 1: Genome-wide association study Manhattan plots for anxiety disorder phenotypes meta-analysed across the GLAD+, QIMR and UKB datasets. *Fear-based disorder GWAS (upper panel) and GAD GWAS (lower panel) results from phenotypes with controls screened for any anxiety disorder and depression. Dashed line red; common genetic variant genome-wide significance threshold ($P < 5 \times 10^{-8}$), black; suggestive significance threshold ($P < 1 \times 10^{-5}$).*

Table 2: Independent genome-wide significant loci associated with anxiety disorder phenotypes

Anxiety disorder	Datasets	Control screening criteria	Locus no.	Region	CHR	Lead SNP	Base pair	A1:A2	Func; gene (kb)	P	OR	N cand. SNPs	Ind. Sig. SNPs	Previous report in GWAS catalog
Fear	QIMR, GLAD+ & UKB	Anx-dep	1	14q24.3	14	rs10047892	75111346	T:C	intergenic; AREL1 (8793)	2.99E-08	1.07	26	rs10047892	EA, worry, systolic blood pressure
		Specific	2	5q31.3	5	rs3996354	143264699	A:G	intergenic; CTB-57H20.1 (56361)	4.71E-08	0.93	20	rs3996354	Novel
	GLAD+ & UKB	Anx-dep	3	1p34.1	1	rs11576254	44835833	T:C	intergenic; ERI3 (14900)	4.70E-08	1.09	64	rs11576254	None at genome-wide significance. At suggestive threshold: metabolite levels, reaction time, anxiety
		Specific	-	-	-	-	-	-	-	-	-	-	-	-
GAD	QIMR, GLAD+ & UKB	Anx-dep	4	9p23	9	rs10120318	11645069	A:T	intergenic; RP11-23D5.1 (368754)	3.07E-13	0.93	301	rs10120318; rs11515172	Neuroticism, depression, wellbeing, any anxiety disorder, GAD symptoms, worry, BMI
			5	6p22.1	6	rs385816	29480224	A:G	intergenic; XXbac-BPG13B8.10 (1890)	1.08E-09	0.92	184	rs385816	SCZ, HDL, BIP, cognitive performance, depression, cardiometabolic & hematological traits, lung cancer, worry, smoking, neuroticism
		Specific	4	9p23	9	rs10960024	11616820	C:G	intergenic; RP11-23D5.1 (340505)	7.17E-13	0.93	301	rs10960024; rs11515172	As locus 4 above
			5	6p22.1	6	rs385816	29480224	A:G	intergenic; XXbac-BPG13B8.10 (1890)	7.32E-09	0.93	184	rs385816	As locus 5 above
			6	2p16.1	2	rs11688767	57988194	A:T	ncRNA intronic; CTD-2026C7.1 (0)	4.20E-08	0.95	19	rs11688767	Cross-disorder, sleep, SCZ, cognitive ability, depression, neuroticism, epilepsy, neuroticism

Table 2: Independent genome-wide significant loci associated with anxiety disorder phenotypes

Anxiety disorder	Datasets	Control screening criteria	Locus no.	Region	CHR	Lead SNP	Base pair	A1:A2	Func; gene (kb)	P	OR	N cand. SNPs	Ind. Sig. SNPs	Previous report in GWAS catalog
GAD	GLAD+ & UKB	Anx-dep	4	9p23	9	rs17189482	11513617	T:G	intergenic; RP11-23D5.1 (237302)	1.24E-11	1.09	295	rs17189482; rs11515172	As locus 4 above
			5	6p22.1	6	rs3117427	29274136	T:C	upstream; OR14J1 (266)	8.70E-09	1.09	219	rs3117427	As locus 5 above
			7	5q34	5	rs3095951	164616460	A:G	intergenic; CTC-340A15.2 (17810)	4.33E-08	1.06	142	rs3095951	Any anxiety disorder, depression, neuroticism, wellbeing, BMI
		Specific	4	9p23	9	rs10960024	11616820	C:G	intergenic; RP11-23D5.1 (340505)	1.63E-11	0.92	250	rs10960024	As locus 4 above

Locus no. = annotated locus number; CHR = chromosome; A1 = effect allele; A2 = noneffect allele; Func gene (kb) = functional consequence of the SNP on the nearest gene with distance in kilobase; OR = odds ratio; N cand. SNPs = Number of genome-wide candidate SNPs that are in LD ($r^2 = 0.6$) with Ind. Sig. SNPs; Ind. Sig. SNPs = independent significant SNPs in the locus. Previous report = GWASs of other phenotypes that reported a genome-wide association with one or more of the candidate SNPs identified in this study (full results shown in **Supplementary Tables 4 & 5** for fear and GAD loci, respectively). EA = educational attainment; BMI = body mass index; SCZ = schizophrenia, HDL = high-density lipoprotein cholesterol, BIP = bipolar disorder.

Gene-level and gene-set association analyses

Fear-based disorders

Three genes were significantly associated with fear-based disorders after Bonferroni correction for the number of genes tested (see **Table 3** for P-value thresholds). The gene *SNX29* on chromosome 16 was significantly associated with fear in all GWASs. The gene *YWHAG* on chromosome 7 was only significantly associated with both fear GWAS in the full meta-analysis, whilst the gene *ER13* on chromosome 1 was only significantly associated with both fear GWAS in the GLAD+ and UKB meta-analysis (**Table 3**). The gene *ER13* includes the SNP-level genome-wide significant locus 1p34.1. No gene sets were significantly associated with fear (Bonferroni correction threshold for 10678 gene sets = $P \leq 4.68 \times 10^{-6}$).

Generalised anxiety disorder

A total of three genes were significantly associated with GAD. All three genes were identified in the GAD_{anx-dep} GWAS in the full meta-analysis. The most significantly associated gene was *TMEM106B* on chromosome 7. This gene was also found in GAD_{specific} GWAS in the full meta-analysis and the GAD_{anx-dep} GWAS in the GLAD+ and UKB meta-analysis. The gene *TRIM31* on chromosome 6 was also associated with GAD_{specific} GWAS in the full meta-analysis. This gene region includes common genetic variants on locus 6p22.1 identified in SNP-level association analyses. The gene *SORCS3* on chromosome 10 was only associated with GAD_{anx-dep} GWAS in the full meta-analysis (**Table 3**). No gene sets were significantly associated with GAD ($P \leq 4.68 \times 10^{-6}$).

Table 3: Gene-level associations with anxiety disorder phenotypes.

Anxiety disorder	Datasets	Control screening criteria	Gene symbol	CHR	N SNPS	Z	P	P Bonf. Threshold (0.05/ N genes)	Previous report in GWAS catalog	Gene description
Fear	QIMR, GLAD+ & UKB	Anx-dep	SNX29	16	1560	5.09	1.77E-07	2.702E-6 (18502 genes)	EA, insomnia, smoking, IQ	Sorting nexin 29
			YWHAG	7	46	4.67	1.49E-06		Blood protein levels. Suggestive significance: SCZ, multiple sclerosis	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein gamma
		Specific	SNX29	16	1560	4.97	3.31E-07	2.702E-6 (18502)	As above	As above
			YWHAG	7	46	4.64	1.75E-06		As above	As above
	GLAD+ & UKB	Anx-dep	SNX29	16	1561	5.17	1.19E-07	2.698E-6 (18531)	As above	As above
			ERI3	1	159	5.05	2.23E-07		EA, alcohol consumption	ERI1 exoribonuclease family member 3
		Specific	SNX29	16	1562	5.07	1.96E-07	2.698E-6 (18531)	As above	As above
			ERI3	1	159	4.62	1.94E-06		As above	As above
GAD	QIMR, GLAD+ & UKB	Anx-dep	TMEM106B	7	125	4.98	3.22E-07	2.702E-6 (18503)	Depression, EA, wellbeing, insomnia, Alzheimer's, neuroticism, GAD, PTSD, HDL	Transmembrane protein 106B
			SORCS3	10	1422	4.92	4.33E-07		EA, depression, blood pressure, externalising behaviours, CAD, insomnia, neuroticism, smoking, ADHD	Sortilin related VPS10 domain containing receptor 3
			TRIM31	6	45	4.81	7.74E-07		Cholesterol, depression, insomnia, IQ	Tripartite motif containing 31
		Specific	TMEM106B	7	125	4.71	1.27E-06	2.702E-6 (18503)	As above	As above
			TRIM31	6	45	4.58	2.31E-06		As above	
	GLAD+ & UKB	Anx-dep	TMEM106B	7	125	4.58	2.27E-06	2.698E-6 (18529)	As above	As above
		Specific	-	-	-	-	-	2.698E-6 (18530)	-	-

Gene-level association analyses conducted in MAGMA. CHR = chromosome; Previous report = GWASs of other phenotypes that report genome-wide significant loci that map onto the gene-level associations identified in this study (**Supplementary Tables 6 & 7**); EA = educational attainment; IQ = general cognitive ability; SCZ = schizophrenia; GAD = generalised anxiety disorder; PTSD = post-traumatic stress disorder; HDL = high-density lipoprotein cholesterol; CAD = coronary artery disease; ADHD = Attention deficit hyperactivity disorder.

SNP-based heritability

We report results here for the most powerful GWAS in the full meta-analyses, as measured by GWAS mean X^2 , and present results for other GWAS in the supplementary. We identified modest liability-scale LDSC h^2_{SNP} for $\text{Fear}_{\text{anx-dep}}$ of 9.6% (SE = 0.006, mean $X^2 = 1.17$) and 10% for $\text{GAD}_{\text{anx-dep}}$ (SE = 0.006, mean $X^2 = 1.25$), assuming a population prevalence of 8.8% and 12.9%, respectively. Population prevalence was based on COPING study sample prevalence (**Supplementary Table 3**). We note that this is not a population-based study, and a proportion of participants were recruited through an inflammatory bowel disease study⁴⁵. However, COPING study anxiety disorder prevalence estimates were within the range of epidemiological studies, which report broad estimates for panic disorder (1-5%), agoraphobia (1-3%), social anxiety disorder (3-13%), specific phobia (8-14%), and GAD (3-13%)^{46,47}. To our knowledge, no previous epidemiological studies have estimated the population prevalence of fear-based disorders as a group. **Supplementary Table 8** shows h^2_{SNP} liability-scale estimates calculated from prevalences 10% higher and lower than COPING study sample prevalences. **Supplementary Table 8** also reports h^2_{SNP} liability-scale estimates for phenotypes with specifically screened controls, and in the meta-analysis excluding QIMR.

Genetic correlation between fear-based disorders and GAD

The genetic correlation between fear and GAD was high ($r_g = 0.85\text{--}0.96$) and significantly different from 0 when analysing both GWAS phenotypes ($\text{fear}_{\text{specific}}\text{--}\text{GAD}_{\text{specific}}$ and $\text{fear}_{\text{anx-dep}}\text{--}\text{GAD}_{\text{anx-dep}}$) in both GWAS meta-analyses datasets (*full meta-analysis* and *GLAD+ and UKB meta-analysis* excluding depression-enriched QIMR dataset). The genetic correlation between fear and GAD was significantly less than 1 when controls were screened specifically and when the QIMR dataset was excluded from the meta-analysis. All other fear–GAD genetic correlations were not significantly different from 1 (**Table 4**; Bonferroni-corrected significance threshold $P \leq 0.0125$).

Table 4: Genetic correlations (r_g) between fear-based disorder and GAD phenotypes

Datasets	Control criteria	r_g	se	Z	P (0)	P (1)
GLAD+ & UKB	Specific	0.85	0.05	15.63	4.90E-55	4.57E-03
	Any anx-dep	0.92	0.03	27.52	9.65E-167	0.0215
QIMR, GLAD+ & UKB	Specific	0.95	0.03	28.77	4.77E-182	0.1282
	Any anx-dep	0.97	0.02	47.03	0.00E+00	0.1808

Genetic correlations estimated using LDSC regression. P (0) - P-value for test of genetic correlation different from 0. P (1) - P-value for test of genetic correlation different from 1.

Genetic correlations with external phenotypes

We calculated genetic correlations between fear and GAD with 345 external phenotypes. Only $\text{fear}_{\text{specific}}$ and $\text{GAD}_{\text{specific}}$ in the GWAS meta-analysis excluding the depression-enriched QIMR dataset were tested for genetic correlations with external phenotypes (as the fear–GAD genetic correlation was significantly different from 1). There was no statistically significant difference in genetic correlations with most external traits across $\text{fear}_{\text{specific}}$ and $\text{GAD}_{\text{specific}}$ (Bonferroni-corrected significance threshold $P \leq 1.45 \times 10^{-4}$; **Figure 2; left panel**). Exceptions were observed for educational attainment, general cognitive ability, which showed significantly stronger negative genetic correlations with $\text{fear}_{\text{specific}}$ than with $\text{GAD}_{\text{specific}}$. Coronary artery disease also showed a significantly higher positive genetic correlation with $\text{fear}_{\text{specific}}$ than with $\text{GAD}_{\text{specific}}$. Conversely, a diagnosis of any bipolar disorder (type I and type II), bipolar disorder type I, anorexia nervosa, and neuroticism showed significantly stronger positive genetic correlations with $\text{GAD}_{\text{specific}}$ than with $\text{fear}_{\text{specific}}$ (**Figure 2; right panel**). Full genetic correlation results with external phenotypes are shown in **Supplementary Table 9**.

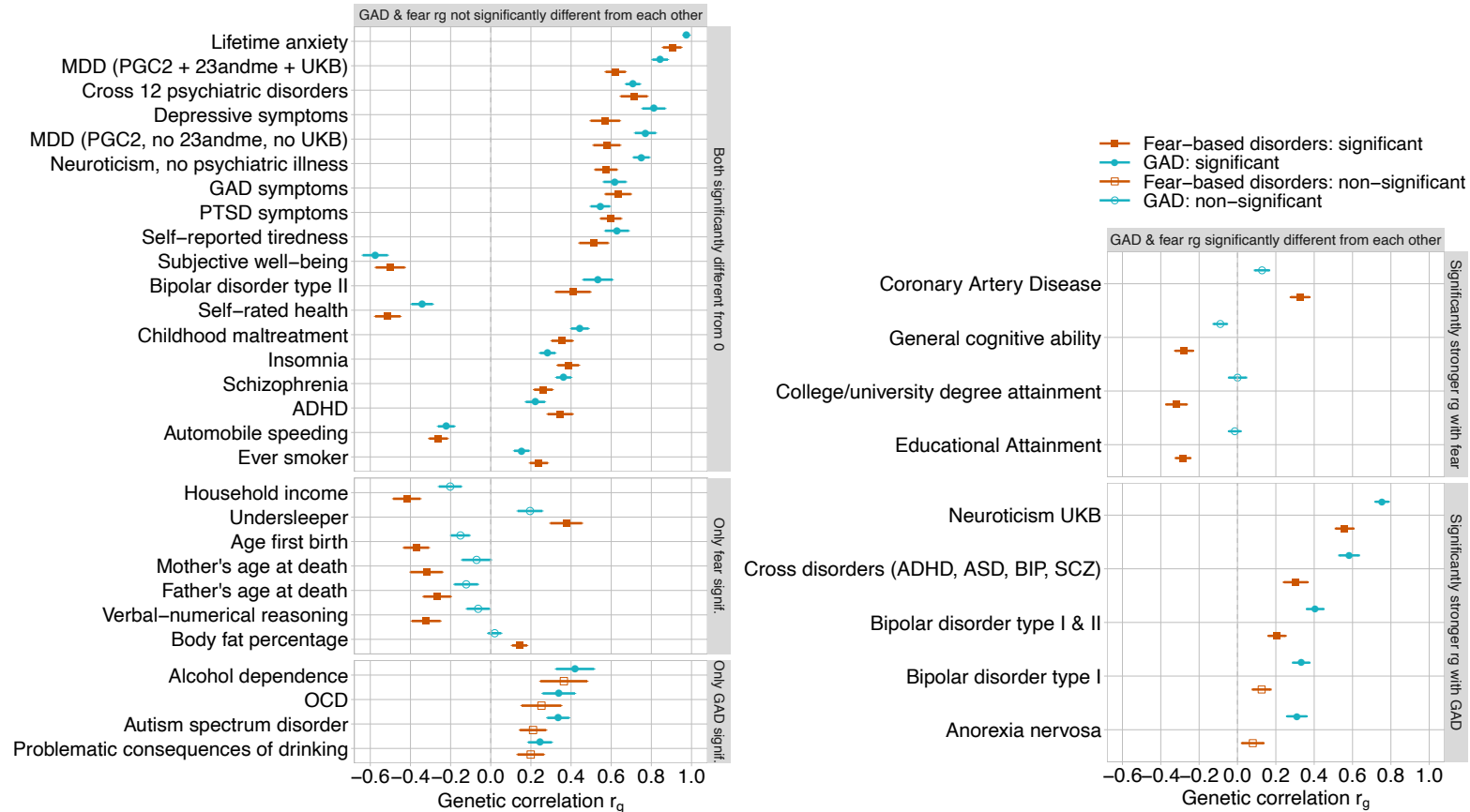


Figure 2: Genetic correlations between anxiety disorder phenotypes and external traits estimated in LDSC regression.

Fear-based disorder GWAS and GAD GWAS results from phenotypes with specifically screened controls, meta-analysed across the UKB and GLAD+ datasets. Genetic correlations with 345 phenotypes were tested (Bonferroni-corrected significance threshold $P \leq 1.45 \times 10^{-4}$). Only traits with genetic correlation P -values significantly different from 0 with at least one of the anxiety disorder phenotypes are shown (after Bonferroni correction for multiple testing). Full results are shown in Supplementary Table 9. External phenotypes were tested for significantly different genetic correlations with fear-based disorders versus GAD using a block jackknife (Bonferroni correction for significance [$P \leq 1.45 \times 10^{-4}$]). Left panel: phenotypes with no statistically significant different genetic correlations between fear-based disorders and GAD, though in the lower two segments, there was a significant association with one trait and not the other. Right panel: phenotypes with statistically significant different genetic correlations between fear-based disorders and GAD. Bars represent standard errors. ADHD = attention deficit hyperactivity disorder; ASD = autism spectrum disorder; BIP = bipolar disorder; GAD = generalised anxiety disorder; MDD = major depressive disorder; OCD = obsessive-compulsive disorder; PGC = Psychiatric Genomics Consortium; PTSD = post-traumatic stress disorder; SCZ = schizophrenia; UKB = UK Biobank.

Discussion

This is the first study to report genome-wide significant loci (three in total) associated with any lifetime fear-based disorder diagnosis, hereafter referred to as fear. In addition, we identified four genome-wide significant loci for lifetime GAD diagnosis, previously identified in other GWAS of anxiety^{15,48} and distress-related traits^{49–51}. The increase in power compared to previous GWAS of fear¹⁹ was achieved by combining detailed and brief diagnostic measures. In agreement with twin studies, we found that fear and GAD were highly genetically correlated with each other and with depression^{9,11,12}. When we excluded the depression-enriched dataset and only screened controls for the specific anxiety disorder being analysed, genetic correlations between these disorders were reduced, suggesting some genetic specificity between them. Differences in common genetic variant-based genetic correlations between fear and GAD with broad domains of other complex traits were explored. Compared with GAD, fear showed significantly stronger negative genetic correlations with general cognitive ability and educational attainment and stronger positive genetic correlations with coronary artery disease. Conversely, neuroticism, bipolar disorder type I and anorexia nervosa were among the traits with significantly higher positive genetic correlations with GAD than fear.

We found independent loci and gene-level associations that support previous GWAS findings of other phenotypes and some novel associations. We identified a novel variant associated with fear that has not been previously reported in GWASs of complex traits. The nearest gene to the intergenic variant is the uncharacterised, non-coding gene *CTB-57H20.1*. Other loci and gene-level associations with fear that we identified have been previously implicated in GWAS of educational attainment⁵², general cognitive ability⁵³, insomnia⁵⁴ and broad depression⁴⁹, including *sorting nexin 29*. The sorting nexin family of proteins is thought to play a role in neuronal function, synaptic plasticity, learning, and memory⁵⁵. One locus associated with fear was previously reported in a worry-neuroticism GWAS in the UK Biobank⁵⁰. No fear loci or gene-level associations overlapped with our GAD results. As GWASs of specific anxiety disorders improve in power, we expect this overlap to increase.

Some of our loci and gene-level associations with GAD were reported in previous UK Biobank GWASs of lifetime anxiety and GAD symptoms^{15,48}, which is expected given the sample overlap. Loci and gene-level associations with GAD were consistently replicated in previous GWASs of other distress-related phenotypes, with all four loci previously identified in depression and neuroticism^{49–51}. This includes the *SORCS3* gene, which is highly expressed in the CA1 hippocampus region and has been implicated in aversive memory extinction in mice⁵⁶. *SORCS3* was also associated with anxiety-like behaviour in mice, mediated by changes in gut microbiome measures suggesting the role of the gut-brain axis⁵⁷. Our finding

warrants further animal modelling and functional genomic interrogation of the role of *SORCS3* in distress liability and translational potential.

The high genetic correlations we observed between fear, GAD, and depression are consistent with previous twin studies^{9,11,12}. Our results are compatible with our previous preliminary GWAS in the UK Biobank¹⁹ but add further detail to the genetic structure of fear and GAD by assessing a broad range of other complex traits. We verified that screening controls specifically for the disorder grouping being analysed improved our ability to distinguish genetic differences between highly genetically correlated disorders (e.g. controls in the fear GWAS were only screened for any fear-based disorder). We found evidence of some genetic differences between fear and GAD once we screened controls specifically and after excluding the depression-ascertained QIMR dataset. This highlights the necessity for careful consideration of how to screen controls when attempting to identify disorder-specific genetics.

Genetic correlations with depression were high and not significantly different between fear and GAD, supporting a common genetic liability factor shared among these disorders^{9,11,12,58}. A common liability factor is often conceptualised as being driven by shared negative affectivity, which is also well captured by measures of the personality trait neuroticism^{22,59}. We found neuroticism had a significantly stronger genetic correlation with GAD than fear. This is in line with twin studies that reported genetic influences on GAD (and depression) were more core to a higher-order dimension of genetic liability to negative affectivity than fear⁶⁰. General negative affectivity also influences fear-based disorders but is less central to them; instead, they are characterised by more specific elements of acute fearful and physiological hyperarousal⁶¹. Our findings partially support a distress-fear genetic distinction. However, we found stronger evidence for an overarching genetic liability to internalising that drives comorbidity across anxiety disorders and depression.

Compared with GAD, fear had stronger negative genetic correlations with educational attainment and general cognitive ability and a higher positive genetic correlation with coronary artery disease (heart disease). These correlations were non-significant with GAD. The specific genetic relationship observed between fear and cognitive-related traits may reflect shared cognitive mechanisms, such as those involved in learning, that are less central to GAD. Associative learning is a key mechanism that distinguishes fear-based disorders from GAD⁶. Our finding that heart disease is uniquely genetically correlated with fear may be explained by genetics playing a bigger part in driving comorbidity between heart disease and fear than with GAD. Core symptoms of fear-based disorders, including experiencing phobic fears and somatic arousal, are linked to a higher risk of heart disease and cardiac fatality^{62–65}. Our findings align with a phenotypic study that revealed fear-based disorders were stronger

predictors of heart disease development than distress disorders⁶⁶. The earlier age of onset of chronically experienced fear-based disorders compared to distress may give rise to longer-term exposure to mechanisms associated with heart disease risk^{1,66,67}. Overlapping genetics with heart disease may underpin shared physiological mechanisms in inflammatory pathways or autonomic dysfunction⁶⁸.

GAD showed significantly stronger positive genetic correlations than fear with a diagnosis of any bipolar disorder (type I or II). Bipolar type II disorder involves episodes of hypomania and is more similar to depression, whereas bipolar type I disorder is characterised by more severe manic episodes⁴. When we analysed bipolar disorder I and II separately, we found GAD and fear had no significantly different genetic correlation with bipolar II, whereas bipolar I was significantly more positively genetically correlated with GAD. This indicates that bipolar II drives down the genetic correlations with fear when analysed together with bipolar I. This aligns with previous studies that found bipolar I had a lower genetic correlation with depression than bipolar II and was more genetically similar to rarer psychiatric illnesses, including schizophrenia^{69,70}. Genetic correlation analyses of the mood disorder spectrum reflect a genetic continuum from manic to depressive clusters. Bipolar I sits on one end of the spectrum, and depressive disorders on the other, with bipolar type II connecting the two disorders⁷¹. Our novel finding adds further fine-grained detail to this genetic structure by highlighting heterogeneity at the internalising pole of the genetic continuum. On the opposite end of the spectrum to bipolar type I lies fear-based disorders, with distress disorders (GAD and depression) sitting between bipolar and fear-based disorders.

We found that fear had a significantly weaker genetic correlation with anorexia nervosa (~0.10) than GAD (~0.30). This is somewhat unexpected given that fear-based disorders have similar rates of comorbidity with anorexia nervosa compared to GAD, thought to be driven by an overlap in genetic susceptibility and fear-conditioning mechanisms⁷²⁻⁷⁴. Few studies have assessed genetic correlations specifically between anorexia nervosa and the individual anxiety disorder subtypes, with analyses restricted to GAD⁷⁵. However, the genetic overlap between other eating disorder phenotypes and fear-based disorders has been reported^{76,77}. One twin study found that a broad eating disorder phenotype had moderate genetic correlations with panic disorder (0.62) and GAD (0.51) and a slightly lower genetic correlation with specific phobias (0.40), agoraphobia (0.38), and social anxiety disorder (0.33)⁷⁷. Thus, further research is required to establish whether genetic correlations with anorexia nervosa differ between individual fear-based disorders. There may also be shared loci that affect traits in different directions, making us unable to detect a *global* genetic correlation. Studies conducting *local* genetic correlation analyses may identify specific shared regions of the

genome supporting shared mechanisms that link fear-based disorders with anorexia nervosa^{78,79}.

Our study is the first to identify fear-based disorder-associated genetic loci. However, findings should be considered in light of limitations that impacted our ability to detect fear-specific genetic influences. First, we were underpowered to assess the fear-based disorders individually. Some twin studies suggest the proportion of disorder-specific genetic influences on individual fear-based disorders varies^{11,12}, with specific phobia being less genetically similar to GAD than panic disorder and agoraphobia¹¹. Second, results may have differed if we had sufficient sample sizes to only use panic disorder diagnosis instead of combining it with panic attack measures. Panic attacks are transdiagnostic and can be included as a specifier for a range of DSM-5 psychiatric disorders⁴. Using self-reports of panic attacks may have elevated the genetic correlations between fear and other psychiatric disorders. However, epidemiological estimates of lifetime panic disorder comorbidity with other mental disorders are also high (80%)⁸⁰. Third, specifically screening controls for a certain anxiety disorder was not possible in the QIMR data as we were limited to using a single-item brief measure of any anxiety disorder. Fourth, although a strength of our study was incorporating detailed measures of anxiety disorders, we could not compare genetic differences between brief versus detailed diagnostic measures. Studies suggest that more detailed measures increase genetic specificity²⁵. Combining detailed with brief measures may have limited genetic specificity and elevated the genetic correlation between fear, GAD, and related disorders such as depression.

Finally, we could not account for the high comorbidity observed across anxiety disorders and depression. We previously found that cases rarely met criteria for only one anxiety disorder and no comorbid MDD (5%) in the GLAD+ dataset⁸¹. Thus, ascertaining data for well-powered GWAS of a single anxiety disorder without comorbidities will be challenging and are less representative of the broader lived experience of these disorders. We did not include heritable comorbid traits as GWAS covariates as this can bias results⁸². Future studies with well-powered GWAS of anxiety subtypes could discern subtype-specific genetics through multivariate genomic structural equation modelling⁸³. As individual-level methods advance^{84,85}, this will enable testing genetic covariances across comorbid disorders and comparing measures in smaller samples than is required for GWAS summary-level modelling.

In summary, our results add further evidence that fear-based disorders and GAD share a high proportion of genetic liability, as well as with depression, while also revealing some genetic differences between them. We identified differences in the genetic relationship between fear and GAD, including those with neuroticism, bipolar type I, anorexia nervosa, general cognitive ability, and coronary artery disease. Our findings add more fine-grained detail to the proposed

hierarchical structure of internalising disorders and partially support fear-based disorders and GAD as separate subfactors of distress and fear. Quantitatively-based dimensional modelling of shared and distinct distress and fear symptoms on the genomic level would further complete this hierarchical structure²². The growth of datasets with detailed phenotyping of all anxiety disorders, such as those used in this study, will be key for further identifying subtype-specific and transdiagnostic genetic factors of the full anxiety disorder spectrum.

References

1. Kessler, R. C. *et al.* Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593–602 (2005).
2. Somers, J. M., Goldner, E. M., Waraich, P. & Hsu, L. Prevalence and incidence studies of anxiety disorders: a systematic review of the literature. *Can. J. Psychiatry* **51**, 100–113 (2006).
3. Craske, M. G. *et al.* Anxiety disorders. *Nat Rev Dis Primers* **3**, 17024 (2017).
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
5. Kessler, R. C. The global burden of anxiety and mood disorders: putting the European Study of the Epidemiology of Mental Disorders (ESEMeD) findings into perspective. *J. Clin. Psychiatry* **68 Suppl 2**, 10–19 (2007).
6. Craske, M. G. & Waters, A. M. Panic Disorder, Phobias, and Generalized Anxiety Disorder. *Annual Review of Clinical Psychology* vol. 1 197–225 Preprint at <https://doi.org/10.1146/annurev.clinpsy.1.102803.143857> (2005).
7. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
8. Ask, H. *et al.* Genetic contributions to anxiety disorders: where we are and where we are heading. *Psychol. Med.* **51**, 2231–2246 (2021).
9. Waszczuk, M. A., Zavos, H. M. S., Gregory, A. M. & Eley, T. C. The phenotypic and genetic structure of depression and anxiety disorder symptoms in childhood, adolescence, and young adulthood. *JAMA Psychiatry* **71**, 905–916 (2014).
10. Roberson-Nay, R., Eaves, L. J., Hettema, J. M., Kendler, K. S. & Silberg, J. L. Childhood separation anxiety disorder and adult onset panic attacks share a common genetic diathesis. *Depress. Anxiety* **29**, 320–327 (2012).
11. Hettema, J. M., Prescott, C. A., Myers, J. M., Neale, M. C. & Kendler, K. S. The structure of genetic and environmental risk factors for anxiety disorders in men and women. *Arch. Gen. Psychiatry* **62**, 182–189 (2005).
12. Tambs, K. *et al.* Structure of genetic and environmental risk factors for dimensional representations of DSM–IV anxiety disorders. *Br. J. Psychiatry* **195**, 301–307 (2009).
13. Kendler, K. S., Prescott, C. A., Myers, J. & Neale, M. C. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Arch. Gen. Psychiatry* **60**, 929–937 (2003).
14. Smoller, J. W. Anxiety Genetics Goes Genomic. *The American journal of psychiatry* vol. 177 190–194 (2020).
15. Purves, K. L. *et al.* A major role for common genetic variation in anxiety disorders. *Mol. Psychiatry* **25**, 3292–3303 (2020).
16. Levey, D. F. *et al.* Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *Am. J. Psychiatry* **177**, 223–232 (2020).
17. Meier, S. M. *et al.* Genetic Variants Associated With Anxiety and Stress-Related Disorders: A Genome-Wide Association Study and Mouse-Model Study. *JAMA Psychiatry* (2019) doi:10.1001/jamapsychiatry.2019.1119.
18. Wendt, F. R. *et al.* Using phenotype risk scores to enhance gene discovery for generalized anxiety disorder and posttraumatic stress disorder. *Mol. Psychiatry* (2022) doi:10.1038/s41380-022-01469-y.
19. Morneau-Vaillancourt, G. *et al.* The genetic and environmental hierarchical structure of anxiety and depression in the UK Biobank. *Depress. Anxiety* **37**, 512–520 (2020).
20. Deckert, J. *et al.* GLRB allelic variation associated with agoraphobic cognitions, increased startle response and fear network activation: a potential neurogenetic pathway to panic disorder. *Mol. Psychiatry* **22**, 1431–1439 (2017).

21. Erhardt, A. *et al.* TMEM132D, a new candidate for anxiety phenotypes: evidence from human and mouse studies. *Mol. Psychiatry* **16**, 647–663 (2011).
22. Watson, D. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry* **21**, 26–54 (2022).
23. Toft, T. *et al.* Mental disorders in primary care: prevalence and co-morbidity among disorders. results from the functional illness in primary care (FIP) study. *Psychol. Med.* **35**, 1175–1184 (2005).
24. Kendler, K. S., Chatzinakos, C. & Bacanu, S.-A. The impact on estimations of genetic correlations by the use of super-normal, unscreened, and family-history screened controls in genome wide case-control studies. *Genet. Epidemiol.* **44**, 283–289 (2020).
25. Cai, N., Choi, K. W. & Fried, E. I. Reviewing the genetics of heterogeneity in depression: operationalizations, manifestations and etiologies. *Hum. Mol. Genet.* **29**, R10–R18 (2020).
26. Flint, J. The genetic basis of major depressive disorder. *Mol. Psychiatry* (2023) doi:10.1038/s41380-023-01957-9.
27. Davies, M. R. *et al.* Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts. *J. Anxiety Disord.* **85**, 102491 (2022).
28. Mitchell, B. L. *et al.* The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes. *Biol. Psychiatry* (2021) doi:10.1016/j.biopsych.2021.10.021.
29. Davis, K. A. S. *et al.* Mental health in UK Biobank - development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open* **6**, e18 (2020).
30. Gigantesco, A. & Morosini, P. Development, reliability and factor analysis of a self-administered questionnaire which originates from the World Health Organization's Composite International Diagnostic Interview – Short Form (CIDI-SF) for assessing mental disorders. *Clinical Practice and Epidemiology in Mental Health* vol. 4 8 Preprint at <https://doi.org/10.1186/1745-0179-4-8> (2008).
31. Byrne, E. M. *et al.* Cohort profile: the Australian genetics of depression study. *BMJ Open* **10**, e032580 (2020).
32. Davies, M. R. *et al.* The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety. *Behav. Res. Ther.* **123**, 103503 (2019).
33. Young, K. S. *et al.* Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK. *Psychol. Med.* 1–14 (2022).
34. Olsen, C. M. *et al.* Cohort profile: the QSkin Sun and Health Study. *Int. J. Epidemiol.* **41**, 929–929i (2012).
35. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
36. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* **53**, 1097–1103 (2021).
37. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).
38. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
39. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
40. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).

41. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* vol. 47 291–295 Preprint at <https://doi.org/10.1038/ng.3211> (2015).
42. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
43. Hübel, C. *et al.* Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **180**, 428–438 (2019).
44. Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. *Nat. Genet.* **50**, 906–908 (2018).
45. Davies, H. L. *et al.* 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. *Int. J. Eat. Disord.* (2022).
46. Bandelow, B. & Michaelis, S. Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin. Neurosci.* **17**, 327–335 (2015).
47. Leray, E. *et al.* Prevalence, characteristics and comorbidities of anxiety disorders in France: Results from the ‘Mental Health in General Population’ Survey (MHGP). *Eur. Psychiatry* **26**, 339–345 (2011).
48. Thorp, J. G. *et al.* Symptom-level modelling unravels the shared genetic architecture of anxiety and depression. *Nat Hum Behav* **5**, 1432–1442 (2021).
49. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019).
50. Nagel, M. *et al.* Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat. Genet.* **50**, 920–927 (2018).
51. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
52. Okbay, A. *et al.* Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat. Genet.* **54**, 437–449 (2022).
53. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat. Commun.* **9**, 2098 (2018).
54. Watanabe, K. *et al.* Genome-wide meta-analysis of insomnia prioritizes genes associated with metabolic and psychiatric pathways. *Nat. Genet.* **54**, 1125–1132 (2022).
55. Vieira, N., Rito, T., Correia-Neves, M. & Sousa, N. Sorting Out Sorting Nexins Functions in the Nervous System in Health and Disease. *Mol. Neurobiol.* **58**, 4070–4106 (2021).
56. Breiderhoff, T. *et al.* Sortilin-related receptor SORCS3 is a postsynaptic modulator of synaptic depression and fear extinction. *PLoS One* **8**, e75006 (2013).
57. Jin, X. *et al.* Gut microbiome partially mediates and coordinates the effects of genetics on anxiety-like behavior in Collaborative Cross mice. *Sci. Rep.* **11**, 270 (2021).
58. Mosing, M. A. *et al.* Genetic and environmental influences on the co-morbidity between depression, panic disorder, agoraphobia, and social phobia: a twin study. *Depress. Anxiety* **26**, 1004–1011 (2009).
59. Hettema, J. M., Neale, M. C., Myers, J. M., Prescott, C. A. & Kendler, K. S. A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am. J. Psychiatry* **163**, 857–864 (2006).
60. Kendler, K. S. & Myers, J. The boundaries of the internalizing and externalizing genetic spectra in men and women. *Psychol. Med.* **44**, 647–655 (2014).
61. Watson, D. Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *J. Abnorm. Psychol.* **114**, 522–536 (2005).

62. Watkins, L. L. *et al.* Phobic anxiety and increased risk of mortality in coronary heart disease. *Psychosom. Med.* **72**, 664–671 (2010).
63. Albert, C. M., Chae, C. U., Rexrode, K. M., Manson, J. E. & Kawachi, I. Phobic anxiety and risk of coronary heart disease and sudden cardiac death among women. *Circulation* **111**, 480–487 (2005).
64. Nabi, H. *et al.* Psychological and somatic symptoms of anxiety and risk of coronary heart disease: the health and social support prospective cohort study. *Biol. Psychiatry* **67**, 378–385 (2010).
65. Roest, A. M., Heideveld, A., Martens, E. J., de Jonge, P. & Denollet, J. Symptom dimensions of anxiety following myocardial infarction: associations with depressive symptoms and prognosis. *Health Psychol.* **33**, 1468–1476 (2014).
66. Roest, A. M. *et al.* Fear and distress disorders as predictors of heart disease: A temporal perspective. *J. Psychosom. Res.* **96**, 67–75 (2017).
67. Lijster, J. M. de *et al.* The Age of Onset of Anxiety Disorders. *Can. J. Psychiatry* **62**, 237–246 (2017).
68. Celano, C. M., Daunis, D. J., Lokko, H. N., Campbell, K. A. & Huffman, J. C. Anxiety Disorders and Cardiovascular Disease. *Curr. Psychiatry Rep.* **18**, 101 (2016).
69. Mallard, T. T. *et al.* Multivariate GWAS of psychiatric disorders and their cardinal symptoms reveal two dimensions of cross-cutting genetic liabilities. *Cell Genom* **2**, (2022).
70. Mullins, N. *et al.* Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021).
71. Coleman, J. R. I. *et al.* The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biol. Psychiatry* (2019) doi:10.1016/j.biopsych.2019.10.015.
72. Meier, S. M. *et al.* Diagnosed Anxiety Disorders and the Risk of Subsequent Anorexia Nervosa: A Danish Population Register Study. *Eur. Eat. Disord. Rev.* **23**, 524–530 (2015).
73. Kaye, W. H., Bulik, C. M., Thornton, L., Barbarich, N. & Masters, K. Comorbidity of anxiety disorders with anorexia and bulimia nervosa. *Am. J. Psychiatry* **161**, 2215–2221 (2004).
74. Godart, N. T., Flament, M. F., Perdereau, F. & Jeammet, P. Comorbidity between eating disorders and anxiety disorders: a review. *Int. J. Eat. Disord.* **32**, 253–270 (2002).
75. Dellava, J. E., Kendler, K. S. & Neale, M. C. Generalized anxiety disorder and anorexia nervosa: evidence of shared genetic variation. *Depress. Anxiety* **28**, 728–733 (2011).
76. Kendler, K. S. *et al.* The structure of the genetic and environmental risk factors for six major psychiatric disorders in women. Phobia, generalized anxiety disorder, panic disorder, bulimia, major depression, and alcoholism. *Arch. Gen. Psychiatry* **52**, 374–383 (1995).
77. Kendler, K. S. *et al.* The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *Am. J. Psychiatry* **168**, 29–39 (2011).
78. Werme, J., van der Sluis, S. & de Leeuw, C. A. LAVA: An integrated framework for local genetic correlation analysis. *bioRxiv* 2020.12.31.424652 (2021) doi:10.1101/2020.12.31.424652.
79. Ronald, A. & Pain, O. A Revolution Is Brewing in How We Understand the Shared Genetic Causes of Psychiatric Disorders. *The American journal of psychiatry* vol. 179 791–793 (2022).
80. de Jonge, P., Roest, A. M., Lim, C. C. W., Levinson, D. & Scott, K. M. Panic Disorder and Panic Attacks. in *Mental Disorders Around the World: Facts and Figures from the WHO World Mental Health Surveys* 93–105 (Cambridge University Press, 2018).

81. Davies, M. R. *et al.* Factors associated with anxiety disorder comorbidity. *J. Affect. Disord.* **323**, 280–291 (2023).
82. Aschard, H., Vilhjálmsdóttir, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am. J. Hum. Genet.* **96**, 329–339 (2015).
83. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513–525 (2019).
84. de Hoyos, L. *et al.* Structural models of genome-wide covariance identify multiple common dimensions in autism. *bioRxiv* (2022) doi:10.1101/2022.10.21.22281213.
85. Pritikin, J. N., Neale, M. C., Prom-Wormley, E. C., Clark, S. L. & Verhulst, B. GW-SEM 2.0: Efficient, Flexible, and Accessible Multivariate GWAS. *Behav. Genet.* **51**, 343–357 (2021).

Chapter 3. Genetic decomposition of the heritable component of reported childhood maltreatment

This chapter is presented as a manuscript of an accepted paper in *Biological Psychiatry: Global Open Science*. It is an exact copy of this manuscript. Supplementary materials for this chapter, as detailed in the text, are included in **Appendix B**.

In press citation: **ter Kuile, A. R.**, Hübel, C., Cheesman, R., Coleman, J. R. I., Peel, A. J., Levey, D. F., Stein, M. B., Gelernter, J., Rayner, C., Eley, T. C., & Breen, G. (In press). Genetic decomposition of the heritable component of reported childhood maltreatment. *Biological Psychiatry Global Open Science*. <https://doi.org/10.1016/j.bpsgos.2023.03.003>

Authors: **Abigail R. ter Kuile**^{1,2}, Christopher Hübel^{1,2,3}, Rosa Cheesman⁸, Jonathan R. I. Coleman^{1,2}, Alicia J. Peel¹, Daniel F. Levey^{4,5}, Murray B. Stein^{6,7}, Joel Gelernter^{4,5}, Christopher Rayner¹, Thalia C. Eley^{1,2}, Gerome Breen^{1,2}

1. Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
2. National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust, London, UK
3. National Centre for Register-based Research, Aarhus Business and Social Sciences, Aarhus University, Aarhus, Denmark
4. Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA
5. Veterans Affairs Connecticut Healthcare System, West Haven, CT, USA
6. Department of Psychiatry and School of Public Health, University of California San Diego, La Jolla, CA, USA
7. VA San Diego Healthcare System, San Diego, CA, USA
8. PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway

Abstract

Background: Decades of research have shown that environmental exposures, including self-reports of trauma, are partly heritable. Heritable characteristics may influence exposure to and interpretations of environmental factors. Identifying heritable factors associated with self-reported trauma could improve our understanding of vulnerability to exposure and the interpretation of life events.

Methods: We used genome-wide association study summary statistics of childhood maltreatment, defined as reporting of abuse (emotional, sexual and physical) and neglect (emotional and physical) (N=185,414). We calculated genetic correlations (r_g) between reported childhood maltreatment and 576 traits to identify phenotypes that might explain the heritability of reported childhood maltreatment, retaining those with $|r_g| > 0.25$. We specified multiple regression models using genomic structural equation modelling to detect residual genetic variance in childhood maltreatment after accounting for genetically correlated traits.

Results: In two separate models, the shared genetic component of twelve health and behavioural traits and seven psychiatric disorders accounted for 59% and 56% of heritability due to common genetic variants (h^2_{SNP}) of childhood maltreatment, respectively. Genetic influences on the h^2_{SNP} of childhood maltreatment were generally accounted for by a shared genetic component across traits. The exceptions to this were general risk tolerance, subjective well-being, post-traumatic stress disorder and autism spectrum disorder, identified as independent contributors to its h^2_{SNP} . These four traits alone were sufficient to explain 58% of the h^2_{SNP} of childhood maltreatment.

Conclusions: We identified putative traits that reflect the h^2_{SNP} of childhood maltreatment. Elucidating the mechanisms underlying these associations may improve trauma prevention and posttraumatic intervention strategies.

Introduction

Traumatic events, namely those perceived as physically or emotionally threatening and violating, are associated with various adverse outcomes, including psychopathology¹⁻⁴. Decades of behavioural genetics research has shown that reported trauma exposures, like many environmental measures and behavioural traits, are partly heritable^{5,6}. Twin studies estimate 6-62% of the variance in reporting different types of trauma is attributable to genetics⁷⁻¹². Interpersonal assaultive traumas (e.g., physical and sexual assault) have higher heritability than non-interpersonal or non-assaultive traumas (e.g., accidents)^{7,9,13}. In relation to these observations, stressful life events (SLEs) dependent on one's behaviour (e.g., fights) are more heritable than those that are independent (e.g., natural disasters), with the latter occurring more often due to chance¹⁴. However, being at higher genetic risk for reported trauma does not signify that an individual is genetically predestined to experience trauma. Furthermore, a large proportion of the total phenotypic variability of reported trauma is not attributable to genetics. The environment itself may be harmful, or a perpetrator may exploit those in vulnerable circumstances^{11,15}. However, environmental risk factors are generally unstable, idiosyncratic, and thus, unpredictable and challenging to examine¹⁶. Exploring traits genetically related to reported trauma in different environmental contexts may provide a framework for social research to help determine trauma risk factors and protect vulnerable individuals⁶.

Heritable behavioural characteristics may contribute to the likelihood of experiencing certain events. Personality traits, such as openness to experience and antisocial behaviour, are phenotypically and genetically correlated with reporting interpersonal assaultive trauma¹⁷. Such partially heritable characteristics may contribute to the heritability of reported trauma through gene-environment correlation (rGE), whereby the environment reflects an individual's genetic propensities via three different processes¹⁸. Passive rGE occurs when a relative's genotype, such as parental genetic variation contributing to risk-taking behaviours, shapes the child's environment and potentially creates an unsafe home^{19,20}. The environment that the parent creates and the parental genotype are correlated as the child receives both from their biological parents. Thus, parental environmental effects may be captured in genetic analyses of offspring traits¹⁸. Evocative rGE arises when an individual's genotype shapes how others engage with them. For example, a child's behavioural difficulties may evoke verbal and physical discipline due to the carer's expectations of how a child should behave¹⁸. Active rGE involves an individual's genetic disposition to, for example, risk-taking modifying and selecting their environment^{18,21}, leading to differing risks of exposure to adverse environments.

Correlations between genetic factors and retrospective reports of trauma may also, in part, be driven by heritable characteristics influencing the subjective interpretation, willingness to disclose and recollection of events^{22,23}. Genetic research has largely relied on retrospective self-reports of trauma exposure which may be more susceptible to genetically influenced perceptions and recollection of events, as opposed to more objective measures prospectively recorded closer to the time of exposure (e.g., court records, caregiver reports)²³. Memory, emotional regulation and interpretation biases are partly heritable^{24–27} and are associated with retrospective reporting of trauma in early life²⁸. Individual differences in subjective experiences are partly influenced by genetics^{29,30}. Subjective appraisal of trauma is important for posttraumatic psychopathology, which is more strongly associated with retrospective self-reports of trauma than objective court records³¹. Individual, partially heritable differences in personality traits such as neuroticism and agreeableness may explain the discrepancy between retrospective and prospective measures of trauma^{31–33}. Furthermore, the consistency and frequency of self-reports are impacted by individual factors involved in the willingness to disclose a traumatic event, such as perception of stigma, fear of negative consequences, or pre-existing relationships with the perpetrator^{34–37}. Lack of disclosure is a barrier to therapeutic and legal interventions³⁶. Thus, a better understanding of the heritable factors that impact the retrospective report of trauma experiences could help improve posttraumatic support.

In sum, the influences on retrospectively reported trauma are complex and difficult to disentangle. A range of heritable traits may be involved. Heritability and genetic correlations between traits can be estimated using genome-wide association study (GWAS) summary statistics³⁸. The proportion of heritability explained by common genetic variants (h^2_{SNP}) ranges from 6–9% for reported interpersonal trauma during childhood^{21,39} to 18% during childhood and adulthood combined⁴⁰. This accounts for a large proportion of the reported twin heritabilities estimated at 20–62%^{7,9–12}. Reported trauma shows genetic correlations with psychiatric disorders, current mental state, personality traits, lifestyle factors, and sociodemographic traits^{21,39,40}. However, these studies did not analytically explain the extent to which the h^2_{SNP} of reported traumas reflects genetic correlations with these complex traits. Identifying specific traits that explain a large proportion of h^2_{SNP} can guide follow-up analyses in assessing certain characteristics involved in rGE and/or the subjective experience of trauma.

We hypothesised that genetic variants associated with relevant behavioural and cognitive traits would overlap with those associated with reported trauma, such that no residual genetic variance of reported trauma would remain after accounting for genetically correlated traits. Twin studies have used multivariate structural equation modelling (SEM) to examine the residual genetic variance of life event measures after accounting for genetically correlated

traits^{41,42}. To our knowledge, multivariate SEM has not been used to explore the extent to which specific heritable characteristics capture the heritability of reported trauma. Twin studies are limited in assessing only a moderate number of traits and environmental measures in the same individuals, which may be particularly challenging in the case of more severe environmental exposures such as trauma⁴³. In contrast, the multivariate SEM extension to GWAS summary statistics⁴⁴ allows the inclusion of many more traits measured in *different* individuals. Here, we decompose the h^2_{SNP} of reported trauma using genomic multiple regression with the Genomic SEM R package⁴⁴. Our primary aim was to measure the amount of residual genetic variance of reported trauma that remains after accounting for genetically correlated traits. Our secondary aim was to identify the traits contributing to the h^2_{SNP} of reported trauma from hundreds of complex traits that were systematically assessed.

Methods and Materials

Samples and measures

We used summary statistics from the largest published GWAS of reported trauma as of 2021, on childhood maltreatment²¹. This GWAS built on our previous work (Coleman *et al.*, 2020), extending it to assess childhood maltreatment specifically, and included 185,414 participants predominantly of European ancestry from five datasets: the UK Biobank⁴⁵, the Adolescent Brain Cognitive Development Study⁴⁶, the Avon Longitudinal Study of Parents and Children^{47,48}, Generation R^{49,50} and the Psychiatric Genomics Consortium³⁹. Childhood maltreatment was defined as reports of emotional, sexual, and physical abuse, and emotional and physical neglect. Most traumas (91.5%) were retrospectively self-reported ($N = 169,766$); however, a small proportion (8.5%) were reported prospectively by a parent or caregiver ($N = 15,651$). The genetic correlation between retrospective and prospective childhood maltreatment was previously reported as 0.72 (SE = 0.36; $P = 0.05$)²¹. Further methodological details can be found elsewhere²¹. In the original publication, the h^2_{SNP} of the continuous meta-analysed phenotype of childhood maltreatment was 0.08 (SE = 0.01) using linkage disequilibrium score (LDSC) regression²¹. We also analysed GWAS summary statistics from Coleman *et al.*⁴⁰ of a retrospectively reported lifetime trauma phenotype that more broadly captures trauma occurring in both childhood and adulthood in the UK Biobank^{40,45} (**Supplementary Methods**).

Bivariate genetic correlations

To identify traits associated with the genetic component of reported trauma, we used bivariate LDSC regression^{51,52} to measure the genetic correlations (r_g) between reported trauma and a wide range of complex traits. We tested 576 traits from GWAS summary statistics for an r_g with reported trauma. We excluded the major histocompatibility complex (MHC) region from our analyses^{51,53}. We considered traits for downstream analyses in Genomic SEM if they met the following criteria: a $|r_g|$ with reported trauma of > 0.25 and $|Z|$ statistic ≥ 5 , a GWAS mean X^2 value > 1.02 and an h^2_{SNP} Z statistic ≥ 5 . These thresholds were based on recommendations by the software developers^{54,55}. All traits that met these criteria were also statistically significant after Bonferroni correction for multiple testing ($\alpha = 0.05 / \text{the number of traits}$; $P \leq 8.68 \times 10^{-5}$), which was less stringent than our selected threshold $r_g |Z|$ statistic ≥ 5 (equivalent to $P \leq 5.73 \times 10^{-7}$). The criteria were stringent to restrict the number of traits included in downstream analyses to well-powered GWAS with potentially larger genetic contributions to the h^2_{SNP} of reported trauma.

Genomic structural equation modelling

To decompose the h^2_{SNP} of reported trauma, we used the Genomic SEM R package version 0.0.5 (<https://github.com/MichelNivard/GenomicSEM/wiki>)⁴⁴. Genomic SEM is a multivariate extension of LDSC that constructs covariance matrices from h^2_{SNP} and r_g calculated by LDSC. GWAS samples can overlap for Genomic SEM as the sampling covariance matrices adjust for potential sample overlap. All GWAS summary statistics in our analyses were based on individuals drawn from European ancestries.

We fitted fully saturated genomic multiple regression models. This approach simultaneously regressed the outcome (i.e., reported childhood maltreatment) on various explanatory variables, which modelled genetic correlations between each explanatory variable. This was informative in two ways. First, we estimated the residual genetic variance of reported trauma not explained by the genetics of the explanatory variables. Second, we estimated the unique contribution of each explanatory variable to the genetic component of reported trauma independent of other explanatory variables (i.e. conditional genetic association, termed b_g). We selected explanatory variables based on results from bivariate LDSC regression. We introduced one explanatory variable at a time, iteratively, from the most highly correlated trait with reported trauma to the least correlated trait. For all models, we used the default diagonally weighted least-squares estimator, in which the precision of genetic covariances (e.g. due to GWAS sample size) is considered.

We identified a practical limit of ≤ 11 explanatory variables in a fully saturated multiple regression model. Standard errors increased as more explanatory variables were added (> 11 explanatory variables resulted in standard errors > 1). Therefore, we had to fit two models for categorically distinct traits identified in LDSC regression analysis: model 1) health and behavioural traits and model 2) psychiatric disorders.

Results

Results from analysing reported childhood maltreatment were highly similar to that of reported lifetime trauma. We therefore present results for the more highly powered GWAS of childhood maltreatment (power quantified by h^2_{SNP} Z statistic = 18.7; mean X^2 value = 1.27) and report our findings for reported lifetime trauma (h^2_{SNP} Z statistic = 15.6; mean X^2 value = 1.22) in the **Supplementary Results** and **Supplementary Tables 7-11**.

Traits genetically correlated with reported childhood maltreatment

Figure 1 shows the 18 bivariate genetic correlations with childhood maltreatment ($|r_g| > 0.25$ and $r_g |Z|$ statistic ≥ 5). Genetic correlations with all 576 traits are summarised in **Supplementary Table 1**. After filtering for sufficiently powered GWAS, 18 traits were genetically correlated with childhood maltreatment with $|r_g| > 0.25$ and $r_g |Z|$ statistic ≥ 5 (**Figure 1; Supplementary Table 2**). The pairwise genetic correlations amongst the health and behavioural traits and psychiatric disorders modelled in Genomic SEM are shown in **Supplementary Tables 3L** and **4H**, respectively. The most well-powered GWAS was retained in cases where pairwise $|r_g|$ in each category was not significantly different from one (calculated using the chi-squared distribution function and $[(|r_g|-1)/\text{se}]^2$ in R v. 4.1.1)^{40,56}.

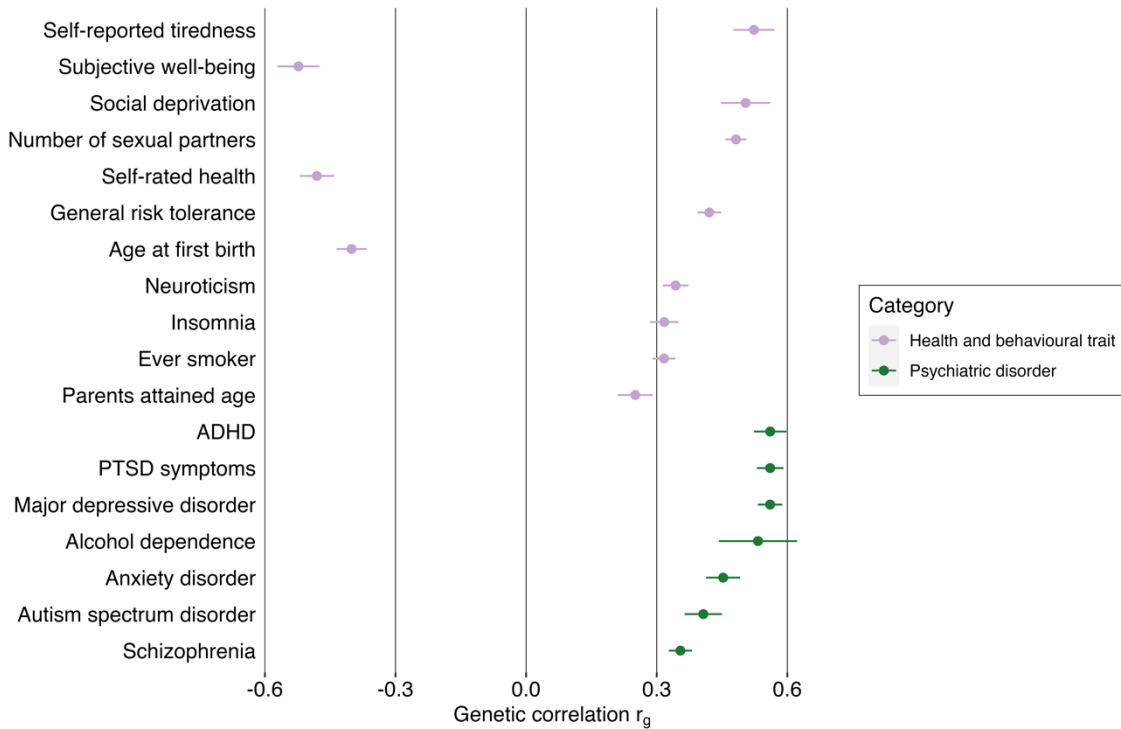


Figure 1. Top bivariate genetic correlations (r_g) between reported childhood maltreatment and various heritable traits.

We calculated the correlations using linkage disequilibrium score (LDSC) regression. We tested correlations with 576 traits. Only traits with an $|r_g| > 0.25$ and $r_g |Z|$ statistic ≥ 5 , sufficiently powered with a mean X^2 value > 1.02 and a common genetic variant-based heritability Z statistic ≥ 5 are shown. Bars represent standard errors. ADHD = attention deficit hyperactivity disorder; PTSD = post-traumatic stress disorder.

Genomic multiple regression

Health and behavioural traits

As noted above, due to being able to include a maximum of 11 explanatory variables at once, we specified separate models for health and behavioural traits and psychiatric disorders. The path diagram in **Figure 2A** shows the results for eleven health and behavioural traits simultaneously specified as explanatory variables of the genetic component of childhood maltreatment. Heritability can be defined as the proportion of variance between individuals for a given trait or disorder that is accounted for by genetic factors. The residual genetic variance is the amount of heritable variation of childhood maltreatment that is unexplained by the genetic factors of the explanatory variables in the model. The genetic influences on these health and behavioural traits explained 59% of the h^2_{SNP} of childhood maltreatment (one minus the residual genetic variance of 0.41 ± 0.07 ; $P = 7.76 \times 10^{-9}$).

Once controlling for the genetic influences of the other explanatory variables, the majority of conditional genetic associations between reported childhood maltreatment and health and behavioural traits were non-significant ($P > 0.05$; dashed lines in **Figure 2A & Supplementary Table 3K**). Therefore, most of the genetic correlations between childhood maltreatment and health and behavioural traits were shared with the other health and behavioural traits in the model. The shared genetic component across these traits may drive the direction of some bivariate genetic correlations, as several traits showed conditional associations with childhood maltreatment in the opposite direction (e.g. reported health $r_g = -0.48$; $b_g = 0.16$). However, two traits were uniquely genetically associated with childhood maltreatment, over and above their genetic correlations with the other traits: subjective well-being ($b_g = -0.47 \pm 0.18$; $P = 0.01$) and general risk tolerance (the reporting of how comfortable you are with taking risks; $b_g = 0.31 \pm 0.06$; $P = 9.96 \times 10^{-7}$). Models where each health trait was introduced iteratively are shown in **Supplementary Tables 3A-K**.

Psychiatric disorders

When taking into account the genetic influences of all nine psychiatric disorders in a genomic multiple regression model (**Figure 2B**), the residual genetic variance of reported childhood maltreatment was 0.44 ± 0.08 ($P = 9.23 \times 10^{-9}$). Thus, 56% (calculated as $1 - 0.44$) of the h^2_{SNP} of childhood maltreatment was explained by genetic components of these psychiatric disorders. In this model, most psychiatric disorders shared their genetic overlap with childhood maltreatment, as indicated by non-significant conditional associations after accounting for shared genetics ($P > 0.05$; **Figure 2B** and **Supplementary Table 4G**). Autism spectrum disorder (ASD) ($b_g = 0.25 \pm 0.11$, $P = 0.02$) and post-traumatic stress disorder (PTSD) symptoms ($b_g = 0.29 \pm 0.09$, $P = 1.40 \times 10^{-3}$) had significant independent associations with childhood maltreatment independent of other disorders. **Supplementary Tables 4A-G** show genomic multiple regression results when psychiatric disorders were included in iterative stages.

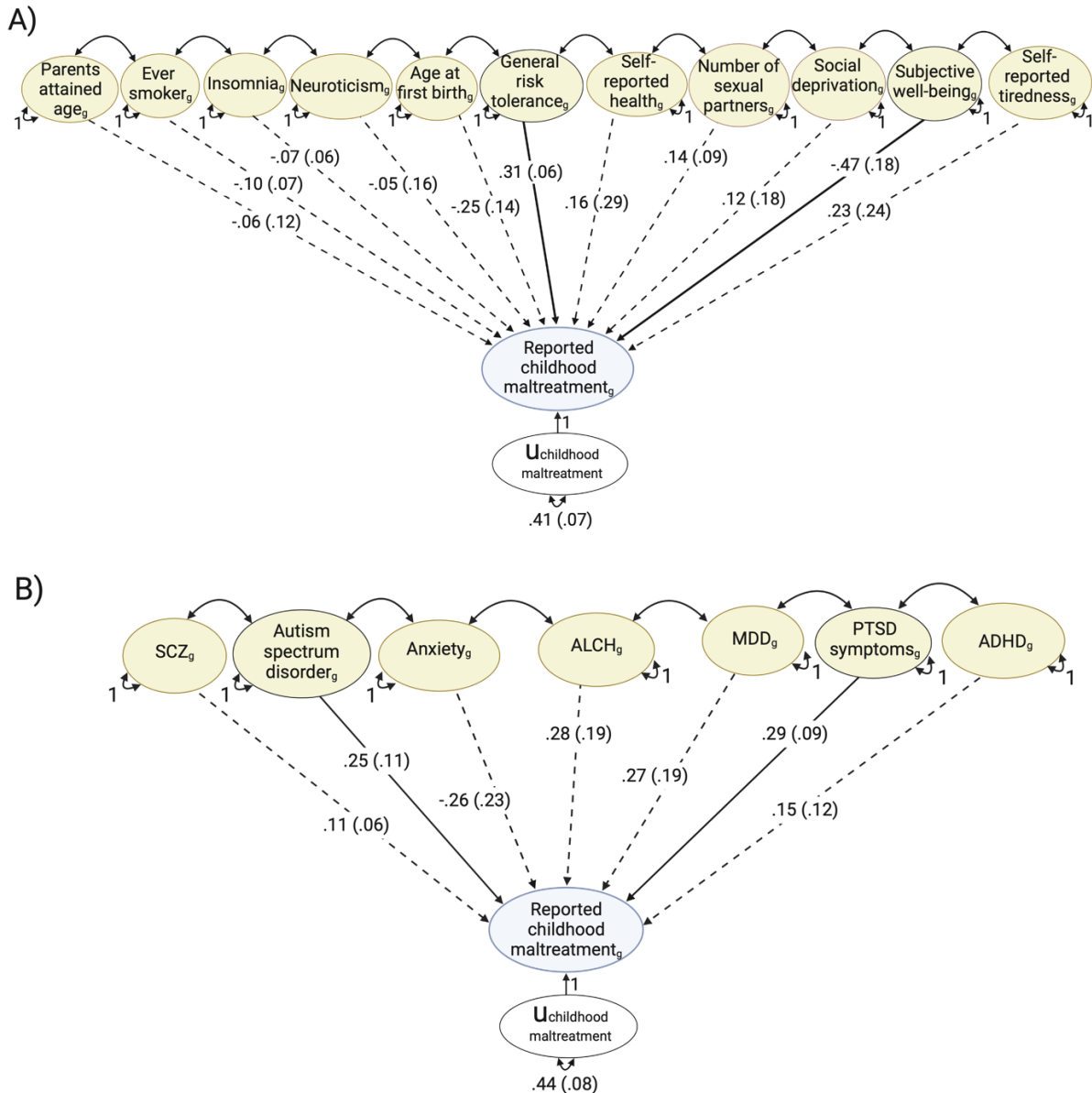


Figure 2. Path diagrams representing results from genomic multiple regression analyses.

Models were specified in Genomic SEM. The genetic component of A) eleven health and behavioural traits B) nine psychiatric disorders are the simultaneously fitted explanatory variables of the genetic component of reported childhood maltreatment using a weighted least-squares estimator. Single-headed arrows are conditional genetic associations ($b_g \pm SE$) between the explanatory variables and childhood maltreatment independent of the genetic influences of the other explanatory variables. A solid line indicates that the conditional genetic association is significant, and a dashed line indicates the conditional genetic association is non-significant. Double-headed arrows connecting explanatory variables represent genetic correlations; for simplicity, these values are not shown here but are in Supplementary Tables 3L and 4H. Double-headed arrows connecting the genetic component of childhood maltreatment to itself is the residual genetic variance of childhood maltreatment ($u_{\text{maltreatment}} \pm SE$) that is unexplained by the genetic influence of either the psychiatric disorders or the health and behavioural traits. ADHD = attention deficit hyperactivity disorder; ALCH = alcohol dependence; MDD = major depressive disorder; PTSD = post-traumatic stress disorder; SCZ = schizophrenia.

Model of independently contributing health and psychiatric traits

To explore the contribution of the independently genetically associated traits (i.e., ASD, PTSD, general risk tolerance, and subjective well-being), as a sensitivity analysis, we specified the genetic component of these four traits as explanatory variables of the genetic component of reported childhood maltreatment simultaneously (**Supplementary Figure 1 & Supplementary Table 5**). The residual genetic variance of reported childhood maltreatment was 0.42 ± 0.06 ($P = 1.24 \times 10^{-12}$). This means 58% of the h^2_{SNP} of childhood maltreatment was explained by the genetic component of ASD, PTSD, general risk tolerance, and subjective well-being. In a model with two additional traits (ADHD and self-reported tiredness), the residual genetic variance of childhood maltreatment only decreased by ~2% (0.40 ± 0.06 ; $P = 2.11 \times 10^{-12}$; **Supplementary Results & Supplementary Table 6**).

Discussion

Using genomic multiple regression in genomic SEM, we identified traits that accounted for ~60% of the h^2_{SNP} of reported trauma. Health and behavioural traits together accounted for 59% of h^2_{SNP} variance, whilst the model exploring psychiatric disorders explained 56%. In both models, a shared genetic component was observed across traits except subjective well-being, general risk tolerance, PTSD symptoms and ASD, which were independently associated with childhood maltreatment. Together, these latter four traits alone were sufficient to explain 58% of the h^2_{SNP} of reported childhood maltreatment.

We found similar results for retrospective lifetime trauma, which included adult trauma. We could not directly compare adulthood and childhood trauma as the lifetime trauma phenotype included both measures. Our findings suggest that the exact timing of trauma does not strongly affect the proportion of genetic variance accounted for by health and psychiatric traits. However, replication in non-overlapping datasets and appropriate trauma measures are required to make strong conclusions about differences between adulthood and childhood trauma.

Independently associated traits, in addition to the genetic components shared across health and psychiatric traits, may be involved in rGE, and/or the reporting of such environments as traumatic. This raises the question of which processes explain these associations and could inform strategies to minimise risk and consequences of trauma.

General risk tolerance is measured by endorsing a willingness to take risks, broadly capturing risk-taking behaviours⁵⁷. rGE may explain the genetic contribution of risk tolerance to the genetic component of reported trauma, consistent with a previous study that found environmental adversities mediate the association between genetic propensity for risk-taking and reported childhood maltreatment²³. A child may passively inherit a parent's genetic propensity for risk-taking behaviours, such as substance use, and be exposed to an environment where the child may be neglected or abused^{20,23}. An individual's own genetic propensity for risk-taking may increase exposure to potentially adverse environments^{17,21,58}. However, active rGE is less prominent during childhood than passive rGE and, therefore, a less plausible explanation for our findings⁵⁹. Alternatively, risk tolerance may capture behaviours associated with disclosure of trauma. A barrier to disclosing trauma includes perceiving it as a risk and fearing negative consequences³⁴. As such, individuals who take fewer risks may also be less inclined to disclose traumas. Further research is required to elucidate the mechanisms that explain the genetic association between general risk tolerance and reported trauma.

Subjective well-being captures an individual's cognitive evaluation of life satisfaction and positive affectivity^{60,61}. Our findings may reflect the role of such cognitions in the perception, recall, and thus reporting of trauma exposure. As found with life events, those with a genetic propensity to positive subjective well-being may be less likely to report trauma retrospectively, while the converse may occur with negative subjective well-being³⁰. A positive disposition could bias towards a more positive recall of experiences^{32,33}. This could explain why some individuals with objective records of experiencing childhood maltreatment do not retrospectively self-report maltreatment³¹. Such individuals are also less likely to develop posttraumatic psychopathology³¹. Positive emotional affectivity may contribute to resiliency by countering the adverse effects of stressful experiences due to a lack of subjective valence⁶². Conversely, if negative affectivity indicates greater sensitivity to trauma and vulnerability to psychopathology, this could have implications for screening for psychopathology risk following trauma⁶³.

The independent genetic association between ASD and childhood maltreatment is supported by several previous studies^{23,64,65}. Family-based polygenic score analyses exploring genetic differences between siblings suggest a greater risk of childhood maltreatment in those with ASD is partly attributable to evocative and active rGE²¹. Difficulty processing social cues may place an individual at greater risk of harmful environmental situations such as exploitation by a perpetrator^{64,66}. Furthermore, individuals with ASD may experience a broader range of life experiences as traumatic⁶⁷. An association between the polygenic score for ASD and trauma has been consistently found with retrospectively but not prospectively reported trauma^{23,64,65,68,69}. One study found the association between a polygenic score for ASD and retrospectively reported trauma was independent of rGE²³. Together, these findings suggest the importance of subjective trauma interpretation in ASD. Future research should determine the specific heritable components of ASD related to the subjective experience of and exposure to trauma, and the potential for screening for posttraumatic symptoms in ASD to provide appropriate support⁶⁷.

As trauma exposure is necessary for a PTSD diagnosis, the unique genetic association between reported trauma and PTSD symptoms could be explained by trauma exposure increasing the risk of PTSD. However, there are plausible reverse or bidirectional mechanisms. This includes passive rGE, whereby parental genetic predisposition to PTSD may act to increase the risk of trauma exposure in the child, as suggested by studies of PTSD and parenting⁷⁰. However, evidence suggests the association between a higher genetic risk for PTSD and increased self-report of childhood trauma could be explained by subjective interpretation processes and not rGE²³. This is supported by previous findings that PTSD

polygenic scores are not associated with objectively assessed trauma exposure severity⁷¹. Except for ASD and PTSD, genetic associations between reported childhood maltreatment and the other psychiatric disorders were explained by genetic factors shared across all other psychiatric disorders included in the model. However, GWAS often use brief phenotypic measures to achieve sufficient power, which may impact our ability to detect disorder-specific genetic influences⁷². Shared genetics may underlie transdiagnostic psychological mechanisms, such as those involved in the subjective experience of trauma, which is more robustly associated with psychiatric disorders than objective measures of trauma³¹. Our findings support targeting transdiagnostic pathways to reduce the general risk of psychopathology following trauma^{73,74}.

Our modelling approach has several limitations. First, the model is fully saturated, and we could not objectively estimate which model best fits the data. Second, we were limited by the number of traits that could be fitted in one model. However, sensitivity analyses did not indicate that all of the residual genetic variance of childhood maltreatment could be explained if all available genetically correlated traits were accounted for in one model. Third, our estimates are based on lower bounds of the total genetic variance explained by common genetic variants that can be, as captured by h^2_{SNP} , estimated from summary statistics and may differ from results using more advanced methods or individual-level genetic data^{40,75}.

Further research is needed to establish the role of traits in terms of whether they are associated with the risk of exposure or the interpretation and recollection of events. Preliminary evidence suggests that the h^2_{SNP} of prospective trauma is lower than retrospectively reported trauma²¹. Thus, genetic factors involved in the retrospective reporting of trauma may have a greater impact on the h^2_{SNP} of trauma than traits involved in the exposure of events. The residual h^2_{SNP} of trauma may reflect traits involved in memory recall^{17,28,33} that lack adequately powered GWASs. Alternatively, unaccounted for parental traits involved in passive rGE, such as antisocial behaviour contributing to an unsafe environment, may partly explain the residual genetic variance. Disentangling the genetic associations with vulnerability to exposure and those with the subjective experience of trauma will be important for distinguishing whether factors are relevant to trauma prevention or posttraumatic interventions.

In summary, we systematically examined traits genetically correlated with reported trauma, implicating possible mechanisms that partly explain the h^2_{SNP} of trauma. Potentially, indirect genetic effects regulating behaviour and cognition are associated with trauma exposure and/or retrospectively self-reporting trauma. We emphasise that our findings do not suggest that an individual is ever at fault or responsible for their exposure to trauma. Furthermore, an h^2_{SNP} of

reported trauma does not mean some individuals are genetically determined to experience trauma. Most of the phenotypic variance of reported trauma is explained by the environment, which is malleable and can be modified into a more supportive and protective environment to mitigate vulnerabilities. For example, if a genetic propensity for ASD, PTSD, general risk tolerance, and subjective well-being reflect genetic risk to reported trauma, more social support may protect such individuals and alleviate adverse posttraumatic effects. However, our findings are correlational, not necessarily causal, and better delineation of the processes involved is needed. Future studies could assess the specific role of these traits in large family-based datasets using within-family designs⁴³. Disentangling passive from evocative and active rGE that may explain our trait-specific associations could have implications for prevention strategies. As GWASs increase in power, the direction of causal relationships or testing the types of pleiotropy could be explored through Mendelian randomisation techniques^{21,76}. Such approaches are crucial to understanding vulnerability to trauma exposure and the subjective interpretation of trauma.

Acknowledgements

We thank the contribution of participants from all cohorts who have shared their life experiences. We gratefully acknowledge the scientists involved in the construction of all cohorts and those who provided the genetic summary results used in this study. We thank Dr A. Grotzinger for his guidance and contribution to the interpretation of our findings using genomic multiple regression with the Genomic SEM R package. This study represents years of independent research funded by the National Institute for Health and Care Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care or King's College London. C.H. is supported by Lundbeck Foundation (Grant no. R276-2018-4581). The authors 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 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 (TR130505). Figure 2 and Supplementary Figure 2 were created with BioRender.com.

Conflict of interest

G.B. has received honoraria, research or conference grants and consulting fees from Illumina, Otsuka, and COMPASS Pathways Ltd. The remaining authors have nothing to disclose.

Author contributions: CRediT author statement

Abigail R. ter Kuile: Conceptualization, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualisation. **Christopher Hübel:** Conceptualization, Formal analysis, Data Curation, Writing - Review & Editing, Visualisation. **Rosa Cheesman:** Conceptualization, Writing - Review & Editing. **Jonathan R. I. Coleman:** Conceptualization, Data Curation, Writing - Review & Editing. **Alicia Peel:** Conceptualization, Writing - Review & Editing. **Daniel F. Levey:** Data Curation, Writing - Review & Editing. **Murray B. Stein:** Data Curation, Writing - Review & Editing. **Joel Gelernter:** Data Curation, Writing - Review & Editing. **Christopher Rayner:** Formal analysis, Writing - Review & Editing. **Thalia C. Eley:** Conceptualization,

Writing - Original Draft, Writing - Review & Editing, Supervision. **Gerome Breen:** Conceptualization, Writing - Original Draft, Writing - Review & Editing, Supervision.

Data availability

Data availability of all GWAS summary statistics used for analyses can be found in **Supplementary Table 12**. Access to Million Veteran Programme PTSD GWAS summary statistics can be requested through bdGaP ([phs001672.v7.p1](#)).

References

1. Scott, K. M. *et al.* Associations between lifetime traumatic events and subsequent chronic physical conditions: a cross-national, cross-sectional study. *PLoS One* **8**, e80573 (2013).
2. Keyes, K. M. *et al.* Potentially traumatic events and the risk of six physical health conditions in a population-based sample. *Depress. Anxiety* **30**, 451–460 (2013).
3. Turner, R. J., Jay Turner, R. & Lloyd, D. A. Lifetime Traumas and Mental Health: The Significance of Cumulative Adversity. *Journal of Health and Social Behavior* vol. 36 360 Preprint at <https://doi.org/10.2307/2137325> (1995).
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
5. Kendler, K. S. & Baker, J. H. Genetic influences on measures of the environment: a systematic review. *Psychol. Med.* **37**, 615–626 (2007).
6. Plomin, R. & Asbury, K. Nature and Nurture: Genetic and Environmental Influences on Behavior. *Ann. Am. Acad. Pol. Soc. Sci.* **600**, 86–98 (2005).
7. Stein, M. B., Jang, K. L., Taylor, S., Vernon, P. A. & Livesley, W. J. Genetic and environmental influences on trauma exposure and posttraumatic stress disorder symptoms: a twin study. *Am. J. Psychiatry* **159**, 1675–1681 (2002).
8. Lyons, M. J. *et al.* Do genes influence exposure to trauma? A twin study of combat. *Am. J. Med. Genet.* **48**, 22–27 (1993).
9. Sartor, C. E. *et al.* Common heritable contributions to low-risk trauma, high-risk trauma, posttraumatic stress disorder, and major depression. *Arch. Gen. Psychiatry* **69**, 293–299 (2012).
10. Richmond-Rakerd, L. S. *et al.* Common genetic contributions to high-risk trauma exposure and self-injurious thoughts and behaviors. *Psychol. Med.* **49**, 421–430 (2019).
11. Pezzoli, P., Antfolk, J., Hatoum, A. S. & Santtila, P. Genetic Vulnerability to Experiencing Child Maltreatment. *Front. Genet.* **10**, 852 (2019).
12. Jang, K. L., Vernon, P. A., John Livesley, W., Stein, M. B. & Wolf, H. Intra- and extra-familial influences on alcohol and drug misuse: a twin study of gene-environment correlation. *Addiction* vol. 96 1307–1318 Preprint at <https://doi.org/10.1046/j.1360-0443.2001.969130710.x> (2001).
13. Jang, K. L., Vernon, P. A., Livesley, W. J., Stein, M. B. & Wolf, H. Intra- and extra-familial influences on alcohol and drug misuse: A twin study of gene-environment correlation. *Addiction* **96**, 1307–1318 (2001).
14. Kendler, K. S., Karkowski, L. M. & Prescott, C. A. The assessment of dependence in the study of stressful life events: validation using a twin design. *Psychol. Med.* **29**, 1455–1460 (1999).
15. Averdijk, M. Victim Selection. *The Oxford Handbook of Offender Decision Making* **6**, 301 (2017).
16. von Stumm, S. & d'Apice, K. From Genome-Wide to Environment-Wide: Capturing the Environome. *Perspect. Psychol. Sci.* 1745691620979803 (2021).
17. Jang, K. L., Stein, M. B., Taylor, S., Asmundson, G. J. G. & Livesley, W. J. Exposure to traumatic events and experiences: aetiological relationships with personality function. *Psychiatry Res.* **120**, 61–69 (2003).
18. Jaffee, S. R. & Price, T. S. Gene-environment correlations: a review of the evidence and implications for prevention of mental illness. *Mol. Psychiatry* **12**, 432–442 (2007).
19. Knafo, A. & Jaffee, S. R. Gene-environment correlation in developmental psychopathology. *Dev. Psychopathol.* **25**, 1–6 (2013).
20. Dube, S. R. *et al.* Growing up with parental alcohol abuse: exposure to childhood abuse, neglect, and household dysfunction. *Child Abuse Negl.* **25**, 1627–1640 (2001).

21. Warrier, V. *et al.* Gene–environment correlations and causal effects of childhood maltreatment on physical and mental health: a genetically informed approach. *The Lancet Psychiatry* **8**, 373–386 (2021).
22. Das, A. Genes, Childhood Trauma, and Late Life Depressive Symptoms. *J. Aging Health* **31**, 1503–1524 (2019).
23. Peel, A. J. *et al.* Genetic and early environmental predictors of adulthood self-reports of trauma. *Br. J. Psychiatry* 1–8 (2022).
24. Cloninger, C. R., Garcia, D. & Pluess, M. The heritability and development of positive affect and emotionality. *Genetics of psychological well-being—the role of heritability and genetics in positive psychology* 97–113 (2015).
25. Kremen, W. S. *et al.* Genetics of verbal working memory processes: a twin study of middle-aged men. *Neuropsychology* **21**, 569–580 (2007).
26. Eley, T. C. *et al.* In the face of uncertainty: a twin study of ambiguous information, anxiety and depression in children. *J. Abnorm. Child Psychol.* **36**, 55–65 (2008).
27. Lau, J. Y. F., Rijdsdijk, F. & Eley, T. C. I think, therefore I am: a twin study of attributional style in adolescents. *J. Child Psychol. Psychiatry* **47**, 696–703 (2006).
28. Goodman, G. S., Quas, J. A. & Ogle, C. M. Child maltreatment and memory. *Annu. Rev. Psychol.* **61**, 325–351 (2010).
29. Boutwell, B. B. On variability & human consciousness. *Heliyon* **4**, e00905 (2018).
30. Wootton, R. E., Davis, O. S. P., Mottershaw, A. L., Wang, R. A. H. & Haworth, C. M. A. Genetic and environmental correlations between subjective wellbeing and experience of life events in adolescence. *Eur. Child Adolesc. Psychiatry* **26**, 1119–1127 (2017).
31. Danese, A. & Widom, C. S. Objective and subjective experiences of child maltreatment and their relationships with psychopathology. *Nat Hum Behav* **4**, 811–818 (2020).
32. Baldwin, J. R., Reuben, A., Newbury, J. B. & Danese, A. Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment. *JAMA Psychiatry* vol. 76 584 Preprint at <https://doi.org/10.1001/jamapsychiatry.2019.0097> (2019).
33. Reuben, A. *et al.* Lest we forget: comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *J. Child Psychol. Psychiatry* **57**, 1103–1112 (2016).
34. Pokharel, B., Hegadoren, K. & Papathanassoglou, E. Factors influencing silencing of women who experience intimate partner violence: An integrative review. *Aggress. Violent Behav.* **52**, 101422 (2020).
35. Ullman, S. E. Social reactions to child sexual abuse disclosures: a critical review. *J. Child Sex. Abus.* **12**, 89–121 (2003).
36. Paine, M. L. & Hansen, D. J. Factors influencing children to self-disclose sexual abuse. *Clin. Psychol. Rev.* **22**, 271–295 (2002).
37. Goodman-Brown, T. B., Edelstein, R. S., Goodman, G. S., Jones, D. P. H. & Gordon, D. S. Why children tell: a model of children’s disclosure of sexual abuse. *Child Abuse Negl.* **27**, 525–540 (2003).
38. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* (2019) doi:10.1038/s41576-019-0137-z.
39. Dalvie, S. *et al.* Genomic influences on self-reported childhood maltreatment. *Transl. Psychiatry* **10**, 38 (2020).
40. Coleman, J. R. I. *et al.* Genome-wide gene–environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Mol. Psychiatry* (2020) doi:10.1038/s41380-019-0546-6.
41. Saudino, K. J., Pedersen, N. L., Lichtenstein, P., McClearn, G. E. & Plomin, R. Can personality explain genetic influences on life events? *J. Pers. Soc. Psychol.* **72**, 196–206 (1997).

42. McAdams, T. A., Gregory, A. M. & Eley, T. C. Genes of experience: explaining the heritability of putative environmental variables through their association with behavioural and emotional traits. *Behav. Genet.* **43**, 314–328 (2013).
43. Friedman, N. P., Banich, M. T. & Keller, M. C. Twin studies to GWAS: there and back again. *Trends Cogn. Sci.* **25**, 855–869 (2021).
44. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513–525 (2019).
45. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
46. Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
47. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
48. Boyd, A. *et al.* Cohort profile: the ‘children of the 90s’—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
49. Hofman, A. *et al.* Growth, development and health from early fetal life until young adulthood: the Generation R Study. *Paediatr. Perinat. Epidemiol.* **18**, 61–72 (2004).
50. Kooijman, M. N. *et al.* The Generation R Study: design and cohort update 2017. *Eur. J. Epidemiol.* **31**, 1243–1264 (2016).
51. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
52. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
53. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* vol. 33 272–279 Preprint at <https://doi.org/10.1093/bioinformatics/btw613> (2017).
54. Nivard, M. G. Minimum sample size, SNP-h² and rG? *Genomic SEM users google group* <https://groups.google.com/g/genomic-sem-users/c/Wn1jptU2VcY/m/9CWxfQbvAAAJ> (2019).
55. Bulik-Sullivan, B. LD Score Regression GitHub Wiki FAQ. (Github).
56. Team, R. C. R Core Team (2019) RA Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. *References-Scientific Research Publishing*.
57. Karlsson Linnér, R. *et al.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat. Genet.* **51**, 245–257 (2019).
58. Ryan, J., Chaudieu, I., Ancelin, M.-L. & Saffery, R. Biological underpinnings of trauma and post-traumatic stress disorder: focusing on genetics and epigenetics. *Epigenomics* **8**, 1553–1569 (2016).
59. Scarr, S. & McCartney, K. How people make their own environments: A theory of genotype --> environment effects. *Child Dev.* **54**, 424 (1983).
60. Okbay, A. *et al.* Genetic associations with subjective well-being also implicate depression and neuroticism. *Nat. Genet.* (2016).
61. Diener, E., Oishi, S. & Lucas, R. E. Personality, culture, and subjective well-being: emotional and cognitive evaluations of life. *Annu. Rev. Psychol.* **54**, 403–425 (2003).
62. Tugade, M. M. & Fredrickson, B. L. Resilient individuals use positive emotions to bounce back from negative emotional experiences. *J. Pers. Soc. Psychol.* **86**, 320–333 (2004).
63. Baldwin, J. R. *et al.* Population vs Individual Prediction of Poor Health From Results of Adverse Childhood Experiences Screening. *JAMA Pediatr.* **175**, 385–393 (2021).

64. Warrier, V. & Baron-Cohen, S. Childhood trauma, life-time self-harm, and suicidal behaviour and ideation are associated with polygenic scores for autism. *Mol. Psychiatry* **26**, 1670–1684 (2019).
65. Ratanatharathorn, A. *et al.* Polygenic risk for autism, attention-deficit hyperactivity disorder, schizophrenia, major depressive disorder, and neuroticism is associated with the experience of childhood abuse. *Mol. Psychiatry* **26**, 1696–1705 (2021).
66. Sevelever, M., Roth, M. E. & Gillis, J. M. Sexual abuse and offending in autism spectrum disorders. *Sex. Disabil.* **31**, 189–200 (2013).
67. Rumball, F., Happé, F. & Grey, N. Experience of Trauma and PTSD Symptoms in Autistic Adults: Risk of PTSD Development Following DSM-5 and Non-DSM-5 Traumatic Life Events. *Autism Res.* (2020) doi:10.1002/aur.2306.
68. Sallis, H. M. *et al.* Genetic liability to schizophrenia is associated with exposure to traumatic events in childhood. *Psychol. Med.* **51**, 1814–1821 (2021).
69. Baldwin, J. R. *et al.* A genetically informed Registered Report on adverse childhood experiences and mental health. *Nat Hum Behav* (2022) doi:10.1038/s41562-022-01482-9.
70. Christie, H., Hamilton-Giachritsis, C., Alves-Costa, F., Tomlinson, M. & Halligan, S. L. The impact of parental posttraumatic stress disorder on parenting: a systematic review. *Eur. J. Psychotraumatol.* **10**, 1550345 (2019).
71. Waszczuk, M. A. *et al.* Polygenic prediction of PTSD trajectories in 9/11 responders. *Psychol. Med.* **52**, 1981–1989 (2022).
72. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
73. McLaughlin, K. A., Colich, N. L., Rodman, A. M. & Weissman, D. G. Mechanisms linking childhood trauma exposure and psychopathology: a transdiagnostic model of risk and resilience. *BMC Med.* **18**, 96 (2020).
74. Baldwin, J. R. *et al.* Childhood Maltreatment and Mental Health Problems: A Systematic Review and Meta-Analysis of Quasi-Experimental Studies. *Am. J. Psychiatry* appiajp20220174 (2023).
75. Evans, L. M. *et al.* Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat. Genet.* **50**, 737–745 (2018).
76. Pingault, J.-B., Cecil, C. A. M., Murray, J., Munafò, M. R. & Viding, E. Causal Inference in Psychopathology: A Systematic Review of Mendelian Randomisation Studies Aiming to Identify Environmental Risk Factors for Psychopathology. *Psychopathology Review* **a4**, 4–25 (2017).

Chapter 4. Genetic influences on self-reported outcomes following psychological treatment for depression and anxiety disorders

This chapter is a manuscript that will be submitted for peer-review following approval from co-authors. Supplementary materials for this chapter, as detailed in the text, are included in **Appendix C**.

Authors: Alicia J. Peel^{1*}, **Abigail R. ter Kuile**^{1,2*}, Christopher Rayner^{1,2}, Kristen Hopkins^{3,4}, Brittany L. Mitchell³, Sarah E. Medland³, Sang Hyuck Lee¹, Brett N. Adey^{1,2}, Chérie Armour⁵, Joshua E. J. Buckman^{6,7}, Enda M. Byrne⁸, Jonathan R. I. Coleman^{1,2}, Andrea Danese^{1,9,10}, Molly R. Davies^{1,2}, Helena L. Davies¹, Ian B. Hickie¹¹, Colette Hirsch^{1,2,12}, Matthew Hotopf^{1,2,12}, Christopher Hübel^{1,2,13}, Ian R. Jones¹⁴, Gursharan Kalsi^{1,2}, Georgina Krebs^{1,12}, Nicholas G. Martin³, Andrew M. McIntosh¹⁵, Jessica Mundy^{1,2}, Kirstin L. Purves^{1,2}, Henry Rogers^{1,2}, Megan Skelton^{1,2}, Daniel J. Smith¹⁵, David Veale^{1,2,12}, James T. R. Walters¹⁴, Naomi R. Wray^{8,16}, Gerome Breen^{1,2} & Thalia C. Eley^{1,2}

*Joint first authors

1. Institute of Psychiatry, Psychology and Neuroscience, King's College London, Denmark Hill, Camberwell, London, UK
2. National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust, London, UK
3. QIMR Berghofer Medical Research Institute, Brisbane, Australia
4. School of Biomedical Sciences, Faculty of Medicine, The University of Queensland, Brisbane, Australia
5. Research Centre for Stress, Trauma, and Related Conditions (STARC), School of Psychology, Queen's University Belfast (QUB), Belfast, Northern Ireland, UK
6. Research Department of Clinical, Educational and Health Psychology, University College London, Gower Street, London, UK
7. iCope – Camden and Islington Psychological Therapies Services, Camden & Islington NHS Foundation Trust, Finsbury Health Centre, Pine Street, London, UK
8. Queensland Brain Institute, The University of Queensland, Brisbane, Australia
9. Department of Child & Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
10. National and Specialist CAMHS Trauma, Anxiety, and Depression Clinic, South London and Maudsley NHS Foundation Trust, London, UK
11. Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia
12. South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital, Monks Orchard Road, Beckenham, Kent, UK
13. National Centre for Register-based Research, Aarhus Business and Social Sciences, Aarhus University, Aarhus, Denmark
14. MRC Centre for Neuropsychiatric Genetics and Genomics, National Centre for Mental Health, Cardiff University, Cardiff, UK
15. Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK
16. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Australia

Abstract

Background: Common genetic variation is anticipated to contribute to psychological treatment outcomes. However, statistical power to identify genetic associations has been limited by challenges in recruiting large samples with prospective treatment outcome data. We aimed to investigate genetic influences on outcomes following psychological treatment for depression and anxiety disorders using retrospective self-reports.

Methods: We conducted a genome-wide association study meta-analysis of retrospectively self-reported outcomes following psychological treatment for depression or anxiety, combining measures across two datasets (total N=15,131; ~75% reported positive outcomes). In a subset of individuals (N=4,439), predictive models for self-reported treatment outcomes were estimated using elastic net regularisation and evaluated using nested 10-fold cross-validation. Polygenic scores of related traits were included as genetic predictors. Models incorporating genetic predictors were compared to models of known sociodemographic/clinical predictors to assess whether genetic influences on psychiatric, cognitive and related traits contributed to prediction over and above existing predictors.

Results: The single nucleotide polymorphism (SNP)-based heritability of self-reported treatment outcomes was not significantly different from zero. No common genetic variants were associated with this phenotype at the genome-wide significance level. Elastic net models containing polygenic scores explained 0.4% of the variance in self-reported treatment outcomes. Incorporating genetic predictors did not significantly improve prediction beyond existing sociodemographic and clinical predictors.

Conclusions: Identifying the common genetic variants associated with retrospectively self-reported treatment outcomes requires larger sample sizes. A small proportion of the variance explained by polygenic scores supports the hypothesis that self-reported treatment outcomes has a heritable component. Further research is needed to validate single-item self-reported measures assessing psychological treatment outcomes.

Introduction

Depression and anxiety disorders are highly prevalent, affecting up to 29% of the global population at some point in their lifetime^{1,2}. The UK Improving Access to Psychological Therapies (IAPT) programme, which provides psychological treatment for mild-moderate depression and anxiety, receives over 1.25 million referrals annually, with approximately 65% of referred individuals receiving treatment³. The IAPT treatments include guided self-help, counselling, group therapies and couples therapy, with Cognitive Behavioural Therapy (CBT) being the primary form of treatment⁴. However, treatment outcomes are variable, and many individuals do not experience improvement in symptoms⁵⁻⁸. A meta-analysis of 228 randomised trials of psychological therapy for depression found that only 41% of patients demonstrated treatment response, defined as a 50% reduction in depressive symptoms⁹. Furthermore, only one-third of patients reached criteria for remission, described as falling below a clinical cut-off for symptom scores. Response rates are consistent across different forms of psychological treatment⁹ and for the treatment of anxiety disorders¹⁰. In addition to variable outcomes, access to psychological therapies is limited by the cost of treatment and a shortage of trained healthcare professionals^{11,12}. There is a strong rationale for investigating who is most likely to benefit from psychological treatment given the limited resources available.

Several clinical, treatment and patient characteristics have been associated with differential outcomes following psychological treatment for anxiety and depression. The term outcomes is used here to broadly refer to any measure capturing a lack of improvement in symptoms, including a reduction in symptoms and complete remission of case status. Clinical characteristics including longer duration of illness¹³⁻¹⁶, earlier age of onset¹⁷, greater functional impairment¹⁸⁻²⁰, greater symptom severity^{21,22} and comorbid anxiety, depressive or personality disorders^{13,23-25} are consistently associated with poor outcomes. Treatment factors such as lower patient engagement²⁶, poor treatment adherence²¹ and previous courses of treatment^{13,27} are also associated with unfavourable outcomes. Associations with sociodemographic characteristics are less well-established. However, evidence suggests that being single²⁸, not being in full-time employment^{19,29,30}, having lower educational attainment²⁹, and experiencing childhood maltreatment¹⁷ are associated with worse prognosis.

Although numerous clinical, treatment and patient factors are associated with psychological treatment outcomes, the prediction of individual response remains poor³¹⁻³³. Identifying novel sources of information to improve the prediction of treatment outcomes is essential for developing models with greater utility. One possible avenue is to investigate genetic factors associated with treatment outcomes. Pharmacological treatments, such as antidepressants,

have yielded promising results in genetic investigations, demonstrating that common genetic variants explain ~13% of the variance in remission^{34,35}. Response to psychological therapies is also anticipated to have a genetic component, given that outcomes are influenced by learning behaviours, interpersonal factors and personality characteristics^{21,26,36}, all of which are partly heritable^{37–39}. However, in the field of ‘therapygenetics’, studies investigating genetic influences on psychological treatment outcomes have not found genome-wide significant associations between common genetic variants and symptom change during treatment, nor have they identified a significant common genetic variant-based heritability estimate^{40,41}. The primary explanation for these findings is that the sample sizes in these studies (both under 3,000 individuals) were insufficient to detect the many common genetic variants that each explain a small amount of variance in treatment outcomes⁴². Achieving the statistical power required for detecting a genetic basis of psychological treatment outcomes is challenging due to the financial and time constraints of collecting prospective symptom data throughout treatment for the thousands of individuals required for well-powered genome-wide association studies (GWAS).

To progress towards identifying genetic factors associated with treatment outcomes, sample size challenges must be overcome. One approach is to use ‘minimal’ or ‘brief’ phenotyping, whereby a single self-reported questionnaire item is used to capture the trait of interest⁴³. This strategy has been used in GWAS of psychiatric phenotypes, such as depression, to achieve sufficient sample sizes to detect genome-wide significant loci (e.g. Howard et al.⁴⁴). However, brief phenotyping has been criticised for capturing qualitatively different traits to more detailed clinical assessments^{45,46}. Nonetheless, briefly assessed depression phenotypes have high genetic correlations with clinically-derived phenotypes for major depressive disorder. ‘Broad’ depression, which was assessed using two single-item self-reported measures on help-seeking for “nerves, anxiety, tension or depression”, had a genetic correlation of 0.87 with clinically-derived major depressive disorder^{47,48}. Similarly, depression assessed through self-reported clinical diagnosis or treatment history had a genetic correlation of 0.85^{48,49}. Although brief measures may not detect disorder-specific genetic influences as effectively as detailed phenotypic measures⁴⁶, the high genetic correlation observed across phenotyping strategies indicates a similar underlying genetic architecture is captured.

A brief phenotyping approach can be applied to psychological treatment responses. Retrospectively self-reported outcomes following treatment can be used instead of prospective symptom data during treatment. In phenotypic analyses, outcomes assessed using retrospective self-reports display a similar pattern of association with sociodemographic and clinical factors as prospective indicators of therapy response⁵⁰. Participants with a lifetime

experience of depression and/or anxiety were asked to retrospectively report how much their symptoms and day-to-day functioning had improved following a previous course of therapy. Responses were assessed using a scale adapted from the Global Rating of Change scale: much worse (-2), a little worse (-1), no change (0), a little better (+1), much better (+2)⁵¹. Having university-level education and using additional therapeutic activities were associated with favourable self-reported outcomes, while the number of disorder episodes, personality disorder symptoms, being male, and receiving therapy for the first time were associated with less favourable self-reported outcomes⁵⁰. Effect sizes were consistent with previous literature, with odds ratios ranging from 0.89 to 1.45. These results indicate that retrospective self-reports capture some aspects of treatment outcomes that overlap with responses measured using traditional approaches. This provides support for the use of retrospectively self-reported outcomes as an alternative to traditional prospective data collection. Combining data from studies that have utilised this simple brief phenotyping approach could enable reaching large enough sample sizes to identify genetic variants associated with treatment outcomes. The associated genetic variants can then be summed to create a polygenic score that captures genetic propensity to favourable treatment outcomes, which could be used in predicting psychological treatment response.

A second route to identifying the genetic basis of treatment outcomes is to incorporate polygenic scores for traits related to treatment response in prediction models, capitalising on pre-existing well-powered GWAS. Previous studies found that higher polygenic scores for autism spectrum disorder and lower polygenic scores for depression and intelligence were associated with remission in depressive symptoms following internet-based CBT (iCBT)^{52,53}. Genetic predictors also independently contributed to the prediction of iCBT outcomes alongside demographic and clinical factors⁵³, indicating that genetic propensity for relevant psychiatric or cognitive traits provides additional information in predicting treatment outcomes. However, evidence is limited for the clinical utility of genetic influences in predicting outcomes for in-person psychological treatments. Similar to GWAS, progress in developing treatment models incorporating genetics has been hindered by the challenge of collecting large samples with both polygenic scores and psychological treatment outcome data available.

We aimed to investigate the genetic influences on retrospectively self-reported outcomes in individuals who have received psychological therapy for anxiety and depressive disorders. Two approaches were used to address this aim. First, we performed a genome-wide association study meta-analysis of retrospectively self-reported single-item questions assessing outcomes following psychological therapy for depression/anxiety. We used this approach to assess whether this phenotype has a heritable basis that can be captured by

common genetic variants, also known as single-nucleotide polymorphisms (SNP-based heritability). Second, we used elastic net regression with nested 10-fold cross-validation to build and evaluate models predicting retrospectively self-reported treatment outcomes from genetic, sociodemographic and clinical factors. Using this approach, we aimed to determine whether genetic influences on psychiatric, cognitive and related traits contribute to the prediction of retrospectively self-reported outcomes beyond known clinical and demographic predictors.

Methods

Phenotypic definitions and genome-wide association analyses were pre-registered on the Open Science Framework (accessible at <https://osf.io/h7d9a/>). Analysis plans for prediction modelling were formulated following pre-registration and prior to commencing analyses.

Samples

This study included data collected in 15,131 genotyped individuals from two datasets. The first dataset was created by combining data from two studies within the National Institute for Health and Care Research (NIHR) BioResource; the Genetic Links to Anxiety and Depression Study (GLAD)⁵⁴ and the COVID-19 Psychiatry and Neurological Genetics Study (COPING)⁵⁵, forming the GLAD+ dataset. The second dataset comprised the Australian Genetics of Depression Study (AGDS)⁵⁶. We included participants of predominantly European-associated genetic ancestry from these samples who met the following three criteria. First, they indicate lifetime experience of a depressive or anxiety disorder based on symptom reports or self-reported diagnoses. Second, they reported receiving any type of psychological therapy for a depressive or anxiety disorder and have completed at least one measure of self-reported treatment outcomes. Third, they provided genetic data that passed quality control criteria.

The GLAD+ dataset

The Genetic Links to Anxiety and Depression (GLAD) Study

The GLAD study aims to facilitate research investigating genetic and environmental risk factors for depression and anxiety disorders by recruiting participants into the NIHR Mental Health BioResource⁵⁴. Recruitment is open to individuals in the UK aged 16+ with lifetime anxiety or depressive disorders. Individuals were initially recruited through a widespread media campaign. Ongoing recruitment is supported by social media advertising and recruitment through NHS Trusts and GP practices. Participants sign up through the GLAD study website (<https://gladstudy.org.uk>), provide informed consent, and complete online, self-report questionnaires including detailed symptom reports. Eligibility is assessed from questionnaire responses using self-reported diagnoses and diagnostic algorithms. Eligible participants are sent a saliva DNA kit in the post to their preferred address and are able to take part in additional optional questionnaires. Saliva samples were genotyped using the UK Biobank v2 Axiom array⁵⁷. Recruitment began in September 2018 and is ongoing. See Davies et al.⁵⁴ for a full description of the GLAD recruitment and data collection procedures. The GLAD study was approved by the London Fulham Research Ethics Committee (REC

reference: 18/LO/1218) following a full review by the committee.

The COVID-19 Psychiatry and Neurological Genetics (COPING) Study

The NIHR BioResource is a resource of UK volunteers who have provided medical, clinical, and genetic data and have consented to re-contact for future research. Volunteers were recruited through several initiatives across the country, including NHS blood transfusion services and disorder-specific clinics and research studies. In April 2020, participants from the NIHR BioResource (<https://bioresource.nihr.ac.uk/>) were invited to take part in the COPING study, which aimed to assess mental health and well-being throughout the COVID-19 pandemic⁵⁵. Participants that consented to take part in COPING completed a baseline survey, similar to that in GLAD, details of which are given elsewhere⁵⁸. COPING study participants are made up of those from several genotyped subsamples within the NIHR BioResource, including the Inflammatory Bowel Disease BioResource ($N = 237$); NHS blood and transplant studies, including INTERVAL ($N = 659$), COMPARE ($N = 253$), and STRategies to Improve Donor Experiences (STRIDES; $N = 57$); and the Research Tissue Bank - Generic ($N = 663$). The NIHR BioResource was approved as a Research Tissue Bank by the East of England, Cambridge Central Committee (REC reference: 17/EE/0025). Ethical approval for the COPING study was obtained by the South West, Central Bristol Research Ethics Committee (REC reference: 20/SW/0078). The combined GLAD and COPING participants used in this analysis are referred to as GLAD+ participants from here on.

Australian Genetics of Depression Study (AGDS) dataset

The Australian Genetics of Depression Study aims to contribute to research assessing the role of genetic variation in the risk of depression and therapeutic response to antidepressants. Recruitment was open to adults across Australia who “had been treated by a doctor, psychiatrist or psychologist for depression”. Individuals were recruited via an open media campaign and a targeted mailout based on recent antidepressant prescription history. Participants registered through the AGDS website (<https://www.geneticsofdepression.org.au>), where they provided informed consent and completed an online self-report questionnaire assessing mental health diagnoses and antidepressant responses. Registered participants were mailed a saliva DNA kit to their preferred address. Recruitment began in September 2016 and is ongoing. See Byrne et al.⁵⁶ for a full description of the AGDS recruitment and data collection procedures. The AGDS was reviewed and approved by QIMR Berghofer Medical Research Institute's Human Research Ethics Committee.

Measures

Self-reported treatment outcomes

Self-reported outcome following psychological therapy for depression and anxiety was defined using two retrospectively reported single-item questions assessing *benefit from therapy* and *improvement in symptoms/functioning* (**Table 1**).

Self-reported benefit from therapy

Within the GLAD+ baseline survey, participants completed self-report modules of the Composite International Diagnostic Interview - Short Form (CIDI-SF) adapted for major depression, generalised anxiety disorder, social anxiety, agoraphobia, panic disorder and specific phobia⁵⁹. Within this, participants were asked, 'Did you ever try or are currently trying the following for these problems?'. Those who selected the answer 'Psychotherapy or other talking therapy more than once' for depression or 'Psychotherapy' for anxiety were presented with further questions, including 'Did you find psychotherapy or other talking therapy helpful?', which has the response options 'Yes', 'No', 'Don't know' and 'Prefer not to answer'. Answers to this item are used to identify those with *self-reported benefit from therapy*, with binary options 'Yes' and 'No'.

For individuals within GLAD+ who responded to *self-reported benefit from therapy* questions in relation to multiple disorders, responses were selected hierarchically. Responses from the major depression module were preferentially selected (94% of responses) to correspond as closely as possible with participants from AGDS, who were all ascertained on the basis of a lifetime depression diagnosis. When unavailable, responses were selected in relation to generalised anxiety disorder (4%), panic disorder (2%), agoraphobia (<1%), specific phobia (<1%), then social anxiety (<1%). The proportion of participants meeting diagnostic criteria for each disorder is given in **Supplementary Table S1**. As participants were asked to report all psychological therapies that they attended more than once, we did not differentiate by the type of treatment received. The proportion of participants reporting each type of treatment is given in **Supplementary Table S2**.

Self-reported improvement in symptoms/functioning

Following completion of the sign-up questionnaires, participants from the GLAD study and AGDS were invited to take part in an optional follow-up survey about therapy outcomes. The survey was made available to all participants on the study websites, and active GLAD participants were sent an email to invite them to take part in a self-report questionnaire about lifetime treatment history⁵⁰. As part of this questionnaire, participants were asked if they had

ever tried psychological therapy for help with anxiety or depression. In the GLAD study, participants were asked about ‘Talking therapies’ in line with the terminology used in UK IAPT services. In AGDS, participants were asked about ‘Psychological treatment or Counselling’. Those who endorsed this item were asked their age at the time of therapy, the main diagnosis that the treatment was for, additional diagnoses, whether therapy was one-to-one or group, the type of therapy, concurrent medication and improvement in symptoms for the most recent and second most recent times they tried therapy. Improvement in symptoms was assessed with the item ‘How much did your symptoms and day to day functioning improve?’, which had Likert scale response options of ‘Much better’, ‘A little better’, ‘No change’, ‘A little worse’ and ‘Much worse’. Answers to this item are used to identify those with *self-reported improvement in symptoms/functioning*. For individuals who reported two courses of therapy, we selected the response relating to their most recent course.

Due to the small sample size across each of the five ordinal categories, a binary definition of the *improvement in symptoms/functioning* phenotype was created, with ‘Much better’ and ‘A little better’ response options denoting ‘Improvement’, and ‘No change’, ‘A little worse’ and ‘Much worse’ responses denoting ‘No improvement’. This binary variable was used to supplement data from the CIDI-SF, being combined with *self-reported benefit from therapy* to create one phenotype representing general self-reported outcomes following therapy. This phenotype is referred to as *self-reported treatment outcomes*.

Table 1. Number of participants from the GLAD+ and AGDS datasets with data on each measures of self-reported outcomes following therapy (Total N = 15,131)

Phenotype	GLAD+		AGDS
	GLAD	COPING	
<i>Self-reported benefit from therapy</i> - adapted CIDI-SF modules	3,206	1,869	-
<i>Self-reported improvement in symptoms/functioning</i> - optional lifetime treatment history questionnaire	2,461	-	7,595

Sociodemographic and clinical predictors of self-reported treatment outcomes

Sociodemographic and clinical predictors previously associated with retrospectively self-reported improvement in symptoms/functioning in the GLAD+ dataset were selected⁵⁰. As participants retrospectively reported outcomes of past treatments, these variables were chosen to reflect the entire period of illness rather than the specific episode in which treatment was sought. We only included associated predictors assessed in the adapted CIDI-SF modules that formed part of the GLAD+ baseline survey to retain a sample size sufficient for prediction modelling. This resulted in three predictors being taken forward. Two were

sociodemographic predictors: self-reported biological sex (male/female) and education level dichotomised from six categorical responses using the median category (University degree: yes/no). The clinical predictor, number of episodes of illness, was assessed as a continuous variable. Although personality disorder symptoms had been previously associated with retrospectively self-reported improvement in symptoms/functioning⁵⁰, this variable was not included due to a lack of consistent measurement across the GLAD and COPING studies.

Aim 1: Genome-wide association study meta-analysis of retrospectively self-reported treatment outcomes

Genotyping quality control and imputation

Participants from the NIHR BioResource dataset were genotyped using the Affymetrix UK Biobank Axiom Array v1 and v2. Genotype data quality control was performed separately in the GLAD and COPING datasets. Genetic variants were excluded if they were rare (with a minor allele frequency < 1%), were missing > 5% per variant, or displayed Hardy-Weinberg equilibrium deviation ($P < 10^{-10}$). Individuals were excluded with genetic variant missingness > 5%, genetic duplication across both samples, undetermined genetic sex, and outliers of identify by descent (indicating genetic contamination). Genotyping and genetic quality control performed in the AGDS dataset is described in detail elsewhere⁶⁰. In both the GLAD+ and AGDS datasets, imputation was performed using the TOPMed imputation panel (Version R2 on GRC38), and variants were filtered to quality info metric > 0.3. Due to small sample size of non-European genetic ancestry participants, analyses in both datasets were restricted to European-associated ancestry clusters selected through principal component analysis (PCA).

Genome-wide association study

For our first aim, we conducted GWAS of single-item questions assessing retrospectively self-reported treatment outcomes in each sample. Logistic mixed model analyses were performed using the software REGENIE for analyses in the GLAD+ dataset and SAIGE in the AGDS dataset. Logistic regression score tests were applied to account for unbalanced case-control ratios, with approximate Firth logistic regression test used in REGENIE and the saddle point approximation test used in SAIGE^{61,62}. The first 10 principal components and genotyping batches were included as covariates. Results were filtered on MAF > 0.01 and R^2 imputation quality metric > 0.3.

GWAS meta-analysis

We meta-analysed the GWAS summary statistics from each sample to obtain a GWAS of overall self-reported treatment outcomes following psychological therapy. We used the software METAL to perform an inverse-variance weighted meta-analysis⁶³. We restricted genetic variants for downstream analyses to those that overlapped across samples.

Gene-level association analysis

Gene-level association analyses were performed with MAGMA v1.08 using the FUMA platform to aggregate the joint effect of genetic variant associations to the level of the gene^{64,65}. A Bonferroni significance threshold was applied by adjusting the alpha-value (0.05) for the number of genes tested for each phenotype.

Estimation of SNP-based heritability

We used GCTA-GREML to estimate the SNP-based heritability of self-reported treatment outcomes using individual-level genotype data⁶⁶. Analyses in the GLAD+ and AGDS datasets were conducted separately. In each dataset, a genomic relatedness matrix was created using unpruned genotype data and adjusted for sample relatedness based on a threshold of > 0.05 and variants were filtered with a minor allele frequency > 0.01 . We included the first 10 principal components and six genotyping batches as covariates. To estimate the SNP-based heritability of self-reported treatment outcomes using summary-level genetic data, Linkage Disequilibrium Score (LDSC) regression was used with default parameters applied⁶⁷. Genetic variants were filtered to those within the HapMap 3 reference panel. SNP-based heritability estimates were transformed to the liability scale, assuming a population prevalence of 41% for response to psychological therapy⁹.

Aim 2: Prediction of retrospectively self-reported treatment outcomes using genetic, sociodemographic and clinical predictors

Creation of polygenic scores

Prediction analyses were conducted in a subset of the GLAD+ dataset in which we had access to both baseline sociodemographic/clinical predictors and individual genetic data ($N = 4,439$). Polygenic scores were derived using the GenoPrep pipeline, a reference-standardised framework using a common set of genetic variants, linkage disequilibrium estimates, and allele frequency estimates to aid reproducibility⁶⁸. We created scores using MegaPRS⁶⁹, which

applies the LDAK heritability model⁷⁰. MegaPRS has good predictive utility in comparison to other common PRS methods, including *P*-value thresholding and clumping (pT+clump), SBLUP, lassosum, LDpred and SBayesR⁶⁸. We selected the optimal MegaPRS model using the pseudo-validation approach, which estimates the best set of effect size distribution parameters out of 236 that were tested without the need for an external validation sample^{68,71}.

We created 27 polygenic scores representing genetic predisposition to a range of psychiatric, cognitive and socio-demographic traits related to psychological treatment outcomes. A list of previously published GWAS summary statistics and trait SNP-based heritabilities (h^2_{SNP}) corresponding to each polygenic score is given in **Supplementary Table S3**. To ensure polygenic scores were derived from GWAS with good statistical power, we only selected traits with an LDSC SNP-based heritability z-score greater than 5, calculated as the SNP-based heritability estimate divided by its standard error⁷².

Prediction modelling

We created three prediction models to assess whether the inclusion of polygenic scores improved the prediction of retrospectively self-reported treatment outcomes over and above previously associated sociodemographic/clinical predictors. Prior to analyses, the first 10 principal components and genotyping batches were regressed out of the polygenic scores, and the residuals were used as predictors. All predictors were scaled and centred based on the mean and standard deviation of the sample, using the 'scale' function in R. Model development and evaluation was carried out in R using the Model Builder format from the GenoPred pipeline^{68,71}, which incorporates functions from the 'glmnet' and 'caret' packages^{73,74}.

Prediction models were created using logistic elastic net regression, which incorporates the ridge regression penalty into Least Absolute Shrinkage and Selection Operator (LASSO) regularisation to reduce model complexity and prevent overfitting⁷⁵. LASSO regularisation applies the L1 penalty, which shrinks the size of all coefficients based on the sum of the absolute values, resulting in some coefficients being reduced to zero and dropped from the model. This method allows for variable selection, typically retaining only one predictor from groups of highly correlated variables⁷⁶. In ridge regression, an L2 penalty is applied to the sum of *squared* coefficients, shrinking the value of all coefficients without dropping any from the model. This method may be more suited to models with correlated predictors⁷⁷. By combining these penalties, shrinkage is applied to coefficients to enable variable selection while allowing correlated predictors to be retained in the model. The tuning parameter lambda (λ) determines the degree of the penalty, with values closer to 1 indicating greater shrinkage. Alpha (α)

determines the relative weightings of the L1 and L2 penalties, with values of alpha closer to 1 favouring the L1 penalty⁷³.

Parameter tuning and predictive utility was determined using nested 10-fold cross-validation⁷⁸. Nested 10-fold cross-validation occurs in loops two levels (for an illustration, see **Supplementary Figure S1**). At the outer level, data are randomly split into training and test sets. The model is developed in 80% of the sample (the “training” set), and the performance of the model is evaluated in the remaining 20% (the “test” set). Nested within the training set, ten further folds are used at the inner level to tune the λ parameter through cross-validation, which determines the degree of the penalty by maximising the predictive ability of the model in “unseen” data. In nine of the inner training data folds, 100 values of λ are assessed to identify the value with the greatest reduction in prediction error for the tenth “validation” fold. This is repeated ten times, with each inner fold of the training data acting as the validation fold once. Through this process, the optimal value of λ is identified in the training set. The performance of this model is then evaluated at the outer level in the unused 20% test set originally held out of the data. Observed outcome values are regressed on the values predicted by the model to determine the proportion of variance in the outcome explained by the prediction model (R^2). The process occurs iteratively, with distinct 20% chunks of the original data, each acting as the independent test set once. The power of this nested approach comes from the full sample contributing to both parameter tuning via cross-validation and model performance evaluation. As each individual in the data forms part of the test set once, they each have one predicted value of the outcome, estimated from the model derived from the inner cross-validation that they were held out of.

Three models were created and compared: (1) genetic predictors only, (2) sociodemographic/clinical predictors only and (3) sociodemographic/clinical and genetic predictors. The relative predictive utility of the models was assessed by comparing correlations between the observed and predicted values of the outcome, using the Hotelling-Williams test as implemented in the ‘psych’ R package⁷⁹; P-values were calculated under a two-sided test. As we pre-selected a limited number of previously associated sociodemographic/clinical predictors, this analysis aimed not to create a model with the highest possible predictive utility but to assess whether the inclusion of genetic predictors increased predictive power over and above known demographic and clinical predictors. Hence we have refrained from interpreting the predictive utility of models including pre-selected sociodemographic/clinical predictors, which may have been prone to overfitting, and have focused primarily on model comparisons.

Results

Sample descriptives

Descriptive statistics for each sample are given in **Table 2**. Both samples were on average middle-aged and were primarily female. The most common education level across samples was having a University degree or higher, with most participants being employed. In both samples, ~75% of participants self-reported positive outcomes following treatment.

Table 2. Descriptive table for GLAD+ and AGDS datasets, total N = 15,131

Study	GLAD+	AGDS
Sample size	7,536	7,595
Age	43.82 (15.1)	47.95 (14.9)
Sex		
Female	5932 (78.7%)	5848 (77.0%)
Male	1603 (21.3%)	1745 (23.0%)
Missing	1	2
Education		
University degree or higher ¹	4339 (57.6%)	5062 (66.7%)
A-levels/AS-levels or equivalent ²	1514 (20.1%)	468 (6.2%)
NVQ/HND/HNC or equivalent ³	607 (8.1%)	1675 (22.1%)
GCSEs/CSEs/O-levels or equivalent ⁴	824 (10.9%)	366 (4.8%)
None of the above	140 (1.9%)	21 (0.2%)
Missing	112	3
Employment status		
In paid employment or self-employed	4663 (61.9%)	4262 (56.2%)
Full or part-time student	575 (7.6%)	452 (6.0%)
Unable to work because of sickness or disability	739 (9.8%)	515 (6.8%)
Retired	907 (12.0%)	1186 (15.6%)
Other (looking after home and/or family, doing unpaid or voluntary work)	337 (4.5%)	412 (5.4%)
Unemployed	252 (3.3%)	251 (3.3%)
None of the above	47 (0.6%)	505 (6.7%)
Missing	16	12
Retrospectively self-reported outcomes following psychological treatment		
Reported benefit/improvement	5693 (75.5%)	5715 (75.2%)
No reported benefit/improvement	1843 (24.5%)	1880 (24.8%)

Descriptives given as Mean (SD) or N (%)

¹Qualifications are taken one year earlier in Australia, so the first year of University is typically equivalent to Year 13 in the UK, for a 4 year bachelor's degree

²Years 11-12 in Australia

³Certificate or diploma, a vocational qualification from TAFE (Technical and Further Education)

⁴Years 8-10 in Australia

Aim 1: Genome-wide association study meta-analysis of retrospectively self-reported treatment outcomes

In both datasets, no significant genetic variants were identified at the level of genome-wide significance ($P < 5 \times 10^{-8}$) (**Supplementary Figures S2 and S3**). We then meta-analysed the results from the GWAS in both samples in METAL ($N = 15,131$; 75% reported positive treatment outcomes). Approximately 6 million genetic variants were shared across the GLAD+ and AGDS datasets. No genetic variants were identified as associated with positive treatment outcomes at the level of genome-wide significance (**Figure 1**). We found little evidence for confounding, with the quantile-quantile plot of the meta-analysed GWAS result showing no inflation (**Figure 1**).

Gene-level association analysis

No genes were significantly associated with self-reported treatment outcomes after Bonferroni correction for the number of genes tested in each GWAS (GLAD+: $\alpha = 0.05 / 18581$ genes = $P < 2.69 \times 10^{-6}$; AGDS: 18656 genes = $P < 2.68 \times 10^{-6}$; Meta-analysis: 18844 genes = $P < 2.65 \times 10^{-6}$).

SNP-based heritability

We did not identify an SNP-based heritability of retrospectively self-reported treatment outcomes that was significantly different from zero using LDSC regression and GCTA-GREML (**Table 3** and **Table 4**). GCTA-GREML power calculations⁸⁰ indicate that the sample size in each dataset (in which individual-level data was available) had a minimum of 80% power to detect an SNP-based heritability of 25%, assuming a population prevalence of 41% for response to psychological therapy⁹. As we did not detect a significant SNP-based heritability, this suggests that the SNP-based heritability of this phenotype is less than 25%.

Table 3. LDSC SNP-based heritability estimates (using summary-level data) in GLAD+ and AGDS datasets, total N = 15,131

Dataset	h^2_{SNP} Observed	h^2_{SNP} SE	Z	h^2_{SNP} Liability
GLAD+	0.04	0.07	0.57	0.08
AGDS	-0.05	0.06	-0.87	-0.11
Meta-analysis	0.01	0.03	0.34	0.02

Table 4. GCTA-GREML SNP-based heritability estimates (using individual-level data) in GLAD+ and AGDS datasets

Dataset	h^2_{SNP} Liability	h^2_{SNP} SE	P-value	N
GLAD+	0.003	0.13	0.49	7442
AGDS	0.027	0.06	0.32	7595

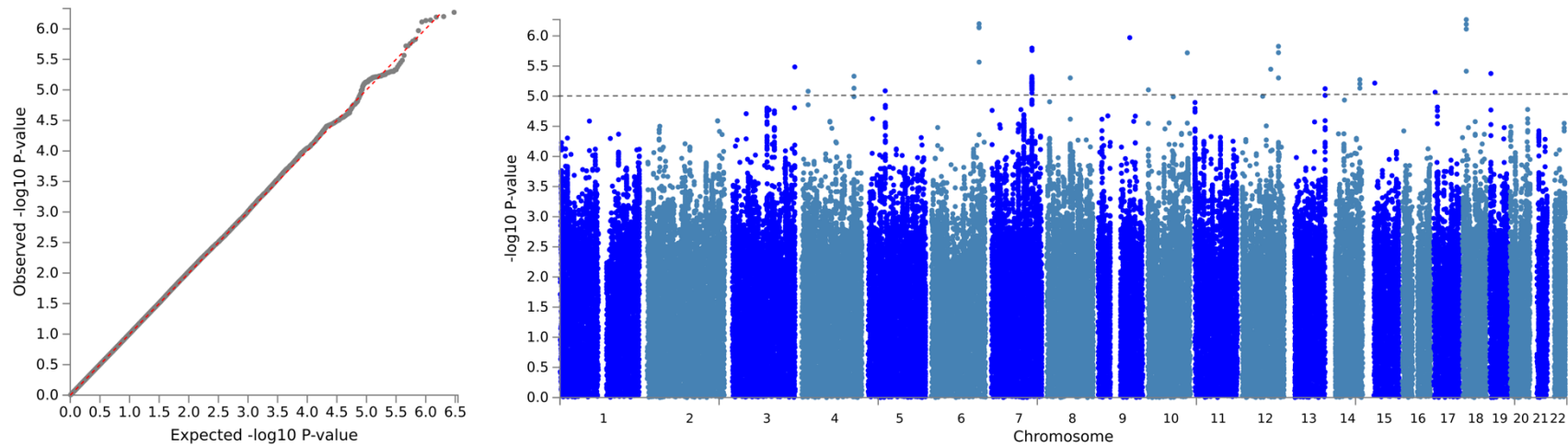


Figure 1: Q-Q plot (left) and Manhattan plot (right) of associations from a meta-analysis of retrospectively self-reported treatment outcomes following psychological therapy (N = 15,131; 75% reported positive outcomes).

In the Manhattan plot, dashed line indicates suggestive threshold ($P < 1 \times 10^{-5}$) for genetic variants associated with retrospectively self-reported treatment outcomes.

Aim 2: Prediction of retrospectively self-reported treatment outcomes using genetic, sociodemographic and clinical predictors

These analyses included a sub-sample of participants from the GLAD+ dataset who had provided both genetic data for polygenic score creation and data on sociodemographic and clinical predictors (N = 4,439). Univariable associations with each predictor are provided in **Table 5**. Education level, number of episodes and polygenic scores for ADHD, self-reported health and age at first birth were significantly associated with retrospectively self-reported treatment outcomes at Bonferroni corrected threshold ($P \leq 0.0017$).

Elastic net models were estimated using nested 10-fold cross-validation to assess whether the inclusion of polygenic predictors improved the prediction of retrospectively self-report treatment outcomes over and above previously associated sociodemographic/clinical predictors. Variance explained on the liability scale (R^2) was determined using a prevalence of 41% for response to psychological therapy⁹.

Polygenic scores explained 0.4% of the variance in retrospectively self-reported treatment outcomes (**Table 6**; $P = 0.007$). Sociodemographic and clinical predictors previously associated with retrospectively self-reported treatment outcomes in this sample (sex, number of episodes, University level education) explained 3.5% of the variance in retrospectively self-reported treatment outcomes (**Table 6**; $P = 2.1 \times 10^{-16}$). Including polygenic scores in the model of sociodemographic/clinical predictors did not significantly improve the predictive utility of the model, as assessed by the correlation between the observed and predicted phenotype (R; P -value of difference in R = 0.409; **Supplementary Table S4**). Coefficients for the predictors retained in the model estimated using the optimal λ tuning parameter derived from 10-fold cross-validation are given in **Supplementary Table S5**.

Table 5. Univariable associations between predictors and retrospectively self-reported treatment outcomes in the GLAD+ dataset (N = 4,439)

Predictor	Univariable β (SE)	P-value	R^2_{Observed}	$R^2_{\text{Liability}}$
<i>Sociodemographic</i>				
Sex	0.027 (0.04)	0.456	0.000	0.000
University degree*	0.172 (0.04)	3.2x10-6	0.005	0.012
<i>Clinical</i>				
Number of episodes*	-0.278 (0.04)	4.3x10-14	0.013	0.031
<i>Genetic (polygenic score)</i>				
ADHD*	-0.124 (0.04)	0.001	0.003	0.006
Self-reported health*	0.121 (0.04)	0.001	0.002	0.006
Age at first birth*	0.119 (0.04)	0.001	0.002	0.005
Body mass index	-0.084 (0.04)	0.024	0.001	0.003
Depressive symptoms	-0.085 (0.04)	0.023	0.001	0.003
Educational attainment	0.097 (0.04)	0.009	0.002	0.004
Daily alcohol use	0.012 (0.04)	0.751	0.000	0.000
Anorexia nervosa	0.005 (0.04)	0.884	0.000	0.000
Anxiety	0.031 (0.04)	0.405	0.000	0.000
Autism	-0.024 (0.04)	0.511	0.000	0.000
Bipolar disorder	0.027 (0.04)	0.468	0.000	0.000
Sleep duration	0.072 (0.04)	0.053	0.001	0.002
Cross disorder	-0.07 (0.04)	0.061	0.001	0.002
Major depressive disorder (PGC2, no 23andMe)	-0.036 (0.04)	0.335	0.000	0.000
Broad depression	0.017 (0.04)	0.640	0.000	0.000
Diabetes Type 2	0.004 (0.04)	0.912	0.000	0.000
Household income	0.031 (0.04)	0.410	0.000	0.000
Insomnia	-0.071 (0.04)	0.055	0.001	0.002
Intelligence	0.046 (0.04)	0.218	0.000	0.001
Memory	-0.022 (0.04)	0.553	0.000	0.000
Neuroticism	0.024 (0.04)	0.511	0.000	0.000
Obsessive compulsive disorder	-0.027 (0.04)	0.471	0.000	0.000
Physical activity	0.023 (0.04)	0.543	0.000	0.000
Post-traumatic stress disorder	-0.045 (0.04)	0.223	0.000	0.001
General risk tolerance	0.03 (0.04)	0.426	0.000	0.000
Schizophrenia	0.008 (0.04)	0.836	0.000	0.000
Ever smoker	-0.047 (0.04)	0.207	0.000	0.001

*Bold indicates significance at $p < .05$, *indicates significance at bonferroni corrected threshold $p < 0.0017$ ($0.05 / 30$ tests). β = logistic regression coefficient denotes the expected change in log odds of having the outcome per unit change in the predictor. Hence, one standard deviation increase in a polygenic score increases the odds of reporting favourable treatment outcomes by a factor of e^β . SE = standard error. R^2_{Observed} = Variance explained on the observed scale. $R^2_{\text{Liability}}$ = Variance explained on the liability scale.*

Table 6. Models predicting self-reported outcomes following psychological therapy for depression/anxiety using genetic, sociodemographic and clinical predictors in the GLAD+ dataset ($N = 4,439$)

Model	R	R SE	R P -value	R^2
Genetic predictors	0.041	0.015	0.007	0.004 (0.4%)
Sociodemographic and clinical predictors	0.123	0.015	2.1×10^{-16}	0.035 (3.5%)
Sociodemographic, clinical and genetic predictors	0.115	0.015	1.3×10^{-14}	0.031 (3.1%)

R = Pearson correlation between model predictions and outcome. R^2 = Variance explained on the liability scale.

Sensitivity analyses

To assess the impact of combining the binary phenotype from the CIDI-SF with the ordinal option treatment history scale, we ran sensitivity analyses with a more homogenous phenotype using only responses from the depression and generalised anxiety CIDI-SF modules in the GLAD+ dataset. Using this phenotype did not influence the significance of the SNP-based heritability estimate using GCTA-GREML ($h^2_{\text{Liability}} = 0.36$; $h^2 \text{ SE} = 0.26$; $P = 0.087$; $N = 4,280$), and did not change the pattern of the results from the prediction model (R for genetic model = 0.044; R for sociodemographic/clinical model = 0.122; R for full model = 0.112; P -value of difference in R = 0.334; $N = 3,843$).

Two sensitivity analyses were run to assess whether current student status was confounding associations with education level in the prediction analysis. First, participants who reported being a full-time or part-time student or who were below the age of 25 were excluded ($N = 532$). The aim of this analysis was to create a subsample of participants in which the University degree variable captured likely lifetime educational attainment. These exclusions did not change the pattern of the results (R for genetic model = 0.034; R for sociodemographic/clinical model = 0.108; R for full model = 0.094; P -value of difference in R = 0.232; $N = 3,907$). Second, education level was removed as a predictor. This also did not change the pattern of the results (R for genetic model = 0.041; R for sociodemographic/clinical model = 0.110; R for full model = 0.107; P -value of difference in R = 0.716; $N = 4,439$).

Discussion

This study presents the largest genetic investigation of psychological treatment outcomes to date (N = 15,131). Two approaches were used to investigate genetic influences on retrospectively self-reported outcomes following psychological treatment for anxiety and depression. First, we performed a genome-wide association study (GWAS) meta-analysis of self-reported treatment outcomes to assess whether this phenotype has a heritable basis that common genetic variants can capture (i.e. SNP-based heritability). Second, we evaluated whether genetic influences on psychiatric, cognitive and related traits improve the prediction of self-reported treatment outcomes over and above known sociodemographic and clinical predictors.

No common genetic variants associated with self-reported treatment outcomes reached genome-wide significance in our GWAS meta-analysis. The benefit of brief phenotyping approaches may only be seen in much larger scale analyses, such that the sample size outweighs the imprecision of the phenotype. Previous genomic analyses using prospective indicators of treatment outcomes did not find genome-wide significant associations in samples of 980 and 2,724 individuals, respectively^{40,41}. Power calculations based on these findings indicated that sample sizes of ~4500 individuals would be necessary to achieve 80% power to detect a SNP-based heritability of 20%, when treatment outcomes are defined as a continuous measure of change in symptom severity from start to end of treatment. However, brief phenotyping approaches result in a lower SNP-based heritability than symptom-based measures. For depression, lifetime cases defined using diagnostic algorithms have a SNP-based heritability of 26%. In comparison, cases derived from single-item measures assessing self-reported diagnoses or seeking treatment for depressive symptoms have lower SNP-based heritabilities of 11% and 13%, respectively⁸¹. These results are attributed to brief phenotyping approaches capturing only a proportion of the genetic liability to depression^{81,82}. Hence, it is reasonable to assume that the SNP-based heritability of self-reported outcomes following psychological therapy would be lower than that of prospective indicators of treatment outcomes.

We had approximately 80% power to estimate a SNP-based heritability of 25% in GCTA-GREML. Estimates were calculated separately in each of the two datasets due to the availability of individual-level genetic data, with a 3:1 case-control ratio in 7,536 and 7,596 individuals in the GLAD+ and AGDS datasets, respectively. Therefore, in line with brief phenotyping of other psychiatric traits, the results of this study indicate that the SNP-based heritability of self-reported outcomes following psychological therapy as a binary measure is lower than 25%. To detect a SNP-based heritability of 10%, these datasets had only

approximately 20% power. To achieve 80% power, individual-level genetic data on 18,600 participants would be needed, corresponding to 2.5 times the size of the current GLAD+ dataset. These sample sizes are not inconceivable in psychiatric genetic research, in which sample sizes with individual-level genetic data and brief measures on depression have exceeded 300,000^{44,45}. With the ongoing recruitment of individuals to studies collecting self-reported treatment outcome data, we anticipate such sample sizes will not take too long to reach.

These conclusions, however, raise questions over the clinical importance of identifying genetic influences that only amount to 10-24% (or less) of the SNP-based heritability of self-reported treatment outcomes. The aim of detecting these variants would be to create a polygenic score that captures genetic propensity to favourable outcomes following psychological treatment. This polygenic score could be utilised in predicting treatment outcomes to help allocate resources to those who are most likely to benefit or provide additional support to those with low propensity to benefit from standard psychological therapy alone. The utility of this polygenic score can be hypothesised from comparisons to related traits. For depression, a polygenic score derived from a GWAS with a SNP-based heritability estimate of 9% explained 1.5-3.2% of the variance (R^2) in diagnoses in independent samples⁴⁸. To put this into context of existing predictors, in our sample the strongest clinical predictor, number of disorder episodes, explained 3.1% of the variance and the strongest demographic predictor, University-level education, explained 1.2% of the variance. These estimates indicate that a polygenic score of self-reported treatment outcomes could explain a comparable proportion of the variance in treatment outcomes to existing sociodemographic/clinical predictors.

Currently, prediction models incorporating the best known clinical and sociodemographic predictors explain ~17% of the variance in outcomes assessed broadly across psychological and antidepressant treatments for depression in primary care³¹. However, the predictive power of models assessing psychological treatment outcomes specifically are less well quantified. If a polygenic score of psychological treatment outcomes could contribute an additional 1.5% to the variance explained above these existing predictors, this would represent a 9% increase in predictive utility. One additional advantage of polygenic scores in comparison to clinical predictors is that the genetic liability that they capture is consistent across life. Hence, once genetic scores are obtained for an individual, this information can be used across numerous instances without any additional burden of repeated data collection.

Consistent with previous models predicting remission following iCBT for depression, we found that higher polygenic scores for neurodevelopmental disorders and lower scores for cognitive traits were associated with poorer self-reported treatment outcomes in the univariable

models^{52,53}. In addition to the six phenotypes that have previously been investigated, which included major depressive disorder, bipolar disorder, ADHD, autism spectrum disorder, intelligence and educational attainment⁵², we tested associations with polygenic scores for a range of other psychiatric, personality, health and demographic traits. We also found significant positive associations with polygenic scores for self-reported health and age at first birth, which are compatible with the shared polygenic nature of these traits with phenotypes associated with treatment outcomes, including educational attainment, ADHD and depression symptoms^{83,84}. Interestingly, both traits also display strong phenotypic and genetic relationships with longevity, indicating a role of psychological well-being in general health status throughout life^{85,86}. As these associations are based on univariable analyses, they may be inflated as they are not mutually adjusted for related traits that share genetic overlap. Nevertheless, polygenic scores for ADHD, depressive symptoms, broad depression, educational attainment, intelligence, self-reported health and age at first birth were all retained in the optimal model derived from 10-fold cross-validation, indicating that they independently contribute to the prediction of self-reported treatment outcomes (**Supplementary Table S5**). Before strong conclusions can be drawn about the overall predictive utility of these models, external validation in independent samples is essential³³.

Despite these associations, the inclusion of polygenic scores did not significantly improve the prediction of self-reported treatment outcomes over previously associated sociodemographic/clinical predictors. This is in contrast with the prediction of remission following iCBT, in which genetic, demographic and clinical predictors displayed comparable importance to prediction⁵³. There are several possible explanations for these differences. First, the definition of outcomes predicted in these two models differed, and therefore heritable traits may not contribute to their classification in the same way. For iCBT, remission was defined as a score ≤ 10 following treatment on the Montgomery-Åsberg Depression Rating Scale-Self report (MADRS-S)⁸⁷. This phenotype is more objective than self-reported treatment outcomes, which may be more greatly influenced by external factors, such as the therapist-client relationship, that influence the overall perception of treatment as beneficial, independently of whether or not symptoms improved.

A second plausible explanation is that the inclusion of several treatment types in the current study masked associations with genetic predictors. As well as including both individual and group therapies, treatment types included CBT, counselling, guided self-help, relationship or family therapy and internet-based therapies (**Supplementary Table S2**). This contrasts with past research exclusively assessing outcomes of iCBT, which followed a rigid structure with regulated module content, duration and therapist contact⁸⁸. Although different forms of

psychological treatment are found, on average, to be equally effective⁸⁹, they may target different mechanisms or processes. For example, CBT is centred around cognitive restructuring and behavioural exposure with a focus on changing maladaptive thoughts and behaviours to improve symptoms and functioning. Conversely, counselling involves a discussion between the therapist and client about life events, emotions, reactions and behaviours, with the therapist challenging the client to understand themselves better and develop coping strategies⁹⁰. Hence, the heritable traits that influence the outcomes of one treatment may have no association, or even the opposite association, with other treatment types. Consistent with this possibility, CBT has been shown to be more effective than counselling for individuals with shorter duration of depression, higher expectations of treatment, and who were taking concurrent antidepressants⁹¹. It is plausible that combining these different treatment types limited the ability to identify associations with heritable traits associated with the outcomes of one treatment specifically.

A final possibility is that the differences in the predictive utility of polygenic scores arise from differences in the samples across these two studies. Previous studies finding a meaningful contribution of genetic factors to the prediction of iCBT outcomes have included individuals receiving treatment as part of the public psychiatric care services in Stockholm, including those referred by their general practitioner or through online self-referrals⁵². These individuals were likely to represent a range of presentations of depression, as demonstrated through differing rates of severity indicators. Only 23% of the sample had a comorbid disorder, 57.5% were currently taking psychotropic medication, and approximately half reported recurrent episodes. This is in contrast to the datasets included in the current study, the majority of whom volunteered to be part of research on the basis of their lifetime experience of depression or anxiety. In the GLAD study and AGDS, participants broadly display similar profiles. The majority of participants have a comorbid psychiatric disorder, over 90% report taking antidepressant medication, and almost all participants report numerous episodes. This highlights that these datasets are enriched for those with severe and recurrent psychiatric disorders, who are likely to have similarly high genetic predisposition to depression, anxiety and related traits. This results in less variation between those who report favourable and negative treatment outcomes, reducing power to identify factors associated with different outcomes. This would suggest that datasets recruited on the basis of lifetime disorder may not be the most suited to this type of research, and their use in combination with less selective naturalistic samples of individuals seeking treatment may provide greater power to identify predictors of outcomes.

In addition to increasing sample size, there are several key considerations for future research relating to the limitations of existing work in this field. First, this study used a combination of two self-reported single-item questions, assessing *benefit from therapy* and *improvement in symptoms/functioning*. This approach was taken to maximise sample size in the GLAD+ dataset and facilitate meta-analysing with AGDS. However, these two types of questions may capture different aspects of treatment response. Using a consistent item in future may help to reduce heterogeneity and increase power. Additionally, the brief phenotyping items used have not been validated against prospective measures of symptom improvement or change in case status. Validation of these items may help to elucidate the brief phenotyping assessments that most closely resemble traditionally assessed treatment outcomes. Second, ~75% of participants in the current sample reported positive treatment outcomes. The imbalance of this “case-control” ratio may indicate that those with perceived negative outcomes are not well represented in these samples. This may also reflect the self-selecting nature of participants who have volunteered to engage with research about their experiences of depression and anxiety. Additionally, imbalanced ratios can result in problems applying statistical methods⁹². More balanced samples may arise from greater recruitment from naturalistic treatment settings or could be achieved by oversampling those with negative outcomes.

Third, our ability to assess whether polygenic scores increase the prediction of self-reported treatment outcomes over and above known predictors was limited by the availability of specific sociodemographic and clinical variables in these datasets. Other factors previously associated with treatment outcomes, such as baseline symptoms, functional impairment and treatment factors, were not included in this analysis due to either the absence of their assessment or inconsistent assessments across studies. Although this limitation is unlikely to have altered the conclusions that polygenic scores did not significantly improve prediction, it does impact the ability to estimate the maximum variance explained in psychological treatment outcomes by a best fitting model including all identified predictors. In addition, clinical predictors were selected to reflect the entire period of illness rather than the specific episode for which treatment was sought⁵⁰. Hence, it is possible that the lifetime assessment of the number of episodes confounded associations with self-reported treatment outcomes. Nevertheless, this association is also demonstrated when outcomes are assessed using traditional prospective measures of symptom change¹⁷. Finally, issues of diversity and generalisability are evident in all genetic investigations, including only participants of predominantly European ancestry. These issues are particularly pertinent in studies of treatment outcomes, given the disparities in access and receipt of mental health care for minoritised groups in the UK⁹³. Research centred around improving outcomes based only on European-associated genetic ancestry

data may exacerbate these inequalities in treatment seeking and receipt, which already disproportionately affect minoritised groups.

These analyses have several implications for future investigations of genetic influences on psychological treatment outcomes. The results indicate that consistent with brief phenotyping approaches of related psychiatric traits, identifying the SNP-based heritable component of retrospectively self-reported treatment outcomes will require larger sample sizes than were reached through the current meta-analysis. To achieve this, increased collection of data relating to psychological treatment outcomes as part of studies that assess mental health history will be vital. Polygenic scores of related psychiatric, cognitive and demographic traits explained a small proportion of variance in self-reported treatment outcomes. This supports the hypothesis that a heritable component to self-reported treatment outcomes exists, warranting further investigation in larger samples. Future work should focus on validation of different items for briefly phenotyped psychological treatment outcomes and investigate whether polygenic scores have greater predictive utility to differentiate between positive and negative outcomes within specific treatment types.

References

1. Steel, Z. *et al.* The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int. J. Epidemiol.* **43**, 476–493 (2014).
2. Kessler, R. C. *et al.* Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593–602 (2005).
3. Wakefield, S. *et al.* Improving Access to Psychological Therapies (IAPT) in the United Kingdom: A systematic review and meta-analysis of 10-years of practice-based evidence. *Br. J. Clin. Psychol.* **60**, 1–37 (2021).
4. Clark, D. M. Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annu. Rev. Clin. Psychol.* **14**, 159–183 (2018).
5. Cuijpers, P., Karyotaki, E., de Wit, L. & Ebert, D. D. The effects of fifteen evidence-supported therapies for adult depression: A meta-analytic review. *Psychother. Res.* **30**, 279–293 (2020).
6. Cipriani, A. *et al.* Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults With Major Depressive Disorder: A Systematic Review and Network Meta-Analysis. *Focus* **16**, 420–429 (2018).
7. Slee, A. *et al.* Pharmacological treatments for generalised anxiety disorder: a systematic review and network meta-analysis. *Lancet* **393**, 768–777 (2019).
8. Hunot, V., Churchill, R., de Lima, M. S. & Teixeira, V. Psychological therapies for generalised anxiety disorder. *Cochrane Database of Systematic Reviews* Preprint at <https://doi.org/10.1002/14651858.cd001848.pub3> (2006).
9. Cuijpers, P. *et al.* The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatr. Scand.* **144**, 288–299 (2021).
10. Loerinc, A. G. *et al.* Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clin. Psychol. Rev.* **42**, 72–82 (2015).
11. Marcus, S. C. & Olfson, M. National trends in the treatment for depression from 1998 to 2007. *Arch. Gen. Psychiatry* **67**, 1265–1273 (2010).
12. Blane, D., Williams, C., Morrison, J., Wilson, A. & Mercer, S. Cognitive behavioural therapy: why primary care should have it all. *British Journal of General Practice* vol. 63 103–104 Preprint at <https://doi.org/10.3399/bjgp13x663235> (2013).
13. Buckman, J. E. J. *et al.* The contribution of depressive ‘disorder characteristics’ to determinations of prognosis for adults with depression: an individual patient data meta-analysis. *Psychol. Med.* **51**, 1068–1081 (2021).
14. Fournier, J. C. *et al.* Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *J. Consult. Clin. Psychol.* **77**, 775–787 (2009).
15. Hölzel, L., Härter, M., Reese, C. & Kriston, L. Risk factors for chronic depression--a systematic review. *J. Affect. Disord.* **129**, 1–13 (2011).
16. Lorenzo-Luaces, L., Rodriguez-Quintana, N. & Bailey, A. J. Double trouble: Do symptom severity and duration interact to predicting treatment outcomes in adolescent depression? *Behav. Res. Ther.* **131**, 103637 (2020).
17. Buckman, J. E. J. *et al.* Risk factors for relapse and recurrence of depression in adults and how they operate: A four-phase systematic review and meta-synthesis. *Clin. Psychol. Rev.* **64**, 13–38 (2018).
18. Delgadillo, J., Huey, D., Bennett, H. & McMillan, D. Case complexity as a guide for psychological treatment selection. *J. Consult. Clin. Psychol.* **85**, 835–853 (2017).
19. Delgadillo, J., Moreea, O. & Lutz, W. Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. *Behav. Res. Ther.* **79**, 15–22 (2016).

20. Saunders, R., Cape, J., Fearon, P. & Pilling, S. Predicting treatment outcome in psychological treatment services by identifying latent profiles of patients. *J. Affect. Disord.* **197**, 107–115 (2016).
21. El Alaoui, S. *et al.* Predicting Outcome in Internet-Based Cognitive Behaviour Therapy for Major Depression: A Large Cohort Study of Adult Patients in Routine Psychiatric Care. *PLoS One* **11**, e0161191 (2016).
22. Hollon, S. D. *et al.* Effect of Cognitive Therapy With Antidepressant Medications vs Antidepressants Alone on the Rate of Recovery in Major Depressive Disorder. *JAMA Psychiatry* vol. 71 1157 Preprint at <https://doi.org/10.1001/jamapsychiatry.2014.1054> (2014).
23. Angstman, K. B. *et al.* Personality Disorders in Primary Care: Impact on Depression Outcomes Within Collaborative Care. *J. Prim. Care Community Health* **8**, 233–238 (2017).
24. Dennehy, E. B., Marangell, L. B., Martinez, J., Balasubramani, G. K. & Wisniewski, S. R. Clinical and functional outcomes of patients who experience partial response to citalopram: secondary analysis of STAR*D. *J. Psychiatr. Pract.* **20**, 178–187 (2014).
25. Newton-Howes, G., Tyrer, P. & Johnson, T. Personality disorder and the outcome of depression: meta-analysis of published studies. *Br. J. Psychiatry* **188**, 13–20 (2006).
26. Renaud, J., Russell, J. J. & Myhr, G. Predicting who benefits most from cognitive-behavioral therapy for anxiety and depression. *J. Clin. Psychol.* **70**, 924–932 (2014).
27. DeRubeis, R. J. *et al.* The Personalized Advantage Index: translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS One* **9**, e83875 (2014).
28. Buckman, J. E. J. *et al.* Role of age, gender and marital status in prognosis for adults with depression: An individual patient data meta-analysis. *Epidemiol. Psychiatr. Sci.* **30**, e42 (2021).
29. Chekroud, A. M. *et al.* Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry* **3**, 243–250 (2016).
30. Skelton, M. *et al.* Trajectories of depression and anxiety symptom severity during psychological therapy for common mental health problems. *Psychol. Med.* 1–11 (2022).
31. Buckman, J. E. J. *et al.* Predicting prognosis for adults with depression using individual symptom data: a comparison of modelling approaches. *Psychol. Med.* 1–11 (2021).
32. O'Driscoll, C. *et al.* The importance of transdiagnostic symptom level assessment to understanding prognosis for depressed adults: analysis of data from six randomised control trials. *BMC Med.* **19**, 109 (2021).
33. Chekroud, A. M. *et al.* The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry* **20**, 154–170 (2021).
34. Tansey, K. E. *et al.* Contribution of common genetic variants to antidepressant response. *Biol. Psychiatry* **73**, 679–682 (2013).
35. Pain, O. *et al.* Identifying the Common Genetic Basis of Antidepressant Response. *Biol Psychiatry Glob Open Sci* **2**, 115–126 (2022).
36. Mojtabai, R. Nonremission and time to remission among remitters in major depressive disorder: Revisiting STAR* D. *Depress. Anxiety* (2017).
37. Bralten, J. *et al.* Genetic underpinnings of sociability in the general population. *Neuropsychopharmacology* **46**, 1627–1634 (2021).
38. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
39. Luciano, M. *et al.* Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism. *Nat. Genet.* **50**, 6–11 (2018).
40. Rayner, C. *et al.* A genome-wide association meta-analysis of prognostic outcomes following cognitive behavioural therapy in individuals with anxiety and depressive disorders. *Transl. Psychiatry* **9**, 150 (2019).

41. Coleman, J. R. I. *et al.* Genome-wide association study of response to cognitive-behavioural therapy in children with anxiety disorders. *Br. J. Psychiatry* **209**, 236–243 (2016).
42. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
43. Sanchez-Roige, S. & Palmer, A. A. Emerging phenotyping strategies will advance our understanding of psychiatric genetics. *Nat. Neurosci.* **23**, 475–480 (2020).
44. Howard, D. M. *et al.* Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.* **9**, 1470 (2018).
45. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
46. Flint, J. The genetic basis of major depressive disorder. *Mol. Psychiatry* (2023) doi:10.1038/s41380-023-01957-9.
47. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668 (2018).
48. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *BioRxiv* 433367 (2019).
49. Hyde, C. L. *et al.* Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics* vol. 48 1031–1036 Preprint at <https://doi.org/10.1038/ng.3623> (2016).
50. Rayner, C. *et al.* 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* **22**, 719 (2022).
51. Button, K. S. *et al.* Minimal clinically important difference on the Beck Depression Inventory - II according to the patient's perspective. *Psychological Medicine* vol. 45 3269–3279 Preprint at <https://doi.org/10.1017/s0033291715001270> (2015).
52. Andersson, E. *et al.* Genetics of response to cognitive behavior therapy in adults with major depression: a preliminary report. *Mol. Psychiatry* **24**, 484–490 (2019).
53. Wallert, J. *et al.* Predicting remission after internet-delivered psychotherapy in patients with depression using machine learning and multi-modal data. *medRxiv* 2021.04.30.21256367 (2021).
54. Davies, M. R. *et al.* The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety. *Behav. Res. Ther.* **123**, 103503 (2019).
55. Young, K. S. *et al.* Depression, anxiety and PTSD symptoms before and during the COVID-19 pandemic in the UK. *Psychol. Med.* 1–31 (2022).
56. Byrne, E. M. *et al.* Cohort profile: the Australian genetics of depression study. *BMJ Open* **10**, e032580 (2020).
57. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
58. Davies, M. R. *et al.* Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts. *J. Anxiety Disord.* **85**, 102491 (2022).
59. Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B. & Wittchen, H.-U. The World Health Organization composite international diagnostic interview short-form (CIDI-SF). *Int. J. Methods Psychiatr. Res.* **7**, 171–185 (1998).
60. Mitchell, B. L. *et al.* The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes. *Biol. Psychiatry* (2021) doi:10.1016/j.biopsych.2021.10.021.
61. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* **53**, 1097–1103 (2021).
62. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).

63. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
64. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
65. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
66. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
67. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* vol. 47 291–295 Preprint at <https://doi.org/10.1038/ng.3211> (2015).
68. Pain, O. *et al.* Evaluation of polygenic prediction methodology within a reference-standardized framework. *PLoS Genet.* **17**, e1009021 (2021).
69. Zhang, Q., Privé, F., Vilhjálmsson, B. & Speed, D. Improved genetic prediction of complex traits from individual-level data or summary statistics. *Nat. Commun.* **12**, 4192 (2021).
70. Speed, D., Holmes, J. & Balding, D. J. Evaluating and improving heritability models using summary statistics. *Nat. Genet.* **52**, 458–462 (2020).
71. Pain, O. & Lewis, C. M. Using Local Genetic Correlation Improves Polygenic Score Prediction Across Traits. *bioRxiv* 2022.03.10.483736 (2022) doi:10.1101/2022.03.10.483736.
72. Krapohl, E. *et al.* Multi-polygenic score approach to trait prediction. *Mol. Psychiatry* (2017).
73. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).
74. Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **28**, (2008).
75. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Series B Stat. Methodol.* **67**, 301–320 (2005).
76. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. (Springer Science & Business Media, 2013).
77. James, G., Witten, D., Hastie, T. & Tibshirani, R. An Introduction to Statistical Learning. *Springer Texts in Statistics* Preprint at <https://doi.org/10.1007/978-1-4614-7138-7> (2013).
78. Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society: Series B (Methodological)* vol. 36 111–133 Preprint at <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x> (1974).
79. Steiger, J. H. Tests for comparing elements of a correlation matrix. *Psychological Bulletin* vol. 87 245–251 Preprint at <https://doi.org/10.1037/0033-2909.87.2.245> (1980).
80. Visscher, P. M. *et al.* Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLoS Genet.* **10**, e1004269 (2014).
81. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nature Genetics* vol. 52 437–447 Preprint at <https://doi.org/10.1038/s41588-020-0594-5> (2020).
82. Kendler, K. S. *et al.* Shared and specific genetic risk factors for lifetime major depression, depressive symptoms and neuroticism in three population-based twin samples. *Psychol. Med.* **49**, 2745–2753 (2019).
83. Mills, M. C. *et al.* Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. *Nat Hum Behav* **5**, 1717–1730 (2021).
84. Harris, S. E. *et al.* Molecular genetic contributions to self-rated health. *Int. J. Epidemiol.* **46**, 994–1009 (2017).
85. Wuorela, M. *et al.* Self-rated health and objective health status as predictors of all-cause mortality among older people: a prospective study with a 5-, 10-, and 27-year follow-up. *BMC Geriatr.* **20**, 120 (2020).

86. Shadyab, A. H. *et al.* Maternal Age at Childbirth and Parity as Predictors of Longevity Among Women in the United States: The Women's Health Initiative. *American Journal of Public Health* vol. 107 113–119 Preprint at <https://doi.org/10.2105/ajph.2016.303503> (2017).
87. Zimmerman, M., Posternak, M. A. & Chelminski, I. Derivation of a definition of remission on the Montgomery–Asberg depression rating scale corresponding to the definition of remission on the Hamilton rating scale for depression. *Journal of Psychiatric Research* vol. 38 577–582 Preprint at <https://doi.org/10.1016/j.jpsychires.2004.03.007> (2004).
88. Hedman, E. *et al.* Effectiveness of Internet-based cognitive behaviour therapy for depression in routine psychiatric care. *J. Affect. Disord.* **155**, 49–58 (2014).
89. Cuijpers, P., van Straten, A., Andersson, G. & van Oppen, P. Psychotherapy for Depression in Adults: A Meta-Analysis of Comparative Outcome Studies. *FOCUS* vol. 8 75–75 Preprint at <https://doi.org/10.1176/foc.8.1.foc75> (2010).
90. Pybis, J., Saxon, D., Hill, A. & Barkham, M. The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: evidence from the 2 UK National Audit of psychological therapies. *BMC Psychiatry* **17**, 215 (2017).
91. Delgadillo, J. & Gonzalez Salas Duhne, P. Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. *J. Consult. Clin. Psychol.* **88**, 14–24 (2020).
92. Dai, X., Fu, G., Zhao, S. & Zeng, Y. Statistical Learning Methods Applicable to Genome-Wide Association Studies on Unbalanced Case-Control Disease Data. *Genes* **12**, (2021).
93. Rayner, C. *et al.* Sociodemographic factors associated with treatment-seeking and treatment receipt: cross-sectional analysis of UK Biobank participants with lifetime generalised anxiety or major depressive disorder. *BJPsych Open* **7**, (2021).

Chapter 5. General discussion

General overview of findings

The overarching aim of this thesis was to use genomic data to better understand genetic contributions to the risk of anxiety and depression and psychological treatment outcomes. **Chapter 2** examined genomic risk factors associated with the clinical heterogeneity of anxiety disorders. **Chapter 3** explored genomic factors associated with vulnerability to experiencing trauma, which is an important socio-environmental risk factor for anxiety and depression. Finally, **Chapter 4** assessed genomic influences on outcomes of positive environmental exposures (i.e. psychological treatment for anxiety or depression). The findings from these studies are now discussed in relation to three key themes that emerged from the results and how these relate to current challenges in anxiety and depression genomics. General limitations and future avenues for genetics research on the risk and treatment of common mental disorders are then discussed.

Chapter 2 investigated the genomic risk factors associated with fear-based disorders versus GAD, two key subtypes that reflect clinical heterogeneity in the anxiety disorder spectrum. Combining detailed and brief diagnostic measures of anxiety disorders improved the power to refine the genomic relationship between fear and GAD and other complex traits. Genetic correlation results supported the utilisation of both a higher-order internalising grouping and lower-order fear and distress subgroupings in research on genetic risk factors. Case ascertainment and control screening criteria influenced the detection of genetic differences between fear and GAD, highlighting the challenge of handling high comorbidity levels in anxiety GWAS. Genetic correlations between fear and GAD were high and did not significantly differ in genetic correlation with several psychiatric traits, including depression, PTSD, OCD, and bipolar disorder type II. The exceptions were neuroticism, bipolar disorder type I and anorexia nervosa, where GAD showed stronger positive genetic correlations with these traits than with fear. This underscores the value of subtype-specific GWAS in understanding the genetic structure across different psychopathologies. Genetic correlations with environmental risk factors, including reported childhood maltreatment, did not differ between fear and GAD. This may indicate shared genetic risk factors, such as heritable behaviours, in gene-environment interplay. Fear and GAD had significant negative genetic correlations with one risk-taking behaviour (automobile speeding) but not with other risk-taking behaviours. Gene-environment interplay was also indicated by an overlap in genome-wide significant loci and genetic correlations with sociodemographic-related traits. Of these phenotypes, fear had stronger negative genetic correlations with educational attainment and general cognitive ability

than GAD. In summary, these findings highlight potential mechanisms linking the similarities and differences between fear and distress by assessing differential genetic correlations across anxiety subtypes and a wide range of complex traits. The GWAS summary statistics from this thesis can be used in future analyses to further explore mechanisms of transdiagnostic and subtype-specific genetic influences on the anxiety disorder spectrum.

Chapters 3 and 4 investigated the genetic component of retrospectively self-reported measures of environmental experiences. **Chapter 3** expands upon previous research demonstrating a significant SNP-based heritability of self-reported trauma, a major socio-environmental risk factor for anxiety and depression. Genomic multiple regression modelling of the heritable component of self-reported childhood maltreatment indicated the role of traits involved in the subjective appraisal of reporting adversity, such as subjective well-being. Risk-taking behaviour, PTSD and ASD were also identified as key traits that capture the SNP-based heritability of reported trauma. These results highlight the value of using self-report measures of adversity to understand the subjective experience of negative exposures and vulnerability to adversity. In contrast, **Chapter 4** did not find a significant SNP-based heritability for brief, retrospectively self-reported outcomes following psychological treatment for anxiety and depression. Furthermore, although some polygenic scores showed significant univariable associations with self-reported treatment outcomes, adding multiple polygenic scores to prediction models did not significantly improve prediction beyond clinical and sociodemographic predictors. However, polygenic scores jointly explained a small, yet significant, proportion of the variance in self-reported psychological treatment outcomes. This suggests that a significant SNP-based heritability may be detected with larger sample sizes. The lack of power to detect a genetic component may reflect noise and heterogeneity associated with the phenotyping strategy to assess psychological treatment outcomes. Brief instead of detailed phenotyping and subjective self-reports rather than objective measures were used. Moreover, the grouping of treatment outcomes of all five anxiety disorders and depression together may have affected noise and heterogeneity. In **Chapter 3**, anxiety and depression overlapped in their genetic influences on reported trauma, possibly capturing a broad internalising trauma-associated psychopathology. Anxiety and depression genetic influences on reported trauma were also shared with other psychiatric disorders included in the multiple regression models, suggesting trauma-related risk factors may reflect the genetics of psychopathology more broadly.

Across **Chapters 3 and 4**, some traits emerged as having genetic associations with both reported adversity and psychological treatment outcomes. Three traits had significant univariable polygenic score associations with psychological treatment outcomes in **Chapter**

4. These traits also had significant bivariate genetic correlations with reported trauma in **Chapter 3** (although this was not independent of other traits included in the multiple regression models). Among these were ADHD, which had one of the strongest positive genetic correlations with reported trauma (**Chapter 3**). A higher polygenic score for ADHD was associated with *worse* psychological treatment outcomes (**Chapter 4**). Sociodemographic-related traits also showed genetic associations in both chapters. In **Chapter 3**, self-reported health, and age at first offspring birth were negatively genetically correlated with reported trauma, while social deprivation and parents' attained age (i.e. lifespan) showed positive genetic correlations. In **Chapter 4**, polygenic scores for higher self-reported health and older age at first birth were associated with *better* psychological treatment outcomes. These findings could potentially suggest some overlap in genetic influences on traits involved in gene-environment interplay in reporting adversity and psychological treatment outcomes.

Three key themes emerge from the broader context of these findings (**Figure 1**). First, in **theme 1**, the strengths and limitations of GWAS using less detailed phenotyping and self-reported measures of environmental exposures are discussed. The effects of the phenotyping strategy can be seen in downstream analyses, such as the genetic structure observed across psychiatric disorders and influences on psychosocial factors. In **theme 2**, the implications of these findings are discussed in the context of the genetic structure of psychopathology and the extent to which disorders can be grouped to increase the statistical power of GWAS. Psychosocial influences may also impact the genetic structures of psychopathology. Finally, in **theme 3**, the possible psychosocial influences involved in gene-environment interplay mechanisms are discussed.

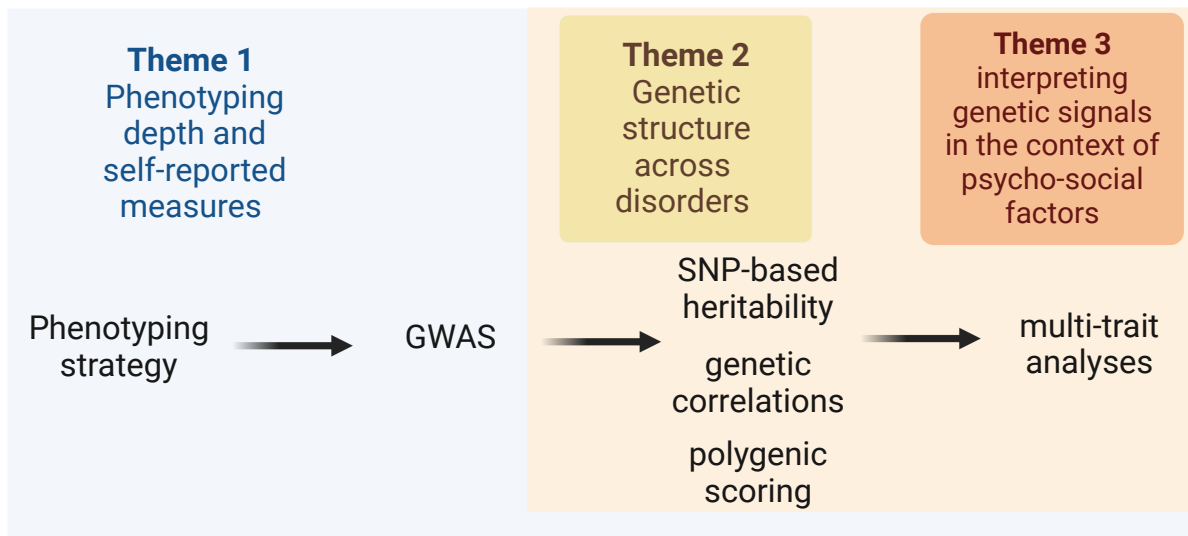


Figure 1: An illustration of the emergence of three key themes based on the findings of this thesis and the way they relate to various stages of GWAS.

Theme 1: Phenotyping depth and the utility of self-reported measures

The phenotyping strategy can have broad implications for interpreting GWAS findings. As outlined in **Chapter 1**, a growing discussion in psychiatric genomics is the balance between the depth of phenotyping and GWAS sample size and biases of self-reported measures. Based on the findings from this thesis and the wider literature, this theme discusses the extent to which brief phenotyping can be used instead of detailed phenotyping or subjective self-report measures in place of objective measures.

In **Chapter 2**, we combined brief and detailed anxiety disorder measures to increase GWAS power to detect genomic influences on clinical heterogeneity. This may have increased broad genetic influences compared to using detailed phenotyping alone. However, our findings likely reflect a balance in specificity and power by combining these two measures. While we could still detect genetic differences between fear and GAD using this approach, notably with neuroticism, the ability to detect distress-fear genetic specificity did not extend to differences in genetic correlations with depression. If we had the sample size to use only the more detailed phenotyping, this may have increased the chance of detecting distress-fear genetic differences. In relation to this, a twin study found that broadly defined depression measures capture approximately a third of the genetic influences identified on more strictly defined, interview-based lifetime MDD¹. This highlights how the precision and detail of phenotyping impact genetic specificity. The SNP-based heritability of more detailed, stringent measures of depression is also higher than briefer, lightly phenotyped measures^{2–5}. A GWAS of MDD in

AGDS using detailed phenotyping reported one locus in a relatively small sample and increased discovery by 23 loci once meta-analysing with the PGC broad depression GWAS. This underscores how GWAS of more heritable, detailed phenotypes (as used in **Chapter 2**) may increase statistical power⁵.

In **Chapter 3**, retrospective self-reports of trauma were used to capture exposure to negative experiences. The findings in the chapter indicate associations with traits likely to influence the subjective appraisal of trauma, such as subjective well-being. Previous research indicates that measures capturing the subjective appraisal of trauma are more strongly associated with psychopathology than measures of objective exposure⁶. Thus, the experience may matter more than the exposure. As subjective appraisal is key for psychopathology, this suggests that using self-reports has value in genomics research of trauma⁷. Of note, the psychiatric disorders identified in **Chapter 3** as independently genetically associated with self-reported trauma (PTSD and ASD) were replicated in an external dataset, also using retrospective self-reports to define trauma⁷. This contrasts with a study of prospectively measured trauma, that found polygenic scores for other psychiatric disorders (ADHD, depression, schizophrenia) to be more strongly associated with exposure than ASD⁸. Thus, the method used to assess trauma phenotypes may potentially determine which psychosocial components and mechanisms can be detected. A future direction in this area of research is to disentangle genetic influences on the subjective appraisal of trauma from influences on vulnerability to exposure. For example, genetic correlation results from GWAS using subjective trauma measures could be compared with those using objective measures.

In contrast, for the studies of genetic factors of psychological treatment outcomes, prospectively assessed, objective, and/or detailed measures may be more appropriate than retrospective self-reports. In **Chapter 4**, analyses were underpowered to detect a significant SNP-based heritability using brief, retrospective self-reports. Furthermore, polygenic scores did not significantly improve the prediction of treatment outcomes, in contrast to a study in a smaller sample that used more objective measures of psychological treatment outcomes⁹. Objective measures may prove more useful in future GWAS to detect a significant SNP-based heritability of psychological treatment outcomes. They are more likely to reduce noise (e.g. less variability in understanding the question) and ensures true exposure to the intervention. Indeed, in pharmacogenomics, measuring drug side effects has shown promise as a phenotyping strategy in GWAS¹⁰, with the added benefit of ensuring objective exposure (i.e. the drug intervention was given). Furthermore, clinically ascertained measures on antidepressant response have more power to detect a significant SNP-based heritability than retrospective self-reports of treatment outcomes^{11–13}. As discussed above in relation to

Chapter 2, this is consistent with phenotyping strategies for depression, where clinically ascertained phenotypes show higher heritability estimates than brief, retrospective self-reports^{2,14}.

The lack of power in **Chapter 4** does not imply that genetic analyses of self-reported positive experiences more broadly are less valuable than objective measures. Twin studies indicate that heritability estimates of the perceived quality of social support may be higher than more objective measures of the quantity of social support¹⁵. Similar to findings on trauma (**Chapter 3**), perceived quality (i.e. the experience) of social support is also more important than quantity (i.e. objective exposure) of social support for protecting against poor mental health. This highlights the importance of subjective appraisal in positive experiences¹⁵. In the case of psychological intervention (**Chapter 4**), the noise that comes with retrospective self-reporting may outweigh the benefits of capturing subjective perceptions of psychological treatment.

In summary, the findings of these chapters partially support the use of brief, retrospective self-reported measures in GWAS. In the context of measures aiming to capture clinical outcomes (e.g. disorder diagnoses, treatment outcomes), objective, detailed measures are preferred, although these may often be unavailable at the scale required for GWAS. For phenotypes aiming to capture the experience of environments, regardless of the exposure, brief, retrospective self-reported measures can capture important genetic influences on perceptions of experiences. This thesis also highlights the need to compare detailed and brief measures of clinical outcomes of psychiatric disorders^{2,14}, much like the need to compare objective and subjective measures of environmental experiences^{6,16}. These comparisons are important for understanding the implications of different strategies on GWAS results. This includes the impact on the power of GWAS but also on detecting genetic structures across psychopathologies and the psychosocial components captured in the GWAS signal. The emergence of medical record linkage in genetic datasets, such as the GLAD study, will provide the opportunity for such comparisons.

Theme 2: the genetic structure of internalising disorders and implications for disorder grouping

An important implication of GWAS results is the evidence it provides for the genetic structure of the psychiatric spectrum. As raised in **Chapter 1**, an aim of modelling covariances between disorders is to understand pervasive comorbidity while considering clinical heterogeneity¹⁷. Such findings have had implications for changing diagnostic criteria. For example, researchers have proposed that classifying anxiety and depressive disorders into higher-order internalising

and lower-order distress-fear dimensions are incorporated into diagnostic systems. Such hierarchical models, like the Hierarchical Taxonomy Of Psychopathology (HiTOP), have been put forward to simplify diagnoses, research and treatment of psychiatric disorders¹⁸. These groupings are often used in etiological research and treatment and, to some extent, have supported the hierarchical model¹⁸. Results from GWAS provide molecular genetic evidence that can validate these hierarchical structures. Although this was not a specific aim of this thesis, as dimensional phenotypes were not tested, the empirical chapter findings can be interpreted in the context of these structures. Furthermore, in GWAS, there is a balance between maximising power by grouping disorders to extract shared components and analysing disorders separately to detect disorder-specific factors. Taking these considerations together, this thesis provides recommendations for future analyses to detect anxiety genetic risk factors, understand trauma as a risk factor for psychopathology, the interplay with genetic risk, and the genetic influences on psychological treatment outcomes.

In **Chapter 2**, there was strong evidence to support a higher-order internalising grouping in understanding genetic risk factors. Given the high genetic overlap between fear, GAD, and depression, this grouping can increase the power to extract shared genetic components. Molecular genetic evidence for shared internalising liability is supported by multivariate GWASs^{19,20}. These studies used Genomic SEM to test the genomic factor structure of traits based on genetic correlations. These factors were then used to improve power for loci discovery and elucidate shared biological pathways^{19,20}. By modelling genetic covariances across anxiety and depression and nine other psychiatric disorders, one study found evidence for a transdiagnostic genomic liability 'p-factor'. However, the 'p factor' had limited utility in elucidating shared biology. Instead, four lower-order factors, including an internalising factor, provided more biological insight. Genetic liability to anxiety and depression were core components of this internalising factor¹⁹. A dimensional approach modelling neuroticism, GAD and depressive symptoms found genomic evidence supporting earlier twin study findings for genetic liability to internalising negative affectivity driving anxiety and depression co-occurrence²⁰. This study also found some genetic distinctions between GAD and depressive symptoms²⁰. Thus, these multivariate approaches highlight the utility of examining both higher- and lower-order dimensions.

Although there was considerable genetic overlap between fear and GAD (**Chapter 2**), there was also evidence to support a lower-order distress-fear grouping to understand genetic risk factors. Fear and GAD were differentially genetically correlated with neuroticism, bipolar disorder and anorexia nervosa. By incorporating more specific distress-fear components of internalising, these findings add to previous molecular genetic structures of the mood disorder

spectrum²¹, as depicted in **Figure 2**. Although differential genetic correlations between fear and GAD were not found in many psychiatric disorders, these findings help address remaining questions on placing some disorders in the hierarchical model. The HiTOP model has been revised several times since its conception²². OCD has recently been placed in the fear subfactor^{18,23}. We found no molecular genetic evidence for a significant difference between fear and GAD genetic correlations with OCD. Similarly, PTSD has been placed within distress, and we found no difference in the genetic overlap between fear and GAD. Analyses of OCD and PTSD GWAS subtypes may reveal differential fear-GAD genetic associations²⁴. In summary, **Chapter 2** showed the potential for using more homogenous internalising phenotypes to test and refine the molecular genetic structure of the psychiatric spectrum.

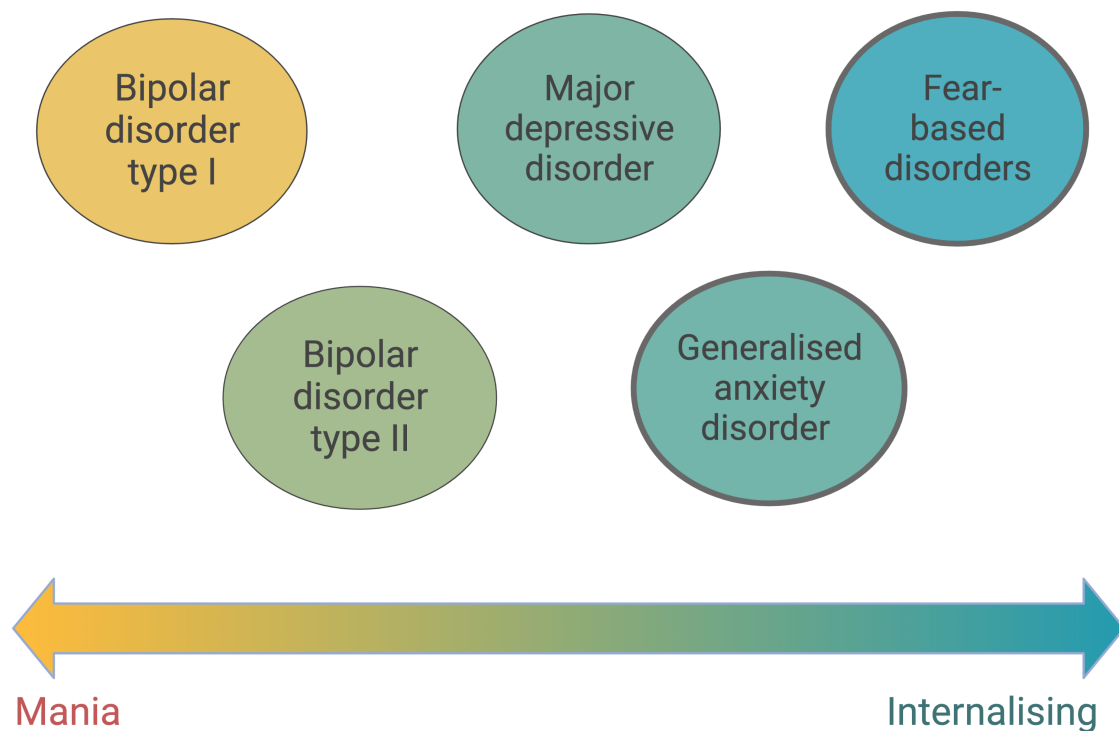


Figure 2: An illustration of the genetic continuum across the mood disorder spectrum to the internalising disorder spectrum, with the addition of fear and GAD.

*The genetic continuum from mania genetics to internalising genetics. The findings from **Chapter 2** add further detail to the findings from Coleman et al.²¹ (see **Chapter 1, Figure 3**), showing distress-fear genetic heterogeneity at the end of the internalising dimension (bold circle borders represent the addition of new findings from **Chapter 2**).*

It is important to account for both shared and unique genetic influences on disorders to differentiate between transdiagnostic and disorder-specific biology. As illustrated in **Chapter 3** with reported trauma, Genomic SEM also enables a way of extracting what is unique to a phenotype by estimating residual genetic variances. This approach could be applied in future

studies using the GWAS summary statistics of fear and GAD from **Chapter 2** to examine transdiagnostic and subtype-specific genetics of internalising. Subtype multivariate modelling would leverage both the higher-order and lower-order dimensions of internalising to increase the power to extract what is common as well as unique. In addition, studies could test the factor structure proposed by the HiTOP model of distress and fear more broadly by incorporating other disorders, such as OCD, PTSD and bipolar disorder subtypes¹⁸. This would provide more granular molecular genetic evidence for the hierarchical structure of internalising disorders and their place on the genetic continuum of psychopathology.

A strength of this thesis was the use of data from both GLAD and AGDS, which have detailed, consistent measures across anxiety and depression, facilitating meta-analysis. Analyses from these datasets also highlighted important issues in case-control ascertainment. The screening criteria used to define cases and controls influenced the ability to detect genetic differences between fear-based disorders and GAD. As such, future studies should consider the effect of ascertainment on detecting subtype-specific influences when designing data collection and analysis. This is especially problematic in studies of anxiety and depression due to their high comorbidity. Indeed, previous GWAS of other psychiatric disorders often broadly screen controls for any psychiatric disorder. Given their high prevalence and comorbidity, this may disproportionately screen controls for anxiety and depression, potentially inflating genetic correlation estimates with internalising disorders²⁵. Publicly available GWAS summary statistics of psychiatric disorders are widely used for downstream analyses, and this should be considered when interpreting results, e.g. the genetic structure of psychopathology. In conclusion, the findings from **Chapter 2** support future research to continue efforts to conduct subtype-specific GWAS. This includes establishing and increasing datasets with detailed measures on the internalising spectrum to achieve the large sample sizes required for disorder-specific GWASs (e.g. the individual fear-based disorders). Such datasets also provide the opportunity to examine dimensional phenotyping, a necessity to adequately test the proposed hierarchical structures of dimensional psychopathology.

The findings from **Chapters 2** and **3** indicate how internalising disorders can be modelled to understand trauma as a shared risk factor and its interplay with genetic risk. In **Chapter 2**, the bivariate genetic correlation with reported trauma did not significantly differ between fear and GAD. As shown in **Chapter 3**, Genomic SEM can be used to model genetic covariation between psychiatric disorders and shared environmental risk factors. The model in **Chapter 3** assessed ‘any’ anxiety disorder instead of looking at fear and GAD separately. When using this phenotype, the genetic relationship between anxiety and reported trauma was shown to be shared with depression. This indicates that similar results may have been observed in

Chapter 3 had fear and GAD been measured separately rather than grouped as ‘any’ anxiety disorder. This suggests that combining anxiety and depression into a broad internalising grouping may be appropriate to understand the internalising factors associated with vulnerability to trauma. This aligns with the HiTOP model, which highlights trauma as a shared transdiagnostic risk factor to the internalising spectrum¹⁸. In contrast to anxiety and depression, findings in **Chapter 3** suggested that some genetic influences of PTSD and ASD on reporting trauma are not shared with the other psychiatric disorders in the model. The potential mechanisms involved in experiencing adverse exposures to anxiety and depressive disorders, PTSD and ASD are discussed below as part of **Theme 3**.

The extent to which internalising disorders can be grouped to detect genetic influences on the psychological treatment outcomes of each disorder remains to be clarified. In **Chapter 4**, the lack of power to detect a genomic component may indicate that grouping all six disorders increased heterogeneity. Future studies may benefit from more homogeneous phenotypes, such as subtype-specific GWAS of treatment outcomes. A significant twin heritability of 15%-36% has been found for fear extinction, a more homogenous phenotype related to the treatment of fear in anxiety disorders^{26,27}. This is in contrast to psychological treatment outcomes for depression and anxiety more broadly, where significant heritability has yet to be found in twin or genome-wide studies. The GWAS dataset in **Chapter 4** largely comprised individuals who had received psychological treatment for depression. We were, therefore, underpowered to conduct fear treatment-specific analyses and detect a SNP-based heritability of response to exposure-based fear treatment. Given the evidence from twin findings of a heritable component of fear extinction, future studies could focus on ascertaining GWAS data on exposure-based fear treatment outcomes. Of note, at the current GWAS sample size, we only had statistical power to detect an SNP-based heritability of 25% for treatment outcomes using individual-level based methods. As twin heritability estimates represent the upper limit of what GWAS data detect²⁸, the SNP-based heritability of fear extinction is likely lower than 25%. Thus, GWAS datasets on exposure-based fear treatment outcomes will also likely need to be larger than the available sample size in **Chapter 4** (~7,550) used for the GWAS of psychological treatment outcomes for depression and anxiety.

Theme 3: interpreting GWAS findings in the context of psychosocial influences

As put forward in **Chapter 1**, environmental effects are not independent of genetics. A portion of the genetic signal from GWAS captures the interplay with the environment²⁹. It has been posited that gene-environment interplay may partly explain the high comorbidity and genetic

covariation observed across psychiatric disorders, contributing to the ‘p-factor’³⁰. For example, genetic influences on experiencing environmental stressors such as trauma may mediate the genetic covariation across psychiatric disorders³⁰. In **Chapter 3**, most genetic influences on reporting trauma were shared across psychiatric disorders. Multiple gene-environment interplay pathways may explain these findings. In this **Theme 3**, plausible heritable characteristics relevant to how individuals engage with the world are discussed, including personality and behavioural traits, cognitive biases, and socioeconomic factors.

Personality, behaviour and cognitive biases

One relevant mechanism to consider here is a gene-environment correlation, whereby heritable influences on shared behavioural and personality traits influence the risk of exposure to adversity, which, in turn, influences the risk of multiple disorders³⁰. As discussed in **Chapter 3**, this may include risk-taking behaviours, which was identified as a key trait that captured the SNP-based heritability of reported trauma. Risk-taking behaviours are considered important risk factors for trauma exposure³¹ and are strongly associated with externalising psychopathology (e.g. ADHD), which may increase vulnerability to later internalising problems³². In **Chapter 2**, most risk-taking behaviours were not significantly associated with fear or GAD. The exception to this was automobile speeding propensity which had a modest, negative genetic correlation with both anxiety disorders. This negative genetic correlation indicates that genetic variants associated with a higher risk of anxiety is associated with lower reporting of automobile speeding. Recent findings indicate the causal effect of genetic risk for externalising disorders (i.e. ADHD) on distress disorders (i.e. PTSD) is partly mediated by genetic influences on risk-taking behaviours, sociodemographic factors, and trauma exposure. However, an independent causal relationship was also found³¹. ADHD is also associated with negative bias recall of adversity³³, which may partly explain its genetic correlation with retrospective self-reported trauma (**Chapter 3**). Thus, other plausible mechanisms linking psychopathologies with adversity include heritable cognitive biases affecting the perception of experiences, which may be independent of gene-environment correlations⁷.

Shared cognitive biases, which, in turn, may increase sensitivity to the environment, may link distress and fear disorders with the adverse effects of traumatic experiences identified in **Chapters 2 and 3**. Cognitive biases impact how people interpret and, thus, respond to their environments. For example, anxiety-related cognitive biases are associated with an increased risk of anxiety disorders, depression and other distress-related disorders, including PTSD^{34–36}. Such biases are heritable and are genetically correlated with anxiety, depression^{37,38}, and environmental sensitivities³⁹. The shared genetic influences between cognitive biases and

environmental sensitivity account for some of the heritability of self-reporting of life events³⁹. Thus, cognitive biases and environmental sensitivities are plausible characteristics that may partly explain how GWAS of distress-related disorders capture the SNP-based heritability of reporting childhood maltreatment in **Chapter 3**.

This thesis also found some evidence to suggest that genetic influences on comorbid psychiatric disorders influence negative experiences (**Chapter 3**) as well as a lack of improvement in anxiety and depressive symptoms from positive experiences, such as psychosocial treatment (**Chapter 4**). In **Chapter 3**, neurodevelopmental disorders (ADHD and ASD) were positively genetically correlated with self-reported trauma. Intriguingly, the polygenic score for ADHD was significantly associated with self-reported negative psychological treatment outcomes in **Chapter 4**. This is consistent with findings from a previous study that reported worse treatment outcomes among those with higher polygenic scores for ASD⁴⁰. Higher polygenic scores for ADHD and ASD are also negatively associated with response to social skills group training⁴¹. This could reflect genetic liability to these disorders influencing symptoms, such as learning or social difficulties, impacting response to psychosocial interventions⁴². If these findings are replicated, they may indicate that a higher genetic propensity for neurodevelopmental disorders is associated with increased adversity and poorer outcomes of psychosocial interventions.

The mechanisms explaining the genetic association between ADHD, ASD, and poor outcomes of both positive (**Chapter 4**) and negative (**Chapter 3**) environmental experiences are unclear. Under the environmental sensitivity framework, hypersensitivity would contribute to the vulnerability of adverse environments and improve responsiveness to positive environments⁴³. Individuals with neurodevelopmental disorders may exhibit hypersensitivity to environmental stimuli⁴³ and difficulties in social information processing⁴². Potentially, a combination of these factors may increase the adverse effects of stressful environments⁴⁴ and negate the positive benefits of supportive social environments, such as psychological treatment. A broader array of interpersonal circumstances may be experienced as stressful in ASD⁴⁴. This may extend to psychological treatment, an inherently social process that may require adaptations to meet the needs of those with neurodevelopmental disorders^{45,46}. Indeed, social cognitive biases, such as interpreting ambiguous social stimuli as negative, are a potential target for adapting psychological treatment in ASD and ADHD^{47,48}.

Although there is considerable overlap between ASD and ADHD, they are still clinically and genetically distinguishable and may influence experiences and treatment outcomes in different ways^{49,50}. Of note, a higher genetic risk for ADHD was more strongly associated with treatment-resistant depression than non-treatment resistance⁵¹. In contrast, a higher

polygenic score for ASD has been associated with an improved response to antidepressants¹¹. Genetic components of ASD and ADHD associated with environmental experiences do not overlap entirely, as indicated in **Chapter 3**. Future research could examine the specific heritable symptom dimensions of neurodevelopmental disorders associated with vulnerability to adversity and response to treatment.

Socioeconomic status

Associations with sociodemographic-related traits are found in all three empirical chapters. This could reflect the role of socioeconomic status (SES) in genetic signals. Genetic influences on characteristics such as general cognitive ability impact educational and economic achievements, in turn influencing a range of phenotypes³⁰. This includes lifestyle and sociodemographic traits that influence the environments we are exposed to, and thus risk and protective factors for physical and mental health outcomes³⁰. Partner selection within socioeconomic groups is common, and offspring born into a higher SES level can inherit both an advantage in environmental circumstances and a genetic propensity to maintain higher a SES level than those born into a lower SES group with greater environmental disadvantages. These mechanisms represent a perpetual cycle in which society rewards those with a higher genetic propensity for educational (i.e. economic) achievements with environmental advantages, increasing the effect of genetics⁵². As such, a complex relationship arises between SES-related traits and influences on genetic correlations across traits, including psychiatric disorders. For example, previous research has shown that when conditioning for genetic influences on SES-related traits, the genetic structuring of psychiatric disorders changes, including among internalising traits^{53,54}. This additional complex relationship should be considered when interpreting the biological, environmental and treatment outcome findings reported in this thesis.

Overlapping genetic influences on educational attainment across psychiatric disorders may indicate gene-environment correlation mechanisms, complicating the interpretation and identification of causal transdiagnostic biological mechanisms^{53,54}. As observed in **Chapter 2**, loci associated with anxiety disorders were previously reported in GWAS of SES-related traits. Global negative genetic correlations were also observed between anxiety disorders and SES-related traits, as found in previous GWAS and epidemiological studies^{55–58}. This chapter showed that fear had a significantly stronger negative genetic correlation with educational attainment and general cognitive ability than GAD. The implications of these findings are not clear and warrant further investigation. This could reflect differences in cognitive mechanisms, gene-environment correlations, or biases in sampling (e.g. participation bias, as discussed further below). Future research could assess how educational attainment and general

cognitive ability differentially influence fear and GAD, and their genetic correlations with other traits. For example, the uncorrelated cognitive and non-cognitive components of educational attainment could be tested for differences in genetic correlations between fear and GAD⁵⁹. Genetic influences on SES-related traits could be partialled out^{53,54}, which may differentially influence fear and GAD and their genetic correlations with other traits, giving a broader picture of how SES may influence the anxiety disorder spectrum.

Findings from **Chapters 3 & 4** may reflect the effect of SES on the risk and resolution of psychopathology through gene-environment correlation mechanisms. In **Chapter 3**, among the top genetic correlations with reported trauma were sociodemographic-related traits associated with lower SES, including social deprivation, poor self-reported health, younger age at first birth and parents' attained age. Genetic influences on socioeconomic status are associated with access to living in safe (or unsafe) neighbourhoods⁶⁰. A genetic propensity for higher SES is linked with migration to safer neighbourhoods, in which exposure to adverse environments is lower, decreasing the risk of experiencing posttraumatic physical and mental health problems⁶⁰. Furthermore, those with a higher SES have better access to healthcare and social support than those with a lower SES^{61,62}. Traits related to socioeconomic advantages are known indicators of treatment outcomes as they influence access to treatment^{9,61–63}. As shown in **Chapter 4**, higher educational level was a significant, non-genetic predictor of better psychological treatment outcomes. The effect of SES may also have been reflected in polygenic signals. Polygenic scores for SES-related traits were associated with better psychological treatment outcomes. This includes higher self-reported health and older age at first birth, which more broadly capture higher SES^{64,65}. Similarly, a higher polygenic score for general cognitive ability was previously associated with better responses to psychological treatment⁹. However, these findings may also reflect spurious associations. Further investigation is required to establish the mechanisms underlying these associations¹⁶. Future studies with sufficient power could examine the influence of SES on reporting adversity and psychological treatment outcomes by conducting genetic analyses across different SES strata. If replicated, these findings could support the need for improving social and educational provisions in socioeconomically disadvantaged areas^{52,60} to minimise the risk and consequences of adverse exposures and improve treatment outcome equity.

General Limitations

Lack of replication and sources of bias

The lack of availability of the measures used in this thesis at the scale required for genomic analyses limited the ability to assess replication in independent datasets. Many genomic

analyses rely on large datasets, often the UK Biobank, which has multiple sources of potential biases, and thus biases may be observed across multiple variables. A large proportion of the publicly available GWAS summary statistics used in this thesis included data from UK Biobank. A strength of this thesis was the use of novel genotyped datasets with rich phenotypic measures of anxiety and depression. However, these datasets were also not free of potential bias.

The UK Biobank has a problem of 'healthy volunteer bias', in which participants generally have better health than the general UK population^{66,67}. Participation across the main studies used in this thesis was also associated with higher SES-related factors, including higher educational achievements. In England and Wales, approximately 34% of the population aged 16 and over have a university-level qualification⁶⁸. In contrast, approximately half of the participants have university-level education in the GLAD study⁶⁹, in the QSkin study⁷⁰, and in the subsample of the UK Biobank that completed the mental health questionnaire⁷¹. Approximately three-quarters of participants in AGDS have university-level qualifications⁷⁰. Therefore, those with a higher education level may be more likely to enrol in genetic studies⁷⁰. Providing genetic data in the GLAD and COPING studies was associated with higher education, older age, and fewer smoking behaviours⁷².

Further sources of participation bias are reflected in the lack of ethnic diversity in these studies. Approximately 96% of participants self-report identifying as White ethnicity in the GLAD study and the UK Biobank mental health questionnaire subset^{69,71}. In contrast, this group comprises 82% of the population in England and Wales⁷³. Due to this lack of diversity, genetic analyses in this thesis were restricted to European-associated genetic ancestry, as discussed in detail below. Improving participation among racially minoritised groups is an active area of ongoing work in the GLAD study. A number of complex systemic and historical factors influence participation in mental health and genetic studies. This includes (but is not limited to) a lack of representation among researchers from racially minoritised groups, which may influence mistrust in research stemming from historical injustices, with fear that genetic research could lead to further racial discrimination^{74–76}. Therefore, the data used in this thesis are missing important socio-cultural environmental factors (e.g. lived experience of racial discrimination) associated with the risk of mental health problems, adversity, and treatment outcomes^{75,77,78}.

Given that a number of heritable traits may be associated with the likelihood of volunteering in these studies (e.g. health-related and SES-related traits), participation bias may have influenced the genetic signal in this thesis²⁹. Recent analyses in the UK Biobank revealed that participation bias has a particularly strong effect on GWAS of socio-behavioural traits⁷⁹. This

bias can deflate the effect size of SNPs in anxiety and depression GWAS and may have impacted the power to detect genome-wide significant loci in **Chapters 2** and **4**. Participation bias can also inflate SNP-based heritability and, thus, a polygenic signal of educational attainment and SES-related traits (**Chapters 3** and **4**). Participation bias can also affect estimates of genetic correlations across several socio-behavioural traits⁷⁹. A downward bias in genetic correlations can arise when two phenotypes influence participation in a study (i.e. collider bias)^{80,81}. For instance, UK Biobank participation bias may underestimate genetic correlations between anxiety and depression and risk-taking, education and physical health⁷⁹, which may have influenced genetic correlation results in **Chapter 2**. The independent genetic association found in **Chapter 3** between reported trauma and general risk tolerance may be partly driven by participation bias. Underestimation of genetic correlations due to participation bias was previously found between risk-taking and several UK Biobank traits modelled in **Chapter 3**, including insomnia, SES-related traits and physical illnesses⁷⁹. Therefore, an important future direction is to replicate findings from this thesis in a population-representative sample⁸², or by using recently established methods that correct for participation bias^{79,83}.

Lack of ancestral diversity

The analyses presented in this thesis were restricted to individuals of PCA-derived European-associated genetic ancestry. Consequently, the findings of this thesis may not be generalisable across genetic ancestry groups. Across the global human population, a continuum of variation exists in structural properties that reflect population history, conferring differences in LD patterning and allele frequencies, as outlined in **Chapter 1, Box 3**. To avoid the occurrence of false-positive associations arising from allele-frequency differences across populations, GWAS are typically limited to groups of individuals of more similar genetic ancestry⁷⁷. Due to the availability of samples, and historical and systemic influences⁸⁴, the most commonly studied from the five major global populations is European-associated ancestry. This has led to the problem of Eurocentric GWAS bias⁸⁵.

Eurocentric GWAS bias is a major problem in psychiatric genetics⁸⁶, and statistical genetics more broadly⁸⁵. In 2021, approximately 86% of all GWASs were based on samples of European-associated genetic ancestry⁷⁵, despite only contributing to 16% of the world population⁸⁵. The potential lack of generalisability of the findings from this thesis to non-European ancestry groups adds to a growing obstacle in health equity⁷⁵. If the Eurocentric GWAS bias continues, translating GWAS findings to clinically benefit those with mental health problems may further health disparities among racially minoritised communities^{85,87}. For example, polygenic score prediction attenuates when applied to genetic ancestry groups that differ from the sample the polygenic score was derived from^{85,88}. Furthermore, the phenotypic

measures used in this thesis are based on Eurocentric beliefs^{78,89,90}. Presentation of disorders, their associated risk factors and treatment outcomes, may therefore differ from non-European populations. This emphasises the need to diversify genetic ancestry in samples and adapt measures to avoid Eurocentric biases at the phenotypic level. Efforts to achieve more globally representative samples for psychiatric genetics research are currently underway⁷⁵. The development of methods to diversify genomic analyses is also rapidly improving^{91–93}.

Future directions

Establishing causality, and disentangling individual from familial genomic effects

A key future avenue to the research conducted in this thesis is establishing mechanisms underlying associations between traits and inferring causality. One powerful approach to assessing causal relationships between two variables is Mendelian Randomization (MR). Results from GWAS summary statistics can be used for MR analyses, where robustly associated genetic variants are used as instrumental variables to test for a causal effect of one variable on another. Both directions of causality can be tested through bidirectional MR⁹⁴. GWASs of anxiety disorders have not yet reached sufficient power for summary-level MR. A number of robust genome-wide significant variants are required to be used as instrumental variables. For example, some methods suggest an analysis of fewer than 30 instruments is limited in power⁹⁵. The release of the first PGC anxiety GWAS summary statistics will be key in providing the tools to test causal relationships with anxiety.

As GWASs of fear and GAD increase in size, bidirectional causal relationships could be tested between fear and GAD, and traits with overlapping genetics, such as the health-related traits identified in **Chapter 2**, including heart disease, insomnia, or comorbid psychiatric disorders. Anxiety often precedes the onset of other psychiatric disorders. MR could be used to test the causality of temporal comorbidity patterns, such as with depression, PTSD and ADHD. The causal relationship across distress-related traits could be examined using multivariable approaches, such as between anxiety, depression and shared negative affectivity (i.e. neuroticism). Multivariable MR takes into account independent and correlated horizontal pleiotropy, estimating the direct effects of one risk factor independent of another and the role of mediation or confounding^{94,96}. As GWAS of environmental measures such as reported trauma grows, this will provide the opportunity to jointly test the direction of the effect of traits in **Chapter 3** and their relationship with one another. Recent applications of such methods have proven promising in testing the causal relation between ADHD and PTSD whilst accounting for various mediating phenotypes³¹. Of note, robust application of MR methods to

identify true causal effects may be challenging for heterogeneous phenotypes such as anxiety and depression. Thus, stratifying GWAS into more homogeneous subgroups may be needed as well as using detailed phenotyping to capture more specific genetic effects⁹⁴. This underscores the importance of GWAS subtyping (**Chapter 2**) and using studies with detailed measures, as in this thesis.

A further future direction in establishing mechanisms of associations is disentangling gene-environment interplay processes. Psychiatric genomics has benefitted from ‘borrowing’ methods from developmental family-based studies and vice versa, driving innovative research^{97,98}. This includes methods that allow for disentangling between genetic influences that reflect *individual* genetic effects (i.e. genetic variation in the individual influences their own trait directly) or *familial* genetic effects (i.e. indirect genetic effects from the parent influences a child trait through the rearing environment)^{16,98}. Familial genetic effects can lead to passive gene-environment correlations. Individual genetic effects can also be mediated by the environment the individual shapes, leading to evocative and active gene-environment correlation⁹⁸. Both individual and familial genetic effects can be tested for causality if the associations are free from biases¹⁶.

Integrating genomic analyses into within-family studies can disentangle individual from familial genomic effects, establish the causality of these influences by minimising biases, and understand the origins of gene-environment interplay⁹⁸. Of relevance to this thesis, GWAS in unrelated individuals aim to detect individual genetic effects on a phenotype (i.e. **Chapter 2**) but inadvertently may detect familial genetic effects (e.g. as suggested in **Chapter 3**) and other sources of bias, including assortative mating. Partners are more likely to choose someone with similar characteristics, including internalising symptoms⁹⁹. These other sources of a genetic signal can impact the interpretation of biological mechanisms underpinning psychiatric disorders, SNP-based heritability estimates, genetic correlations, and MR tests^{100–102}. Within-family GWAS are less affected by such biases and are therefore key for estimating individual genetic effects that more directly influence a phenotype of interest¹⁰². Integrating polygenic scores into a within-family design can elucidate the role of familial genetic effects on child outcomes¹⁰³. Such designs could be applied to understand genomic influences on the risk of trauma exposure (e.g. risk-taking behaviour in **Chapter 3**) and the development of internalising psychopathology (**Chapter 2**). Identifying true environmental influences from parents has implications for prevention strategies that could aim to modify the environment of at-risk families. As within-family genotyped datasets increase in size, these will be key for addressing key gaps in this thesis and the developmental translation of findings more broadly

in anxiety and depression genomics.

Expanding and integrating genome-wide data

A major future direction of anxiety and depression genomics is expanding genome-wide data to the full spectrum of allele frequencies. The GWAS used in this thesis were limited to common genetic variants. Common variants are more likely to have lower penetrance on phenotypes due to selective pressures^{104,105}. Thus, the GWAS signal also reflects the role of natural selection in common traits¹⁰⁶. Rarer variants with larger effect sizes likely also influence common disorders and may better elucidate biological mechanisms due to greater penetrance. However, rare variants are challenging and expensive to capture at the scale required for association analyses^{86,107}. As such, brief phenotyping approaches may be prioritised to attain the large sample sizes required for rare variant analysis of common mental disorders. More progress has been made in identifying rare variants associated with depression than with anxiety^{108–110}. Of note, a recent study indicates that rare variants only contribute to a small proportion of the heritability of complex traits, suggesting that rare variants do not fully explain the missing heritability between GWAS and twin estimates¹⁰⁶. As whole genome sequencing and exome sequencing data become more cost-effective to generate and increase in size, this will progress our understanding of the role of rare variants as a risk factor for common mental disorders and, hopefully, lead to emerging biological insights.

Future biological insights on the risk and treatment of common mental disorders can stem from integrating genomic data with other sources. For example, findings on molecular genetics can be extended to understand brain structure and function by integrating GWAS with neuroimaging data. This has the potential to generate and test novel hypotheses. For example, differences in genetic influences on neurocircuitry could be tested between distress and fear or in understanding differences in the neurobiology of treatment response^{111,112}. Such initiatives include the ENIGMA-anxiety disorder working group, which aims to ascertain neuroimaging data across anxiety disorders¹¹³.

Improvement in the prediction of risk and treatment of common mental disorders will require the integration of multimodal data. For example, the integration of neuroimaging data may improve prediction models of treatment response in anxiety and depression^{111,112,114}. Incorporating more detailed data on environmental risk factors into prediction modelling, such as contextual information on childhood maltreatment, is another promising avenue for treatment response research¹¹⁵. Prediction from polygenic scores may improve once they contain a broader spectrum of allele frequencies¹¹⁶. Integrating biological pathway-based data

into polygenic scores may have greater power for predicting disorder subtypes than SNP-based scores¹¹⁷. Recent methodological developments show that polygenic scores derived from more diverse populations of genetic ancestry can be used to improve predictive power^{118,119}. Large-scale GWAS analyses in non-European genetic ancestry groups are expected to significantly improve the clinical potential for polygenic scores⁸⁸. Furthermore, this will be vital to ensure that progress in predicting risk and treating common mental disorders benefits individuals globally.

Conclusions

The balance between phenotyping depth and sample size and using self-reported versus objective measures likely impact the genetic specificity and psychosocial components of GWAS findings. The genetic structure of internalising disorders provides evidence for the hierarchical classification of anxiety and depressive disorders into higher-order internalising and lower-order distress-fear dimensions. Future genomic studies aiming to detect subtype genetic risk factors should examine both higher-order and lower-order dimensions of internalising to increase the power to extract what is common and unique. The genetic signals from GWAS also capture the interplay with the environment, in which multiple pathways may partly explain the high comorbidity and genetic covariation across psychopathologies. This includes heritable behaviours, cognitive biases, and the impact of neurodevelopmental disorders and socioeconomic factors on responses to adverse environments and psychological treatments.

References

1. Kendler, K. S. *et al.* Shared and specific genetic risk factors for lifetime major depression, depressive symptoms and neuroticism in three population-based twin samples. *Psychol. Med.* **49**, 2745–2753 (2019).
2. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
3. Giannakopoulou, O. *et al.* The Genetic Architecture of Depression in Individuals of East Asian Ancestry: A Genome-Wide Association Study. *JAMA Psychiatry* **78**, 1258–1269 (2021).
4. Levey, D. F. *et al.* Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat. Neurosci.* **24**, 954–963 (2021).
5. Mitchell, B. L. *et al.* The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes. *Biol. Psychiatry* (2021) doi:10.1016/j.biopsych.2021.10.021.
6. Danese, A. & Widom, C. S. Objective and subjective experiences of child maltreatment and their relationships with psychopathology. *Nat Hum Behav* **4**, 811–818 (2020).
7. Peel, A. J. *et al.* Genetic and early environmental predictors of adulthood self-reports of trauma. *Br. J. Psychiatry* 1–8 (2022).
8. Baldwin, J. R. *et al.* A genetically informed Registered Report on adverse childhood experiences and mental health. *Nat Hum Behav* (2022) doi:10.1038/s41562-022-01482-9.
9. Wallert, J. *et al.* Predicting remission after internet-delivered psychotherapy in patients with depression using machine learning and multi-modal data. *Transl. Psychiatry* **12**, 357 (2022).
10. McInnes, G., Yee, S. W., Pershad, Y. & Altman, R. B. Genomewide Association Studies in Pharmacogenomics. *Clin. Pharmacol. Ther.* **110**, 637–648 (2021).
11. Pain, O. *et al.* Identifying the Common Genetic Basis of Antidepressant Response. *Biol Psychiatry Glob Open Sci* **2**, 115–126 (2022).
12. Li, Q. S., Tian, C., Hinds, D. & 23andMe Research Team. Genome-wide association studies of antidepressant class response and treatment-resistant depression. *Transl. Psychiatry* **10**, 360 (2020).
13. Tansey, K. E. *et al.* Contribution of common genetic variants to antidepressant response. *Biol. Psychiatry* **73**, 679–682 (2013).
14. Flint, J. The genetic basis of major depressive disorder. *Mol. Psychiatry* (2023) doi:10.1038/s41380-023-01957-9.
15. Wang, R. A. H., Davis, O. S. P., Wootton, R. E., Mottershaw, A. & Haworth, C. M. A. Social support and mental health in late adolescence are correlated for genetic, as well as environmental, reasons. *Sci. Rep.* **7**, 13088 (2017).
16. Pingault, J.-B. *et al.* Research Review: How to interpret associations between polygenic scores, environmental risks, and phenotypes. *J. Child Psychol. Psychiatry* **63**, 1125–1139 (2022).
17. Waszczuk, M. A., Kotov, R., Ruggero, C., Gamez, W. & Watson, D. Hierarchical structure of emotional disorders: From individual symptoms to the spectrum. *J. Abnorm. Psychol.* **126**, 613–634 (2017).
18. Watson, D. *et al.* Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry* **21**, 26–54 (2022).
19. Grotzinger, A. D. *et al.* Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat. Genet.* **54**, 548–559 (2022).

20. Thorp, J. G. *et al.* Symptom-level modelling unravels the shared genetic architecture of anxiety and depression. *Nat Hum Behav* **5**, 1432–1442 (2021).
21. Coleman, J. R. I. *et al.* The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biol. Psychiatry* (2019) doi:10.1016/j.biopsych.2019.10.015.
22. Kotov, R. *et al.* The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *J. Abnorm. Psychol.* **126**, 454–477 (2017).
23. Forbes, M. K. *et al.* Principles and procedures for revising the Hierarchical Taxonomy of Psychopathology. (2023) doi:10.31234/osf.io/xr48p.
24. Levin-Aspenson, H. F., Watson, D., Ellickson-Larew, S., Stanton, K. & Stasik-O'Brien, S. M. Beyond Distress and Fear: Differential Psychopathology Correlates of PTSD Symptom Clusters. *J. Affect. Disord.* **284**, 9–17 (2021).
25. Kendler, K. S., Chatzinakos, C. & Bacanu, S.-A. The impact on estimations of genetic correlations by the use of super-normal, unscreened, and family-history screened controls in genome wide case-control studies. *Genet. Epidemiol.* **44**, 283–289 (2020).
26. Purves, K. L. *et al.* Evidence for distinct genetic and environmental influences on fear acquisition and extinction. *Psychol. Med.* 1–9 (2021).
27. Hettema, J. M., Annas, P., Neale, M. C., Kendler, K. S. & Fredrikson, M. A twin study of the genetics of fear conditioning. *Arch. Gen. Psychiatry* **60**, 702–708 (2003).
28. Barry, C.-J. S. *et al.* How to estimate heritability: a guide for genetic epidemiologists. *Int. J. Epidemiol.* (2022) doi:10.1093/ije/dyac224.
29. Abdellaoui, A. & Verweij, K. J. H. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nat Hum Behav* **5**, 686–694 (2021).
30. Avinun, R. The E Is in the G: Gene-Environment-Trait Correlations and Findings From Genome-Wide Association Studies. *Perspect. Psychol. Sci.* **15**, 81–89 (2020).
31. Wendt, F. R. *et al.* The Relationship of Attention-Deficit/Hyperactivity Disorder With Posttraumatic Stress Disorder: A Two-Sample Mendelian Randomization and Population-Based Sibling Comparison Study. *Biol. Psychiatry* **93**, 362–369 (2023).
32. Adler, L. A., Kunz, M., Chua, H. C., Rotrosen, J. & Resnick, S. G. Attention-deficit/hyperactivity disorder in adult patients with posttraumatic stress disorder (PTSD): is ADHD a vulnerability factor? *J. Atten. Disord.* **8**, 11–16 (2004).
33. Vrijen, J. N. *et al.* ADHD symptoms in healthy adults are associated with stressful life events and negative memory bias. *Atten. Defic. Hyperact. Disord.* **10**, 151–160 (2018).
34. Taylor, S., Koch, W. J., Woody, S. & McLean, P. Anxiety sensitivity and depression: How are they related? *J. Abnorm. Psychol.* **105**, 474–479 (1996).
35. Schmidt, N. B., Zvolensky, M. J. & Maner, J. K. Anxiety sensitivity: prospective prediction of panic attacks and Axis I pathology. *J. Psychiatr. Res.* **40**, 691–699 (2006).
36. Olatunji, B. O., Armstrong, T., Fan, Q. & Zhao, M. Risk and resiliency in posttraumatic stress disorder: Distinct roles of anxiety and disgust sensitivity. *Psychol. Trauma* **6**, 50–55 (2014).
37. Eley, T. C., Gregory, A. M., Clark, D. M. & Ehlers, A. Feeling anxious: a twin study of panic/somatic ratings, anxiety sensitivity and heartbeat perception in children. *J. Child Psychol. Psychiatry* **48**, 1184–1191 (2007).
38. Zavos, H. M. S., Rijdsdijk, F. V., Gregory, A. M. & Eley, T. C. Genetic influences on the cognitive biases associated with anxiety and depression symptoms in adolescents. *J. Affect. Disord.* **124**, 45–53 (2010).
39. Peel, A. J. *et al.* A multivariate genetic analysis of anxiety sensitivity, environmental sensitivity and reported life events in adolescents. *J. Child Psychol. Psychiatry* **64**, 289–298 (2023).
40. Andersson, E. *et al.* Genetics of response to cognitive behavior therapy in adults with major depression: a preliminary report. *Mol. Psychiatry* **24**, 484–490 (2019).
41. Li, D. *et al.* The influence of common polygenic risk and gene sets on social skills group training response in autism spectrum disorder. *NPJ Genom Med* **5**, 45 (2020).

42. Schmidt, N. B. & Vereenoghe, L. Interpersonal Cognitive Biases in Children and Young People with Neurodevelopmental Disorders: A Systematic Review. *Current Developmental Disorders Reports* **8**, 219–235 (2021).
43. Greven, C. U. *et al.* Sensory Processing Sensitivity in the context of Environmental Sensitivity: A critical review and development of research agenda. *Neurosci. Biobehav. Rev.* **98**, 287–305 (2019).
44. Rumball, F., Happé, F. & Grey, N. Experience of Trauma and PTSD Symptoms in Autistic Adults: Risk of PTSD Development Following DSM-5 and Non-DSM-5 Traumatic Life Events. *Autism Res.* (2020) doi:10.1002/aur.2306.
45. Cooper, K., Loades, M. E. & Russell, A. J. Adapting Psychological Therapies for Autism - Therapist Experience, Skills and Confidence. *Res. Autism Spectr. Disord.* **45**, 43–50 (2018).
46. Gallant, C., Roudbarani, F., Ibrahim, A., Maddox, B. B. & Weiss, J. A. Clinician Knowledge, Confidence, and Treatment Practices in Their Provision of Psychotherapy to Autistic Youth and Youth with ADHD. *J. Autism Dev. Disord.* (2022) doi:10.1007/s10803-022-05722-9.
47. Schmidt, N. B. & Vereenoghe, L. Inclusiveness of cognitive bias modification research toward children and young people with neurodevelopmental disorders: A systematic review. *Int J Dev Disabil* **68**, 86–101 (2022).
48. Miranda, A., Berenguer, C., Roselló, B., Baixauli, I. & Colomer, C. Social Cognition in Children with High-Functioning Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder. Associations with Executive Functions. *Front. Psychol.* **8**, 1035 (2017).
49. Mattheisen, M. *et al.* Identification of shared and differentiating genetic architecture for autism spectrum disorder, attention-deficit hyperactivity disorder and case subgroups. *Nat. Genet.* **54**, 1470–1478 (2022).
50. Antshel, K. M. & Russo, N. Autism Spectrum Disorders and ADHD: Overlapping Phenomenology, Diagnostic Issues, and Treatment Considerations. *Curr. Psychiatry Rep.* **21**, 34 (2019).
51. Fabbri, C. *et al.* Genetic and clinical characteristics of treatment-resistant depression using primary care records in two UK cohorts. *Mol. Psychiatry* **26**, 3363–3373 (2021).
52. Abdellaoui, A., Dolan, C. V., Verweij, K. J. H. & Nivard, M. G. Gene-environment correlations across geographic regions affect genome-wide association studies. *Nat. Genet.* **54**, 1345–1354 (2022).
53. Wendt, F. R. *et al.* Multivariate genome-wide analysis of education, socioeconomic status and brain phenome. *Nat Hum Behav* **5**, 482–496 (2021).
54. Marees, A. T. *et al.* Genetic correlates of socio-economic status influence the pattern of shared heritability across mental health traits. *Nat Hum Behav* **5**, 1065–1073 (2021).
55. Mwinyi, J. *et al.* Anxiety Disorders are Associated with Low Socioeconomic Status in Women but Not in Men. *Womens. Health Issues* **27**, 302–307 (2017).
56. de Graaf, R., ten Have, M., Tuithof, M. & van Dorsselaer, S. First-incidence of DSM-IV mood, anxiety and substance use disorders and its determinants: results from the Netherlands Mental Health Survey and Incidence Study-2. *J. Affect. Disord.* **149**, 100–107 (2013).
57. Sareen, J., Afifi, T. O., McMillan, K. A. & Asmundson, G. J. G. Relationship between household income and mental disorders: findings from a population-based longitudinal study. *Arch. Gen. Psychiatry* **68**, 419–427 (2011).
58. Levey, D. F. *et al.* Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *Am. J. Psychiatry* **177**, 223–232 (2020).
59. Demange, P. A. *et al.* Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).
60. Abdellaoui, A. *et al.* Genetic correlates of social stratification in Great Britain. *Nature Human Behaviour*. Preprint at (2019).

61. Delgadillo, J., Asaria, M., Ali, S. & Gilbody, S. On poverty, politics and psychology: the socioeconomic gradient of mental healthcare utilisation and outcomes. *Br. J. Psychiatry* **209**, 429–430 (2016).
62. Saxon, D. *et al.* Psychotherapy provision, socioeconomic deprivation, and the inverse care law. *Psychother. Res.* **17**, 515–521 (2007).
63. Rayner, C. *et al.* 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* **22**, 719 (2022).
64. van Roode, T., Sharples, K., Dickson, N. & Paul, C. Life-Course Relationship between Socioeconomic Circumstances and Timing of First Birth in a Birth Cohort. *PLoS One* **12**, e0170170 (2017).
65. Alvarez-Galvez, J. *et al.* The impact of socio-economic status on self-rated health: study of 29 countries using European social surveys (2002-2008). *Int. J. Environ. Res. Public Health* **10**, 747–761 (2013).
66. Stamatakis, E. *et al.* Is Cohort Representativeness Passé? Poststratified Associations of Lifestyle Risk Factors with Mortality in the UK Biobank. *Epidemiology* **32**, 179–188 (2021).
67. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
68. Waddington, B. Education, England and Wales - Office for National Statistics. <https://www.ons.gov.uk/peoplepopulationandcommunity/educationandchildcare/bulletins/educationenglandandwales/census2021> (2023).
69. Davies, M. R. *et al.* The Genetic Links to Anxiety and Depression (GLAD) Study: Online recruitment into the largest recontactable study of depression and anxiety. *Behav. Res. Ther.* **123**, 103503 (2019).
70. Byrne, E. M. *et al.* Cohort profile: the Australian genetics of depression study. *BMJ Open* **10**, e032580 (2020).
71. Davis, K. A. S. *et al.* Mental health in UK Biobank - development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open* **6**, e18 (2020).
72. Bright, S. J. *et al.* Sociodemographic, mental health, and physical health factors associated with participation within re-contactable mental health cohorts: an investigation of the GLAD Study. *Research Square* (2022) doi:10.21203/rs.3.rs-2367106/v1.
73. Catney, G. *et al.* Ethnic diversification and neighbourhood mixing: A rapid response analysis of the 2021 Census of England and Wales. *Geogr. J.* (2023) doi:10.1111/geoj.12507.
74. Folorunso, O. O. *et al.* Building an intentional and impactful summer research experience to increase diversity in mental health research. *Neuropsychopharmacology* **47**, 2189–2193 (2022).
75. Fatumo, S. *et al.* A roadmap to increase diversity in genomic studies. *Nat. Med.* **28**, 243–250 (2022).
76. Kraft, S. A. *et al.* Beyond Consent: Building Trusting Relationships With Diverse Populations in Precision Medicine Research. *Am. J. Bioeth.* **18**, 3–20 (2018).
77. Peterson, R. E. *et al.* Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* **179**, 589–603 (2019).
78. Faheem, A. 'Not a cure, but helpful' – exploring the suitability of evidence-based psychological interventions to the needs of Black, Asian and Minority Ethnic (BAME) communities. *Cogn. Behav. Ther.* **16**, e4 (2023).
79. Schoeler, T. *et al.* Correction for participation bias in the UK Biobank reveals non-negligible impact on genetic associations and downstream analyses. *bioRxiv* 2022.09.28.509845 (2022) doi:10.1101/2022.09.28.509845.

80. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* (2019) doi:10.1038/s41576-019-0137-z.
81. Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M. & Davey Smith, G. Collider scope: when selection bias can substantially influence observed associations. *Int. J. Epidemiol.* **47**, 226–235 (2018).
82. Pedersen, C. B. *et al.* The iPSYCH2012 case-cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol. Psychiatry* **23**, 6–14 (2017).
83. van Alten, S., Domingue, B. W., Faul, J., Galama, T. & Marees, A. T. Correcting for volunteer bias in GWAS uncovers novel genetic variants and increases heritability estimates. *medRxiv* (2022) doi:10.1101/2022.11.10.22282137.
84. Popejoy, A. B. & Fullerton, S. M. Genomics is failing on diversity. *Nature* vol. 538 161–164 (2016).
85. Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
86. Andreassen, O. A., Hindley, G. F. L., Frei, O. & Smeland, O. B. New insights from the last decade of research in psychiatric genetics: discoveries, challenges and clinical implications. *World Psychiatry* **22**, 4–24 (2023).
87. Bansal, N. *et al.* Understanding ethnic inequalities in mental healthcare in the UK: A meta-ethnography. *PLoS Med.* **19**, e1004139 (2022).
88. Fatumo, S. & Inouye, M. African genomes hold the key to accurate genetic risk prediction. *Nat Hum Behav* (2023) doi:10.1038/s41562-023-01549-1.
89. Ecks, S. The strange absence of things in the ‘culture’ of the DSM-V. *CMAJ* **188**, 142–143 (2015).
90. Madigoe, T., Burns, J., Zhang, M. & Subramaney, U. Towards a culturally appropriate trauma assessment in a South African Zulu community. *Psychol. Trauma* **9**, 274–281 (2017).
91. Ruan, Y. *et al.* Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54**, 573–580 (2022).
92. Zhang, H. *et al.* Novel Methods for Multi-ancestry Polygenic Prediction and their Evaluations in 5.1 Million Individuals of Diverse Ancestry. *bioRxiv* 2022.03.24.485519 (2023) doi:10.1101/2022.03.24.485519.
93. Atkinson, E. G. *et al.* Cross-ancestry genomic research: time to close the gap. *Neuropsychopharmacology* **47**, 1737–1738 (2022).
94. Wootton, R. E., Jones, H. J. & Sallis, H. M. Mendelian randomisation for psychiatry: how does it work, and what can it tell us? *Mol. Psychiatry* **27**, 53–57 (2022).
95. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
96. Rosoff, D. B., Smith, G. D. & Lohoff, F. W. Prescription Opioid Use and Risk for Major Depressive Disorder and Anxiety and Stress-Related Disorders: A Multivariable Mendelian Randomization Analysis. *JAMA Psychiatry* **78**, 151–160 (2021).
97. Friedman, N. P., Banich, M. T. & Keller, M. C. Twin studies to GWAS: there and back again. *Trends Cogn. Sci.* **25**, 855–869 (2021).
98. Cheesman, R., Ayorech, Z., Eilertsen, E. M. & Ystrom, E. Why we need families in genomic research on developmental psychopathology. *JCPP Adv.* (2023) doi:10.1002/jcv2.12138.
99. Mathews, C. Assortative mating in the affective disorders: A systematic review and meta-analysis. *Comprehensive Psychiatry* vol. 42 257–262 Preprint at <https://doi.org/10.1053/comp.2001.24575> (2001).
100. Border, R. *et al.* Cross-trait assortative mating is widespread and inflates genetic correlation estimates. *Science* **378**, 754–761 (2022).
101. Grotzinger, A. D. & Keller, M. C. Potential bias in genetic correlations. *Science* vol. 378 709–710 (2022).

102. Howe, L. J. *et al.* Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nat. Genet.* **54**, 581–592 (2022).
103. Demange, P. A. *et al.* Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. *Nat. Commun.* **13**, 4801 (2022).
104. Michaelson, J. J. Genetic Approaches to Understanding Psychiatric Disease. *Neurotherapeutics* **14**, 564–581 (2017).
105. O'Connor, L. J. *et al.* Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *Am. J. Hum. Genet.* **105**, 456–476 (2019).
106. Weiner, D. J. *et al.* Polygenic architecture of rare coding variation across 394,783 exomes. *Nature* **614**, 492–499 (2023).
107. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
108. Tian, R. *et al.* Whole exome sequencing in the UK Biobank reveals risk gene *SLC2A1* and biological insights for major depressive disorder. *bioRxiv* (2021) doi:10.1101/2021.05.04.21256398.
109. Kendall, K. M. *et al.* Association of Rare Copy Number Variants With Risk of Depression. *JAMA Psychiatry* **76**, 818–825 (2019).
110. Olsson, E. *et al.* Whole-exome DNA sequencing in childhood anxiety disorders identifies rare de novo damaging coding variants. *Depress. Anxiety* **39**, 474–484 (2022).
111. Lueken, U. *et al.* Neurobiological markers predicting treatment response in anxiety disorders: A systematic review and implications for clinical application. *Neurosci. Biobehav. Rev.* **66**, 143–162 (2016).
112. Klumpp, H. & Fitzgerald, J. M. Neuroimaging Predictors and Mechanisms of Treatment Response in Social Anxiety Disorder: an Overview of the Amygdala. *Curr. Psychiatry Rep.* **20**, 89 (2018).
113. Bas-Hoogendam, J. M. *et al.* ENIGMA-anxiety working group: Rationale for and organization of large-scale neuroimaging studies of anxiety disorders. *Hum. Brain Mapp.* **43**, 83–112 (2022).
114. Sajjadi, M. *et al.* Prediction of depression treatment outcome from multimodal data: a CAN-BIND-1 report. *Psychol. Med.* 1–11.
115. Harkness, K. L. *et al.* The Differential Relation of Emotional, Physical, and Sexual Abuse Histories to Antidepressant Treatment Remission and Persistence of Anhedonia in Major Depression: A CAN-BIND-1 Report. *Can. J. Psychiatry* 07067437231156255 (2023).
116. Fusar-Poli, L., Rutten, B. P. F., van Os, J., Aguglia, E. & Guloksuz, S. Polygenic risk scores for predicting outcomes and treatment response in psychiatry: hope or hype? *Int. Rev. Psychiatry* **34**, 663–675 (2022).
117. Choi, S. W. *et al.* PRSet: Pathway-based polygenic risk score analyses and software. *PLoS Genet.* **19**, e1010624 (2023).
118. Naret, O. *et al.* Improving polygenic prediction with genetically inferred ancestry. *HGG Adv* **3**, 100109 (2022).
119. Ishigaki, K. *et al.* Multi-ancestry genome-wide association analyses identify novel genetic mechanisms in rheumatoid arthritis. *Nat. Genet.* **54**, 1640–1651 (2022).

Appendix A. Supplementary materials for Chapter 2

Table of contents

Supplementary Methods	162
<i>Ethical approval.....</i>	<i>162</i>
Genotyping quality control	162
GLAD+.....	162
UK Biobank.....	162
QIMR	163
Sensitivity analyses	163
Supplementary Results	164
<i>GWAS meta-analyses across GLAD+, QIMR and UKB datasets</i>	<i>164</i>
Supplementary Figure 1: Q-Q plots.....	164
Supplementary Figure 2: Manhattan plots.....	165
<i>GWAS meta-analyses across GLAD+ and UKB datasets (excluding QIMR).....</i>	<i>166</i>
Supplementary Figure 3: Q-Q plots.....	166
Supplementary Figure 4: Manhattan plots.....	167
<i>Region plots of genome-wide significant loci.....</i>	<i>168</i>
Fear-based results.....	168
Supplementary Figure 5: Locus on 14q24.3.....	168
Supplementary Figure 6: Locus on 5q31.3.....	168
Supplementary Figure 7: Locus on 1p34.1.....	169
GAD results	169
Supplementary Figure 8: Locus on 9p23.....	169
Supplementary Figure 9: Locus on 6p22.1.....	170
Supplementary Figure 10: Locus on 2p16.1.....	170
Supplementary Figure 11: Locus on 5q34.....	171
<i>GWAS within each dataset.....</i>	<i>171</i>
Supplementary Figure 12: Q-Q plots for GAD and fear GWASs in the GLAD+ dataset	172
Supplementary Figure 13: Q-Q plots for GAD and fear GWASs in the QIMR dataset	173
Supplementary Figure 14: Manhattan plots for GAD and fear GWASs in the QIMR dataset	174
Supplementary Figure 15: Q-Q plots for GAD and fear GWASs in the UKB dataset..	175
Supplementary Figure 16: Manhattan plots for GAD GWASs in the UKB dataset.....	176
Supplementary Tables	177
Supplementary Table 1: Summary of available measures in each study for genome-wide association analysis.....	177
Supplementary table 2: Self-report diagnoses defined by responses to a single self-report item	178
Supplementary Table 3: Population prevalence estimates calculated from COPING sample prevalence.	178
Supplementary Table 4: Independent genome-wide significant loci associated with fear-based disorders, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.	179

Supplementary Table 5: Independent genome-wide significant loci associated with generalised anxiety disorder, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.	179
Supplementary Table 6: Gene-level associations with fear-based disorders, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.	187
Supplementary Table 7: Gene-level associations with GAD, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.	189
Supplementary Table 8: LDSC regression SNP-based heritability (h^2_{SNP}) estimates of GAD and fear-based disorder phenotypes	198
Supplementary Table 9: Genetic correlations (r_g) between anxiety disorder phenotypes and 345 external traits estimated in LDSC regression.	200
Supplementary Table 10: SNP-based heritability (h^2_{SNP}) estimates of GAD and fear-based disorders screened using detailed versus brief diagnostic measures.	210

Supplementary Methods

Ethical approval

The GLAD Study was approved by the London Fulham Research Ethics Committee on 21st August 2018 following a full review by the committee (REC reference: 18/LO/1218). Ethical approval for the COPING Study was granted following a full review by the South West Central Bristol Research Ethics Committee (REC reference: 20/SW/0078) on 27th April 2020. The AGDS and QSkin cohorts were reviewed and approved by the QIMR Berghofer Medical Research Institute's Human Research Ethics Committee. Analyses in the UK Biobank were performed under UK Biobank Project Application 82087. The UK Biobank has ethical approval from the North West Multi-centre Research Ethics Committee (MREC) as a Research Tissue Bank (RTB) approval.

Genotyping quality control

GLAD+

In both the GLAD and COPING studies, genotyping was performed using the Affymetrix UK Biobank Axiom Array v1 and v2. Participants were defined as having European-associated ancestry based on genomic principal component analysis (PCA) conducted in Plink. Genotyping quality control was performed on the PCA-defined European-associated ancestry subsample. Pre-imputed variants that were rare ($MAF < 1\%$), not present in $> 95\%$ of participants, and showed deviation from the Hardy-Weinberg equilibrium (HWE; $p < 1 \times 10^{-10}$) were removed. Individuals were excluded if they had $> 95\%$ of variants missing, a genetic duplicate across the two studies, a mismatch between self-reported and chromosomal sex and were outliers of identity by descent, indicating sample contamination (> 3 standard deviations from the mean).

UK Biobank

Pre-processing and collection of genotype data was conducted centrally as described elsewhere¹. Participants were genotyped using the Affymetrix UK BiLEVE Axiom and Affymetrix UK Biobank Axiom array. Participants were defined as having European-associated ancestry based on genomic PCA four-means clustering. In the PCA-defined European-associated ancestry dataset, pre-imputed variants with $MAF < 1\%$, missing in $> 98\%$ of individuals, and significant result from the HWE deviation test ($p < 1 \times 10^{-8}$). Participants were

removed if they had > 98% of variants missing and a mismatch between genetic and phenotypic sex.

QIMR

Participants in the AGDS and QIMR studies were genotyped using the Illumina Global Screening Array v2. Genotyping quality control was performed in a PCA-defined European-associated ancestry subsample, as described in Mitchell et al 2022². Briefly, prior to imputation variants were removed with MAF < 1%, high missingness > 95% and deviation from the HWE ($p < 1 \times 10^{-6}$).

Sensitivity analyses

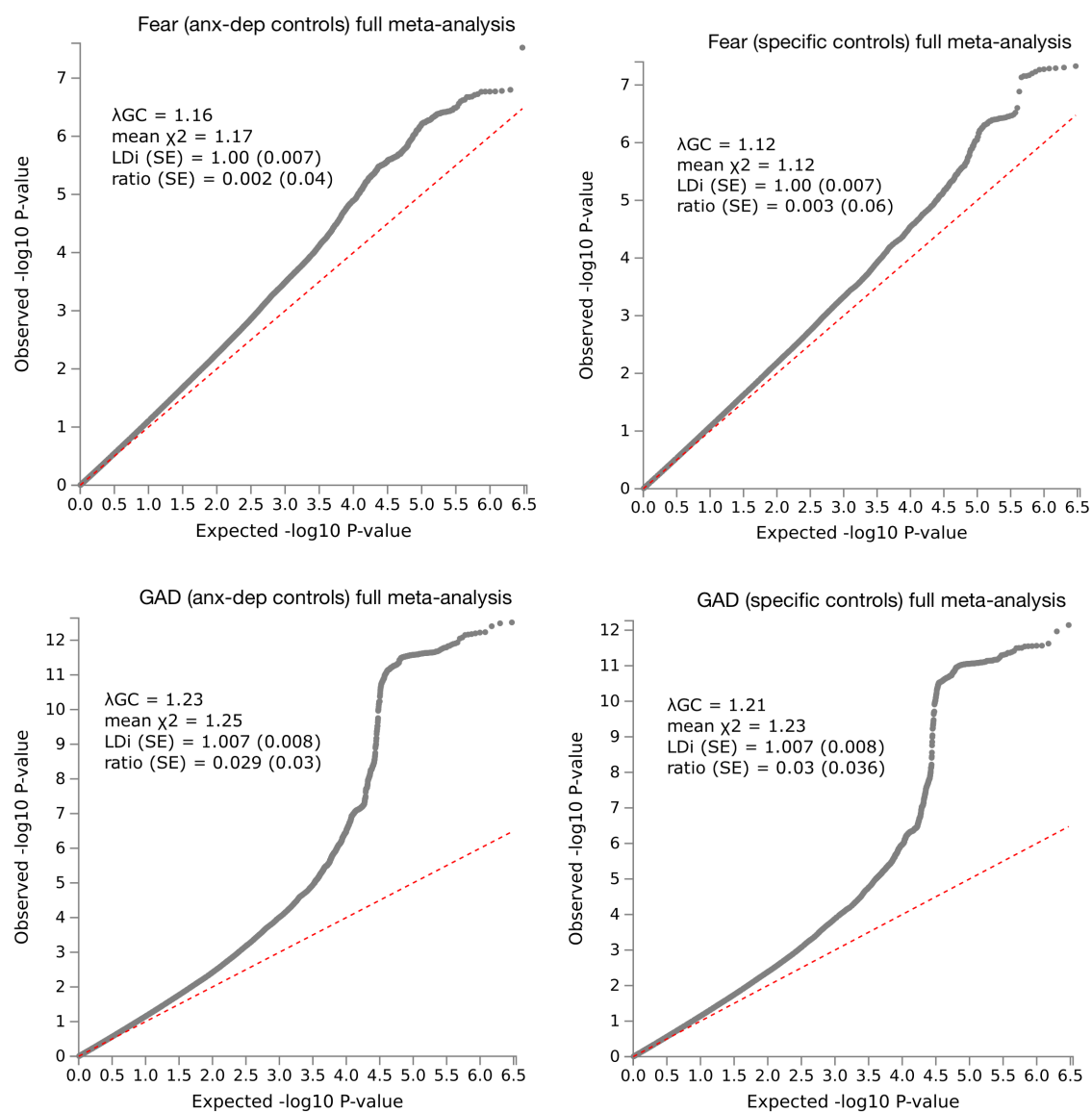
As a sensitivity analysis, we estimated the h^2_{SNP} of GAD and the fear-based disorder where we screened cases and controls using only i) detailed symptom-based measures ii) brief self-report measures. Analyses were restricted to the GLAD+ dataset, where both diagnostic measures were available to screen cases and controls for all five anxiety disorders. Controls were specifically screened for the disorder being analysed. GCTA-GREML was used to estimate h^2_{SNP} on the liability scale using individual-level data³. Sample sizes were unpowered to estimate h^2_{SNP} in LDSC regression.

Supplementary Results

GWAS meta-analyses across GLAD+, QIMR and UKB datasets

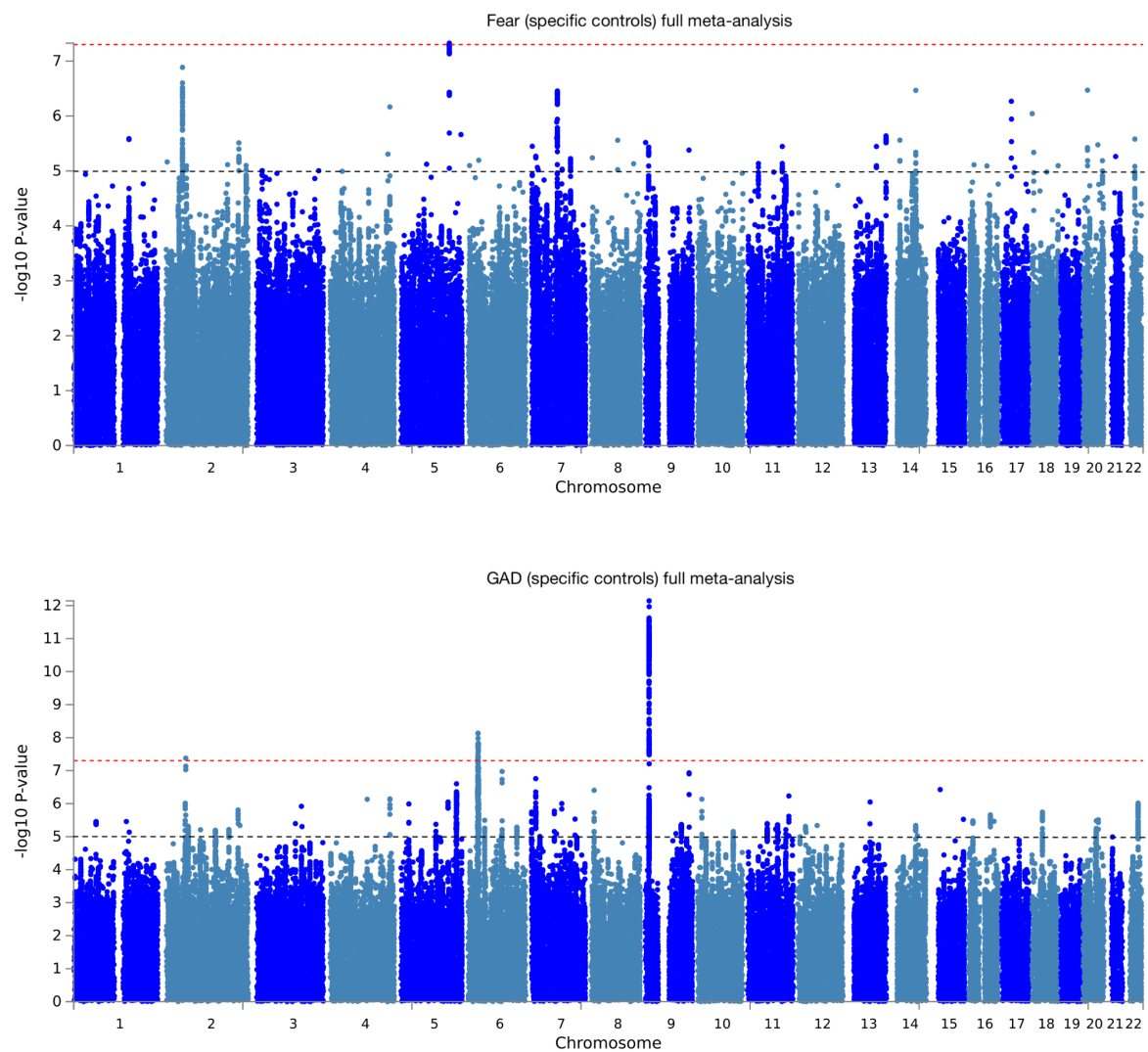
Supplementary Figure 1: Q-Q plots

Q-Q plots of GWAS of fear and GAD meta-analysed across the GLAD+, QIMR and UKB datasets.



Supplementary Figure 2: Manhattan plots

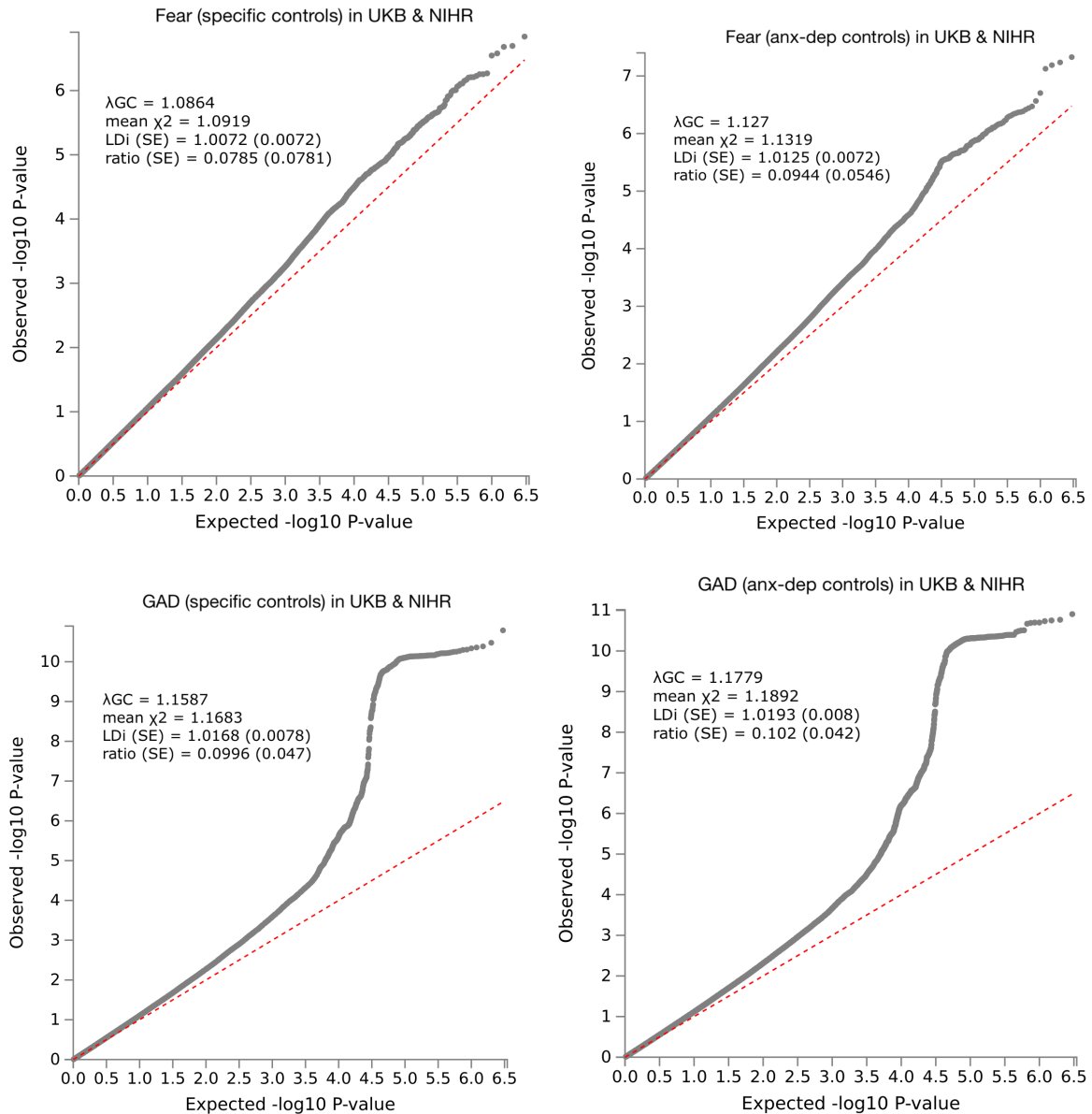
Manhattan plots for GWAS of GAD and fear with specifically screened controls meta-analysed across the GLAD+, QIMR and UKB datasets.



GWAS meta-analyses across GLAD+ and UKB datasets (excluding QIMR)

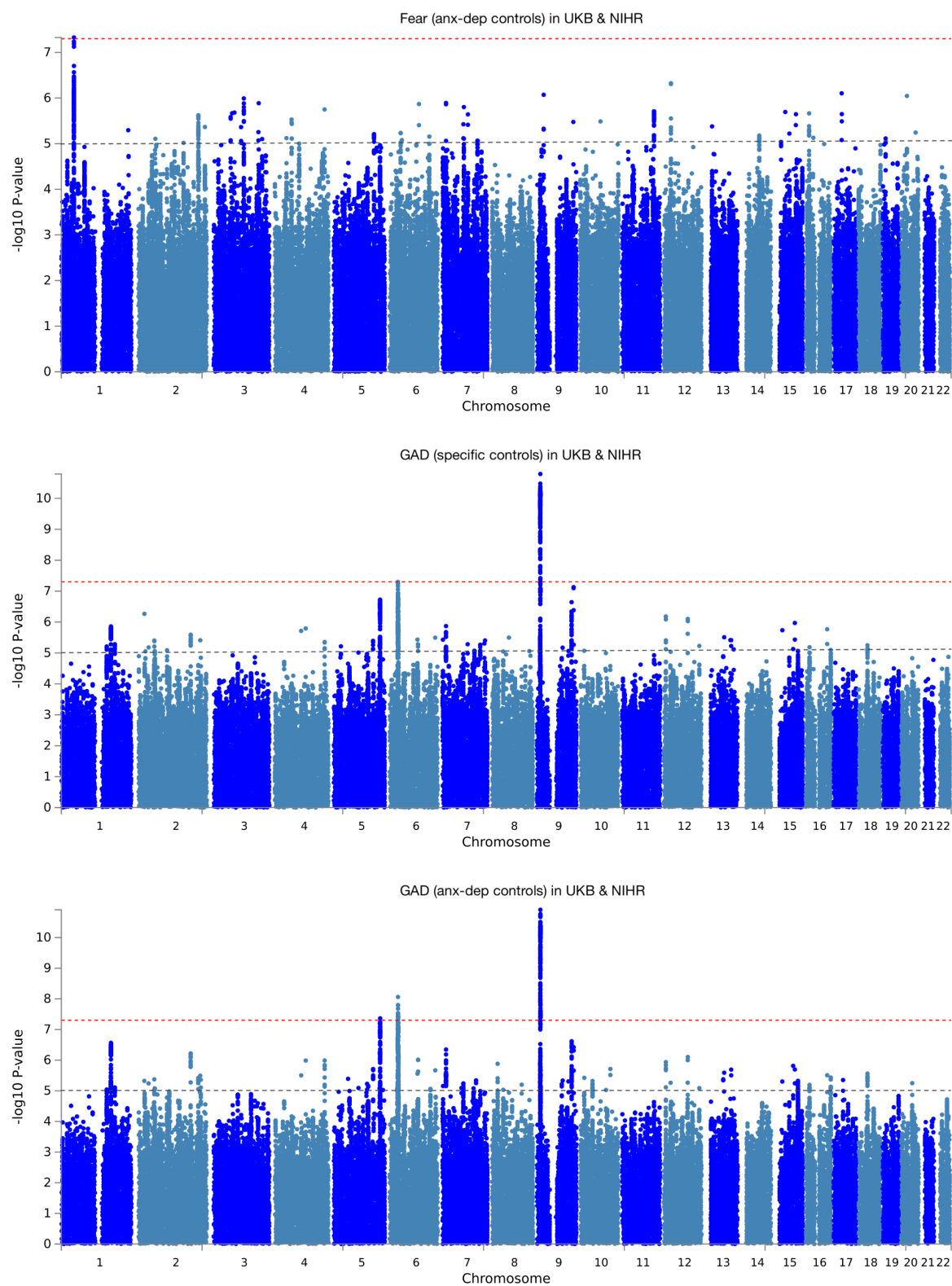
Supplementary Figure 3: Q-Q plots

Q-Q plots for GWAS of GAD and fear meta-analysed across the GLAD+ (NIHR) and UKB datasets.



Supplementary Figure 4: Manhattan plots

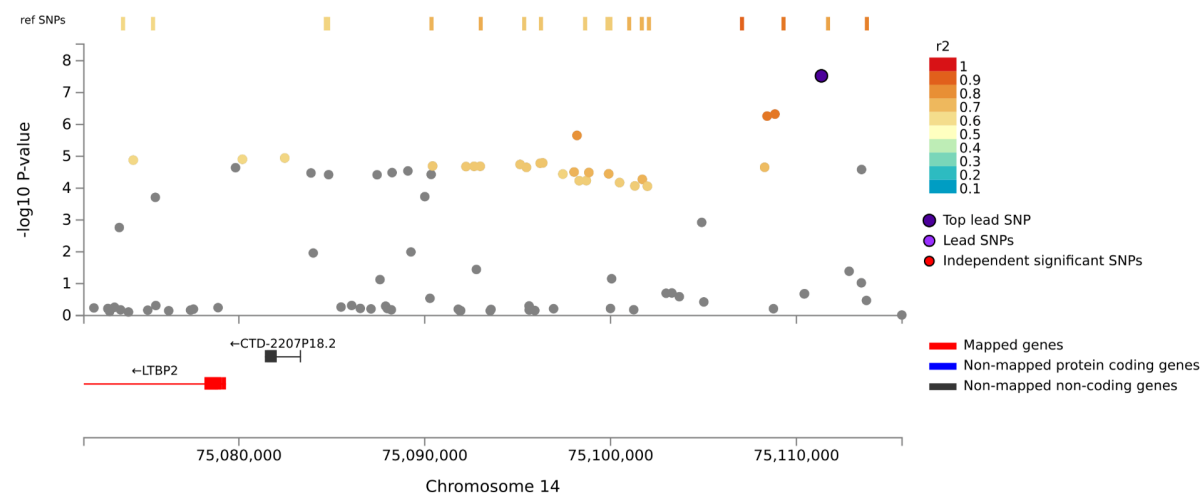
Manhattan plots for GWAS of GAD and fear meta-analysed across the GLAD+ and UKB datasets. No loci reached genome-wide significance in the fear specific GWAS.



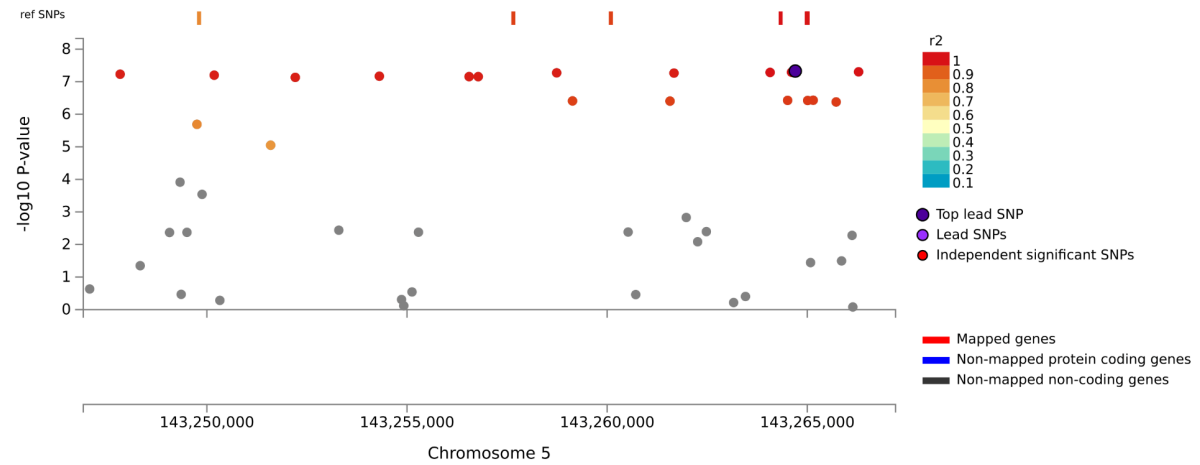
Region plots of genome-wide significant loci

Fear-based results

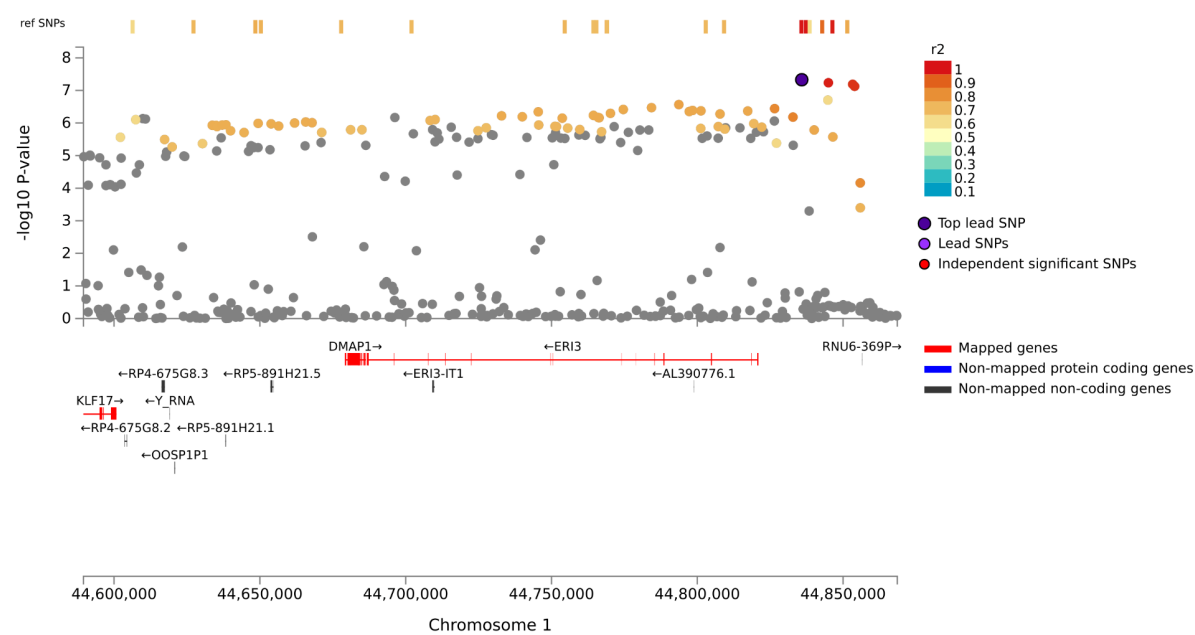
Supplementary Figure 5: Locus on 14q24.3



Supplementary Figure 6: Locus on 5q31.3



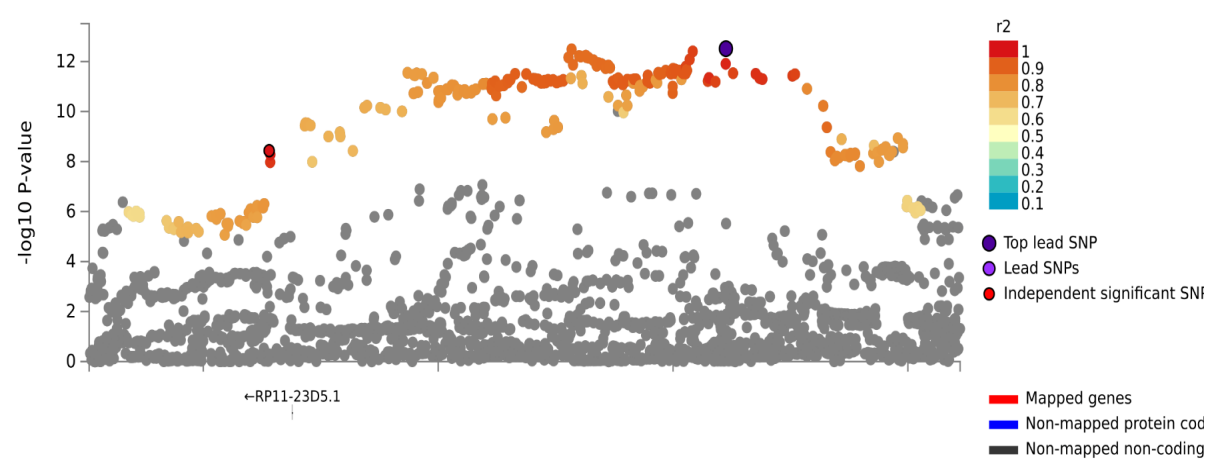
Supplementary Figure 7: Locus on 1p34.1



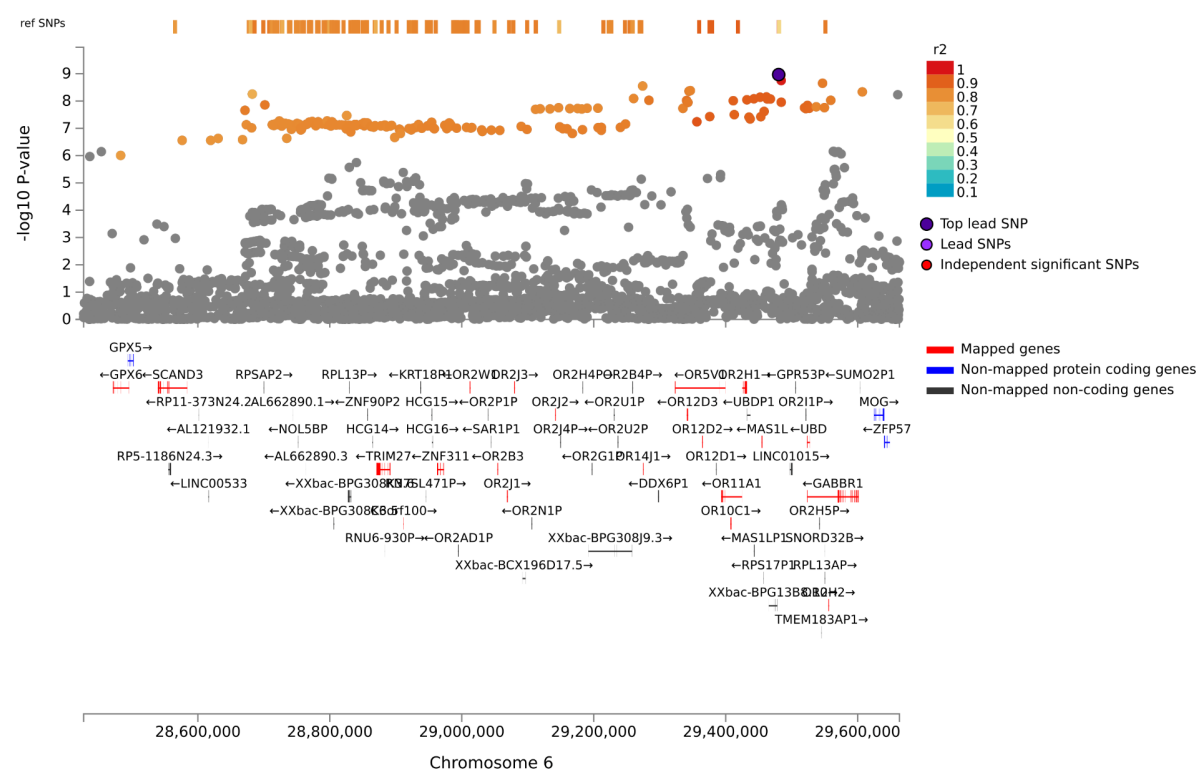
GAD results

Supplementary Figure 8: Locus on 9p23

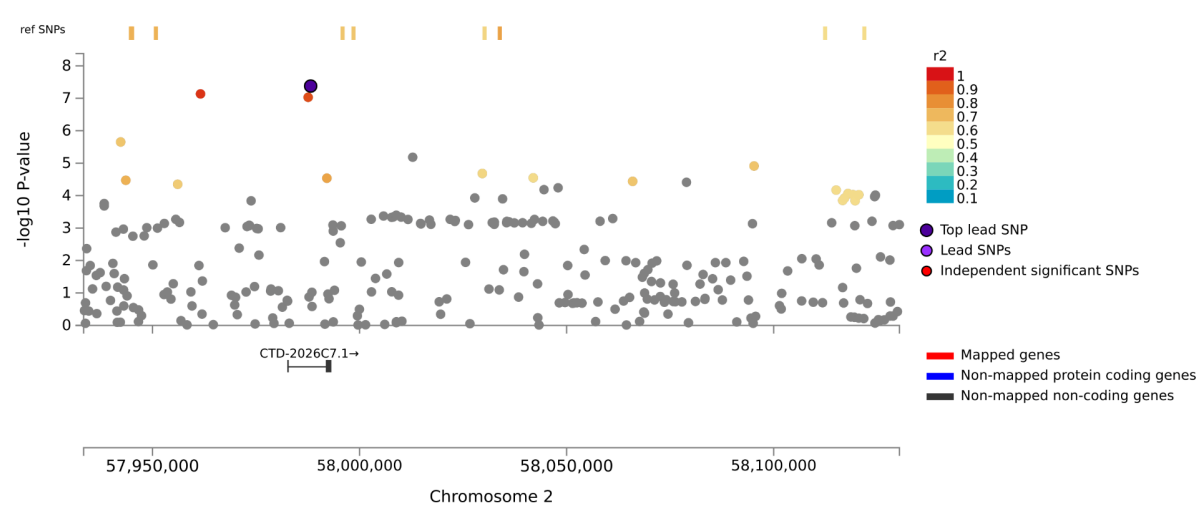
Regional plot for the 9p23 locus shown for the most well-powered GAD GWAS (anx-dep controls in the full meta-analysis).



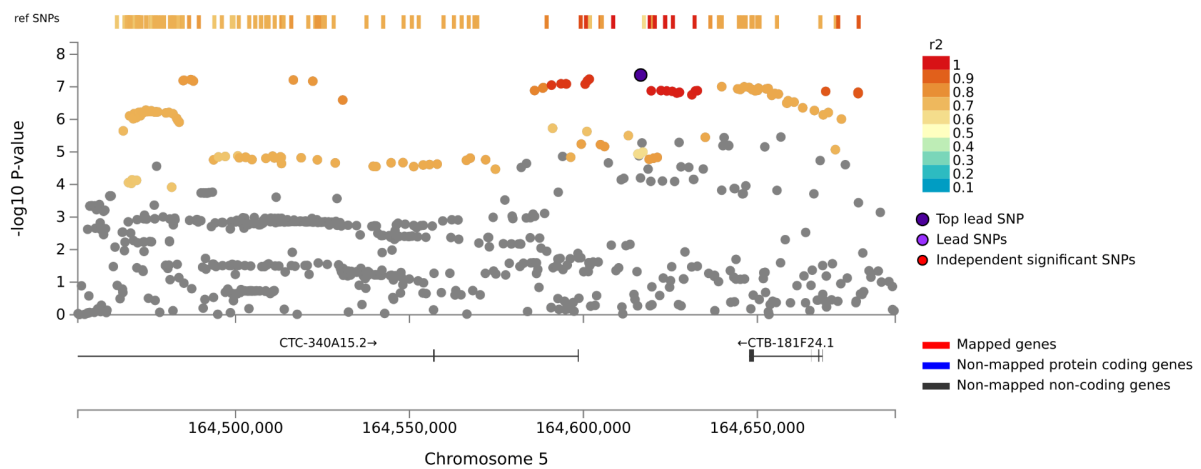
Supplementary Figure 9: Locus on 6p22.1
Regional plot for the 6p22.1 locus shown for the most well-powered GAD GWAS (anx-dep controls in the full meta-analysis).



Supplementary Figure 10: Locus on 2p16.1



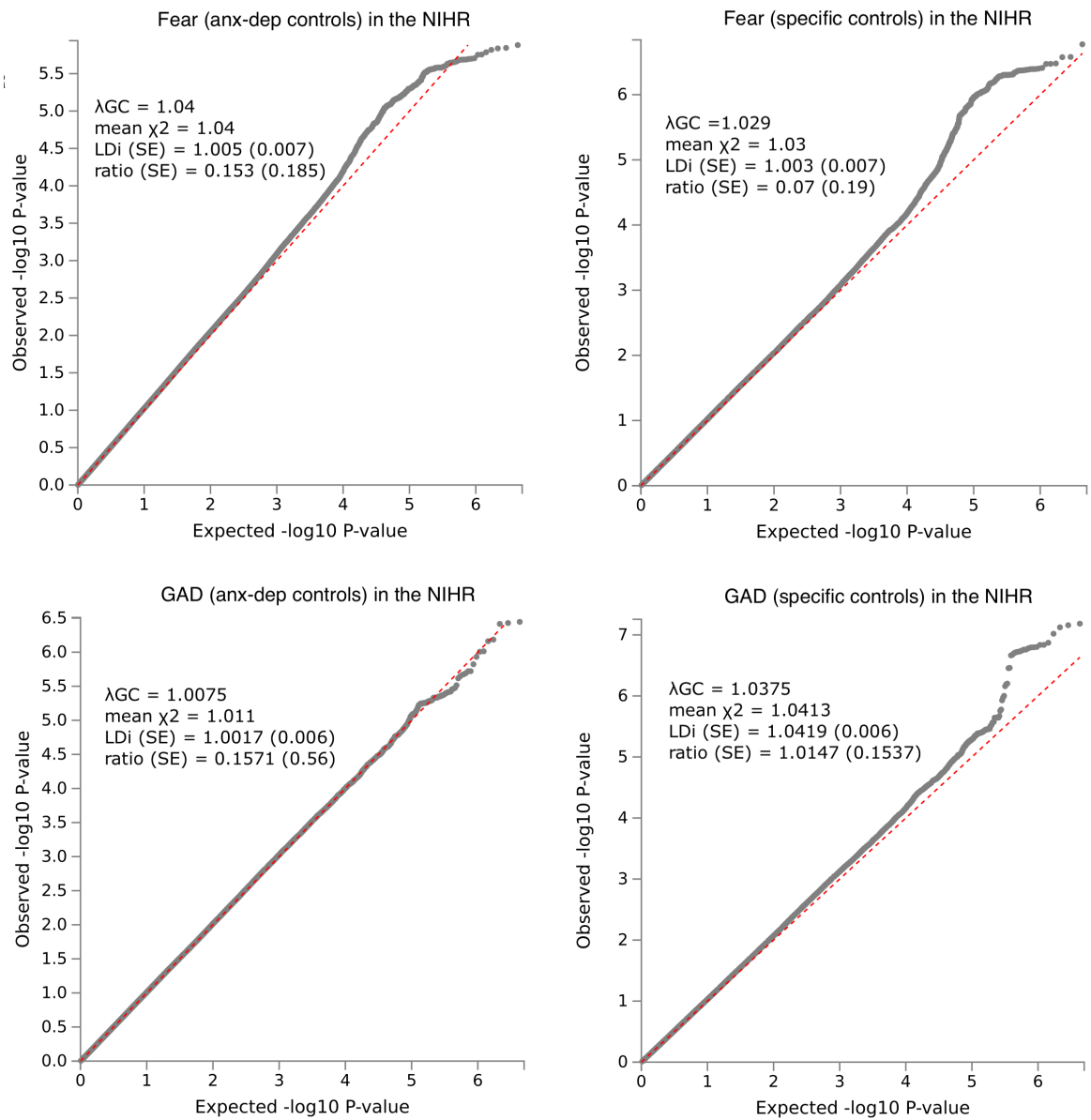
Supplementary Figure 11: Locus on 5q34



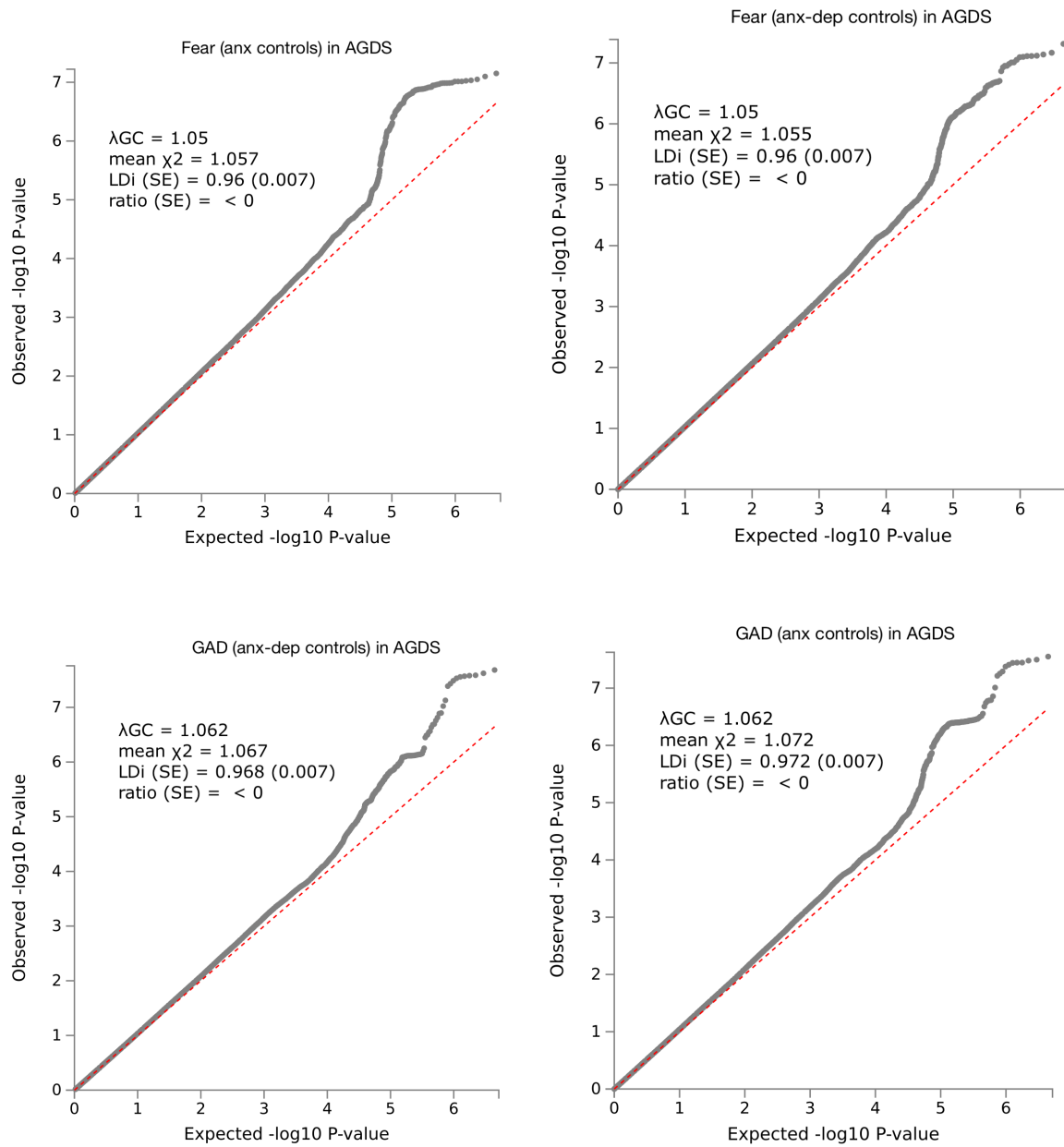
GWAS within each dataset

No genome-wide significant loci were identified for GAD and fear GWASs in the GLAD+ dataset. Manhattan plots in which SNPs exceeded genome-wide significance are shown in **Supplementary Figures 14 & 15** for GWAS in the QIMR dataset and UKB dataset, respectively. In the QIMR dataset, the lead SNP in the fear_{anx-dep} GWAS was intergenic variant rs80182515 on 10q21.1. No loci reached genome-wide significance in the fear_{anx} GWAS (where controls were screened for any anxiety disorder and not depression). The lead SNP in the QIMR GAD_{specific} and GAD_{anx} GWAS was a long non-coding RNA intronic variant rs61547418 on 16p13.3, which mapped onto the gene RBFOX1, also previously reported in GWAS of MDD in QIMR². The lead SNP in the UKB GAD_{specific} and GAD_{anx-dep} GWAS was rs17189482 on 9p23, which was genome-wide significant in the GAD meta-analyses. No loci reached genome-wide significance in UK Biobank fear GWASs.

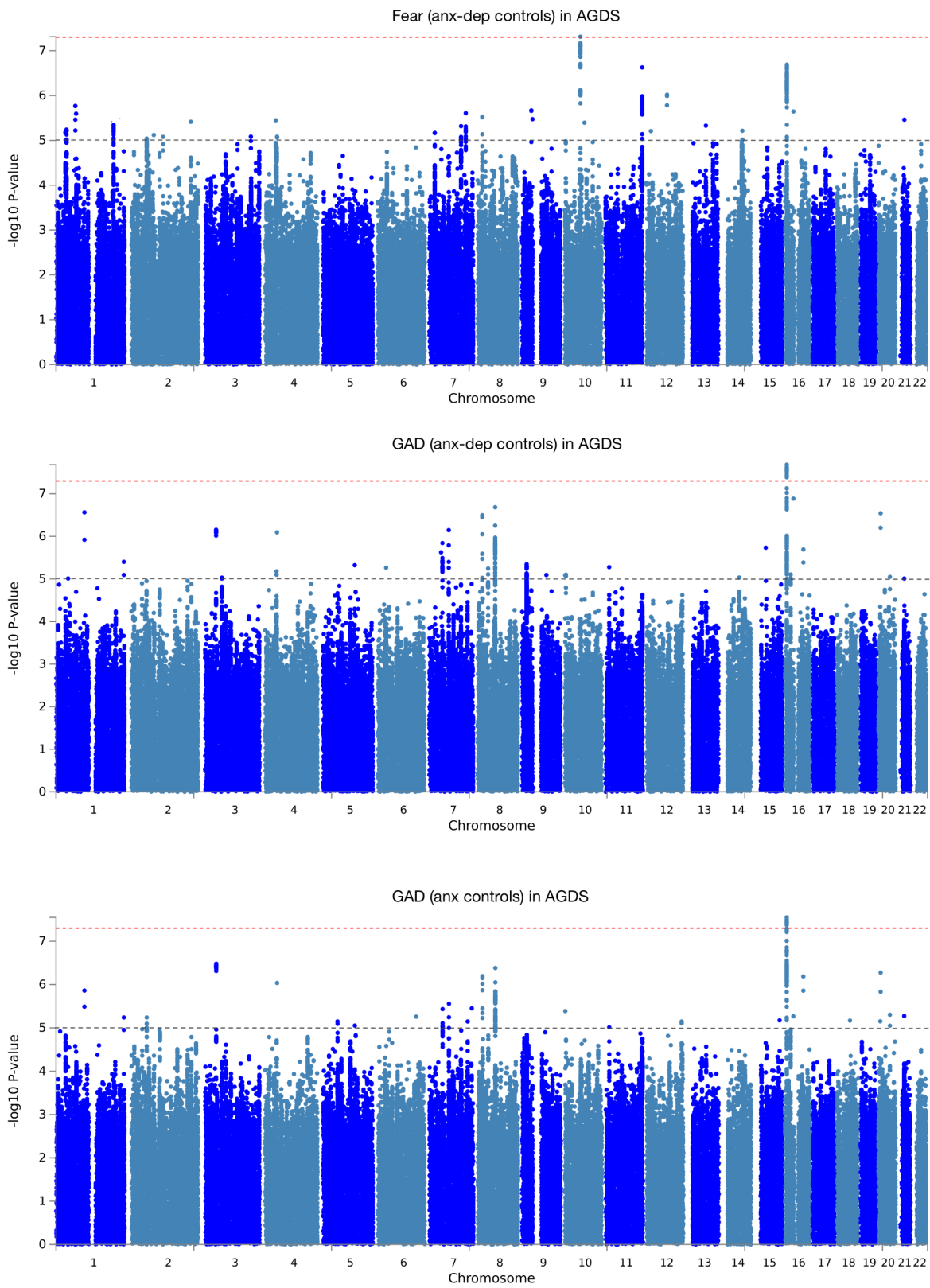
Supplementary Figure 12: Q-Q plots for GAD and fear GWASs in the GLAD+ dataset



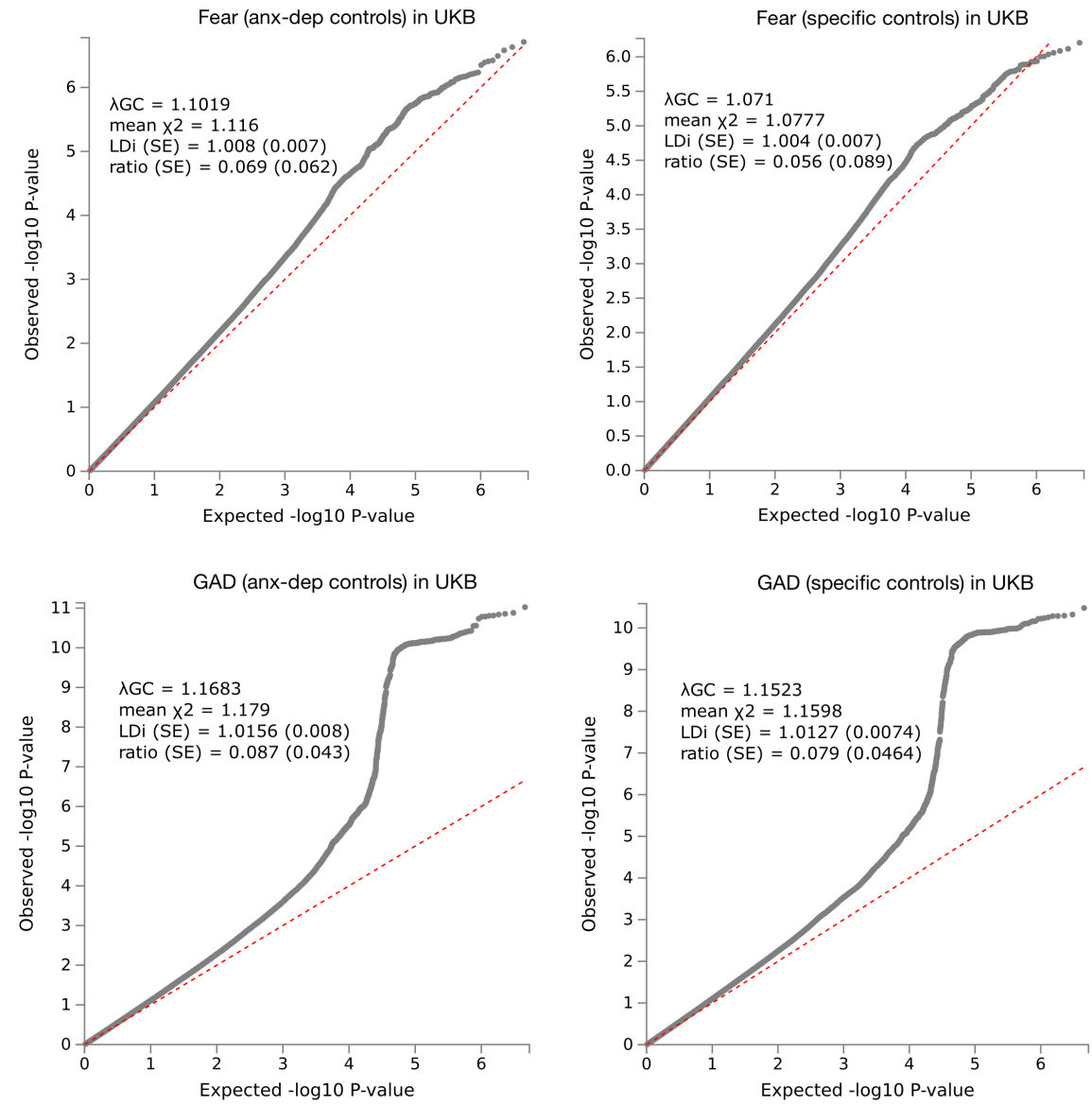
Supplementary Figure 13: Q-Q plots for GAD and fear GWASs in the QIMR dataset



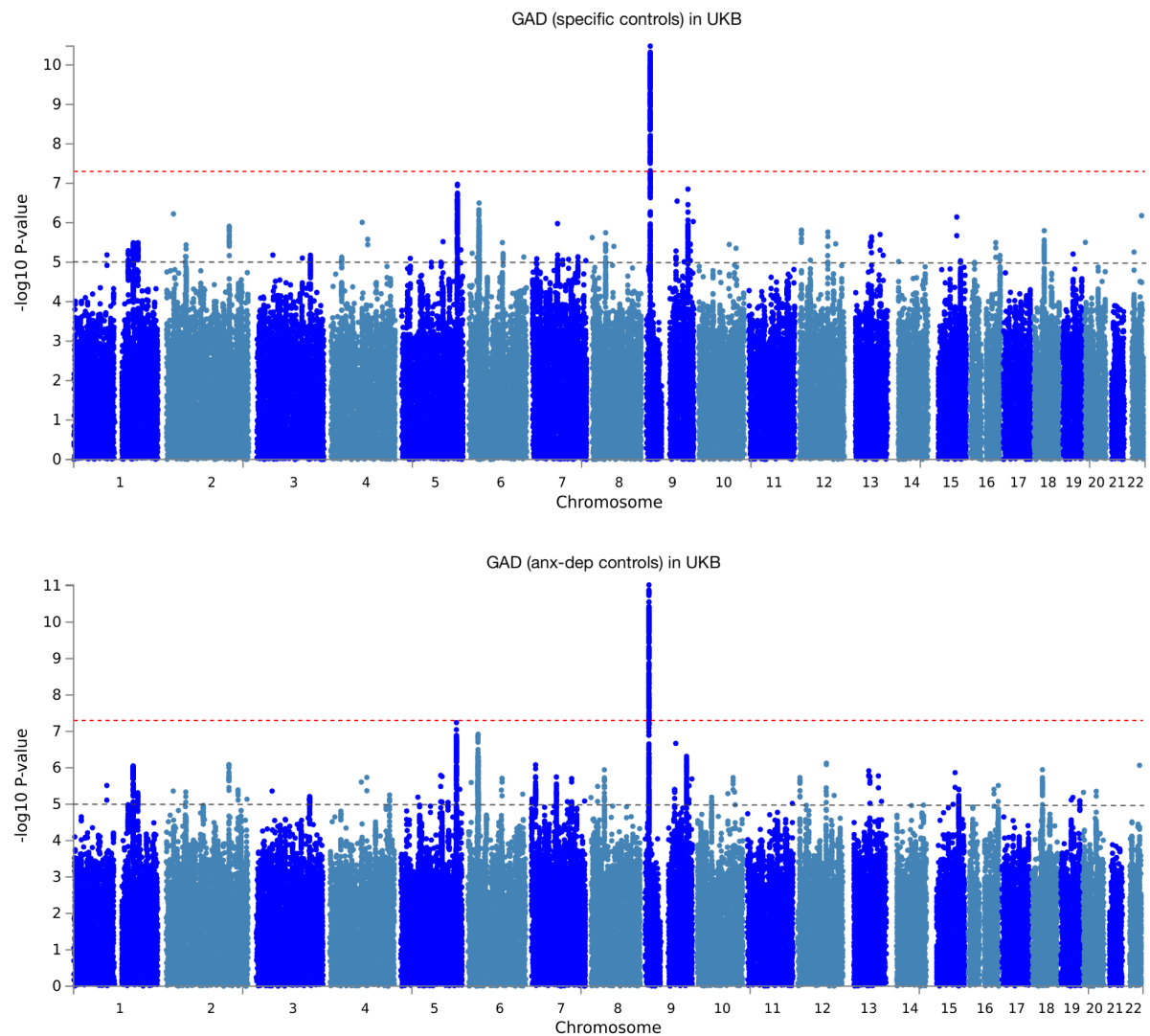
Supplementary Figure 14: Manhattan plots for GAD and fear GWASs in the QIMR dataset



Supplementary Figure 15: Q-Q plots for GAD and fear GWASs in the UKB dataset.



Supplementary Figure 16: Manhattan plots for GAD GWASs in the UKB dataset



References

1. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
2. Mitchell, B. L. *et al.* The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes. *Biol. Psychiatry* (2021) doi:10.1016/j.biopsych.2021.10.021.
3. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).

Supplementary Tables

Supplementary Table 1: Summary of available measures in each study for genome-wide association analysis.										
Study sample	Measure									
	Detailed phenotyping: symptom-based diagnoses					Brief phenotyping: self-report diagnoses				
	GAD	Panic disorder	Agoraphobia	Social phobia	Specific phobia	GAD	Panic disorder and/or panic attacks	Agoraphobia	Social phobia	Specific phobia
GLAD	x	x	x	x	x	x	x	x	x	x
COPING	x	x	x	x	x	x	x	x	x	x
AGDS	x	x	x	x	x	x	x	x	x	x
Qskin*										
UKB	x					x	x	x	x	x
* Only available measure in Qskin is brief, self-report diagnosis of "anxiety", used to screen control for any anxiety disorder.										
x = the availability of measures in each study; GLAD, Genetic Links to Anxiety and Depression Study; COPING, COVID-19 Psychiatric and Neurological Genetics Study; AGDS, Australian Genetics of Depression Study; UKB, UK Biobank, GAD; generalised anxiety disorder										

Supplementary table 2: Self-report diagnoses defined by responses to a single self-report item

Study sample and item phrasing	Response options for screening using self-report diagnoses				
	GAD	Panic	Agoraphobia	Social phobia	Specific phobia
GLAD & COPING: <i>Have you ever been diagnosed with one or more of the following mental health problems by a professional, even if you don't have it currently?</i>	Anxiety, nerves or generalised anxiety disorder	Panic disorder or Panic attacks	Agoraphobia	Social anxiety or social phobia	Specific phobia (e.g. phobia of flying)
AGDS: <i>Have you ever been diagnosed with any of the following?</i>	Anxiety disorder (Generalised anxiety disorder)	Panic disorder	Agoraphobia	Social anxiety disorder (also known as Social phobia)	Specific phobia (e.g. animals, heights, storms, blood, injection/ injury, flying, enclosed spaces)
QSkin: <i>Have you ever been diagnosed with, experienced or been treated for any of the following conditions?</i>	Anxiety Note: only used for screening controls, not cases				
UKB: <i>Have you been diagnosed with one or more of the following mental health problems by a professional, even if you don't have it currently?</i>	Anxiety, nerves or generalised anxiety disorder	Panic attacks	Agoraphobia	Social anxiety or social phobia	Any other phobia (e.g. disabling fear of heights or spiders)
GLAD, Genetic Links to Anxiety and Depression Study; COPING, COVID-19 Psychiatric and Neurological Genetics Study; AGDS, Australian Genetics of Depression Study; UKB, UK Biobank, GAD; generalised anxiety disorder					

Supplementary Table 3: Population prevalence estimates calculated from COPING sample prevalence.

Phenotype	N COPING cases	N COPING total	Population prevalence = COPING sample prevalence	10%	Population prevalence = COPING sample prevalence + 10%	Population prevalence = COPING sample prevalence - 10%
GAD	3210	24903	0.129	0.013	0.142	0.116
Fear	2198	24903	0.088	0.009	0.097	0.079

Sample prevalence was calculated in the full COPING study (not restricted to individuals with genotype data available).

Supplementary Table 4: Independent genome-wide significant loci associated with fear-based disorders, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.

GWASs of other phenotypes are shown if they reported a genome-wide association with one or more of the candidate SNPs identified in this study.

Region	Locus no.	chr	Trait	PMID	First Auth	P	bp	snp	Mapped Gene	Strongest	Context
Region 14q24.3											
14q24.3	1	14	Systolic blood pressure	30224653	Evangelou E	3.00E-13	75074316	rs11159091	LTBP2	rs11159091-A	intron
14q24.3	1	14	Systolic blood pressure	30578418	Giri A	2.00E-11	75098313	rs11159096	LTBP2 - AREL1	rs11159096-A	regulatory region
14q24.3	1	14	Worry	29942085	Nagel M	1.00E-08	75109195	rs888415	LTBP2 - AREL1	rs888415-A	regulatory region
14q24.3	1	14	Educational attainment (MTAG)	30038396	Lee JJ	1.00E-08	75113682	rs7154563	LTBP2 - AREL1	rs7154563-T	intergenic
14q24.3	1	14	Worry	29942085	Nagel M	5.00E-08	75099883	rs11159097	LTBP2 - AREL1	rs11159097-C	intergenic
Region 1p34.1											
1p34.1	3	1	Anxiety	33859377	Thorp JG	9.00E-08	44751710	rs325163	ERI3	rs325163-T	intron
1p34.1	3	1	Metabolite levels	23823483	Rhee EP	3.00E-06	44602159	rs6662888	KLF17 - KLF18	rs6662888-?	intergenic
1p34.1	3	1	Reaction time	29844566	Davies G	3.00E-06	44701277	rs7525929	ERI3	rs7525929-?	intron
1p34.1	3	1	Reaction time	29844566	Davies G	3.00E-06	44826511	rs12078192	ERI3 - RNU6-369P	rs12078192-?	intergenic
1p34.1	3	1	Metabolite levels	23823483	Rhee EP	7.00E-06	44630318	rs6700522	OOSP1P1 - DMAP1	rs6700522-?	intergenic

Supplementary Table 5: Independent genome-wide significant loci associated with generalised anxiety disorder, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.

GWASs of other phenotypes are shown if they reported a genome-wide association with one or more of the candidate SNPs identified in this study.

Region	Locus no.	chr	Trait	PMID	FirstAuth	P	bp	snp	MappedGene	Strongest	Context
Region 9p23											
9p23	4	9	Neuroticism	30643256	Baselmans BML	1.00E-19	11651144	rs10960067		rs10960067-G	intron
9p23	4	9	Depressive symptoms	30643256	Baselmans BML	2.00E-19	11651144	rs10960067		rs10960067-G	intron
9p23	4	9	Well-being spectrum	30643256	Baselmans BML	6.00E-17	11455955	rs10491952		rs10491952-T	intergenic
9p23	4	9	Neuroticism	29255261	Luciano M	2.00E-16	11453149	rs56116032		rs56116032-A	intergenic
9p23	4	9	Neuroticism	29255261	Luciano M	2.00E-16	11651144	rs10960067		rs10960067-A	intron
9p23	4	9	Neuroticism	30643256	Baselmans BML	8.00E-16	11139180	rs10959560		rs10959560-T	intergenic
9p23	4	9	Depressive symptoms	30643256	Baselmans BML	8.00E-16	11139180	rs10959560		rs10959560-T	intergenic
9p23	4	9	Neuroticism	30643256	Baselmans BML	1.00E-15	11400078	rs10959767		rs10959767-G	intron
9p23	4	9	Depressive symptoms	30643256	Baselmans BML	1.00E-15	11400078	rs10959767		rs10959767-G	intron
9p23	4	9	Depression	30718901	Howard DM	1.00E-15	11513019	rs1982277		rs1982277-T	intergenic

9p23	4	9	Positive affect	30643256	Baselmans BML	6.00E-14	11139180	rs10959560		rs10959560-T	intergenic
9p23	4	9	Bipolar disorder or MDD	31926635	Coleman JRI	1.00E-13	11379630	rs10959753		rs10959753-C	intron
9p23	4	9	Anxiety	33859377	Thorp JG	1.00E-13	11553381	rs10959925	AKAP8P1 - JKAMPP1	rs10959925-G	intergenic
9p23	4	9	Neuroticism	29255261	Luciano M	2.00E-13	11786602	rs1565001		rs1565001-C	intron
9p23	4	9	Neuroticism	29255261	Luciano M	3.00E-13	11699270	rs10960103		rs10960103-G	intron
9p23	4	9	Neuroticism	29942085	Nagel M	1.00E-12	11449346	rs62550480		rs62550480-T	intergenic
9p23	4	9	Neuroticism	30595370	Kichaev G	7.00E-12	11453149	rs56116032		rs56116032-?	intergenic
9p23	4	9	Neuroticism	29942085	Nagel M	1.00E-11	11257041	rs11496294		rs11496294-C	intergenic
9p23	4	9	Depressive symptoms	29292387	Turley P	2.00E-11	11557797	rs10809520		rs10809520-T	intergenic
9p23	4	9	Neuroticism (MTAG)	29292387	Turley P	2.00E-11	11557797	rs10809520		rs10809520-T	intergenic
9p23	4	9	Lifetime anxiety disorder	31748690	Purves KL	3.00E-11	11519984	rs10959883		rs10959883-C	intergenic
9p23	4	9	Mood instability	31168069	Ward J	6.00E-11	11453149	rs56116032		rs56116032-?	intergenic
9p23	4	9	Neuroticism	29500382	Nagel M	9.00E-11	11449346	rs62550480		rs62550480-T	intergenic
9p23	4	9	Neuroticism	29255261	Luciano M	1.00E-10	11369213	rs12378446		rs12378446-C	intron
9p23	4	9	Neuroticism	29500382	Nagel M	1.00E-10	11703892	rs56318386		rs56318386-T	intron
9p23	4	9	Neuroticism	27089181	Okbay A	2.00E-10	11699270	rs10960103		rs10960103-C	intron
9p23	4	9	Depressive symptoms	29292387	Turley P	5.00E-10	11557797	rs10809520		rs10809520-T	intergenic
9p23	4	9	General factor of neuroticism	30867560	Hill WD	7.00E-10	11453149	rs56116032		rs56116032-A	intergenic
9p23	4	9	Experiencing mood swings	29500382	Nagel M	1.00E-09	11703892	rs56318386		rs56318386-T	intron
9p23	4	9	Experiencing mood swings	29500382	Nagel M	2.00E-09	11445598	rs10959813	PTPRD - NA	rs10959813-A	intergenic
9p23	4	9	Anxiety symptoms	31748690	Purves KL	4.00E-09	11513617	rs17189482		rs17189482-G	intergenic
9p23	4	9	Depression	29700475	Wray NR	5.00E-09	11544964	rs10959913		rs10959913-T	intergenic
9p23	4	9	Neuroticism	29500382	Nagel M	6.00E-09	11257041	rs11496294		rs11496294-C	intergenic
9p23	4	9	Depression	27089181	Okbay A	6.00E-09	11557797	rs10809520		rs10809520-T	intergenic
9p23	4	9	Mood instability	29187730	Ward J	8.00E-09	11459410	rs10959826		rs10959826-G	intergenic
9p23	4	9	Feeling hurt	29500382	Nagel M	9.00E-09	11256918	rs10117184		rs10117184-A	intergenic
9p23	4	9	Neuroticism	27067015	Smith DJ	9.00E-09	11369213	rs12378446		rs12378446-T	intron
9p23	4	9	Adolescent idiopathic scoliosis	30019117	Liu J	1.00E-08	11701596	rs10809558		rs10809558-?	intron
9p23	4	9	Feeling worry	29500382	Nagel M	2.00E-08	11431990	rs10959797		rs10959797-A	intron
9p23	4	9	Worry	29942085	Nagel M	2.00E-08	11610677	rs62551581		rs62551581-C	intergenic
9p23	4	9	Body mass index	31669095	Zhu Z	3.00E-08	11727896	rs62555664		rs62555664-?	intron
9p23	4	9	Major depressive disorder	31969693	Coleman JRI	4.00E-08	11256041	rs11515172		rs11515172-T	intergenic
9p23	4	9	Feeling worry	29500382	Nagel M	4.00E-08	11610896	rs62551583		rs62551583-A	intergenic
Region 6p22.1											
6p22.1	5	6	Peak expiratory flow	30804560	Shrine N	1.00E-48	28301099	rs7752448	ZSCAN31	rs7752448-A	intron
6p22.1	5	6	Schizophrenia	29483656	Pardinas AF	2.00E-44	29206683	rs3130820	OR2G1P - OR2U1P	rs3130820-?	intron
6p22.1	5	6	Hemoglobin levels	32327693	Oskarsson GR	2.00E-38	29342775	rs3749971	OR12D3, OR5V1	rs3749971-A	missense
6p22.1	5	6	FEV1	30804560	Shrine N	3.00E-35	28301099	rs7752448	ZSCAN31	rs7752448-A	intron

6p22.1	5	6	Schizophrenia	31740837	Lam M	7.00E-30	29144532	rs9257566	OR2J2 - OR2J4P	rs9257566-?	intergenic
6p22.1	5	6	Schizophrenia	30285260	Ikeda M	1.00E-29	29206683	rs150817755	OR2G1P - OR2U1P	rs150817755-T	intron
6p22.1	5	6	Schizophrenia	26198764	Goes FS	2.00E-29	28306671	rs13217619	ZSCAN31	rs13217619-T	intron
6p22.1	5	6	Schizophrenia	26198764	Goes FS	2.00E-28	29244219	rs144447022	OR2U2P - OR14J1	rs144447022-G	intron
6p22.1	5	6	Schizophrenia (MTAG)	32606422	Wu Y	9.00E-28	29244219	rs144447022	OR2U2P - OR14J1	rs144447022-?	intron
6p22.1	5	6	Anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, or Tourette syndrome (pleiotropy)	31835028	Cross-Disorder Group of the Psychiatric Genomics Consortium	1.00E-27	28306671	rs13217619	ZSCAN31	rs13217619-?	intron
6p22.1	5	6	Autism spectrum disorder or schizophrenia	28540026	Anney RIL	3.00E-26	28667669	rs116137698	LINC00533 - RPSAP2	rs116137698-A	intergenic
6p22.1	5	6	Cardiometabolic and hematological traits	27668658	Iotchkova V	3.00E-26	29118747	rs3130725	OR2J3 - OR2J2	rs3130725-G	intergenic
6p22.1	5	6	Schizophrenia	30285260	Ikeda M	1.00E-25	28983274	rs116757206	ZNF311 - OR2AD1P	rs116757206-G	intergenic
6p22.1	5	6	Schizophrenia	30285260	Ikeda M	7.00E-25	29342775	rs114071887	OR5V1, OR12D3	rs114071887-G	missense
6p22.1	5	6	Apolipoprotein A1 levels	32203549	Richardson TG	9.00E-25	28682725	rs9393926	LINC00533 - RPSAP2	rs9393926-G	intergenic
6p22.1	5	6	Lung function (FVC)	30804560	Shrine N	1.00E-22	28301099	rs7752448	ZSCAN31	rs7752448-A	intron
6p22.1	5	6	Major depressive disorder (MTAG)	33479212	Yao X	2.00E-22	28366151	rs2232423	ZSCAN12	rs2232423-A	missense
6p22.1	5	6	Well-being spectrum (multivariate analysis)	30643256	Baselmans BML	6.00E-21	28349698	rs13213152	ZSCAN12	rs13213152-G	3_prime_UTR
6p22.1	5	6	Depressive symptoms	30643256	Baselmans BML	7.00E-21	28349698	rs13213152	ZSCAN12	rs13213152-G	3_prime_UTR
6p22.1	5	6	Sarcoidosis (Lofgren's syndrome vs non-Lofgren's syndrome)	26651848	Rivera NV	1.00E-20	29356331	rs9257809	OR5V1	rs9257809-G	intron
6p22.1	5	6	Sarcoidosis (Lofgren's syndrome vs non-Lofgren's syndrome)	26651848	Rivera NV	1.00E-20	29461730	rs1233491	RPS17P1 - LINC02829	rs1233491-C	intergenic
6p22.1	5	6	HDL cholesterol levels	32203549	Richardson TG	2.00E-20	28992159	rs3129797	ZNF311 - OR2AD1P	rs3129797-C	intergenic
6p22.1	5	6	Sarcoidosis (Lofgren's syndrome vs non-Lofgren's syndrome)	26651848	Rivera NV	7.00E-20	28803291	rs9257248	NOP56P1 - LINC01623	rs9257248-A	intergenic
6p22.1	5	6	Bipolar disorder (MTAG)	32606422	Wu Y	7.00E-20	29244219	rs144447022	OR2U2P - OR14J1	rs144447022-?	intron
6p22.1	5	6	Cognitive performance	30038396	Lee JJ	8.00E-20	28934352	rs148696809	LINC01556 - KRT18P1	rs148696809-T	intergenic
6p22.1	5	6	Blood protein levels	28240269	Suhre K	1.00E-19	29283672	rs3129682	OR14J1 - DDX6P1	rs3129682-T	intergenic
6p22.1	5	6	Sarcoidosis (Lofgren's syndrome vs non-Lofgren's syndrome)	26651848	Rivera NV	4.00E-19	28905791	rs3130895	TRIM27 - LINC01556	rs3130895-A	intergenic
6p22.1	5	6	Neuropsychiatric disorders	33479212	Yao X	5.00E-19	28354835	rs67981811	ZSCAN12	rs67981811-C	3_prime_UTR
6p22.1	5	6	Sarcoidosis (Lofgren's syndrome vs non-Lofgren's syndrome)	26651848	Rivera NV	7.00E-19	29057639	rs3129788	OR2B3 - OR2J1	rs3129788-A	intergenic

6p22.1	5	6	Lung cancer	28604730	McKay JD	7.00E-19	29607101	rs115870917	SUMO2P1 - MOG	rs115870917-C	intergenic
6p22.1	5	6	Body size at age 10	32376654	Richardson TG	2.00E-18	28831611	rs3131336	LINC01623	rs3131336-C	intron
6p22.1	5	6	Sarcoidosis (Lofgren's syndrome vs non-Lofgren's syndrome)	26651848	Rivera NV	5.00E-18	28349698	rs13213152	ZSCAN12	rs13213152-G	3_prime_UTR
6p22.1	5	6	Bipolar disorder or major depressive disorder	31926635	Coleman JRI	7.00E-18	28673998	rs145410455	LINC00533 - RPSAP2	rs145410455-G	intergenic
6p22.1	5	6	Cognitive aspects of educational attainment	33414549	Demange PA	2.00E-17	28411941	rs67381177	ZSCAN23 - COX11P1	rs67381177-G	regulatory_region
6p22.1	5	6	Verbal-numerical reasoning	32895543	de la Fuente J	2.00E-17	28678357	rs7775835	LINC00533 - RPSAP2	rs7775835-?	intergenic
6p22.1	5	6	Lung function (FEV1/FVC)	30804560	Shrine N	7.00E-16	28301099	rs7752448	ZSCAN31	rs7752448-A	intron
6p22.1	5	6	Brain morphology (MOSTest)	32665545	van der Meer D	7.00E-16	29607101	rs3131856	SUMO2P1 - MOG	rs3131856-?	intergenic
6p22.1	5	6	General cognitive ability	29844566	Davies G	5.00E-15	29372323	rs429479	OR5V1	rs429479-A	intron
6p22.1	5	6	Urate levels	31578528	Tin A	7.00E-15	28301099	rs7752448	ZSCAN31	rs7752448-A	intron
6p22.1	5	6	Triglyceride levels	32203549	Richardson TG	1.00E-14	28796071	rs3131337	NOP56P1 - LINC01623	rs3131337-G	intergenic
6p22.1	5	6	Intelligence (MTAG)	29326435	Hill WD	1.00E-13	29244219	rs144447022	OR2U2P - OR14J1	rs144447022-T	intron
6p22.1	5	6	Number of sexual partners	30643258	Karlsson Linner R	2.00E-13	28671343	rs6908726	LINC00533 - RPSAP2	rs6908726-C	intergenic
6p22.1	5	6	Triglycerides	30275531	Klarin D	3.00E-13	28793904	rs3118359	NOP56P1 - LINC01623	rs3118359-T	intergenic
6p22.1	5	6	Aerodigestive squamous cell cancer (pleiotropy)	33667223	Lesueur C	3.00E-13	29415464	rs115482250	OR11A1	rs115482250-A	intron
6p22.1	5	6	Aerodigestive squamous cell cancer (pleiotropy)	33667223	Lesueur C	6.00E-13	28827879	rs116711004	LINC01623	rs116711004-T	Non coding transcript exon
6p22.1	5	6	Depression (broad)	29662059	Howard DM	8.00E-13	29549653	rs1233393	GABBR1	rs1233393-G	intron
6p22.1	5	6	Aerodigestive squamous cell cancer (pleiotropy)	33667223	Lesueur C	1.00E-12	28318331	rs35744819	ZSCAN31, ZKSCAN3	rs35744819-T	intron
6p22.1	5	6	Depression (broad)	29662059	Howard DM	1.00E-12	28345282	rs35017208	ZKSCAN3 - ZSCAN12	rs35017208-T	intergenic
6p22.1	5	6	Lung cancer	28604730	McKay JD	1.00E-12	28796071	rs116461399	NOP56P1 - LINC01623	rs116461399-C	intergenic
6p22.1	5	6	Lung cancer	28604730	McKay JD	1.00E-12	29254982	rs139244745	OR2U2P - OR14J1	rs139244745-G	intron
6p22.1	5	6	Depression (broad)	29662059	Howard DM	1.00E-12	29415464	rs2523443	OR11A1	rs2523443-A	intron
6p22.1	5	6	Lung cancer	28604730	McKay JD	2.00E-12	28281894	rs34661125	PGBD1 - SMIM15P2	rs34661125-A	intergenic
6p22.1	5	6	Pork consumption	32066663	Niarchou M	2.00E-12	28314880	rs35016036	ZSCAN31	rs35016036-?	intron
6p22.1	5	6	Oily fish consumption	32066663	Niarchou M	2.00E-12	28314880	rs35016036	ZSCAN31	rs35016036-?	intron
6p22.1	5	6	Cognitive traits (MTAG)	32895543	de la Fuente J	2.00E-12	29206251	rs3117342	OR2G1P	rs3117342-?	intron
6p22.1	5	6	Depression (broad)	29662059	Howard DM	2.00E-12	29254982	rs9257681	OR2U2P - OR14J1	rs9257681-G	intron
6p22.1	5	6	Aerodigestive squamous cell cancer (pleiotropy)	33667223	Lesueur C	2.00E-12	29254982	rs139244745	OR2U2P - OR14J1	rs139244745-G	intron
6p22.1	5	6	Urate levels	31578528	Tin A	2.00E-12	29372323	rs429479	OR5V1	rs429479-A	intron
6p22.1	5	6	Aerodigestive squamous cell cancer (pleiotropy)	33667223	Lesueur C	3.00E-12	28436145	rs55690788	OR2E1P - GPX6	rs55690788-C	intergenic

6p22.1	5	6	Lung cancer	28604730	McKay JD	3.00E-12	28934352	rs148696809	LINC01556 - KRT18P1	rs148696809-C	intergenic
6p22.1	5	6	Depression (broad)	29662059	Howard DM	3.00E-12	29133642	rs4947263	OR2J3 - OR2J2	rs4947263-T	Non coding transcript exon
6p22.1	5	6	Depressive symptoms (MTAG)	32606422	Wu Y	3.00E-12	29244219	rs144447022	OR2U2P - OR14J1	rs144447022-?	intron
6p22.1	5	6	Squamous cell lung carcinoma	28604730	McKay JD	3.00E-12	29254982	rs139244745	OR2U2P - OR14J1	rs139244745-G	intron
6p22.1	5	6	Depression (broad)	29662059	Howard DM	4.00E-12	28482917	rs13191038	GPX6, GPX6	rs13191038-C	intron
6p22.1	5	6	Help-seeking from a GP	32231276	Cai N	4.00E-12	28795856	rs3135296	NOP56P1 - LINC01623	rs3135296-A	intergenic
6p22.1	5	6	Aerodigestive squamous cell cancer (pleiotropy)	33667223	Lesueur C	4.00E-12	28939283	rs115218101	KRT18P1 - RN7SL471P	rs115218101-C	intergenic
6p22.1	5	6	Lung cancer	28604730	McKay JD	4.00E-12	29148394	rs116826541	OR2J2 - OR2J4P	rs116826541-G	intergenic
6p22.1	5	6	Urate levels	31578528	Tin A	6.00E-12	28839908	rs3118365	LINC01623 - ZNF90P2	rs3118365-A	regulatory_region
6p22.1	5	6	Lung cancer	28604730	McKay JD	7.00E-12	28394680	rs13201681	RNU2-45P - ZSCAN23	rs13201681-T	intergenic
6p22.1	5	6	General cognitive ability	29844566	Davies G	9.00E-12	28678357	rs7775835	LINC00533 - RPSAP2	rs7775835-T	intergenic
6p22.1	5	6	Depression (broad)	29662059	Howard DM	1.00E-11	28630691	rs13194504	LINC00533 - RPSAP2	rs13194504-A	intergenic
6p22.1	5	6	Squamous cell lung carcinoma	28604730	McKay JD	1.00E-11	28796071	rs116461399	NOP56P1 - LINC01623	rs116461399-C	intergenic
6p22.1	5	6	Aerodigestive squamous cell cancer (pleiotropy)	33667223	Lesueur C	1.00E-11	29044701	rs116345369	SAR1AP1	rs116345369-A	Non coding transcript exon
6p22.1	5	6	Estimated glomerular filtration rate in diabetes	31451708	Hellwege JN	1.00E-11	29413003	rs2023463	OR11A1	rs2023463-T	intron
6p22.1	5	6	Chronic obstructive pulmonary disease	33909500	Moll M	2.00E-11	28366151	rs2232423	ZSCAN12	rs2232423-G	missense
6p22.1	5	6	Depression (broad)	29662059	Howard DM	2.00E-11	28954293	rs3129791	HCG15	rs3129791-A	intron
6p22.1	5	6	Receptor for advanced glycosylation end products levels	33067605	Folkersen L	2.00E-11	29274136	rs3117427	OR14J1	rs3117427-T	intron
6p22.1	5	6	Chronic obstructive pulmonary disease x ever smoker interaction (2df)	33106845	Kim W	2.00E-11	29549653	rs1233393	GABBR1	rs1233393-G	intron
6p22.1	5	6	Squamous cell lung carcinoma	28604730	McKay JD	3.00E-11	29133642	rs141600123	OR2J3 - OR2J2	rs141600123-T	Non coding transcript exon
6p22.1	5	6	Chronic obstructive pulmonary disease	33909500	Moll M	3.00E-11	29342775	rs3749971	OR12D3, OR5V1	rs3749971-A	missense
6p22.1	5	6	Chronic obstructive pulmonary disease x current smoker interaction (2df)	33106845	Kim W	3.00E-11	29549653	rs1233393	GABBR1	rs1233393-G	intron
6p22.1	5	6	Depression (broad)	29662059	Howard DM	4.00E-11	28803291	rs9257248	NOP56P1 - LINC01623	rs9257248-T	intergenic
6p22.1	5	6	Squamous cell lung carcinoma	28604730	McKay JD	4.00E-11	28934352	rs148696809	LINC01556 - KRT18P1	rs148696809-C	intergenic
6p22.1	5	6	Squamous cell lung carcinoma	28604730	McKay JD	7.00E-11	28358009	rs13213986	ZSCAN12	rs13213986-A	3_prime_UTR
6p22.1	5	6	Pork consumption	32066663	Niarchou M	7.00E-11	29340743	rs9257800	OR5V1	rs9257800-?	intron
6p22.1	5	6	Oily fish consumption	32066663	Niarchou M	7.00E-11	29340743	rs9257800	OR5V1	rs9257800-?	intron
6p22.1	5	6	Help-seeking from a psychiatrist	32231276	Cai N	1.00E-10	29111775	rs3129120	OR2J3 - OR2J2	rs3129120-T	intergenic
6p22.1	5	6	Fish- and plant-related diet	32066663	Niarchou M	2.00E-10	28301099	rs7752448	ZSCAN31	rs7752448-?	intron

6p22.1	5	6	Squamous cell lung carcinoma	28604730	McKay JD	2.00E-10	28482917	rs114242735	GPX6, GPX6	rs114242735-C	intron
6p22.1	5	6	Cortical thickness (MOSTest)	32665545	van der Meer D	2.00E-10	29607101	rs3131856	SUMO2P1 - MOG	rs3131856-?	intergenic
6p22.1	5	6	Biological sex	33888908	Pirastu N	4.00E-10	29548089	rs926552	GABBR1	rs926552-G	intron
6p22.1	5	6	Lung cancer in ever smokers	28604730	McKay JD	7.00E-10	28304384	rs67340775	ZSCAN31	rs67340775-C	intron
6p22.1	5	6	Cutaneous lupus erythematosus	25827949	Kunz M	8.00E-10	29461730	rs1233491	RPS17P1 - LINC02829	rs1233491-C	intergenic
6p22.1	5	6	Breast cancer	29059683	Michailidou K	1.00E-09	28332141	rs13214023	ZKSCAN3	rs13214023-A	intron
6p22.1	5	6	Blood protein levels	28240269	Suhre K	1.00E-09	29044701	rs3131085	SAR1AP1	rs3131085-A	Non coding transcript exon
6p22.1	5	6	Estimated glomerular filtration rate	31451708	Hellwege JN	1.00E-09	29432173	rs2143668	OR2H1 - UBDP1	rs2143668-A	intergenic
6p22.1	5	6	Lung cancer in ever smokers	28604730	McKay JD	2.00E-09	28411941	rs67381177	ZSCAN23 - COX11P1	rs67381177-C	regulatory_region
6p22.1	5	6	Heel bone mineral density	30048462	Kim SK	2.00E-09	29223493	rs3130825	OR2G1P - OR2U1P	rs3130825-?	intron
6p22.1	5	6	Chronic obstructive pulmonary disease x current smoker interaction (main effect)	33106845	Kim W	2.00E-09	29549653	rs1233393	GABBR1	rs1233393-G	intron
6p22.1	5	6	Chronic obstructive pulmonary disease in non-current smokers	33106845	Kim W	2.00E-09	29549653	rs1233393	GABBR1	rs1233393-G	intron
6p22.1	5	6	Estimated glomerular filtration rate in diabetes	31451708	Hellwege JN	3.00E-09	28868494	rs3132374	HCG14 - TRIM27	rs3132374-C	intergenic
6p22.1	5	6	Intelligence	29942086	Savage JE	3.00E-09	29342775	rs3749971	OR12D3, OR5V1	rs3749971-A	missense
6p22.1	5	6	White matter microstructure (axial diffusivities)	31666681	Zhao B	3.00E-09	29458241	rs1233493	RPS17P1 - LINC02829	rs1233493-?	intergenic
6p22.1	5	6	Barrett's esophagus	22961001	Su Z	4.00E-09	29356331	rs9257809	OR5V1	rs9257809-A	intron
6p22.1	5	6	Asthma and major depressive disorder	31619474	Zhu Z	5.00E-09	28934352	rs148696809	LINC01556 - KRT18P1	rs148696809-?	intergenic
6p22.1	5	6	Breast cancer	29059683	Michailidou K	5.00E-09	29248149	rs3130834	OR2U2P - OR14J1	rs3130834-C	intron
6p22.1	5	6	Breast cancer	29059683	Michailidou K	6.00E-09	29075088	rs3129178	OR2J1 - OR2J3	rs3129178-A	intergenic
6p22.1	5	6	Barrett's esophagus	27527254	Gharahkhani P	6.00E-09	29356331	rs9257809	OR5V1	rs9257809-A	intron
6p22.1	5	6	Barrett's esophagus or Esophageal adenocarcinoma	27527254	Gharahkhani P	6.00E-09	29356331	rs9257809	OR5V1	rs9257809-A	intron
6p22.1	5	6	Lung cancer in ever smokers	28604730	McKay JD	7.00E-09	28619353	rs114385935	LINC00533 - RPSAP2	rs114385935-G	intron
6p22.1	5	6	Positive affect	30643256	Baselmans BML	8.00E-09	28314880	rs35016036	ZSCAN31	rs35016036-T	intron
6p22.1	5	6	Estimated glomerular filtration rate in non-diabetics	31451708	Hellwege JN	9.00E-09	29458241	rs1233493	RPS17P1 - LINC02829	rs1233493-A	intergenic
6p22.1	5	6	Smoking initiation (ever regular vs never regular)	30679032	Brazel DM	1.00E-08	28366151	rs2232423	ZSCAN12	rs2232423-G	missense
6p22.1	5	6	Smoking initiation (ever regular vs never regular)	30679032	Brazel DM	1.00E-08	28366151	rs2232423	ZSCAN12	rs2232423-G	missense
6p22.1	5	6	White matter microstructure (mean diffusivities)	31666681	Zhao B	1.00E-08	29458241	rs1233493	RPS17P1 - LINC02829	rs1233493-?	intergenic

6p22.1	5	6	Smoking initiation (ever regular vs never regular)	30679032	Brazel DM	1.00E-08	29548089	rs926552	GABBR1	rs926552-A	intron
6p22.1	5	6	Depression	29942085	Nagel M	2.00E-08	28678357	rs7775835	LINC00533 - RPSAP2	rs7775835-T	intergenic
6p22.1	5	6	Lung adenocarcinoma	19836008	Landi MT	2.00E-08	28776117	rs4324798	NOP56P1 - LINC01623	rs4324798-A	intergenic
6p22.1	5	6	Worry	29942085	Nagel M	2.00E-08	29260431	rs3117425	OR2U2P - OR14J1	rs3117425-T	Non coding transcript exon
6p22.1	5	6	Myopia (age of diagnosis)	29808027	Tedja MS	3.00E-08	29144532	rs145283874	OR2J2 - OR2J4P	rs145283874-T	intergenic
6p22.1	5	6	Breast cancer	29059683	Michailidou K	3.00E-08	29356331	rs9257809	OR5V1	rs9257809-G	intron
6p22.1	5	6	Waist-to-hip ratio adjusted for BMI	34021172	Christakoudi S	4.00E-08	28734676	rs1233604	RPSAP2 - NOP56P1	rs1233604-A	intergenic
6p22.1	5	6	Major depressive disorder (lifetime)	32231276	Cai N	4.00E-08	29548089	rs926552	GABBR1	rs926552-G	intron
Region 2p16.1											
2p16.1	6	2	Anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, or Tourette syndrome (pleiotropy)	31835028	Cross-Disorder Group of the Psychiatric Genomics Consortium	2.00E-14	57988194	rs11688767	ACTG1P22	rs11688767-?	intron
2p16.1	6	2	Sleep duration	30804565	Jansen PR	1.00E-13	57987593	rs11682175	ACTG1P22	rs11682175-C	intron
2p16.1	6	2	Autism spectrum disorder or schizophrenia	28540026	Anney RIL	1.00E-12	57987593	rs11682175	ACTG1P22	rs11682175-?	intron
2p16.1	6	2	Schizophrenia	26198764	Goes FS	1.00E-12	57987593	rs11682175	ACTG1P22	rs11682175-C	intron
2p16.1	6	2	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	31374203	Lam M	2.00E-12	57988194	rs11688767	ACTG1P22	rs11688767-?	intron
2p16.1	6	2	Schizophrenia	31268507	Periyasamy S	4.00E-12	57987593	rs11682175	ACTG1P22	rs11682175-?	intron
2p16.1	6	2	Schizophrenia (MTAG)	33479212	Yao X	4.00E-12	57987593	rs11682175	ACTG1P22	rs11682175-T	intron
2p16.1	6	2	Schizophrenia (MTAG)	32606422	Wu Y	5.00E-12	57987593	rs11682175	ACTG1P22	rs11682175-?	intron
2p16.1	6	2	Body fat distribution (arm fat ratio)	30664634	Rask-Andersen M	1.00E-11	57961602	rs13011472	NA - ACTG1P22	rs13011472-G	intergenic
2p16.1	6	2	Schizophrenia	25056061	Ripke S	1.00E-11	57987593	rs11682175	ACTG1P22	rs11682175-C	intron
2p16.1	6	2	Depressive symptoms (MTAG)	32606422	Wu Y	1.00E-11	57987593	rs11682175	ACTG1P22	rs11682175-?	intron
2p16.1	6	2	Bipolar disorder or major depressive disorder	31926635	Coleman JRI	2.00E-11	57987593	rs11682175	ACTG1P22	rs11682175-C	intron
2p16.1	6	2	Schizophrenia	30285260	Ikeda M	2.00E-11	57987593	rs11682175	ACTG1P22	rs11682175-C	intron
2p16.1	6	2	Bipolar disorder (MTAG)	32606422	Wu Y	1.00E-09	57987593	rs11682175	ACTG1P22	rs11682175-?	intron
2p16.1	6	2	Depression	29700475	Wray NR	5.00E-09	57987593	rs11682175	ACTG1P22	rs11682175-C	intron
2p16.1	6	2	Epilepsy	30531953	International League Against Epilepsy	8.00E-09	57950346	rs4671319	NA - ACTG1P22	rs4671319-G	intergenic

					Consortium on Complex Epilepsies						
2p16.1	6	2	Lung function (FVC)	30595370	Kichaev G	9.00E-09	57956088	rs13025126	NA - ACTG1P22	rs13025126-?	intergenic
2p16.1	6	2	Positive affect	30643256	Baselmans BML	1.00E-08	57942325	rs7606466	NA - ACTG1P22	rs7606466-G	intergenic
2p16.1	6	2	Major depressive disorder	29728651	Li X	1.00E-08	58041936	rs2717046	ACTG1P22 - VRK2	rs2717046-T	intergenic
2p16.1	6	2	Irritable mood	29500382	Nagel M	1.00E-08	57987593	rs11682175	ACTG1P22	rs11682175-T	intron
2p16.1	6	2	Schizophrenia (MTAG)	32606422	Wu Y	1.00E-08	57961602	rs13011472	NA - ACTG1P22	rs13011472-?	intergenic
2p16.1	6	2	Body shape index	34021172	Christakoudi S	2.00E-08	57988194	rs11688767	ACTG1P22	rs11688767-T	intron
2p16.1	6	2	Neuroticism	29942085	Nagel M	2.00E-08	57988194	rs11688767	ACTG1P22	rs11688767-A	intron
2p16.1	6	2	Neuropsychiatric disorders	33479212	Yao X	2.00E-08	57987593	rs11682175	ACTG1P22	rs11682175-T	intron
2p16.1	6	2	Life satisfaction	30643256	Baselmans BML	3.00E-08	57942325	rs7606466	NA - ACTG1P22	rs7606466-G	intergenic
2p16.1	6	2	Waist circumference adjusted for body mass index	34021172	Christakoudi S	3.00E-08	57942325	rs7606466	NA - ACTG1P22	rs7606466-C	intergenic
2p16.1	6	2	Depression	29942085	Nagel M	3.00E-08	57988194	rs11688767	ACTG1P22	rs11688767-A	intron
Region 5q34											
5q34	7	5	Depression	30718901	Howard DM	2.00E-30	164523472	rs11135349		rs11135349-C	intron
5q34	7	5	Broad depression or major depressive disorder (self-reported)	30626913	Amare AT	8.00E-16	164475774	rs7714851		rs7714851-T	intron
5q34	7	5	Well-being spectrum (multivariate analysis)	30643256	Baselmans BML	1.00E-14	164484948	rs4543289		rs4543289-G	Non coding transcript exon
5q34	7	5	Neuroticism	30643256	Baselmans BML	1.00E-13	164522204	rs6873557		rs6873557-G	intron
5q34	7	5	Life satisfaction	30643256	Baselmans BML	2.00E-11	164477151	rs6896348		rs6896348-G	intron
5q34	7	5	Bipolar disorder or major depressive disorder	31926635	Coleman JRI	3.00E-11	164523472	rs11135349		rs11135349-C	intron
5q34	7	5	Positive affect	30643256	Baselmans BML	6.00E-11	164481879	rs4077277		rs4077277-T	intron
5q34	7	5	Depressive symptoms (MTAG)	29292387	Turley P	6.00E-11	164488849	rs10434704		rs10434704-C	intron
5q34	7	5	Major depressive disorder (MTAG)	33479212	Yao X	2.00E-10	164523472	rs11135349		rs11135349-A	intron
5q34	7	5	Depression	29700475	Wray NR	1.00E-09	164523472	rs11135349		rs11135349-C	intron
5q34	7	5	Depressive symptoms	29292387	Turley P	1.00E-09	164530809	rs10045971		rs10045971-C	intron
5q34	7	5	Body mass index	30239722	Pulit SL	2.00E-09	164557954	rs2861089	NA - LINC01938	rs2861089-A	intron
5q34	7	5	Lifetime anxiety disorder	31748690	Purves KL	3.00E-09	164593555	rs2861139	NA - LINC01938	rs2861139-C	intron
5q34	7	5	Neuroticism	29942085	Nagel M	8.00E-09	164487132	rs7721129		rs7721129-A	intron
5q34	7	5	Major depressive disorder	27479909	Hyde CL	1.00E-08	164484948	rs4543289		rs4543289-T	Non coding transcript exon
5q34	7	5	Neuroticism (MTAG)	29292387	Turley P	1.00E-08	164488849	rs10434704		rs10434704-C	intron
5q34	7	5	Lamb consumption	32066663	Niarchou M	2.00E-08	164548126	rs766120	NA - LINC01938	rs766120-?	intron
5q34	7	5	Depressive symptoms	32606422	Wu Y	4.00E-08	164484948	rs4543289		rs4543289-?	Non coding transcript exon
5q34	7	5	Depression	29942085	Nagel M	4.00E-08	164484948	rs4543289		rs4543289-T	Non coding transcript exon

Supplementary Table 6: Gene-level associations with fear-based disorders, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.

Reported trait	Study accession	Variant and risk allele	P-value	Location	Mapped gene
ERI3					
Educational attainment	GCST90105038	rs1738050-C	1.00E-24	737361:23:00	ERI3
Drinks per week	GCST90243989	rs11808967-G	3.00E-13	738377:17:00	ERI3
Educational attainment (years of education)	GCST006442	rs1778904-C	2.00E-12	737329:36:00	ERI3
Educational attainment (MTAG)	GCST006571	rs1738050-C	4.00E-12	737361:23:00	ERI3
Educational attainment (years of education)	GCST006442	rs1738050-C	1.00E-10	737361:23:00	ERI3
YWHAG					
Heat shock protein beta-1 levels	GCST90161270	rs36119738-TG	3.00E-23	1271811:18:00	HSPB1, YWHAG
Blood protein levels	GCST009696	rs184260005-?	1.00E-13	1272342:45:00	YWHAG
Blood protein levels	GCST009696	rs61536380-?	2.00E-11	1272449:02:00	YWHAG
Schizophrenia	GCST007205	rs6946310-T	3.00E-07	1271786:42:00	HSPB1, YWHAG
Schizophrenia	GCST007201	rs2961033-C	5.00E-07	1272257:26:00	YWHAG
3-hydroxypropylmercapturic acid levels in smokers	GCST002956	rs2908197-?	2.00E-06	1271802:32:00	HSPB1, YWHAG
Theophylline levels in elite athletes	GCST90134212	rs2961047-?	3.00E-06	1271769:19:00	YWHAG, HSPB1
Multiple sclerosis	GCST001096	rs17149161-A	6.00E-06	1272488:52:00	YWHAG
Multiple sclerosis	GCST001096	rs7789940-G	6.00E-06	1272038:53:00	HSPB1, YWHAG
Multiple sclerosis	GCST001096	rs7779014-T	8.00E-06	1272444:49:00	YWHAG
Multiple sclerosis	GCST001096	rs758944-T	8.00E-06	1272073:20:00	HSPB1, YWHAG
SNX29					
Educational attainment	GCST90105038	rs9922788-A	2.00E-22	202517:15:00	SNX29
Educational attainment	GCST90105038	rs12448347-A	1.00E-18	206950:14:00	SNX29
Insomnia	GCST90131901	rs830734-A	7.00E-17	204031:50:00	SNX29
Insomnia	GCST90131901	rs111505982-A	7.00E-17	204307:42:00	SNX29
Educational attainment (years of education)	GCST006442	rs34098770-A	3.00E-16	202415:15:00	SNX29
Educational attainment	GCST90105038	rs34321532-C	4.00E-16	200422:43:00	SNX29
Educational attainment (MTAG)	GCST006571	rs34565142-T	8.00E-16	202307:30:00	SNX29
Educational attainment (years of education)	GCST006442	rs350281-T	1.00E-14	202314:02:00	SNX29
Insomnia	GCST90131901	rs9927276-C	2.00E-14	204370:20:00	SNX29
Smoking initiation	GCST90243985	rs11075066-G	3.00E-14	206688:44:00	SNX29
Total PHF-tau (SNP x SNP interaction)	GCST010340	rs1560104-? x rs2732932-?	3.00E-13	16:12614351 2: 18034351	SNX29 - CPPED1 x KCNS3
Insomnia	GCST90131901	rs60866545-A	3.00E-13	203295:48:00	SNX29
Educational attainment	GCST90105038	rs3851013-A	3.00E-13	202086:26:00	SNX29
Insomnia	GCST90131901	rs830728-A	1.00E-12	203642:37:00	SNX29

Insomnia	GCST90131901	rs2911494-C	2.00E-12	205175:17:00	SNX29
Insomnia	GCST90131902	rs830734-A	3.00E-12	204031:50:00	SNX29
Cigarettes smoked per day	GCST90243976	rs11648938-C	4.00E-12	200020:52:00	SNX29
Insomnia	GCST007988	rs830716-C	9.00E-12	203843:32:00	SNX29
Educational attainment (MTAG)	GCST006571	rs12926226-C	2.00E-11	206934:42:00	SNX29
Educational attainment (years of education)	GCST006442	rs1035578-A	3.00E-11	207307:48:00	SNX29
Total PHF-tau (SNP x SNP interaction)	GCST010340	rs7319335-? x rs7206244-?	5.00E-11	207872:15:00	No mapped genes x SNX29
Insomnia	GCST90131901	rs1701111-A	5.00E-11	203431:47:00	SNX29
Breast milk fatty acid composition (maternal genotype effect)	GCST005586	rs7198595-A	5.00E-11	207866:50:00	SNX29
Intelligence	GCST006250	rs2457192-A	3.00E-10	201742:24:00	SNX29
Insomnia	GCST90131901	rs6498291-A	3.00E-10	206771:14:00	SNX29
Insomnia	GCST90131903	rs11639640-A	3.00E-10	204394:03:00	SNX29
Cognitive performance (MTAG)	GCST006570	rs2457192-A	3.00E-10	201742:24:00	SNX29
Monocyte percentage of white cells	GCST90002394	rs35986807-C	5.00E-10	199889:27:00	SNX29
Insomnia	GCST90131902	rs9936260-A	5.00E-10	204305:57:00	SNX29
Insomnia	GCST90131903	rs1701111-A	6.00E-10	203431:47:00	SNX29
Educational attainment	GCST90100570	rs350201-?	6.00E-10	202095:25:00	SNX29
Educational attainment	GCST90105038	rs7195739-A	7.00E-10	209264:58:00	SNX29
Heel bone mineral density	GCST006979	rs1703492-A	9.00E-10	200734:17:00	SNX29
Insomnia	GCST90131901	rs830719-C	1.00E-09	203813:48:00	SNX29
Insomnia	GCST90131903	rs830733-T	1.00E-09	204167:55:00	SNX29
Attention deficit hyperactivity disorder or autism spectrum disorder or intelligence (pleiotropy)	GCST90134330	rs2457192-?	1.00E-09	201742:24:00	SNX29
Total PHF-tau (SNP x SNP interaction)	GCST010340	rs992131-? x rs1560104-?	3.00E-09	210255:11:00	No mapped genes x SNX29 - CPPED1
Neurofibrillary tangles (SNP x SNP interaction)	GCST010343	rs2485954-? x rs11648938-?	3.00E-09	200020:52:00	No mapped genes x SNX29
Educational attainment (years of education)	GCST007037	rs12599793-?	3.00E-09	202053:50:00	SNX29
Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	GCST008595	rs1035578-?	3.00E-09	207307:48:00	SNX29
3-hydroxypropylmercapturic acid levels in smokers	GCST002956	rs116383923-?	3.00E-09	206371:45:00	SNX29
Insomnia	GCST007988	rs9936260-A	4.00E-09	204305:57:00	SNX29
Insomnia	GCST90131901	rs6498269-A	4.00E-09	202767:18:00	SNX29
Cigarettes smoked per day	GCST90243987	rs11648938-C	4.00E-09	200020:52:00	SNX29
Carotid Intima-media thickness (mean of the maximum cIMT)	GCST90092504	rs190770959-C	6.00E-09	201094:37:00	SNX29
Insomnia	GCST90131902	rs2904423-A	7.00E-09	204318:21:00	SNX29
Free thyroxine concentration	GCST006896	rs8063103-C	8.00E-09	210174:58:00	SNX29, CPPED1
Phospholipid levels in very large HDL	GCST90093012	rs28731373-T	1.00E-08	210282:26:00	CPPED1, SNX29
Insomnia	GCST90131901	rs4780414-A	1.00E-08	202606:15:00	SNX29

Highest math class taken (MTAG)	GCST006568	rs2457192-A	1.00E-08	201742:24:00	SNX29
General cognitive ability	GCST006269	rs2457192-A	1.00E-08	201742:24:00	SNX29
Educational attainment (years of education)	GCST006442	rs387027-A	1.00E-08	199874:06:00	SNX29
Total lipid levels in very large HDL	GCST90093010	rs28731373-T	2.00E-08	210282:26:00	CPPED1, SNX29
Protein quantitative trait loci (liver)	GCST011427	rs58049731-T	2.00E-08	204419:30:00	SNX29
Intelligence (MTAG)	GCST005316	rs350251-A	2.00E-08	201827:55:00	SNX29
Intelligence	GCST006250	rs168407-A	2.00E-08	202045:40:00	SNX29
Height	GCST007841	rs11075041-?	2.00E-08	200164:50:00	SNX29
Free cholesterol levels in large HDL	GCST90092848	rs28731373-T	2.00E-08	210282:26:00	CPPED1, SNX29
Cholesteryl ester levels in very large HDL	GCST90093006	rs28731373-T	2.00E-08	210282:26:00	CPPED1, SNX29
Age at first sexual intercourse	GCST90000047	rs2870488-T	2.00E-08	204292:17:00	SNX29
Adolescent idiopathic scoliosis	GCST006287	rs3902080-?	2.00E-08	208364:38:00	SNX29
Type 2 diabetes	GCST90018926	rs143708136-T	3.00E-08	201537:12:00	SNX29
Smoking initiation (ever regular vs never regular) (MTAG)	GCST007468	rs67119854-G	3.00E-08	206502:29:00	SNX29
LDL cholesterol levels x short total sleep time interaction (2df test)	GCST009365	rs12598569-?	3.00E-08	202156:40:00	SNX29
Insomnia	GCST90131903	rs60866545-A	3.00E-08	203295:48:00	SNX29
Insomnia	GCST90131903	rs2911494-C	3.00E-08	205175:17:00	SNX29
Cholesterol levels in very large HDL	GCST90093004	rs28731373-T	3.00E-08	210282:26:00	CPPED1, SNX29
Monohexosylceramide (d18:1/22:0) levels	GCST90024127	rs140132652-T	4.00E-08	209876:58:00	SNX29, CPPED1
Intelligence	GCST006250	rs72784651-T	4.00E-08	201458:47:00	SNX29

Supplementary Table 7: Gene-level associations with GAD, and their association with previously published GWAS of complex traits, reported in the GWAS Catalog.

Reported trait	Study accession	Variant and risk allele	P-value	Location	Mapped gene
TMEM106B					
Serum levels of protein	GCST90090262	rs1548884-C	1.00E-42	204009:15:00	TMEM106B
Blood protein levels	GCST006585	rs10950398-A	6.00E-27	203761:05:00	TMEM106B
Educational attainment	GCST90105038	rs7810903-A	8.00E-26	197748:52:00	THSD7A, TMEM106B
Depression	GCST007342	rs2043539-A	8.00E-24	203577:54:00	TMEM106B
Differential aging in older adults (frontal cortex)	GCST008795	rs1990622-A	3.00E-23	204076:21:00	VWDE, TMEM106B
Depression	GCST012355	rs10950393-T	3.00E-22	203739:00:00	TMEM106B
Differential aging (frontal cortex)	GCST008800	rs1990622-A	2.00E-19	204076:21:00	VWDE, TMEM106B
Depression	GCST90239706	rs2043539-?	3.00E-16	203577:54:00	TMEM106B
Frontotemporal dementia	GCST006154	rs7791726-G	4.00E-16	204068:43:00	TMEM106B, VWDE
Major depressive disorder	GCST90020227	rs1054169-?	2.00E-15	203904:50:00	TMEM106B
High density lipoprotein cholesterol levels	GCST90239649	rs4721060-?	3.00E-15	203805:46:00	TMEM106B
Depressive symptoms	GCST007340	rs6460895-G	2.00E-14	203555:34:00	TMEM106B
Frontotemporal dementia with GRN mutation	GCST006148	rs7791726-G	2.00E-14	204068:43:00	TMEM106B, VWDE

Educational attainment	GCST90105038	rs73676288-A	3.00E-14	198677:01:00	TMEM106B, THSD7A
Liver enzyme levels (alkaline phosphatase)	GCST90013406	rs6460898-A	6.00E-14	203561:27:00	TMEM106B
Type 2 diabetes	GCST010557	rs13237518-C	2.00E-13	203839:47:00	TMEM106B
Help-seeking from a GP	GCST010011	rs3807866-A	5.00E-13	203519:32:00	TMEM106B
Well-being spectrum (multivariate analysis)	GCST007341	rs7805419-C	9.00E-13	204054:05:00	TMEM106B
Insomnia	GCST90131901	rs4721061-C	2.00E-12	203805:53:00	TMEM106B
Height	GCST007841	rs11509137-?	3.00E-12	203743:13:00	TMEM106B
Depression (broad)	GCST005902	rs3807865-A	7.00E-12	203519:56:00	THSD7A, TMEM106B
Major depressive disorder (MTAG)	GCST012090	rs3807865-A	9.00E-12	203519:56:00	THSD7A, TMEM106B
Depressed affect	GCST006475	rs10950393-C	2.00E-11	203739:00:00	TMEM106B
High density lipoprotein cholesterol levels	GCST90239652	rs1990622-G	2.00E-11	204076:21:00	VWDE, TMEM106B
Smoking initiation	GCST90243985	rs4719302-A	3.00E-11	203440:14:00	THSD7A, TMEM106B
Experiencing mood swings	GCST006944	rs11509880-A	4.00E-11	203711:45:00	TMEM106B
Triglyceride levels	GCST90239661	rs34046032-?	4.00E-11	203814:37:00	TMEM106B
Depression	GCST006477	rs3823612-C	5.00E-11	203659:09:00	TMEM106B
Cortical thickness	GCST90091061	rs5011439-C	5.00E-11	203826:45:00	TMEM106B
Alzheimer's disease	GCST90027158	rs13237518-A	5.00E-11	203839:47:00	TMEM106B
Depressive symptoms (MTAG)	GCST010642	rs3823612-?	5.00E-11	203659:09:00	TMEM106B
Depressive symptoms	GCST010647	rs3823612-?	5.00E-11	203659:09:00	TMEM106B
Neuroticism	GCST006476	rs11509880-A	6.00E-11	203711:45:00	TMEM106B
Neurofibrillary tangles (SNP x SNP interaction)	GCST010343	rs17127497-? x rs17665337-?	6.00E-11	203341:27:00	No mapped genes x N/A - TMEM106B
Bipolar disorder or major depressive disorder	GCST010416	rs4721057-G	7.00E-11	203800:15:00	TMEM106B
Neuroticism	GCST006940	rs11509880-A	9.00E-11	203711:45:00	TMEM106B
Aspartate aminotransferase levels	GCST90018944	rs3807865-A	9.00E-11	203519:56:00	THSD7A, TMEM106B
HDL cholesterol levels	GCST010242	rs10950390-C	1.00E-10	203091:42:00	TMEM106B
Neuroticism	GCST007339	rs62435692-T	1.00E-10	203392:13:00	TMEM106B, THSD7A
Generalized anxiety disorder (phenotype risk score)	GCST90103930	rs12699336-A	1.00E-10	204067:24:00	TMEM106B, VWDE
Neuroticism	GCST010017	rs11509880-A	2.00E-10	203711:45:00	TMEM106B
Type 2 diabetes	GCST010555	rs13237518-C	8.00E-10	203839:47:00	TMEM106B
Feeling tense	GCST006952	rs11509880-A	1.00E-09	203711:45:00	TMEM106B
F-dessert liking (derived food-liking factor)	GCST90094754	rs2043539-A	1.00E-09	203577:54:00	TMEM106B
Vertex-wise cortical thickness	GCST90095131	rs5011439-C	1.00E-09	203826:45:00	TMEM106B
Depression	GCST90058037	rs76854159-A	2.00E-09	203764:13:00	TMEM106B
Major depressive disorder or stroke (pleiotropy)	GCST90095484	rs4721058-?	2.00E-09	203800:50:00	TMEM106B
Depressed affect	GCST006475	rs57506017-?	2.00E-09	203433:39:00	THSD7A, TMEM106B
Post-traumatic stress disorder (phenotype risk score)	GCST90103931	rs4721064-A	2.00E-09	204086:12:00	VWDE, TMEM106B
Aspartate aminotransferase levels	GCST90019497	rs145666097-T	2.00E-09	203244:03:00	TMEM106B
HDL cholesterol	GCST006611	rs3173615-C	2.00E-09	203836:51:00	TMEM106B
Late-onset Alzheimer's disease	GCST90044699	rs5011436-?	3.00E-09	203825:52:00	TMEM106B

HDL cholesterol	GCST90018956	rs397715721-GT	3.00E-09	203548:34:00	TMEM106B
Serum alkaline phosphatase levels	GCST90011900	rs13229988-G	4.00E-09	203810:31:00	TMEM106B
Serum alkaline phosphatase levels	GCST90018942	rs1990620-G	4.00E-09	204080:02:00	TMEM106B, VVDE
Insomnia	GCST90131902	rs6460902-A	5.00E-09	203605:05:00	TMEM106B
Neuroticism	GCST005232	rs57506017-A	6.00E-09	203433:39:00	THSD7A, TMEM106B
Triglycerides	GCST006613	rs3173615-C	7.00E-09	203836:51:00	TMEM106B
General factor of neuroticism	GCST007709	rs7805419-T	1.00E-08	204054:05:00	TMEM106B
Crisps liking	GCST90094747	rs10230959-G	1.00E-08	200664:32:00	THSD7A, TMEM106B
Insomnia	GCST90131901	rs62435692-T	1.00E-08	203392:13:00	TMEM106B, THSD7A
Apolipoprotein A1 levels	GCST010241	rs10950390-C	2.00E-08	203091:42:00	TMEM106B
F-highly palatable foods liking (derived food-liking factor)	GCST90094785	rs2043539-A	2.00E-08	203577:54:00	TMEM106B
F-sauces liking (derived food-liking factor)	GCST90094831	rs2043539-A	2.00E-08	203577:54:00	TMEM106B
Major depressive disorder (probable)	GCST005904	rs5011432-C	2.00E-08	203824:22:00	TMEM106B
Coronary artery disease	GCST005195	rs11509880-A	3.00E-08	203711:45:00	TMEM106B
Coronary artery disease	GCST010866	rs11509880-?	3.00E-08	203711:45:00	TMEM106B
Johnsonella ignava abundance in stool	GCST90032438	rs2356054-T	3.00E-08	201953:33:00	THSD7A, TMEM106B
Coronary artery disease	GCST005196	rs11509880-A	3.00E-08	203711:45:00	TMEM106B
Cigarettes smoked per day in Schizophrenia	GCST012180	rs148253479-?	3.00E-08	203149:56:00	TMEM106B
Depression	GCST005839	rs10950398-A	3.00E-08	203761:05:00	TMEM106B
Coronary artery disease	GCST90132314	rs11509880-?	4.00E-08	203711:45:00	TMEM106B
Triglyceride levels	GCST90239664	rs5011439-C	4.00E-08	203826:45:00	TMEM106B
TRIM31					
High light scatter reticulocyte count	GCST90002385	rs2523990-G	5.00E-71	501830:12:00	TRIM31-AS1, TRIM31
Immature fraction of reticulocytes	GCST90002387	rs115333019-G	8.00E-43	502235:48:00	TRIM40, TRIM31-AS1
Lymphocyte count	GCST90085815	rs3132681-C	1.00E-38	501543:44:00	TRIM31, RNF39
Eosinophil percentage of white cells	GCST004600	rs1111180-C	4.00E-37	501709:25:00	TRIM31, RNF39
Eosinophil percentage of granulocytes	GCST004617	rs1111180-C	2.00E-35	501709:25:00	TRIM31, RNF39
High light scatter reticulocyte percentage of red cells	GCST90002386	rs2523990-G	2.00E-32	501830:12:00	TRIM31-AS1, TRIM31
Neutrophil percentage of granulocytes	GCST004623	rs3132682-C	2.00E-30	501282:51:00	TRIM31, RNF39
High light scatter reticulocyte percentage of red cells	GCST90002386	rs11961190-T	3.00E-29	501811:40:00	TRIM31, TRIM31-AS1
Total cholesterol levels	GCST90239673	rs2844795-?	8.00E-29	501773:50:00	TRIM31, TRIM31-AS1
High light scatter reticulocyte count	GCST90002385	rs115254023-T	8.00E-29	502003:08:00	TRIM31-AS1, TRIM40
Apolipoprotein A1 levels	GCST010241	rs2523978-C	7.00E-28	501938:11:00	TRIM31-AS1, TRIM40
High light scatter reticulocyte count	GCST90002385	rs11961190-T	4.00E-27	501811:40:00	TRIM31, TRIM31-AS1
Lymphocyte count	GCST90085815	rs2023472-A	6.00E-25	501807:27:00	TRIM31, TRIM31-AS1
HDL cholesterol levels	GCST010242	rs2523978-C	4.00E-24	501938:11:00	TRIM31-AS1, TRIM40
Total cholesterol levels	GCST90239676	rs2844795-T	7.00E-24	501773:50:00	TRIM31, TRIM31-AS1
Reticulocyte count	GCST90002405	rs11961190-T	1.00E-23	501811:40:00	TRIM31, TRIM31-AS1
Ebbinghaus illusion (overestimation)	GCST011568	rs6915823-?	1.00E-23	501766:53:00	TRIM31-AS1, TRIM31
Reticulocyte fraction of red cells	GCST90002406	rs11961190-T	2.00E-22	501811:40:00	TRIM31, TRIM31-AS1

Shingles	GCST005001	rs9261431-?	2.00E-19	501983:25:00	TRIM40, TRIM31-AS1
Triglyceride levels	GCST90239661	rs2517601-?	7.00E-19	501777:35:00	TRIM31, TRIM31-AS1
Low density lipoprotein cholesterol levels	GCST010204	rs9378220-?	4.00E-18	501828:38:00	TRIM31-AS1, TRIM31
Waist-to-hip ratio adjusted for BMI	GCST012230	rs1264712-A	9.00E-17	501425:25:00	TRIM31, RNF39
Mumps	GCST005003	rs111537560-?	5.00E-16	501756:31:00	TRIM31
Non-HDL cholesterol levels	GCST90239670	rs9378220-A	2.00E-15	501828:38:00	TRIM31-AS1, TRIM31
Beta-2 microglobulin plasma levels	GCST001863	rs2023472-A	3.00E-15	501807:27:00	TRIM31, TRIM31-AS1
Linoleic acid levels	GCST90092880	rs2844795-C	4.00E-14	501773:50:00	TRIM31, TRIM31-AS1
Autism spectrum disorder or schizophrenia	GCST004521	rs116676919-?	4.00E-14	501884:39:00	TRIM31, TRIM31-AS1
MHC class I polypeptide-related sequence A levels	GCST90101316	rs2523987-C	2.00E-13	501876:16:00	TRIM31, TRIM31-AS1
Depression (broad)	GCST005902	rs2517601-C	5.00E-13	501777:35:00	TRIM31, TRIM31-AS1
Omega-6 fatty acid levels	GCST90092933	rs2844795-C	5.00E-13	501773:50:00	TRIM31, TRIM31-AS1
Mouth ulcers	GCST007839	rs2517664-T	9.00E-13	501764:02:00	TRIM31, TRIM31-AS1
Hepatocellular carcinoma in hepatitis B infection	GCST005746	rs1110446-?	9.00E-13	501725:20:00	TRIM31
Total lipid levels in medium HDL	GCST90092898	rs1264705-G	2.00E-12	501604:17:00	TRIM31, RNF39
Immature fraction of reticulocytes	GCST004628	rs17187854-T	3.00E-12	501635:26:00	RNF39, TRIM31
Insomnia	GCST90131901	rs2844795-T	4.00E-12	501773:50:00	TRIM31, TRIM31-AS1
Estimated glomerular filtration rate in diabetes	GCST008746	rs2517601-C	4.00E-12	501777:35:00	TRIM31, TRIM31-AS1
Insomnia	GCST90131903	rs2844795-T	3.00E-11	501773:50:00	TRIM31, TRIM31-AS1
Human milk oligosaccharide concentration (lacto-N-fucopentaose III) in non-secretors	GCST90027228	rs13214608-?	3.00E-11	501686:36:00	RNF39, TRIM31
Chronic obstructive pulmonary disease	GCST011766	rs2523989-T	1.00E-10	501847:38:00	TRIM31, TRIM31-AS1
Psoriatic arthritis	GCST007043	rs115618145-?	1.00E-10	501714:18:00	RNF39, TRIM31
General cognitive ability	GCST006269	rs1362104-A	2.00E-10	502237:19:00	TRIM31-AS1, TRIM40
Hepatocellular carcinoma in hepatitis B infection	GCST005746	rs1110446-?	3.00E-10	501725:20:00	TRIM31
A body shape index	GCST012231	rs1264712-A	4.00E-10	501425:25:00	TRIM31, RNF39
General cognitive ability	GCST006269	rs2844790-C	4.00E-10	502111:36:00	TRIM31-AS1, TRIM40
Insomnia	GCST90131903	rs3132680-A	5.00E-10	501762:58:00	TRIM31, TRIM31-AS1
Insomnia	GCST90131903	rs9378220-A	6.00E-10	501828:38:00	TRIM31-AS1, TRIM31
Takayasu arteritis	GCST90137503	rs1116221-?	1.00E-09	501731:53:00	TRIM31
F-sweet food liking (derived food-liking factor)	GCST90094854	rs2517664-T	1.00E-09	501764:02:00	TRIM31, TRIM31-AS1
Total fatty acid levels	GCST90092987	rs2844795-C	2.00E-09	501773:50:00	TRIM31, TRIM31-AS1
Total lipid levels in lipoprotein particles	GCST90092989	rs2844795-C	2.00E-09	501773:50:00	TRIM31, TRIM31-AS1
Mean corpuscular hemoglobin	GCST004630	rs78026509-G	2.00E-09	501781:33:00	TRIM31-AS1, TRIM31
Total phospholipid levels in lipoprotein particles	GCST90092991	rs2844795-C	2.00E-09	501773:50:00	TRIM31, TRIM31-AS1
Itch intensity from mosquito bite	GCST004861	rs143211693-?	4.00E-09	501659:37:00	RNF39, TRIM31
Insomnia	GCST90131901	rs6457150-C	5.00E-09	501686:33:00	RNF39, TRIM31
Neuritic plaques (SNP x SNP interaction)	GCST010342	rs11746354-? x rs2517592-?	5.00E-09	502095:20:00	No mapped genes x TRIM31-AS1 - TRIM40
Total cholines levels	GCST90092812	rs2844795-C	6.00E-09	501773:50:00	TRIM31, TRIM31-AS1

Protein quantitative trait loci (liver)	GCST011427	rs1419674-T	7.00E-09	502091:58:00	TRIM40, TRIM31-AS1
Insomnia	GCST90131901	rs9378220-A	8.00E-09	501828:38:00	TRIM31-AS1, TRIM31
General cognitive ability	GCST006269	rs2517664-T	1.00E-08	501764:02:00	TRIM31, TRIM31-AS1
Takayasu arteritis	GCST011936	rs1264697-?	1.00E-08	501669:11:00	TRIM31, RNF39
Free cholesterol to total lipids ratio in small HDL	GCST90092949	rs142292417-T	1.00E-08	501473:16:00	RNF39, TRIM31
Tuberculosis	GCST005006	rs2844790-?	1.00E-08	502111:36:00	TRIM31-AS1, TRIM40
Neurociticism	GCST006940	rs9261418-A	2.00E-08	501820:43:00	TRIM31, TRIM31-AS1
Prostate cancer	GCST006085	rs7767188-A	2.00E-08	501772:39:00	TRIM31, TRIM31-AS1
General cognitive ability	GCST006269	rs3132681-A	2.00E-08	501543:44:00	TRIM31, RNF39
Prostate cancer	GCST002606	rs115457135-A	2.00E-08	501772:39:00	TRIM31-AS1, TRIM31
Type 1 diabetes and autoimmune thyroid diseases	GCST002876	rs2523989-A	2.00E-08	501847:38:00	TRIM31, TRIM31-AS1
Haemorrhoidal disease	GCST90014033	rs1156533-?	2.00E-08	501628:52:00	RNF39, TRIM31
Cholesteryl esters to total lipids ratio in large LDL	GCST90092859	rs2844795-C	3.00E-08	501773:50:00	TRIM31, TRIM31-AS1
Ratio of linoleic acid to total fatty acids	GCST90092881	rs1245372-A	3.00E-08	501309:41:00	RNF39, TRIM31
SORCS3					
Educational attainment	GCST90105038	rs2491365-T	2.00E-41	1746296:58:00	SORCS3
Educational attainment	GCST90105038	rs61867294-A	2.00E-36	1746834:09:00	SORCS3
Smoking initiation	GCST90243968	rs9787523-C	8.00E-27	1745021:42:00	SORCS3
Depression	GCST007342	rs1021363-A	4.00E-23	1747528:01:00	SORCS3
Educational attainment (MTAG)	GCST006571	rs902305-A	2.00E-22	1746217:46:00	SORCS3
Systolic blood pressure	GCST012376	rs181200083-C	1.00E-21	1746030:17:00	SORCS3
Age at first sexual intercourse	GCST90000047	rs3896224-A	3.00E-21	1745144:55:00	SORCS3
Externalizing behaviour (multivariate analysis)	GCST90061435	rs11596214-G	6.00E-21	1744911:14:00	SORCS3
Highest math class taken (MTAG)	GCST006568	rs10786832-T	6.00E-20	1747654:50:00	SORCS3
Major depressive disorder	GCST90020227	rs1021363-?	1.00E-19	1747528:01:00	SORCS3
Educational attainment	GCST90105038	rs11192147-A	1.00E-19	1744231:28:00	SORCS3
Depression	GCST012355	rs11596214-G	2.00E-18	1744911:14:00	SORCS3
Depression	GCST012355	rs7906899-A	2.00E-18	1749572:27:00	SORCS3
Educational attainment	GCST90105038	rs10786823-C	2.00E-18	1747312:41:00	SORCS3
Cardiovascular disease	GCST007072	rs145158522-?	4.00E-18	1747845:01:00	SORCS3
Educational attainment (years of education)	GCST006442	rs790647-A	4.00E-18	1750288:46:00	SORCS3
Diastolic blood pressure	GCST006627	rs191784289-C	5.00E-18	1752263:04:00	SORCS3
Highest math class taken (MTAG)	GCST006568	rs17118088-A	5.00E-18	1750093:10:00	SORCS3
Systolic blood pressure	GCST012376	rs535313355-C	1.00E-17	1743999:42:00	SORCS3
Depression	GCST90239706	rs1021363-?	2.00E-17	1747528:01:00	SORCS3
Diastolic blood pressure	GCST012378	rs181200083-C	6.00E-17	1746030:17:00	SORCS3
Major depressive disorder	GCST90020227	rs17186548-?	2.00E-16	1746529:46:00	SORCS3
Self-reported math ability (MTAG)	GCST006569	rs2864034-A	2.00E-16	1748636:23:00	SORCS3
Educational attainment	GCST90105038	rs4400723-T	2.00E-16	1755812:25:00	SORCS3, RNU6-463P
Medication use (agents acting on the renin-angiotensin system)	GCST90018988	rs74157123-T	4.00E-16	1741791:26:00	SORCS3, LINC02620

Medication use (agents acting on the renin-angiotensin system)	GCST007930	rs74157123-T	4.00E-16	1741791:26:00	SORCS3, LINC02620
Educational attainment	GCST90105038	rs72819955-T	5.00E-16	1754870:26:00	SORCS3, RNU6-463P
Insomnia	GCST90131901	rs9787523-T	7.00E-16	1745021:42:00	SORCS3
Well-being spectrum (multivariate analysis)	GCST007341	rs11599236-C	1.00E-15	1744925:14:00	SORCS3
Educational attainment	GCST90105038	rs72819942-T	2.00E-15	1754592:45:00	RNU6-463P, SORCS3
Anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, or Tourette syndrome (pleiotropy)	GCST009600	rs61867293-?	8.00E-15	1746746:06:00	SORCS3
Systolic blood pressure	GCST012376	rs540369678-T	2.00E-14	1750769:53:00	SORCS3
Educational attainment	GCST90105038	rs1490176-A	2.00E-14	1746684:27:00	SORCS3
Major depressive disorder (MTAG)	GCST012090	rs1021363-A	2.00E-14	1747528:01:00	SORCS3
Medication use (calcium channel blockers)	GCST90018987	rs191572726-C	3.00E-14	1753731:10:00	SORCS3
Medication use (calcium channel blockers)	GCST007929	rs191572726-C	3.00E-14	1753731:10:00	SORCS3
Diastolic blood pressure	GCST012378	rs535313355-C	4.00E-14	1743999:42:00	SORCS3
Depressive symptoms (MTAG)	GCST005323	rs2496022-C	4.00E-14	1747888:40:00	SORCS3
Educational attainment (MTAG)	GCST006571	rs12252384-T	5.00E-14	1750152:03:00	SORCS3
Cognitive performance (MTAG)	GCST006570	rs3896224-A	5.00E-14	1745144:55:00	SORCS3
Smoking initiation (ever regular vs never regular) (MTAG)	GCST007468	rs9787523-C	6.00E-14	1745021:42:00	SORCS3
Insomnia	GCST90131901	rs67700133-A	1.00E-13	1745024:11:00	SORCS3
Depressive symptoms	GCST007340	rs1021363-G	1.00E-13	1747528:01:00	SORCS3
Educational attainment (years of education)	GCST006442	rs11599236-T	1.00E-13	1744925:14:00	SORCS3
Diastolic blood pressure	GCST012378	rs540369678-T	2.00E-13	1750769:53:00	SORCS3
Insomnia	GCST90131901	rs10786831-A	2.00E-13	1747590:13:00	SORCS3
Systolic blood pressure	GCST007087	rs191784289-?	3.00E-13	1752263:04:00	SORCS3
Educational attainment	GCST90105038	rs10884009-C	3.00E-13	1741910:20:00	SORCS3, LINC02620
Depressive symptoms (MTAG)	GCST010642	rs2496024-?	5.00E-13	1747957:18:00	SORCS3
Educational attainment	GCST90105038	rs703482-C	7.00E-13	1752167:24:00	SORCS3
Externalizing behaviour (multivariate analysis)	GCST90061435	rs12356045-G	9.00E-13	1756862:16:00	RNU6-463P, SORCS3
Neuroticism	GCST007339	rs1490176-A	1.00E-12	1746684:27:00	SORCS3
Self-reported math ability	GCST006573	rs2864034-C	1.00E-12	1748636:23:00	SORCS3
Feeling miserable	GCST006943	rs11599236-T	2.00E-12	1744925:14:00	SORCS3
Subjective well-being (MTAG)	GCST005325	rs1961639-G	2.00E-12	1747947:19:00	SORCS3
Positive affect	GCST007338	rs1490176-A	2.00E-12	1746684:27:00	SORCS3
Insomnia	GCST90131903	rs9787523-T	2.00E-12	1745021:42:00	SORCS3
Cognitive performance (MTAG)	GCST006570	rs790647-A	3.00E-12	1750288:46:00	SORCS3
Neuropsychiatric disorders	GCST012088	rs1021363-A	4.00E-12	1747528:01:00	SORCS3
Bipolar disorder or major depressive disorder	GCST010416	rs61867293-C	6.00E-12	1746746:06:00	SORCS3
Diastolic blood pressure	GCST90132904	rs7082671-A	7.00E-12	1750371:08:00	SORCS3
Highest math class taken	GCST006574	rs902305-A	7.00E-12	1746217:46:00	SORCS3
Lifetime smoking index	GCST009096	rs3896224-A	1.00E-11	1745144:55:00	SORCS3

Attention deficit hyperactivity disorder (MTAG)	GCST010644	rs6584649-?	1.00E-11	1749765:46:00	SORCS3
Life satisfaction	GCST007337	rs1490176-A	1.00E-11	1746684:27:00	SORCS3
Attention deficit hyperactivity disorder (MTAG)	GCST012093	rs12265655-T	2.00E-11	1749756:16:00	SORCS3
Pulse pressure	GCST012377	rs181200083-C	2.00E-11	1746030:17:00	SORCS3
Insomnia	GCST90131901	rs113316302-A	2.00E-11	1746372:46:00	SORCS3
Lifetime smoking	GCST90100569	rs3896224-?	2.00E-11	1745144:55:00	SORCS3
Highest math class taken	GCST006574	rs10400054-A	2.00E-11	1750083:55:00	SORCS3
Depressed affect	GCST006475	rs11599236-C	3.00E-11	1744925:14:00	SORCS3
Neuropsychiatric disorders	GCST012088	rs11192230-A	3.00E-11	1748593:47:00	SORCS3
Insomnia	GCST90131903	rs7900775-T	4.00E-11	1747353:05:00	SORCS3
Highest math class taken (MTAG)	GCST006568	rs11192147-A	5.00E-11	1744231:28:00	SORCS3
Neuroticism (MTAG)	GCST005326	rs2451500-T	6.00E-11	1747895:19:00	SORCS3
Attention deficit hyperactivity disorder or autism spectrum disorder or intelligence (pleiotropy)	GCST90134330	rs11591402-?	6.00E-11	1749803:16:00	SORCS3
Depressive symptoms	GCST005324	rs2496022-C	7.00E-11	1747888:40:00	SORCS3
Attention deficit hyperactivity disorder or autism spectrum disorder or intelligence (pleiotropy)	GCST90134330	rs11192193-?	8.00E-11	1746834:55:00	SORCS3
Medication use (diuretics)	GCST90018985	rs145158522-T	9.00E-11	1747845:01:00	SORCS3
Medication use (diuretics)	GCST007928	rs145158522-T	9.00E-11	1747845:01:00	SORCS3
Age of smoking initiation (MTAG)	GCST007462	rs2491383-G	1.00E-10	1747624:42:00	SORCS3
Risk-taking behavior (multivariate analysis)	GCST90239693	rs9787523-?	1.00E-10	1745021:42:00	SORCS3
Treatment or medication use - amlodipine (UKB data field 20003_1140879802)	GCST90042195	rs540369678-T	1.00E-10	1750769:53:00	SORCS3
Age of smoking initiation	GCST90243972	rs142708876-ACTT	2.00E-10	1748020:27:00	SORCS3
Neuroticism	GCST007084	rs2791459-?	2.00E-10	1750458:05:00	SORCS3
Gut microbiota (beta diversity)	GCST008901	rs184389714-T	2.00E-10	1752870:15:00	SORCS3
Educational attainment (years of education)	GCST006442	rs7914674-T	2.00E-10	1749658:37:00	SORCS3
Systolic blood pressure	GCST007267	rs111790405-T	3.00E-10	1744771:27:00	SORCS3
Educational attainment (MTAG)	GCST006571	rs10509790-C	3.00E-10	1755786:52:00	SORCS3
Cognitive performance (MTAG)	GCST006570	rs55680960-C	3.00E-10	1755057:46:00	SORCS3
Highest math class taken (MTAG)	GCST006568	rs12356045-A	3.00E-10	1756862:16:00	RNU6-463P, SORCS3
Smoking initiation (ever regular vs never regular) (MTAG)	GCST007468	rs727341-C	4.00E-10	1752773:00:00	SORCS3
Anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, or Tourette syndrome (pleiotropy)	GCST009600	rs9787523-?	4.00E-10	1745021:42:00	SORCS3
Gastroesophageal reflux disease	GCST90000514	rs1021363-A	5.00E-10	1747528:01:00	SORCS3
Mood instability	GCST008357	rs11599236-?	5.00E-10	1744925:14:00	SORCS3
Neuropsychiatric disorders	GCST012088	rs11599236-T	6.00E-10	1744925:14:00	SORCS3
Smoking initiation (ever regular vs never regular)	GCST007474	rs11192347-A	6.00E-10	1752835:55:00	SORCS3
Schizophrenia	GCST90128471	rs7900775-T	6.00E-10	1747353:05:00	SORCS3
Smoking status	GCST007085	rs3896224-?	7.00E-10	1745144:55:00	SORCS3

Depression	GCST005839	rs61867293-C	7.00E-10	1746746:06:00	SORCS3
Insomnia	GCST90131901	rs61867344-A	9.00E-10	1748741:25:00	SORCS3
Adolescent idiopathic scoliosis	GCST006287	rs10884064-?	9.00E-10	1748565:16:00	SORCS3
Attention deficit hyperactivity disorder or autism spectrum disorder or intelligence (pleiotropy)	GCST90134330	rs3896224-?	9.00E-10	1745144:55:00	SORCS3
Attention deficit hyperactivity disorder (MTAG)	GCST010644	rs11192193-?	1.00E-09	1746834:55:00	SORCS3
Smoking initiation (ever regular vs never regular)	GCST007474	rs9787523-C	1.00E-09	1745021:42:00	SORCS3
Insomnia	GCST90131903	rs67700133-A	1.00E-09	1745024:11:00	SORCS3
Insomnia	GCST90131901	rs11192154-C	2.00E-09	1744568:12:00	SORCS3
Systolic blood pressure	GCST90132903	rs7082671-A	2.00E-09	1750371:08:00	SORCS3
Self-reported math ability (MTAG)	GCST006569	rs4537697-A	2.00E-09	1744330:11:00	SORCS3
Well-being spectrum (multivariate analysis)	GCST007341	rs61867341-A	2.00E-09	1748523:36:00	SORCS3
Age at first birth	GCST90000050	rs755895472-C	2.00E-09	1747619:56:00	SORCS3
Educational attainment (MTAG)	GCST006571	rs11192147-A	3.00E-09	1744231:28:00	SORCS3
Depressive symptoms	GCST007340	rs11192283-C	3.00E-09	1750322:32:00	SORCS3
Neuroticism	GCST007339	rs11192283-C	3.00E-09	1750322:32:00	SORCS3
Educational attainment	GCST90100570	rs790647-?	3.00E-09	1750288:46:00	SORCS3
Depression	GCST006477	rs1961639-A	4.00E-09	1747947:19:00	SORCS3
Depressive symptoms	GCST010647	rs1961639-?	4.00E-09	1747947:19:00	SORCS3
Neuropsychiatric disorders	GCST012088	rs11599313-T	4.00E-09	1746668:49:00	SORCS3
Educational attainment (MTAG)	GCST006571	rs1484246-A	4.00E-09	1753370:46:00	SORCS3
Pulse pressure	GCST012377	rs535313355-C	4.00E-09	1743999:42:00	SORCS3
Highest math class taken (MTAG)	GCST006568	rs703482-C	4.00E-09	1752167:24:00	SORCS3
Schizophrenia	GCST011769	rs7900775-T	4.00E-09	1747353:05:00	SORCS3
Age of smoking initiation	GCST90243986	rs142708876-ACTT	4.00E-09	1748020:27:00	SORCS3
Attention deficit hyperactivity disorder	GCST007543	rs11591402-T	4.00E-09	1749803:16:00	SORCS3
Depression	GCST90058037	rs1021363-A	5.00E-09	1747528:01:00	SORCS3
Autism spectrum disorder (MTAG)	GCST010643	rs6584649-?	6.00E-09	1749765:46:00	SORCS3
Highest math class taken	GCST006574	rs11192147-C	7.00E-09	1744231:28:00	SORCS3
Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	GCST008595	rs11192193-?	7.00E-09	1746834:55:00	SORCS3
Childhood maltreatment	GCST012253	rs3896224-A	7.00E-09	1745144:55:00	SORCS3
General factor of neuroticism	GCST007709	rs2791459-C	7.00E-09	1750458:05:00	SORCS3
General factor of neuroticism	GCST007709	rs11596214-G	7.00E-09	1744911:14:00	SORCS3
General factor of neuroticism	GCST007709	rs12416372-G	7.00E-09	1744991:09:00	SORCS3
Major depressive disorder	GCST006041	rs10786831-G	8.00E-09	1747590:13:00	SORCS3
Depressive symptoms (MTAG)	GCST010642	rs11596241-?	1.00E-08	1746417:38:00	SORCS3
Attention deficit hyperactivity disorder	GCST010649	rs11591402-?	1.00E-08	1749803:16:00	SORCS3
Autism spectrum disorder (MTAG)	GCST012092	rs1021363-A	1.00E-08	1747528:01:00	SORCS3
Adolescent idiopathic scoliosis	GCST006287	rs11192399-?	1.00E-08	1755419:16:00	SORCS3
Educational attainment (years of education)	GCST007037	rs1869165-?	1.00E-08	1750769:52:00	SORCS3

Mental disorder (without autoimmune disease)	GCST011007	rs12265655-?	1.00E-08	1749756:16:00	SORCS3
Total PHF-tau (SNP x SNP interaction)	GCST010340	rs17239258-? x rs703462-?	1.00E-08	4:182341902 10:1 05083666	TENM3 x SORCS3
Depression (broad)	GCST005902	rs1021363-G	1.00E-08	1747528:01:00	SORCS3
Cognitive ability (MTAG)	GCST005141	rs1947988-?	1.00E-08	1753130:48:00	SORCS3
Biological sex	GCST90013473	rs12356045-G	1.00E-08	1756862:16:00	RNU6-463P, SORCS3
Depressed affect	GCST006475	rs1490176-?	2.00E-08	1746684:27:00	SORCS3
General risk tolerance (MTAG)	GCST007325	rs2491392-C	2.00E-08	1746646:26:00	SORCS3
Schizophrenia (MTAG)	GCST010640	rs61867294-?	2.00E-08	1746834:09:00	SORCS3
Schizophrenia (MTAG)	GCST012089	rs61867294-A	2.00E-08	1746834:09:00	SORCS3
Insomnia	GCST90131902	rs113316302-A	2.00E-08	1746372:46:00	SORCS3
Educational attainment	GCST90105038	rs117430256-T	2.00E-08	1742828:42:00	SORCS3, LINC02620
Intelligence (MTAG)	GCST005316	rs1947988-T	2.00E-08	1753130:48:00	SORCS3
Total bone mineral density x gut microbiota (genus Dialister) interaction	GCST90094982	rs78032159-G	2.00E-08	1755058:47:00	RNU6-463P, SORCS3
Life satisfaction	GCST007337	rs11192283-C	2.00E-08	1750322:32:00	SORCS3
Self-reported math ability (MTAG)	GCST006569	rs12356045-A	2.00E-08	1756862:16:00	RNU6-463P, SORCS3
Feeling miserable	GCST006943	rs12359689-A	3.00E-08	1755603:32:00	SORCS3
Intelligence	GCST006250	rs3896224-A	3.00E-08	1745144:55:00	SORCS3
Multisite chronic pain	GCST008512	rs11599236-T	3.00E-08	1744925:14:00	SORCS3
Educational attainment	GCST90105038	rs141468874-A	3.00E-08	1751389:55:00	SORCS3
Help-seeking from a GP	GCST010011	rs11599236-T	3.00E-08	1744925:14:00	SORCS3
Cognitive performance	GCST006572	rs7895991-T	3.00E-08	1753963:26:00	SORCS3
Multisite chronic pain	GCST012332	rs11599236-?	3.00E-08	1744925:14:00	SORCS3
Attention deficit hyperactivity disorder or autism spectrum disorder or intelligence (pleiotropy)	GCST90134330	rs1490182-?	3.00E-08	1747047:51:00	SORCS3
Cognitive aspects of educational attainment	GCST90011875	rs7895991-C	3.00E-08	1753963:26:00	SORCS3
Feeling nervous	GCST006948	rs2930456-C	4.00E-08	1754850:10:00	SORCS3
Cheese consumption	GCST90133000	rs191288851-G	4.00E-08	1750280:38:00	SORCS3
Insomnia	GCST90131901	rs2451469-A	4.00E-08	1747570:40:00	SORCS3
Positive affect	GCST007338	rs11192283-C	4.00E-08	1750322:32:00	SORCS3
Educational attainment	GCST90105038	rs140163394-A	4.00E-08	1752598:00:00	SORCS3
Attention deficit hyperactivity disorder or autism spectrum disorder or intelligence (pleiotropy)	GCST90134330	rs11596241-?	4.00E-08	1746417:38:00	SORCS3
Sib-shared facial trait 902; Facial segment 53; 3D morphology of the chin	GCST90016415	rs7915785-?	4.00E-08	1758730:33:00	SORCS3, RNU6-463P
Neuroticism	GCST007084	rs11599236-?	7.00E-08	1744925:14:00	SORCS3
Neurofibrillary tangles (SNP x SNP interaction)	GCST010343	rs11113836-? x rs1475762-?	7.00E-08	1744208:08:00	No mapped genes x SORCS3
Depression (broad)	GCST005902	rs11599236-C	8.00E-08	1744925:14:00	SORCS3
Neurofibrillary tangles (SNP x SNP interaction)	GCST010343	rs4420651-? x rs2451463-?	9.00E-08	1746603:49:00	No mapped genes x SORCS3

Alzheimer disease and age of onset	GCST003427	rs117792039-A	2.00E-07	1750857:37:00	SORCS3
Trauma exposure	GCST009982	rs3896224-A	3.00E-07	1745144:55:00	SORCS3
Cognitive ability	GCST005142	rs1947988-?	8.00E-07	1753130:48:00	SORCS3
Autism and major depressive disorder (MTAG)	GCST007553	rs61867293-T	1.00E-06	1746746:06:00	SORCS3
General cognitive ability	GCST006269	rs1484246-?	1.00E-06	1753370:46:00	SORCS3
Cis-4-decenoylcarnitine (C10:1) levels in elite athletes	GCST90133911	rs975145-?	2.00E-06	1742102:30:00	LINC02620, SORCS3
Rosacea symptom severity	GCST005790	rs78911933-G	2.00E-06	1744571:20:00	SORCS3
Schizophrenia	GCST003048	rs61867294-A	2.00E-06	1746834:09:00	SORCS3
Depression	GCST003769	rs7074335-T	2.00E-06	1749020:36:00	SORCS3
Help-seeking from a GP	GCST010011	rs10786831-A	2.00E-06	1747590:13:00	SORCS3
Memory T-cell telomere length	GCST90101890	rs117670274-T	3.00E-06	1743522:05:00	SORCS3, LINC02620
Diabetic kidney disease	GCST003098	rs1997066-?	4.00E-06	1750073:25:00	SORCS3
Interleukin-16 levels	GCST004430	rs12765671-G	5.00E-06	1748750:11:00	SORCS3
Metabolite levels	GCST009391	rs6584642-?	6.00E-06	1748871:35:00	SORCS3
Non-alcoholic fatty liver disease activity score in non-alcoholic fatty liver disease	GCST008471	rs61861255-A	6.00E-06	1744878:32:00	SORCS3
Serum bilirubin levels x sex interaction in metabolic syndrome	GCST007016	rs10786848-?	7.00E-06	1752934:11:00	SORCS3
Breast cancer	GCST004988	rs1855581-G	7.00E-06	1743543:46:00	SORCS3
Reaction time	GCST006268	rs771652991-?	8.00E-06	1755096:01:00	SORCS3
Colonoscopy-negative controls vs population controls	GCST005147	rs11818462-?	8.00E-06	1755301:54:00	SORCS3
N-methylpyrrolidone levels in elite athletes	GCST90134229	rs11192362-?	8.00E-06	1753491:51:00	SORCS3
Gut microbiota relative abundance (unassigned genus belonging to family Lachnospiraceae)	GCST90006999	rs11016281-C	9.00E-06	2141704:06:00	SORCS3
General cognitive ability	GCST006269	rs902305-?	9.00E-06	1746217:46:00	SORCS3
Schizophrenia	GCST90128471	rs11192131-C	9.00E-06	1743338:00:00	SORCS3, LINC02620

Supplementary Table 8: LDSC regression SNP-based heritability (h2SNP) estimates of GAD and fear-based disorder phenotypes						
Datasets	Anxiety disorder	Control criteria	GWAS mean χ^2	h2SNP converted to Liability scale		
				Population prevalence = COPING sample prevalence	Population prevalence = COPING sample prevalence + 10%	Population prevalence = COPING sample prevalence - 10%
GLAD+ & UKB	GAD	Specific	1.17	0.074 (0.006)	0.076 (0.006)	0.072 (0.006)
		Any anx-dep	1.19	0.088 (0.007)	0.090 (0.007)	0.085 (0.006)
	Fear	Specific	1.09	0.054 (0.007)	0.056 (0.007)	0.052 (0.006)
		Any anx-dep	1.13	0.084 (0.007)	0.086 (0.008)	0.080 (0.007)
	GAD	Specific	1.23	0.087 (0.006)	0.088 (0.006)	0.083 (0.005)

QIMR, GLAD+ & UKB		Any anx-dep	1.25	0.100 (0.006)	0.103 (0.006)	0.096 (0.006)
		Specific	1.12	0.063 (0.006)	0.064 (0.006)	0.060 (0.005)
	Fear	Any anx-dep	1.17	0.096 (0.006)	0.098 (0.006)	0.092 (0.006)
h2SNP estimated using LDSC regression. Standard errors are presented in parentheses						

Supplementary Table 9: Genetic correlations (r_g) between anxiety disorder phenotypes and 345 external traits estimated in LDSC regression.

$p(0)$ - p -value for test of r_g different from 0. **Bold** - r_g significantly different from 0, or significant jack-knife p value for difference in r_g s ($p < 1.45E-04$, Bonferroni correction threshold for 345 phenotypes).

External Phenotype			GAD					Fear-based disorders					Comparison GAD & fear r_g		
Trait name	PMID	h^2_{SNP} obs z	r_g	se	z	$p(0)$	Sig. $p(0)$	r_g	se	z	$p(0)$	Sig. $p(0)$	z	p	Sig. $p(\text{diff.})$
Educational Attainment (2018)	30038396	34.91	-0.01	0.03	-0.47	6.38E-01	No	-0.28	0.04	-7.56	3.95E-14	Yes	7.35	2.00E-13	Sig. stronger fear
College/university degree attainment	27046643	19.35	0.00	0.04	0.01	9.94E-01	No	-0.32	0.05	-6.01	1.91E-09	Yes	6.23	4.79E-10	Sig. stronger fear
Educational Attainment (female)	27225129	23.24	-0.02	0.04	-0.40	6.89E-01	No	-0.31	0.05	-6.02	1.70E-09	Yes	6.12	9.61E-10	Sig. stronger fear
Educational Attainment (2016)	27225129	29.07	-0.03	0.03	-0.76	4.50E-01	No	-0.31	0.04	-6.94	4.05E-12	Yes	6.10	1.09E-09	Sig. stronger fear
Bipolar disorder (PGC2)	31043756	20.98	0.38	0.05	8.28	1.28E-16	Yes	0.15	0.05	2.90	3.80E-03	No	5.16	2.51E-07	Sig. stronger GAD
Bipolar disorder type I & II	34002096	25.71	0.41	0.04	9.45	3.54E-21	Yes	0.21	0.04	4.69	2.77E-06	Yes	4.80	1.55E-06	Sig. stronger GAD
Educational Attainment (male)	27225129	21.83	-0.03	0.04	-0.80	4.25E-01	No	-0.29	0.05	-5.70	1.23E-08	Yes	4.70	2.59E-06	Sig. stronger fear
Cross disorders (ADHD, ASD, BIP, SCZ)	23453885	14.74	0.58	0.05	11.50	1.30E-30	Yes	0.30	0.06	4.96	6.92E-07	Yes	4.67	3.00E-06	Sig. stronger GAD
Bipolar disorder (PGC1)	21926972	10.90	0.37	0.06	6.25	4.24E-10	Yes	0.03	0.07	0.42	6.77E-01	No	4.66	3.10E-06	Sig. stronger GAD
Coronary Artery Disease (2017)	28714975	15.96	0.13	0.04	3.50	5.00E-04	No	0.33	0.05	6.78	1.17E-11	Yes	-4.60	4.25E-06	Sig. stronger fear
Bipolar disorder type I	34002096	24.18	0.33	0.04	7.83	5.01E-15	Yes	0.13	0.05	2.73	6.30E-03	No	4.35	1.37E-05	Sig. stronger GAD
Age first birth (male)	27798627	10.90	-0.09	0.06	-1.49	1.37E-01	No	-0.43	0.08	-5.32	1.05E-07	Yes	4.15	3.33E-05	Sig. stronger fear
General cognitive ability (2018)	29942086	29.98	-0.09	0.03	-2.61	9.10E-03	No	-0.28	0.05	-6.05	1.43E-09	Yes	4.06	4.87E-05	Sig. stronger fear
Anorexia nervosa (PGC2)	31308545	15.42	0.31	0.05	6.15	7.78E-10	Yes	0.08	0.06	1.43	1.52E-01	No	3.90	9.50E-05	Sig. stronger GAD
BMI (Locke) female	25673413	18.74	-0.06	0.04	-1.54	1.23E-01	No	0.13	0.04	3.04	2.40E-03	No	-3.90	9.52E-05	Sig. stronger fear
Neuroticism UKB	30593698	19.22	0.75	0.03	22.40	3.60E-111	Yes	0.56	0.05	12.19	3.40E-34	Yes	3.83	1.27E-04	Sig. stronger GAD

External Phenotype			GAD						Fear-based disorders				Comparison GAD & fear r_g		
MDD (PGC2 + 23andme, no UKB)	29700475	23.17	0.81	0.04	20.55	7.10E-94	Yes	0.59	0.05	11.87	1.66E-32	Yes	3.82	1.36E-04	Sig. stronger GAD
BMI (EUR Locke)	25673413	18.21	-0.04	0.03	-1.26	2.09E-01	No	0.12	0.04	2.93	3.40E-03	No	-3.79	1.48E-04	Not Sig. diff.
Coronary Artery Disease (mixed ancestry)	26343387	14.55	0.14	0.04	3.44	6.00E-04	No	0.32	0.05	6.23	4.56E-10	Yes	-3.78	1.54E-04	Not Sig. diff.
MDD (PGC2 + 23andme + UKB)	29700475	23.88	0.84	0.04	21.90	2.90E-106	Yes	0.62	0.05	12.50	7.29E-36	Yes	3.76	1.73E-04	Not Sig. diff.
Body fat %	30593698	29.63	0.02	0.03	0.60	5.49E-01	No	0.14	0.04	3.81	1.00E-04	Yes	-3.75	1.80E-04	Not Sig. diff.
BMI (MIX Locke)	25673413	18.24	-0.04	0.03	-1.19	2.33E-01	No	0.12	0.04	3.11	1.90E-03	No	-3.67	2.46E-04	Not Sig. diff.
SCZ & bipolar	24280982	17.92	0.44	0.05	9.35	8.49E-21	Yes	0.21	0.05	3.86	1.00E-04	Yes	3.63	2.81E-04	Not Sig. diff.
Fat mass meta-analysis of males & females	31852892	27.70	-0.01	0.03	-0.32	7.47E-01	No	0.11	0.04	2.82	4.80E-03	No	-3.63	2.83E-04	Not Sig. diff.
Body fat % meta-analysis of males & females	30593698	29.40	0.02	0.03	0.67	5.02E-01	No	0.14	0.04	3.64	3.00E-04	No	-3.57	3.64E-04	Not Sig. diff.
Neuroticism UKB (female)	30593698	17.30	0.73	0.04	18.10	3.49E-73	Yes	0.53	0.05	10.86	1.77E-27	Yes	3.54	3.96E-04	Not Sig. diff.
Age first birth	27798627	17.52	-0.15	0.05	-3.26	1.10E-03	No	-0.37	0.06	-5.90	3.59E-09	Yes	3.49	4.82E-04	Not Sig. diff.
Body fat % (female)	30593698	26.51	0.01	0.04	0.14	8.87E-01	No	0.16	0.04	3.93	8.43E-05	Yes	-3.48	5.02E-04	Not Sig. diff.
Overweight	23563607	16.66	-0.08	0.05	-1.71	8.69E-02	No	0.10	0.05	1.87	6.16E-02	No	-3.44	5.88E-04	Not Sig. diff.
Neuroticism (2016)	27089181	12.38	0.80	0.04	18.72	3.35E-78	Yes	0.58	0.06	10.17	2.71E-24	Yes	3.44	5.92E-04	Not Sig. diff.
Depressive symptoms	27089181	12.24	0.81	0.06	14.21	8.20E-46	Yes	0.57	0.07	7.82	5.32E-15	Yes	3.39	6.90E-04	Not Sig. diff.
BMI	31852892	26.87	-0.01	0.03	-0.20	8.43E-01	No	0.12	0.04	3.08	2.10E-03	No	-3.35	7.95E-04	Not Sig. diff.
Neuroticism UKB (male)	30593698	12.52	0.77	0.04	18.54	9.10E-77	Yes	0.58	0.06	9.59	9.17E-22	Yes	3.27	1.09E-03	Not Sig. diff.
General cognitive ability (2017)	28530673	18.70	-0.05	0.05	-1.12	2.62E-01	No	-0.26	0.06	-4.63	3.66E-06	Yes	3.24	1.19E-03	Not Sig. diff.
Waist circumference all ancestries	25673412	16.12	-0.05	0.04	-1.42	1.55E-01	No	0.11	0.05	2.24	2.54E-02	No	-3.23	1.24E-03	Not Sig. diff.
Verbal-numerical reasoning	27046643	10.78	-0.06	0.06	-1.10	2.69E-01	No	-0.32	0.07	-4.51	6.48E-06	Yes	3.21	1.32E-03	Not Sig. diff.
Fat mass female	31852892	24.76	-0.03	0.04	-0.78	4.35E-01	No	0.11	0.04	2.76	5.70E-03	No	-3.20	1.36E-03	Not Sig. diff.
Body fat % healthy	30593698	24.52	0.03	0.04	0.66	5.11E-01	No	0.15	0.04	3.29	1.00E-03	No	-3.15	1.61E-03	Not Sig. diff.
Weight EU female	23754948	18.24	-0.13	0.04	-2.85	4.30E-03	No	0.05	0.05	1.01	3.12E-01	No	-3.14	1.71E-03	Not Sig. diff.
SCZ (PGC2)	25056061	25.06	0.39	0.04	9.64	5.58E-22	Yes	0.25	0.05	5.51	3.58E-08	Yes	3.12	1.78E-03	Not Sig. diff.
Waist circumference EUR	25673412	16.49	-0.05	0.04	-1.38	1.68E-01	No	0.10	0.05	2.18	2.91E-02	No	-3.12	1.81E-03	Not Sig. diff.
Fat mass	31852892	27.96	-0.01	0.03	-0.38	7.03E-01	No	0.11	0.04	2.93	3.40E-03	No	-3.11	1.88E-03	Not Sig. diff.
Household income, one per household	27818178	9.91	-0.20	0.06	-3.35	8.00E-04	No	-0.42	0.07	-5.89	3.87E-09	Yes	3.09	1.98E-03	Not Sig. diff.
SCZ (PGC2) (mixed ancestry)	25056061	26.09	0.39	0.04	9.79	1.26E-22	Yes	0.25	0.04	5.75	9.18E-09	Yes	3.07	2.11E-03	Not Sig. diff.
Waist circumference EUR female	25673412	16.57	-0.05	0.04	-1.19	2.33E-01	No	0.11	0.05	2.34	1.91E-02	No	-3.06	2.19E-03	Not Sig. diff.
Household income	27818178	10.54	-0.20	0.06	-3.60	3.00E-04	No	-0.42	0.07	-6.16	7.43E-10	Yes	3.04	2.40E-03	Not Sig. diff.
Neuroticism, no psychiatric illness (male)	30593698	7.75	0.79	0.06	14.52	9.46E-48	Yes	0.57	0.08	7.57	3.79E-14	Yes	2.97	2.99E-03	Not Sig. diff.
BMI male	31852892	24.83	-0.01	0.04	-0.25	8.04E-01	No	0.10	0.04	2.18	2.95E-02	No	-2.94	3.30E-03	Not Sig. diff.
Neuroticism, no psychiatric illness	30593698	12.44	0.75	0.04	18.66	1.14E-77	Yes	0.57	0.06	10.41	2.30E-25	Yes	2.94	3.34E-03	Not Sig. diff.
Parents attained age	27015805	12.40	0.02	0.05	0.44	6.62E-01	No	0.19	0.06	3.29	1.00E-03	No	-2.92	3.47E-03	Not Sig. diff.
BMI meta-analysis of males & females	31852892	27.18	-0.01	0.03	-0.25	8.04E-01	No	0.11	0.04	2.95	3.20E-03	No	-2.91	3.58E-03	Not Sig. diff.
Waist-hip ratio EUR	25673412	16.32	-0.01	0.04	-0.17	8.65E-01	No	0.14	0.05	2.66	7.70E-03	No	-2.90	3.70E-03	Not Sig. diff.

External Phenotype			GAD						Fear-based disorders				Comparison GAD & fear r_g		
Mother's age at death	27015805	8.81	-0.07	0.07	-0.97	3.30E-01	No	-0.32	0.08	-4.01	6.11E-05	Yes	2.88	4.01E-03	Not Sig. diff.
Anorexia nervosa (PGC1)	28494655	7.33	0.32	0.07	4.57	4.93E-06	Yes	0.06	0.09	0.61	5.43E-01	No	2.84	4.52E-03	Not Sig. diff.
MDD (PGC2, no 23andme, no UKB)	29700475	16.24	0.77	0.05	14.79	1.78E-49	Yes	0.58	0.07	8.66	4.83E-18	Yes	2.83	4.65E-03	Not Sig. diff.
Fat mass healthy meta-analysis (not alco/smoking-adjust)	31852892	23.27	0.00	0.04	0.03	9.73E-01	No	0.11	0.05	2.36	1.83E-02	No	-2.82	4.80E-03	Not Sig. diff.
Body fat % healthy meta-analysis (not alcohol/smoking-adjust)	30593698	24.69	0.02	0.04	0.62	5.34E-01	No	0.14	0.04	3.20	1.40E-03	No	-2.82	4.86E-03	Not Sig. diff.
BMI female	31852892	24.00	-0.01	0.04	-0.32	7.48E-01	No	0.12	0.04	2.90	3.80E-03	No	-2.81	5.03E-03	Not Sig. diff.
College completion	23722424	13.47	-0.12	0.05	-2.18	2.91E-02	No	-0.32	0.07	-4.55	5.41E-06	Yes	2.80	5.05E-03	Not Sig. diff.
Hip circumference all ancestries female	25673412	16.70	-0.05	0.04	-1.34	1.81E-01	No	0.09	0.05	1.80	7.21E-02	No	-2.76	5.85E-03	Not Sig. diff.
BMI healthy meta-analysis of males & females	31852892	23.39	-0.01	0.04	-0.15	8.77E-01	No	0.11	0.05	2.48	1.33E-02	No	-2.75	6.02E-03	Not Sig. diff.
Waist-hip ratio all ancestries	25673412	16.49	-0.01	0.04	-0.18	8.58E-01	No	0.14	0.05	2.69	7.20E-03	No	-2.72	6.48E-03	Not Sig. diff.
Fat mass healthy meta-analysis of males & females	31852892	23.51	0.00	0.04	0.01	9.93E-01	No	0.10	0.05	2.31	2.11E-02	No	-2.68	7.38E-03	Not Sig. diff.
Body fat % anorexia-free meta-analysis of males & females	30593698	24.44	0.03	0.04	0.67	5.02E-01	No	0.14	0.04	3.12	1.80E-03	No	-2.68	7.43E-03	Not Sig. diff.
Body fat % anorexia-free	30593698	24.65	0.03	0.04	0.73	4.66E-01	No	0.15	0.05	3.25	1.20E-03	No	-2.67	7.70E-03	Not Sig. diff.
Father attained age	27015805	13.45	0.08	0.06	1.34	1.81E-01	No	0.23	0.06	3.93	8.40E-05	Yes	-2.66	7.71E-03	Not Sig. diff.
BMI anorexia-free male	31852892	19.49	-0.03	0.05	-0.56	5.76E-01	No	0.10	0.05	1.98	4.83E-02	No	-2.61	9.08E-03	Not Sig. diff.
Body fat % healthy meta-analysis of males & females	30593698	24.63	0.02	0.04	0.61	5.44E-01	No	0.14	0.04	3.15	1.60E-03	No	-2.59	9.63E-03	Not Sig. diff.
Age at Menarche (marker of puberty in females)	28436984	20.27	-0.09	0.03	-2.69	7.10E-03	No	0.01	0.04	0.17	8.69E-01	No	-2.58	9.77E-03	Not Sig. diff.
BMI healthy	31852892	23.21	0.00	0.04	-0.06	9.52E-01	No	0.11	0.05	2.50	1.25E-02	No	-2.56	1.06E-02	Not Sig. diff.
BMI anorexia-free	31852892	23.07	0.00	0.04	-0.02	9.81E-01	No	0.11	0.05	2.42	1.55E-02	No	-2.55	1.07E-02	Not Sig. diff.
BMI Locke & UKB (updated)	30124842	28.98	-0.02	0.03	-0.52	6.04E-01	No	0.08	0.04	2.11	3.48E-02	No	-2.53	1.14E-02	Not Sig. diff.
Fat mass healthy male	31852892	19.00	-0.01	0.05	-0.11	9.12E-01	No	0.10	0.05	1.93	5.39E-02	No	-2.52	1.19E-02	Not Sig. diff.
BMI healthy male	31852892	19.52	-0.03	0.05	-0.55	5.83E-01	No	0.10	0.05	1.96	4.96E-02	No	-2.50	1.24E-02	Not Sig. diff.
Mother attained age	27015805	9.82	0.04	0.07	0.63	5.30E-01	No	0.24	0.07	3.29	1.00E-03	No	-2.50	1.24E-02	Not Sig. diff.
Obesity class 2	23563607	14.25	-0.01	0.05	-0.25	7.99E-01	No	0.13	0.06	2.20	2.80E-02	No	-2.50	1.25E-02	Not Sig. diff.
Body mass index healthy meta-analysis (not alcohol/smoking-adjust)	30593698	23.18	-0.01	0.04	-0.12	9.07E-01	No	0.12	0.05	2.56	1.05E-02	No	-2.49	1.27E-02	Not Sig. diff.
Obesity class 1	23563607	18.43	-0.04	0.04	-1.01	3.15E-01	No	0.10	0.05	1.97	4.90E-02	No	-2.49	1.28E-02	Not Sig. diff.
Fat mass anorexia-free meta-analysis of males & females	31852892	23.33	0.00	0.04	0.07	9.46E-01	No	0.10	0.05	2.28	2.25E-02	No	-2.49	1.29E-02	Not Sig. diff.
Heavy versus never smoker	28166213	15.46	0.09	0.05	1.75	8.05E-02	No	0.24	0.06	4.14	3.46E-05	Yes	-2.47	1.36E-02	Not Sig. diff.
Extreme BMI	23563607	12.61	-0.11	0.06	-1.73	8.38E-02	No	0.07	0.07	1.02	3.07E-01	No	-2.46	1.39E-02	Not Sig. diff.
Hip circumference EUR	25673412	16.99	-0.07	0.04	-1.86	6.35E-02	No	0.05	0.05	0.97	3.35E-01	No	-2.45	1.41E-02	Not Sig. diff.
BMI anorexia-free meta-analysis of males & females	31852892	23.24	-0.01	0.04	-0.13	8.97E-01	No	0.11	0.05	2.41	1.59E-02	No	-2.45	1.42E-02	Not Sig. diff.
Plateletcrit	27863252	12.24	-0.01	0.03	-0.38	7.07E-01	No	0.09	0.04	2.38	1.75E-02	No	-2.44	1.47E-02	Not Sig. diff.
Ever smoker	30643258	28.66	0.15	0.04	4.07	4.76E-05	Yes	0.24	0.05	5.37	7.75E-08	Yes	-2.43	1.52E-02	Not Sig. diff.

External Phenotype			GAD						Fear-based disorders					Comparison GAD & fear r_g		
Fat mass male	31852892	24.43	0.00	0.04	0.08	9.37E-01	No	0.10	0.04	2.23	2.59E-02	No	-2.42	1.57E-02	Not Sig. diff.	
Fat mass healthy	31852892	23.49	0.00	0.04	0.04	9.66E-01	No	0.10	0.05	2.33	2.00E-02	No	-2.41	1.60E-02	Not Sig. diff.	
Hip circumference EUR female	25673412	16.70	-0.05	0.04	-1.31	1.90E-01	No	0.09	0.05	1.75	7.96E-02	No	-2.40	1.65E-02	Not Sig. diff.	
Self-rated health	27864402	15.67	-0.34	0.05	-6.43	1.29E-10	Yes	-0.51	0.06	-8.12	4.56E-16	Yes	2.38	1.75E-02	Not Sig. diff.	
Social deprivation	27818178	8.04	0.17	0.06	2.68	7.50E-03	No	0.37	0.09	3.95	7.76E-05	Yes	-2.35	1.86E-02	Not Sig. diff.	
Hip circumference all ancestries	25673412	16.80	-0.07	0.04	-1.91	5.68E-02	No	0.05	0.05	0.99	3.25E-01	No	-2.33	1.97E-02	Not Sig. diff.	
Age first birth (female)	27798627	14.57	-0.17	0.05	-3.42	6.00E-04	No	-0.34	0.07	-5.13	2.87E-07	Yes	2.32	2.01E-02	Not Sig. diff.	
Fat mass anorexia-free	31852892	23.31	0.00	0.04	0.11	9.14E-01	No	0.10	0.05	2.29	2.19E-02	No	-2.31	2.09E-02	Not Sig. diff.	
Type-2 Diabetes 2017	28566273	13.24	0.06	0.05	1.34	1.79E-01	No	0.21	0.06	3.60	3.00E-04	No	-2.28	2.29E-02	Not Sig. diff.	
Undersleeper	27494321	10.43	0.20	0.06	3.18	1.50E-03	No	0.38	0.08	4.76	1.92E-06	Yes	-2.27	2.31E-02	Not Sig. diff.	
Body fat % healthy female	30593698	19.07	0.03	0.05	0.68	4.94E-01	No	0.17	0.06	2.99	2.80E-03	No	-2.26	2.41E-02	Not Sig. diff.	
ADHD (male)	29325848	12.06	0.24	0.06	4.39	1.15E-05	Yes	0.40	0.08	4.98	6.35E-07	Yes	-2.24	2.53E-02	Not Sig. diff.	
Mean diameter for HDL particles	27005778	4.12	0.22	0.10	2.27	2.33E-02	No	-0.05	0.11	-0.44	6.59E-01	No	2.21	2.69E-02	Not Sig. diff.	
Lifecourse bone mineral density: -15	29304378	8.18	-0.01	0.07	-0.18	8.61E-01	No	-0.21	0.09	-2.34	1.91E-02	No	2.20	2.79E-02	Not Sig. diff.	
MDD (PGC1)	22472876	5.75	0.80	0.10	7.72	1.20E-14	Yes	0.59	0.11	5.58	2.48E-08	Yes	2.20	2.80E-02	Not Sig. diff.	
SCZ PGC 2018 (mixed ancestry)	31740837	30.04	0.38	0.04	10.79	3.92E-27	Yes	0.28	0.04	7.19	6.44E-13	Yes	2.20	2.80E-02	Not Sig. diff.	
Obesity class 3	23563607	8.42	-0.01	0.08	-0.07	9.48E-01	No	0.18	0.09	2.13	3.31E-02	No	-2.20	2.80E-02	Not Sig. diff.	
Body fat % healthy male	30593698	20.18	0.01	0.05	0.31	7.58E-01	No	0.11	0.05	2.32	2.03E-02	No	-2.19	2.88E-02	Not Sig. diff.	
Lifetime anxiety (female)	31748690	11.65	0.98	0.03	32.26	2.38E-228	Yes	0.85	0.06	14.93	2.20E-50	Yes	2.18	2.93E-02	Not Sig. diff.	
Body fat % male	30593698	25.32	0.04	0.04	0.92	3.56E-01	No	0.12	0.04	2.76	5.80E-03	No	-2.18	2.93E-02	Not Sig. diff.	
BMI (Locke) male	25673413	15.40	-0.03	0.04	-0.79	4.32E-01	No	0.08	0.05	1.74	8.21E-02	No	-2.17	3.00E-02	Not Sig. diff.	
Parents' age at death	27015805	10.44	-0.06	0.07	-0.81	4.17E-01	No	-0.23	0.08	-3.10	1.90E-03	No	2.17	3.01E-02	Not Sig. diff.	
Autism spectrum disorder	30804558	13.25	0.34	0.06	6.15	7.55E-10	Yes	0.21	0.07	3.22	1.30E-03	No	2.16	3.07E-02	Not Sig. diff.	
Schizophrenia	29483656	27.32	0.36	0.04	9.73	2.23E-22	Yes	0.26	0.05	5.66	1.52E-08	Yes	2.16	3.08E-02	Not Sig. diff.	
Cholesterol esters in large HDL	27005778	4.43	0.23	0.10	2.41	1.60E-02	No	-0.02	0.11	-0.18	8.57E-01	No	2.16	3.08E-02	Not Sig. diff.	
Age at Menarche (marker of puberty in females)	25231870	19.98	-0.07	0.04	-1.61	1.07E-01	No	0.03	0.05	0.67	5.02E-01	No	-2.15	3.14E-02	Not Sig. diff.	
Waist-hip ratio all ancestries BMI-adjusted	25673412	12.36	0.02	0.04	0.38	7.04E-01	No	0.11	0.05	2.21	2.71E-02	No	-2.15	3.19E-02	Not Sig. diff.	
Body fat % anorexia-free male	30593698	20.12	0.02	0.05	0.34	7.35E-01	No	0.11	0.05	2.33	1.96E-02	No	-2.13	3.29E-02	Not Sig. diff.	
Body fat % anorexia-free female	30593698	19.01	0.04	0.05	0.77	4.44E-01	No	0.16	0.06	2.93	3.40E-03	No	-2.11	3.48E-02	Not Sig. diff.	
Neuroticism, no psychiatric illness (female)	30593698	10.65	0.71	0.06	13.02	9.56E-39	Yes	0.57	0.06	8.91	5.14E-19	Yes	2.09	3.65E-02	Not Sig. diff.	
Fat mass anorexia-free male	31852892	18.97	0.00	0.05	-0.09	9.27E-01	No	0.10	0.05	1.95	5.08E-02	No	-2.06	3.93E-02	Not Sig. diff.	
Lifecourse bone mineral density: meta-analysis	29304378	12.26	-0.04	0.04	-1.03	3.02E-01	No	-0.14	0.05	-2.91	3.60E-03	No	2.05	4.01E-02	Not Sig. diff.	
Insomnia	30804565	21.95	0.28	0.04	7.39	1.48E-13	Yes	0.39	0.05	7.20	6.09E-13	Yes	-2.04	4.14E-02	Not Sig. diff.	
High FEV1 versus average FEV1 never smokers	28166213	5.85	0.05	0.08	0.58	5.65E-01	No	0.28	0.10	2.70	6.90E-03	No	-2.01	4.48E-02	Not Sig. diff.	
Lean mass (fat free mass)	30593698	24.18	-0.07	0.03	-2.49	1.27E-02	No	-0.01	0.03	-0.35	7.26E-01	No	-2.00	4.55E-02	Not Sig. diff.	
Lean mass (fat free mass) male	30593698	21.16	-0.06	0.03	-2.02	4.37E-02	No	0.00	0.04	-0.07	9.48E-01	No	-2.00	4.57E-02	Not Sig. diff.	
Osteoporosis: femoral neck bone mineral density	26367794	7.14	-0.30	0.12	-2.57	1.01E-02	No	-0.06	0.10	-0.59	5.55E-01	No	-2.00	4.59E-02	Not Sig. diff.	

External Phenotype			GAD						Fear-based disorders						Comparison GAD & fear r_g		
ADHD	30478444	15.95	0.22	0.05	4.67	3.09E-06	Yes	0.35	0.06	5.52	3.41E-08	Yes	-1.98	4.79E-02	Not Sig. diff.		
Body fat % (Lu. et al)	26833246	12.66	0.05	0.06	0.89	3.73E-01	No	0.19	0.07	2.75	5.90E-03	No	-1.97	4.89E-02	Not Sig. diff.		
Father's age at death	27015805	10.06	-0.12	0.06	-2.08	3.80E-02	No	-0.27	0.07	-3.90	9.54E-05	Yes	1.96	4.96E-02	Not Sig. diff.		
Primary Biliary Cirrhosis	26394269	6.11	0.13	0.07	1.97	4.94E-02	No	0.28	0.08	3.75	2.00E-04	No	-1.94	5.23E-02	Not Sig. diff.		
Low FEV1 versus high FEV1 never smokers	28166213	12.95	0.00	0.05	0.03	9.76E-01	No	-0.12	0.06	-1.85	6.39E-02	No	1.94	5.24E-02	Not Sig. diff.		
Systemic Lupus Erythematosus	26502338	5.73	0.11	0.09	1.15	2.52E-01	No	0.29	0.09	3.35	8.00E-04	No	-1.90	5.69E-02	Not Sig. diff.		
Hemoglobin concentration	27863252	11.11	0.03	0.04	0.69	4.89E-01	No	0.12	0.05	2.49	1.27E-02	No	-1.90	5.72E-02	Not Sig. diff.		
Heart Rate	23583979	8.48	0.11	0.05	2.17	2.97E-02	No	0.24	0.07	3.47	5.00E-04	No	-1.90	5.79E-02	Not Sig. diff.		
Concentration of large HDL particles	27005778	4.66	0.19	0.09	2.09	3.64E-02	No	-0.01	0.10	-0.06	9.54E-01	No	1.88	6.08E-02	Not Sig. diff.		
Fat mass healthy female	31852892	17.69	0.00	0.05	0.02	9.83E-01	No	0.10	0.06	1.80	7.13E-02	No	-1.83	6.69E-02	Not Sig. diff.		
Lifetime anxiety	31748690	14.52	0.98	0.02	63.72	0.00E+00	Yes	0.90	0.05	19.60	1.72E-85	Yes	1.82	6.91E-02	Not Sig. diff.		
Age at Menopause	26414677	8.58	-0.06	0.05	-1.12	2.61E-01	No	-0.18	0.06	-2.88	3.90E-03	No	1.80	7.20E-02	Not Sig. diff.		
serum creatinine, non diabetes mellitus)	26831199	10.89	-0.04	0.05	-0.76	4.45E-01	No	0.07	0.06	1.05	2.92E-01	No	-1.79	7.28E-02	Not Sig. diff.		
Total lipids in large HDL	27005778	4.61	0.20	0.09	2.13	3.29E-02	No	-0.01	0.10	-0.11	9.09E-01	No	1.79	7.35E-02	Not Sig. diff.		
Severe early onset obesity vs controls	30677029	7.61	0.25	0.08	3.35	8.00E-04	No	0.08	0.09	0.93	3.54E-01	No	1.78	7.56E-02	Not Sig. diff.		
ADHD (mixed ancestry)	30478444	16.29	0.23	0.05	4.74	2.10E-06	Yes	0.34	0.06	5.54	3.00E-08	Yes	-1.77	7.72E-02	Not Sig. diff.		
BMI healthy female	31852892	17.43	0.01	0.05	0.22	8.27E-01	No	0.11	0.06	1.98	4.75E-02	No	-1.76	7.78E-02	Not Sig. diff.		
Triglycerides	20686565	5.36	0.05	0.04	1.22	2.22E-01	No	0.15	0.06	2.72	6.40E-03	No	-1.76	7.79E-02	Not Sig. diff.		
Waist circumference EUR male	25673412	12.48	-0.07	0.05	-1.24	2.16E-01	No	0.06	0.06	0.89	3.76E-01	No	-1.76	7.90E-02	Not Sig. diff.		
Creatinine	27005778	4.52	0.01	0.09	0.15	8.79E-01	No	-0.19	0.11	-1.72	8.47E-02	No	1.75	8.04E-02	Not Sig. diff.		
Waist circumference all ancestries BMI-adjusted	25673412	16.12	-0.04	0.04	-1.03	3.04E-01	No	0.03	0.04	0.70	4.84E-01	No	-1.74	8.22E-02	Not Sig. diff.		
Phospholipids in large HDL	27005778	4.22	0.17	0.10	1.82	6.88E-02	No	-0.02	0.11	-0.15	8.83E-01	No	1.73	8.42E-02	Not Sig. diff.		
Lean mass (fat free mass) meta-analysis males & females	30593698	24.33	-0.08	0.03	-2.79	5.20E-03	No	-0.02	0.03	-0.63	5.29E-01	No	-1.72	8.48E-02	Not Sig. diff.		
High FEV1 versus average FEV1 heavy smokers	28166213	5.42	-0.06	0.07	-0.79	4.31E-01	No	-0.23	0.11	-2.18	2.91E-02	No	1.72	8.63E-02	Not Sig. diff.		
Total body less head bone mineral density children	28743860	8.86	0.08	0.07	1.09	2.78E-01	No	0.22	0.09	2.61	9.00E-03	No	-1.69	9.10E-02	Not Sig. diff.		
Lean mass (fat free mass) healthy male	30593698	17.07	-0.04	0.04	-1.09	2.76E-01	No	0.02	0.04	0.34	7.33E-01	No	-1.68	9.30E-02	Not Sig. diff.		
Male pattern baldness	28196072	7.07	0.01	0.04	0.29	7.73E-01	No	-0.07	0.05	-1.32	1.87E-01	No	1.65	1.00E-01	Not Sig. diff.		
Fat mass anorexia-free female	31852892	17.41	0.01	0.05	0.11	9.10E-01	No	0.10	0.06	1.74	8.27E-02	No	-1.64	1.00E-01	Not Sig. diff.		
Platelet count	27863252	12.75	0.00	0.03	-0.13	9.00E-01	No	0.07	0.04	1.68	9.35E-02	No	-1.63	1.02E-01	Not Sig. diff.		
PTSD EUR female	28439101	4.24	0.25	0.11	2.35	1.88E-02	No	0.04	0.13	0.28	7.81E-01	No	1.62	1.05E-01	Not Sig. diff.		
Consumption (AUDIT-C)	30336701	15.89	-0.07	0.05	-1.30	1.92E-01	No	-0.16	0.06	-2.59	9.70E-03	No	1.62	1.06E-01	Not Sig. diff.		
Type 2 Diabetes 2012	22885922	9.48	-0.01	0.07	-0.11	9.15E-01	No	0.12	0.08	1.52	1.28E-01	No	-1.61	1.08E-01	Not Sig. diff.		
eGFRcrea (serum creatinine)	26831199	11.18	-0.02	0.05	-0.46	6.43E-01	No	0.06	0.06	1.00	3.20E-01	No	-1.61	1.08E-01	Not Sig. diff.		
Waist circumference EUR BMI-adjusted	25673412	16.25	-0.04	0.04	-0.94	3.49E-01	No	0.03	0.05	0.70	4.87E-01	No	-1.59	1.13E-01	Not Sig. diff.		
Hematocrit	27863252	11.04	0.03	0.04	0.97	3.34E-01	No	0.11	0.05	2.39	1.71E-02	No	-1.58	1.14E-01	Not Sig. diff.		
Lean mass (fat free mass) female	30593698	23.30	-0.10	0.03	-3.01	2.60E-03	No	-0.04	0.04	-1.14	2.54E-01	No	-1.58	1.14E-01	Not Sig. diff.		

External Phenotype			GAD					Fear-based disorders					Comparison GAD & fear r_g		
Eosinophil % of white cells	27863252	9.42	-0.09	0.04	-2.47	1.35E-02	No	-0.01	0.05	-0.22	8.23E-01	No	-1.56	1.19E-01	Not Sig. diff.
AUDIT total score	30336701	16.87	0.01	0.05	0.13	8.99E-01	No	-0.09	0.06	-1.50	1.34E-01	No	1.56	1.19E-01	Not Sig. diff.
Lifecourse bone mineral density: 45-60	29304378	9.69	-0.09	0.07	-1.36	1.74E-01	No	-0.21	0.08	-2.71	6.60E-03	No	1.56	1.19E-01	Not Sig. diff.
Visceral adipose tissue volume adj. for BMI female	27918534	4.83	0.16	0.10	1.59	1.13E-01	No	-0.01	0.11	-0.05	9.58E-01	No	1.55	1.20E-01	Not Sig. diff.
Waist-hip ratio EUR BMI-adjusted	25673412	12.42	0.02	0.04	0.42	6.73E-01	No	0.11	0.05	2.19	2.82E-02	No	-1.55	1.22E-01	Not Sig. diff.
Celiac Disease	20190752	5.71	0.02	0.08	0.26	7.97E-01	No	0.17	0.10	1.72	8.60E-02	No	-1.54	1.23E-01	Not Sig. diff.
Cholesterol in large HDL	27005778	4.08	0.19	0.10	1.96	4.97E-02	No	0.01	0.11	0.06	9.50E-01	No	1.54	1.23E-01	Not Sig. diff.
Bipolar disorder type II	34002096	8.20	0.53	0.07	7.28	3.42E-13	Yes	0.41	0.09	4.70	2.59E-06	Yes	1.53	1.27E-01	Not Sig. diff.
Rheumatoid Arthritis 2014	24390342	5.82	-0.02	0.05	-0.39	6.93E-01	No	0.07	0.06	1.05	2.96E-01	No	-1.52	1.28E-01	Not Sig. diff.
Childhood maltreatment	33740410	17.17	0.44	0.04	10.20	2.04E-24	Yes	0.36	0.05	6.75	1.53E-11	Yes	1.51	1.31E-01	Not Sig. diff.
Mean corpuscular hemoglobin MCH	27863252	8.33	0.01	0.03	0.22	8.25E-01	No	0.08	0.04	1.91	5.65E-02	No	-1.50	1.34E-01	Not Sig. diff.
Type 2 diabetes dominance deviation model	26961502	4.17	0.06	0.08	0.75	4.55E-01	No	0.21	0.10	1.99	4.69E-02	No	-1.48	1.40E-01	Not Sig. diff.
Eosinophil % of granulocytes	27863252	9.39	-0.09	0.04	-2.67	7.50E-03	No	-0.02	0.05	-0.33	7.41E-01	No	-1.47	1.41E-01	Not Sig. diff.
Weight EU male	23754948	15.23	-0.02	0.05	-0.34	7.33E-01	No	0.06	0.06	1.08	2.81E-01	No	-1.47	1.41E-01	Not Sig. diff.
Total cholesterol	20686565	5.81	0.04	0.04	1.02	3.08E-01	No	0.14	0.06	2.46	1.41E-02	No	-1.47	1.43E-01	Not Sig. diff.
Longest 10% survival	27015805	8.00	-0.05	0.07	-0.76	4.50E-01	No	-0.18	0.08	-2.22	2.64E-02	No	1.45	1.46E-01	Not Sig. diff.
Autism spectrum disorder worldwide	28540026	9.02	0.21	0.08	2.73	6.30E-03	No	0.10	0.08	1.19	2.33E-01	No	1.43	1.53E-01	Not Sig. diff.
Tourette Syndrome	30818990	8.03	0.18	0.07	2.60	9.30E-03	No	0.31	0.08	3.76	2.00E-04	No	-1.42	1.56E-01	Not Sig. diff.
Neutrophil % of granulocytes	27863252	9.45	0.09	0.04	2.60	9.40E-03	No	0.00	0.05	0.08	9.39E-01	No	1.42	1.57E-01	Not Sig. diff.
Mean corpuscular volume	27863252	9.01	0.02	0.03	0.65	5.17E-01	No	0.09	0.04	2.15	3.17E-02	No	-1.41	1.59E-01	Not Sig. diff.
Neutrophil % of white cells	27863252	11.26	0.08	0.04	2.10	3.57E-02	No	0.01	0.04	0.27	7.84E-01	No	1.41	1.59E-01	Not Sig. diff.
Self-reported tiredness	28322280	11.46	0.63	0.06	10.66	1.51E-26	Yes	0.51	0.07	7.10	1.24E-12	Yes	1.39	1.65E-01	Not Sig. diff.
BMI anorexia-free female	31852892	17.13	0.01	0.05	0.29	7.70E-01	No	0.10	0.06	1.85	6.42E-02	No	-1.39	1.65E-01	Not Sig. diff.
Fasting glucose (BMI-adjusted)	22581228	5.43	-0.06	0.07	-0.91	3.65E-01	No	0.04	0.08	0.57	5.69E-01	No	-1.36	1.74E-01	Not Sig. diff.
Insomnia (2017)	28604731	9.42	0.31	0.06	5.33	9.83E-08	Yes	0.43	0.09	4.83	1.35E-06	Yes	-1.36	1.75E-01	Not Sig. diff.
Reproductive behavior: No. Children Ever Born Female	27798627	9.57	0.06	0.07	0.91	3.64E-01	No	0.16	0.09	1.79	7.37E-02	No	-1.35	1.76E-01	Not Sig. diff.
Persistent thin individuals vs severe early onset obesity	30677029	8.85	0.17	0.07	2.41	1.57E-02	No	0.06	0.08	0.73	4.66E-01	No	1.35	1.76E-01	Not Sig. diff.
Atopic dermatitis	26482879	4.22	0.06	0.08	0.77	4.39E-01	No	0.18	0.10	1.84	6.52E-02	No	-1.33	1.84E-01	Not Sig. diff.
Physical activity (objectively-measured) male	34753499	7.72	-0.11	0.07	-1.65	9.88E-02	No	-0.23	0.08	-2.72	6.50E-03	No	1.30	1.94E-01	Not Sig. diff.
Sum eosinophil basophil counts	27863252	8.91	-0.07	0.03	-2.18	2.93E-02	No	0.00	0.05	0.03	9.77E-01	No	-1.29	1.98E-01	Not Sig. diff.
Reproductive behavior: Number of Children Ever Born	27798627	14.06	0.08	0.06	1.46	1.45E-01	No	0.17	0.07	2.35	1.88E-02	No	-1.25	2.11E-01	Not Sig. diff.
Height female 2013	23754948	14.55	-0.11	0.04	-2.63	8.50E-03	No	-0.04	0.05	-0.88	3.78E-01	No	-1.23	2.19E-01	Not Sig. diff.
Tobacco: CPD (cigaretts per day)	20418890	4.22	0.00	0.08	-0.03	9.77E-01	No	0.14	0.11	1.25	2.13E-01	No	-1.22	2.21E-01	Not Sig. diff.
Childhood Intelligence	23358156	5.90	-0.20	0.08	-2.54	1.10E-02	No	-0.37	0.12	-3.01	2.60E-03	No	1.22	2.21E-01	Not Sig. diff.
Rheumatoid Arthritis 2014 EUR	24390342	5.89	-0.02	0.05	-0.31	7.57E-01	No	0.06	0.06	1.01	3.15E-01	No	-1.22	2.23E-01	Not Sig. diff.
Physical activity (objectively-measured)	34753499	13.98	-0.07	0.05	-1.36	1.74E-01	No	-0.16	0.07	-2.44	1.48E-02	No	1.22	2.24E-01	Not Sig. diff.

External Phenotype			GAD						Fear-based disorders					Comparison GAD & fear r_g		
Physical activity (objectively-measured) female	34753499	10.36	-0.06	0.06	-1.02	3.08E-01	No	-0.16	0.07	-2.29	2.21E-02	No	1.22	2.24E-01	Not Sig. diff.	
Concentration of chylomicrons & largest VLDL particles	27005778	4.26	-0.03	0.10	-0.30	7.67E-01	No	0.12	0.11	1.12	2.61E-01	No	-1.21	2.28E-01	Not Sig. diff.	
Subjective well-being	27089181	11.71	-0.58	0.06	-9.22	3.06E-20	Yes	-0.50	0.07	-6.90	5.12E-12	Yes	-1.17	2.41E-01	Not Sig. diff.	
Lean mass (fat free mass) anorexia-free male	30593698	17.06	-0.04	0.04	-1.11	2.67E-01	No	0.02	0.04	0.37	7.11E-01	No	-1.17	2.41E-01	Not Sig. diff.	
Osteoporosis: lumbar spine bone mineral density	26367794	7.10	-0.28	0.12	-2.40	1.64E-02	No	-0.14	0.10	-1.40	1.61E-01	No	-1.16	2.45E-01	Not Sig. diff.	
Autism spectrum disorder EUR	28540026	7.82	0.20	0.08	2.34	1.95E-02	No	0.11	0.09	1.26	2.09E-01	No	1.15	2.51E-01	Not Sig. diff.	
BMI childhood: 3-10 years	26604143	10.30	-0.03	0.06	-0.44	6.61E-01	No	0.05	0.06	0.77	4.40E-01	No	-1.15	2.52E-01	Not Sig. diff.	
Insomnia (female)	28604731	4.99	0.31	0.08	3.89	9.97E-05	Yes	0.44	0.11	4.01	6.10E-05	Yes	-1.12	2.61E-01	Not Sig. diff.	
Low FEV1 versus high FEV1 heavy smokers	28166213	10.56	0.08	0.05	1.51	1.30E-01	No	0.15	0.07	2.29	2.22E-02	No	-1.11	2.66E-01	Not Sig. diff.	
Fasting insulin (BMI-adjusted)	22581228	6.50	0.03	0.08	0.35	7.28E-01	No	0.14	0.10	1.38	1.68E-01	No	-1.10	2.73E-01	Not Sig. diff.	
Hip circumference all ancestries male	25673412	11.84	-0.12	0.05	-2.33	1.97E-02	No	-0.05	0.06	-0.81	4.20E-01	No	-1.09	2.74E-01	Not Sig. diff.	
Birth Length	25281659	7.28	0.07	0.08	0.80	4.22E-01	No	-0.03	0.10	-0.34	7.38E-01	No	1.08	2.80E-01	Not Sig. diff.	
Hip circumference EUR male	25673412	12.13	-0.12	0.05	-2.28	2.24E-02	No	-0.05	0.06	-0.79	4.32E-01	No	-1.08	2.82E-01	Not Sig. diff.	
Cross-disorder: autism subset	23453885	8.04	0.13	0.08	1.71	8.69E-02	No	0.03	0.09	0.38	7.02E-01	No	1.07	2.85E-01	Not Sig. diff.	
Oversleeper	27494321	4.51	0.04	0.09	0.42	6.73E-01	No	0.15	0.11	1.39	1.65E-01	No	-1.01	3.11E-01	Not Sig. diff.	
Eosinophil count	27863252	8.79	-0.07	0.03	-2.09	3.63E-02	No	0.00	0.05	-0.02	9.81E-01	No	-1.00	3.17E-01	Not Sig. diff.	
Granulocyte % of myeloid white cells	27863252	8.68	0.01	0.04	0.19	8.53E-01	No	-0.04	0.05	-0.82	4.15E-01	No	0.99	3.24E-01	Not Sig. diff.	
Sleep Duration	27494321	10.73	-0.11	0.06	-1.74	8.11E-02	No	-0.19	0.08	-2.52	1.17E-02	No	0.97	3.34E-01	Not Sig. diff.	
Chronotype	27494321	16.62	-0.02	0.05	-0.35	7.25E-01	No	-0.07	0.06	-1.17	2.43E-01	No	0.96	3.37E-01	Not Sig. diff.	
Lean mass (fat free) anorexia-free	30593698	20.55	-0.05	0.03	-1.40	1.62E-01	No	-0.01	0.04	-0.34	7.37E-01	No	-0.96	3.39E-01	Not Sig. diff.	
Triglycerides in very large VLDL	27005778	4.29	0.01	0.10	0.08	9.35E-01	No	-0.08	0.11	-0.72	4.70E-01	No	0.95	3.42E-01	Not Sig. diff.	
Monocyte % of white cells	27863252	8.26	0.02	0.04	0.46	6.49E-01	No	0.04	0.05	0.95	3.41E-01	No	-0.95	3.44E-01	Not Sig. diff.	
Insomnia male	28604731	6.98	0.28	0.08	3.31	9.00E-04	No	0.36	0.10	3.61	3.00E-04	No	-0.94	3.49E-01	Not Sig. diff.	
Problematic consequences of drinking	30336701	12.68	0.25	0.06	4.18	2.98E-05	Yes	0.20	0.07	3.03	2.50E-03	No	0.93	3.51E-01	Not Sig. diff.	
Concentration of medium VLDL particles	27005778	4.35	0.05	0.08	0.58	5.61E-01	No	0.14	0.10	1.39	1.66E-01	No	-0.93	3.54E-01	Not Sig. diff.	
Lifecourse bone mineral density: 30-45	29304378	6.32	-0.14	0.08	-1.75	7.95E-02	No	-0.05	0.10	-0.47	6.39E-01	No	-0.91	3.65E-01	Not Sig. diff.	
PTSD symptoms	33510476	19.05	0.55	0.05	11.36	6.69E-30	Yes	0.60	0.05	11.56	6.84E-31	Yes	-0.89	3.75E-01	Not Sig. diff.	
Lean mass (fat free mass) healthy	30593698	20.52	-0.05	0.03	-1.42	1.56E-01	No	-0.02	0.04	-0.37	7.14E-01	No	-0.87	3.82E-01	Not Sig. diff.	
OCD	28761083	7.13	0.34	0.08	4.21	2.57E-05	Yes	0.25	0.10	2.56	1.06E-02	No	0.87	3.86E-01	Not Sig. diff.	
Autism 2015	26998691	8.94	0.14	0.08	1.86	6.35E-02	No	0.06	0.09	0.73	4.68E-01	No	0.86	3.88E-01	Not Sig. diff.	
Birth weight EUR	27680694	14.81	0.02	0.04	0.53	5.93E-01	No	-0.03	0.05	-0.52	6.02E-01	No	0.86	3.91E-01	Not Sig. diff.	
Adiponectin	22479202	4.88	0.08	0.08	1.07	2.87E-01	No	0.00	0.10	0.03	9.77E-01	No	0.86	3.91E-01	Not Sig. diff.	
Lifetime cannabis use	27023175	5.61	0.17	0.09	1.94	5.23E-02	No	0.07	0.12	0.59	5.56E-01	No	0.86	3.91E-01	Not Sig. diff.	
Reaction time	27046643	13.47	0.11	0.06	2.02	4.39E-02	No	0.17	0.06	2.67	7.50E-03	No	-0.85	3.94E-01	Not Sig. diff.	
Lymphocyte % of white cells	27863252	12.68	-0.07	0.04	-1.88	6.05E-02	No	-0.02	0.04	-0.47	6.42E-01	No	-0.83	4.07E-01	Not Sig. diff.	
Red cell distribution width	27863252	8.53	-0.04	0.05	-0.81	4.17E-01	No	-0.01	0.04	-0.33	7.42E-01	No	-0.83	4.08E-01	Not Sig. diff.	
ADHD female	29325848	4.65	0.26	0.10	2.54	1.10E-02	No	0.17	0.12	1.39	1.63E-01	No	0.81	4.15E-01	Not Sig. diff.	

External Phenotype			GAD					Fear-based disorders					Comparison GAD & fear r_g		
HbA1c	20858683	5.33	-0.04	0.08	-0.45	6.51E-01	No	-0.12	0.10	-1.17	2.42E-01	No	0.81	4.19E-01	Not Sig. diff.
Cholesterol esters in large VLDL	27005778	4.94	0.06	0.09	0.67	5.00E-01	No	0.14	0.10	1.46	1.44E-01	No	-0.78	4.34E-01	Not Sig. diff.
Childhood obesity	22484627	9.64	0.02	0.06	0.35	7.29E-01	No	-0.04	0.07	-0.59	5.52E-01	No	0.77	4.44E-01	Not Sig. diff.
Lifecourse bone mineral density: 60+	29304378	9.06	0.01	0.06	0.22	8.28E-01	No	-0.05	0.06	-0.73	4.63E-01	No	0.76	4.47E-01	Not Sig. diff.
Alcohol dependence (mixed ancestry)	30482948	5.45	0.44	0.09	4.85	1.21E-06	Yes	0.38	0.10	3.62	3.00E-04	No	0.73	4.64E-01	Not Sig. diff.
Crohn's disease	26192919	9.03	0.07	0.05	1.42	1.55E-01	No	0.13	0.06	2.02	4.36E-02	No	-0.73	4.68E-01	Not Sig. diff.
Granulocyte count	27863252	8.37	0.05	0.04	1.32	1.87E-01	No	0.02	0.05	0.40	6.90E-01	No	0.73	4.68E-01	Not Sig. diff.
Ratio: visceral-to-subcutaneous adipose tissue vol. female	27918534	4.23	0.17	0.11	1.51	1.30E-01	No	0.07	0.13	0.50	6.20E-01	No	0.72	4.69E-01	Not Sig. diff.
Daily alcohol use	27911795	6.65	-0.04	0.08	-0.48	6.31E-01	No	-0.12	0.09	-1.31	1.90E-01	No	0.72	4.71E-01	Not Sig. diff.
Birth weight	23202124	6.13	0.04	0.08	0.56	5.78E-01	No	-0.02	0.09	-0.26	7.95E-01	No	0.72	4.72E-01	Not Sig. diff.
Fasting insulin (age- & sex-adjusted)	20081858	6.22	-0.03	0.09	-0.31	7.60E-01	No	0.04	0.10	0.38	7.04E-01	No	-0.71	4.81E-01	Not Sig. diff.
Sum basophil neutrophil counts	27863252	8.35	0.06	0.04	1.45	1.47E-01	No	0.02	0.05	0.53	5.94E-01	No	0.70	4.83E-01	Not Sig. diff.
Fasting glucose (age- & sex-adjusted)	20081858	5.07	-0.05	0.07	-0.73	4.63E-01	No	-0.01	0.09	-0.11	9.16E-01	No	-0.68	4.96E-01	Not Sig. diff.
Total body lean mass adulthood	28724990	8.44	-0.11	0.07	-1.61	1.07E-01	No	-0.16	0.09	-1.85	6.46E-02	No	0.65	5.17E-01	Not Sig. diff.
Birth weight all ancestries	27680694	14.89	0.02	0.04	0.42	6.75E-01	No	-0.03	0.05	-0.64	5.25E-01	No	0.65	5.18E-01	Not Sig. diff.
Inflammatory bowel diseases (ulcerative colitis + Crohn's disease)	26192919	10.02	0.10	0.05	2.09	3.69E-02	No	0.15	0.06	2.37	1.79E-02	No	-0.64	5.23E-01	Not Sig. diff.
Neutrophil count	27863252	8.34	0.06	0.04	1.53	1.26E-01	No	0.02	0.05	0.49	6.25E-01	No	0.64	5.23E-01	Not Sig. diff.
Visceral adipose tissue volume adj. for BMI	27918534	5.56	0.10	0.09	1.11	2.68E-01	No	0.04	0.10	0.38	7.07E-01	No	0.62	5.34E-01	Not Sig. diff.
urinary albumin-to-creatinine ratio, non diabetes mellitus	26631737	5.19	-0.11	0.09	-1.22	2.22E-01	No	-0.03	0.12	-0.25	8.04E-01	No	-0.62	5.36E-01	Not Sig. diff.
Total cholesterol in HDL	20686565	5.32	0.04	0.05	0.82	4.11E-01	No	-0.02	0.06	-0.26	7.92E-01	No	0.60	5.46E-01	Not Sig. diff.
Offspring birth weight	29309628	9.84	-0.09	0.06	-1.50	1.32E-01	No	-0.04	0.08	-0.51	6.11E-01	No	-0.60	5.46E-01	Not Sig. diff.
Lean mass (fat free) anorexia-free meta-analysis of males & females	30593698	21.26	-0.06	0.03	-1.62	1.05E-01	No	-0.04	0.04	-0.92	3.56E-01	No	-0.60	5.48E-01	Not Sig. diff.
UACR (urinary albumin-to-creatinine ratio)	26631737	5.28	-0.08	0.09	-0.86	3.88E-01	No	0.00	0.11	-0.04	9.72E-01	No	-0.59	5.53E-01	Not Sig. diff.
Total lipids in medium VLDL	27005778	4.22	0.09	0.09	1.07	2.83E-01	No	0.16	0.10	1.50	1.33E-01	No	-0.59	5.55E-01	Not Sig. diff.
Leptin (BMI-adjusted)	26833098	5.68	0.14	0.08	1.90	5.80E-02	No	0.09	0.11	0.86	3.90E-01	No	0.59	5.56E-01	Not Sig. diff.
Automobile speeding	30643258	25.40	-0.22	0.04	-5.37	7.89E-08	Yes	-0.26	0.05	-5.63	1.81E-08	Yes	0.58	5.59E-01	Not Sig. diff.
Appendicular (arms & legs) lean mass	28724990	9.53	-0.11	0.07	-1.52	1.28E-01	No	-0.15	0.09	-1.73	8.45E-02	No	0.58	5.62E-01	Not Sig. diff.
Height (2018)	30124842	22.59	-0.05	0.02	-2.31	2.11E-02	No	-0.05	0.03	-1.59	1.12E-01	No	-0.57	5.68E-01	Not Sig. diff.
Schizophrenia-vs-bipolar	24280982	8.32	0.06	0.06	0.89	3.73E-01	No	0.11	0.08	1.26	2.09E-01	No	-0.57	5.68E-01	Not Sig. diff.
Sum neutrophil eosinophil counts	27863252	8.55	0.05	0.04	1.12	2.64E-01	No	0.02	0.05	0.37	7.14E-01	No	0.56	5.72E-01	Not Sig. diff.
High light scatter reticulocyte count	27863252	9.36	0.12	0.04	3.04	2.30E-03	No	0.11	0.05	2.27	2.31E-02	No	0.56	5.77E-01	Not Sig. diff.
Tobacco: smoking cessation	20418890	5.73	-0.01	0.09	-0.09	9.29E-01	No	-0.06	0.11	-0.59	5.53E-01	No	0.56	5.79E-01	Not Sig. diff.
White blood cell count	27863252	10.55	0.03	0.04	0.71	4.75E-01	No	0.01	0.05	0.32	7.53E-01	No	0.53	5.94E-01	Not Sig. diff.

External Phenotype			GAD						Fear-based disorders						Comparison GAD & fear r_g		
Visceral adipose tissue volume	27918534	5.49	0.04	0.09	0.44	6.59E-01	No		0.09	0.11	0.83	4.06E-01	No		-0.53	5.96E-01	Not Sig. diff.
Mean corpuscular hemoglobin concentration MCHC	27863252	8.94	-0.02	0.04	-0.61	5.40E-01	No		0.01	0.05	0.11	9.16E-01	No		-0.52	6.04E-01	Not Sig. diff.
Lean mass (fat free mass) healthy meta	30593698	21.24	-0.06	0.03	-1.68	9.36E-02	No		-0.04	0.04	-0.95	3.42E-01	No		-0.50	6.16E-01	Not Sig. diff.
Lymphocyte count	27863252	12.36	-0.03	0.03	-0.94	3.45E-01	No		0.00	0.04	0.10	9.19E-01	No		-0.50	6.17E-01	Not Sig. diff.
Glaucoma (primary angle closure)	27064256	7.08	-0.09	0.07	-1.35	1.78E-01	No		-0.06	0.08	-0.74	4.62E-01	No		-0.49	6.22E-01	Not Sig. diff.
Alcohol dependence	30482948	5.60	0.42	0.10	4.42	9.65E-06	Yes		0.36	0.12	3.12	1.80E-03	No		0.48	6.31E-01	Not Sig. diff.
Subcutaneous adipose tissue volume	27918534	5.99	0.01	0.09	0.11	9.10E-01	No		0.07	0.10	0.66	5.09E-01	No		-0.48	6.31E-01	Not Sig. diff.
Extreme Height	23563607	11.28	-0.04	0.05	-0.74	4.59E-01	No		-0.01	0.06	-0.23	8.21E-01	No		-0.48	6.34E-01	Not Sig. diff.
Total lipids in large VLDL	27005778	4.61	0.08	0.09	0.83	4.08E-01	No		0.12	0.11	1.13	2.58E-01	No		-0.47	6.40E-01	Not Sig. diff.
Hip circumference EUR male BMI-adjusted	25673412	12.23	-0.08	0.05	-1.57	1.15E-01	No		-0.11	0.06	-1.88	6.07E-02	No		0.46	6.45E-01	Not Sig. diff.
Myeloid white cell count	27863252	8.90	0.06	0.04	1.41	1.58E-01	No		0.03	0.05	0.54	5.91E-01	No		0.44	6.59E-01	Not Sig. diff.
Concentration of very large VLDL particles	27005778	4.28	0.06	0.09	0.58	5.63E-01	No		0.11	0.11	0.96	3.39E-01	No		-0.44	6.61E-01	Not Sig. diff.
Concentration of small VLDL particles	27005778	4.28	0.12	0.08	1.44	1.49E-01	No		0.16	0.10	1.64	1.02E-01	No		-0.44	6.62E-01	Not Sig. diff.
Low FEV1 versus average FEV1 never smokers	28166213	6.35	0.08	0.10	0.78	4.36E-01	No		0.11	0.10	1.12	2.61E-01	No		-0.42	6.72E-01	Not Sig. diff.
Reproductive behavior: No. of Children Ever Born Male	27798627	6.67	0.12	0.08	1.40	1.61E-01	No		0.16	0.11	1.54	1.23E-01	No		-0.41	6.85E-01	Not Sig. diff.
Daily alcohol use male	27911795	5.44	-0.04	0.08	-0.51	6.11E-01	No		-0.09	0.11	-0.87	3.85E-01	No		0.40	6.86E-01	Not Sig. diff.
Memory	27046643	8.13	0.19	0.06	3.17	1.50E-03	No		0.17	0.08	2.07	3.84E-02	No		0.38	7.03E-01	Not Sig. diff.
Asthma	20860503	4.71	0.01	0.09	0.15	8.78E-01	No		0.07	0.12	0.61	5.44E-01	No		-0.37	7.10E-01	Not Sig. diff.
Platelet distribution width	27863252	8.24	0.01	0.04	0.12	9.01E-01	No		0.03	0.04	0.58	5.63E-01	No		-0.36	7.18E-01	Not Sig. diff.
GAD symptoms	31906708	13.54	0.62	0.06	11.02	2.99E-28	Yes		0.64	0.06	9.90	4.08E-23	Yes		-0.36	7.22E-01	Not Sig. diff.
Pubertal growth: diff. in height age 8 & adult in females	23449627	4.04	0.14	0.10	1.36	1.73E-01	No		0.09	0.14	0.65	5.13E-01	No		0.35	7.27E-01	Not Sig. diff.
HOMA-IR: insulin resistance (age- & sex-adjusted)	20081858	5.62	0.01	0.10	0.06	9.53E-01	No		0.05	0.11	0.50	6.15E-01	No		-0.34	7.37E-01	Not Sig. diff.
Phospholipids in very large VLDL	27005778	4.02	0.03	0.10	0.33	7.39E-01	No		0.01	0.11	0.09	9.30E-01	No		0.32	7.46E-01	Not Sig. diff.
Mean platelet volume	27863252	7.22	0.00	0.03	-0.10	9.22E-01	No		0.02	0.04	0.53	5.97E-01	No		-0.32	7.47E-01	Not Sig. diff.
Osteoporosis: heel bone mineral density	28869591	10.20	-0.05	0.04	-1.24	2.13E-01	No		-0.04	0.04	-0.91	3.61E-01	No		-0.32	7.52E-01	Not Sig. diff.
Total body lean mass childhood	28743860	8.36	-0.16	0.07	-2.36	1.83E-02	No		-0.13	0.09	-1.43	1.53E-01	No		-0.31	7.55E-01	Not Sig. diff.
Hip circumference all ancestries BMI-adjusted	25673412	17.26	-0.06	0.04	-1.63	1.03E-01	No		-0.07	0.05	-1.51	1.32E-01	No		0.31	7.55E-01	Not Sig. diff.
Total lipids in very large VLDL	27005778	4.87	0.04	0.09	0.41	6.85E-01	No		0.01	0.11	0.09	9.27E-01	No		0.31	7.56E-01	Not Sig. diff.
Fat-free/lean mass healthy meta-analysis (not alcohol/smoking-adjust)	30593698	21.17	-0.06	0.03	-1.67	9.50E-02	No		-0.04	0.04	-0.94	3.48E-01	No		-0.31	7.57E-01	Not Sig. diff.
HOMA-B: beta-cell function (age- & sex-adjusted)	20081858	6.42	0.04	0.09	0.44	6.57E-01	No		0.07	0.10	0.66	5.07E-01	No		-0.31	7.60E-01	Not Sig. diff.
HbA1c EUR	28898252	8.39	-0.05	0.05	-0.96	3.37E-01	No		-0.02	0.06	-0.39	6.98E-01	No		-0.30	7.62E-01	Not Sig. diff.
Monocyte count	27863252	8.51	0.03	0.04	0.93	3.50E-01	No		0.05	0.05	1.06	2.91E-01	No		-0.29	7.74E-01	Not Sig. diff.
Lean mass (fat free mass) healthy female	30593698	18.69	-0.08	0.04	-1.93	5.39E-02	No		-0.09	0.05	-1.78	7.48E-02	No		0.28	7.80E-01	Not Sig. diff.
Height male 2013	23754948	15.05	-0.02	0.04	-0.53	5.94E-01	No		0.00	0.05	-0.06	9.56E-01	No		-0.26	7.92E-01	Not Sig. diff.
Hip circumference EUR BMI-adjusted	25673412	17.12	-0.06	0.04	-1.57	1.17E-01	No		-0.07	0.05	-1.47	1.42E-01	No		0.26	7.97E-01	Not Sig. diff.
Coronary Artery Disease (EUR, South & East Asian)	26343387	8.45	0.06	0.05	1.19	2.33E-01	No		0.07	0.07	1.00	3.18E-01	No		-0.25	8.05E-01	Not Sig. diff.

External Phenotype			GAD						Fear-based disorders						Comparison GAD & fear r_g		
Total lipids in chylomicrons & largest VLDL particles	27005778	4.35	0.01	0.10	0.08	9.40E-01	No		0.05	0.12	0.46	6.44E-01	No		-0.25	8.06E-01	Not Sig. diff.
Leptin (not BMI-adjusted)	26833098	6.61	0.10	0.07	1.43	1.54E-01	No		0.13	0.09	1.43	1.53E-01	No		-0.24	8.08E-01	Not Sig. diff.
Pubertal growth: height at age 10 in girls & 12 in boys	23449627	9.11	-0.17	0.06	-2.63	8.50E-03	No		-0.17	0.09	-1.97	4.85E-02	No		0.23	8.15E-01	Not Sig. diff.
Visceral adipose tissue volume adj. for BMI male	27918534	4.07	0.03	0.10	0.24	8.10E-01	No		0.07	0.13	0.56	5.76E-01	No		-0.23	8.17E-01	Not Sig. diff.
number of sexual partners (risky behaviour)	30643258	26.78	0.09	0.04	2.19	2.84E-02	No		0.10	0.05	2.25	2.46E-02	No		-0.23	8.17E-01	Not Sig. diff.
Hip circumference all ancestries male BMI-adjusted	25673412	12.34	-0.09	0.05	-1.68	9.20E-02	No		-0.11	0.06	-1.90	5.79E-02	No		0.23	8.18E-01	Not Sig. diff.
Sitting height ratio (SHR) Sitting/total height EUR male	25865494	4.91	-0.05	0.10	-0.53	5.93E-01	No		-0.07	0.11	-0.63	5.32E-01	No		0.23	8.20E-01	Not Sig. diff.
Total lipids in small VLDL	27005778	4.20	0.12	0.08	1.52	1.28E-01	No		0.16	0.10	1.60	1.10E-01	No		-0.22	8.25E-01	Not Sig. diff.
Subcutaneous adipose tissue volume female	27918534	4.44	0.15	0.10	1.44	1.49E-01	No		0.14	0.13	1.11	2.69E-01	No		0.22	8.29E-01	Not Sig. diff.
Sitting height ratio (SHR) Sitting/total height EUR	25865494	7.75	-0.05	0.08	-0.65	5.17E-01	No		-0.06	0.09	-0.72	4.72E-01	No		0.21	8.31E-01	Not Sig. diff.
Extreme Waist-Hip Ratio	23563607	5.94	0.03	0.08	0.41	6.79E-01	No		0.05	0.10	0.50	6.19E-01	No		-0.21	8.32E-01	Not Sig. diff.
Basophil % of granulocytes	27863252	6.67	0.00	0.05	0.08	9.40E-01	No		0.01	0.06	0.16	8.75E-01	No		-0.20	8.40E-01	Not Sig. diff.
Height 2014	25282103	19.82	-0.06	0.03	-2.28	2.24E-02	No		-0.06	0.04	-1.51	1.30E-01	No		-0.20	8.45E-01	Not Sig. diff.
Infant Head Circumference	22504419	5.09	-0.28	0.10	-2.89	3.80E-03	No		-0.24	0.13	-1.88	6.08E-02	No		-0.19	8.49E-01	Not Sig. diff.
Reticulocyte fraction of red cells	27863252	8.85	0.10	0.04	2.71	6.70E-03	No		0.11	0.04	2.60	9.40E-03	No		-0.18	8.58E-01	Not Sig. diff.
Extraversion (item response theory)	26362575	6.65	-0.08	0.08	-0.91	3.65E-01	No		-0.10	0.09	-1.11	2.68E-01	No		0.18	8.61E-01	Not Sig. diff.
Red blood cell count	27863252	9.93	0.01	0.04	0.18	8.56E-01	No		0.02	0.04	0.51	6.13E-01	No		-0.17	8.62E-01	Not Sig. diff.
Pubertal growth: height age 12 in boys	23449627	4.73	-0.08	0.09	-0.95	3.41E-01	No		-0.10	0.11	-0.96	3.39E-01	No		0.17	8.65E-01	Not Sig. diff.
Cholesterol esters in medium VLDL	27005778	4.34	0.13	0.08	1.57	1.16E-01	No		0.14	0.10	1.46	1.44E-01	No		-0.15	8.82E-01	Not Sig. diff.
Ulcerative Colitis	26192919	8.29	0.10	0.06	1.72	8.60E-02	No		0.13	0.07	1.89	5.89E-02	No		-0.15	8.82E-01	Not Sig. diff.
Hip circumference all ancestries female BMI-adjusted	25673412	13.97	-0.05	0.04	-1.24	2.14E-01	No		-0.05	0.05	-0.82	4.10E-01	No		0.14	8.87E-01	Not Sig. diff.
Lifetime anxiety (male)	31748690	5.27	0.98	0.07	14.41	4.43E-47	Yes		0.99	0.10	9.60	8.16E-22	Yes		-0.14	8.89E-01	Not Sig. diff.
Free cholesterol in large VLDL	27005778	4.45	0.05	0.09	0.53	5.94E-01	No		0.04	0.10	0.37	7.12E-01	No		0.14	8.91E-01	Not Sig. diff.
High light scatter % of red cells	27863252	9.56	0.11	0.04	2.97	2.90E-03	No		0.10	0.05	2.26	2.38E-02	No		0.14	8.93E-01	Not Sig. diff.
Ratio of visceral-to-subcutaneous adipose tissue vol.	27918534	5.77	0.13	0.09	1.48	1.39E-01	No		0.13	0.10	1.26	2.09E-01	No		0.14	8.92E-01	Not Sig. diff.
Basophil % of white cells	27863252	6.08	0.03	0.05	0.59	5.53E-01	No		0.02	0.06	0.34	7.37E-01	No		0.13	8.93E-01	Not Sig. diff.
PTSD	31594949	5.83	0.65	0.09	7.23	4.72E-13	Yes		0.67	0.12	5.72	1.06E-08	Yes		-0.13	8.96E-01	Not Sig. diff.
General risk tolerance (UKB + 10 replication cohorts)	30643258	22.70	0.04	0.05	0.95	3.43E-01	No		0.05	0.05	0.85	3.94E-01	No		-0.13	8.99E-01	Not Sig. diff.
Immature fraction of reticulocytes	27863252	8.29	0.11	0.04	2.60	9.40E-03	No		0.09	0.05	1.62	1.06E-01	No		0.12	9.05E-01	Not Sig. diff.
Major depressive disorder (CONVERGE)	26176920	4.54	-0.29	0.09	-3.45	6.00E-04	No		-0.28	0.11	-2.48	1.33E-02	No		-0.12	9.05E-01	Not Sig. diff.
Common epilepsies	25087078	4.78	0.08	0.10	0.84	3.99E-01	No		0.08	0.11	0.70	4.87E-01	No		-0.11	9.14E-01	Not Sig. diff.
Basophil count	27863252	7.27	0.04	0.05	0.76	4.47E-01	No		0.03	0.06	0.53	6.00E-01	No		-0.09	9.29E-01	Not Sig. diff.
Pubertal growth: diff. in height age 8 & adult males & females	23449627	5.97	0.08	0.08	0.89	3.73E-01	No		0.08	0.10	0.73	4.66E-01	No		-0.08	9.34E-01	Not Sig. diff.
Pubertal growth: height age 10 in girls	23449627	7.54	-0.20	0.08	-2.41	1.60E-02	No		-0.21	0.10	-2.16	3.10E-02	No		0.08	9.35E-01	Not Sig. diff.
Ratio of visceral-to-subcutaneous adipose tissue vol. adj. for BMI	27918534	5.40	0.12	0.09	1.34	1.79E-01	No		0.12	0.11	1.11	2.69E-01	No		0.08	9.38E-01	Not Sig. diff.

External Phenotype			GAD						Fear-based disorders						Comparison GAD & fear r_g		
Total cholesterol in large VLDL	27005778	4.15	0.04	0.09	0.45	6.53E-01	No		0.03	0.11	0.32	7.52E-01	No		-0.07	9.41E-01	Not Sig. diff.
Cross 12 psychiatric disorders	36471075	25.80	0.71	0.04	19.84	1.30E-87	Yes		0.71	0.07	10.72	8.63E-27	Yes		-0.06	9.49E-01	Not Sig. diff.
Smoking initiation	20418890	12.27	0.26	0.06	4.27	1.94E-05	Yes		0.24	0.08	2.97	3.00E-03	No		0.06	9.52E-01	Not Sig. diff.
Lean mass (fat free mass) anorexia-free female	30593698	18.59	-0.07	0.04	-1.83	6.71E-02	No		-0.09	0.05	-1.76	7.79E-02	No		-0.05	9.57E-01	Not Sig. diff.
Primary sclerosing cholangitis	27992413	4.48	-0.09	0.08	-1.03	3.03E-01	No		-0.08	0.09	-0.86	3.88E-01	No		0.05	9.63E-01	Not Sig. diff.
Triglycerides in very small VLDL	27005778	4.10	0.15	0.08	1.86	6.32E-02	No		0.16	0.11	1.49	1.37E-01	No		-0.03	9.73E-01	Not Sig. diff.
Sitting height ratio EUR female	25865494	4.44	-0.05	0.11	-0.52	6.06E-01	No		-0.06	0.12	-0.49	6.24E-01	No		-0.02	9.87E-01	Not Sig. diff.
Reticulocyte count	27863252	8.84	0.10	0.04	2.82	4.80E-03	No		0.11	0.04	2.61	9.10E-03	No		-0.01	9.91E-01	Not Sig. diff.
Persistent thin individuals versus controls	30677029	4.33	0.01	0.10	0.06	9.54E-01	No		0.01	0.12	0.05	9.64E-01	No		0.00	9.97E-01	Not Sig. diff.
Hip circumference EUR female BMI-adjusted	25673412	13.80	-0.05	0.04	-1.15	2.50E-01	No		-0.04	0.06	-0.80	4.23E-01	No		0.00	9.99E-01	Not Sig. diff.

Supplementary Table 10: SNP-based heritability (h^2_{SNP}) estimates of GAD and fear-based disorders screened using detailed versus brief diagnostic measures.

Anxiety disorder	Diagnostic measure	Cases	Controls	h^2	se	P
GAD	Detailed	10826	7986	0.037	0.037	0.152
	Brief	15257	9160	0.107	0.031	4.59E-05
Fear	Detailed	12261	6621	0.08	0.036	0.0116
	Brief	5523	10156	0.049	0.031	0.0481

h^2_{SNP} estimated in GCTA-GREML on the liability-scale.

Appendix B. Supplementary materials for Chapter 3

Table of contents

Supplementary methods	212
<i>Sample and measures.....</i>	<i>212</i>
Supplementary results	213
<i>Sensitivity analyses</i>	<i>213</i>
Supplementary Figure 1. Path diagram representing results from genomic multiple regression analyses of independently contributing traits to the genetic component of childhood maltreatment.	214
<i>Self-reported lifetime trauma</i>	<i>214</i>
Supplementary references	215
Supplementary Tables	216
<i>Supplementary tables for reported childhood maltreatment.....</i>	<i>216</i>
Supplementary Table 1: LD Score genetic correlation (rg) results for all phenotypes with reported childhood maltreatment.	216
Supplementary Table 2: LD Score genetic correlation (rg) results for genetically correlated phenotypes with reported childhood maltreatment rg > 0.25 and rg Z statistic ≥ 5, a heritability Z statistics ≥ 5, mean chi-square > 1.02.....	222
Supplementary table 3: A-K Genomic SEM multiple regression models where each psychological, social and health trait is introduced in iterative stages, in order of the most highly genetically correlated trait with reported childhood maltreatment.	224
Supplementary table 4: A-G Genomic SEM multiple regression models where each Psychiatric disorder is introduced in iterative stages, in order of the most highly genetically correlated trait with reported childhood maltreatment.	228
Supplementary table 5: Genomic SEM genomic multiple regression analyses of independently contributing traits selected based on results shown in figure 2A/ supplementary table 3K and figure 2B/ supplementary table 4G.	230
Supplementary table 6 of sensitivity analysis: Genomic SEM genomic multiple regression analyses of independently contributing traits (in figure 3/ supplementary table 5) and two additional traits that broadly capture the shared genetic associations observed across	

the psychiatric disorders (as captured by ADHD) and health traits (as captured by self-reported tiredness).	232
<i>Supplementary tables for self-reported lifetime trauma</i>	233
Link to access to Supplementary Tables 7-8: LD Score genetic correlation (rg) results between self-reported trauma and 576 phenotypes.	233
Supplementary table 9: A-K Genomic SEM multivariable models where each psychological, social and health trait is introduced in iterative stages, in order of the most highly genetically correlated trait with self-reported lifetime trauma.	233
Supplementary table 10: A-G Genomic SEM multivariable models where each psychiatric disorder is introduced in iterative stages, in order of the most highly genetically correlated trait with self-reported lifetime trauma.	238
Supplementary table 11: Genomic SEM genomic multiple regression analyses of independently contributing traits to self-reported lifetime trauma, selected based on results shown in supplementary table 9K and supplementary table 10G.	240
Link access to Supplementary Table 12: GWAS summary statistics data availability information.	240

Supplementary methods

Sample and measures

Self-reported lifetime trauma in the UK Biobank

We also analysed published GWAS summary statistics of a lifetime trauma phenotype that more broadly captures trauma occurring in both childhood and adulthood in the UK Biobank^{1,2}. The lifetime trauma GWAS included 35,269 individuals who retrospectively reported trauma and 63,451 individuals who reported no traumas¹. Participants were predominantly of European ancestry and answered the online mental health questionnaire³. Participants were considered reporters of trauma if they self-reported at least two types of lifetime trauma and were excluded from analyses if they only reported one. These trauma types were selected based on their association in the UK Biobank with major depressive disorder (MDD) and needed to show an odds ratio >2.5 for selection. These pre-selected self-reported traumas occurred in childhood (i.e., felt loved less than often, felt hated by a family member, sexually abused), adulthood (i.e., partner physical violence, belittlement, and sexual interference) and one-lifetime item (i.e., victim of sexual assault). Individuals were considered non-reporters of trauma if they did not self-report any pre-selected traumas. Further methodological details can be found in the original publication¹. The h^2_{SNP} of this self-reported lifetime trauma definition using LDSC was 0.18 (s.e = 0.01) on the liability scale¹. We also calculated the genetic correlation between reported childhood maltreatment and retrospectively reported lifetime

trauma, and tested if the genetic correlation was significantly different from 1 (calculated using the chi-squared distribution function and $[(|r_g|-1)/se]^2$ in R v. 4.1.1)^{1,4}.

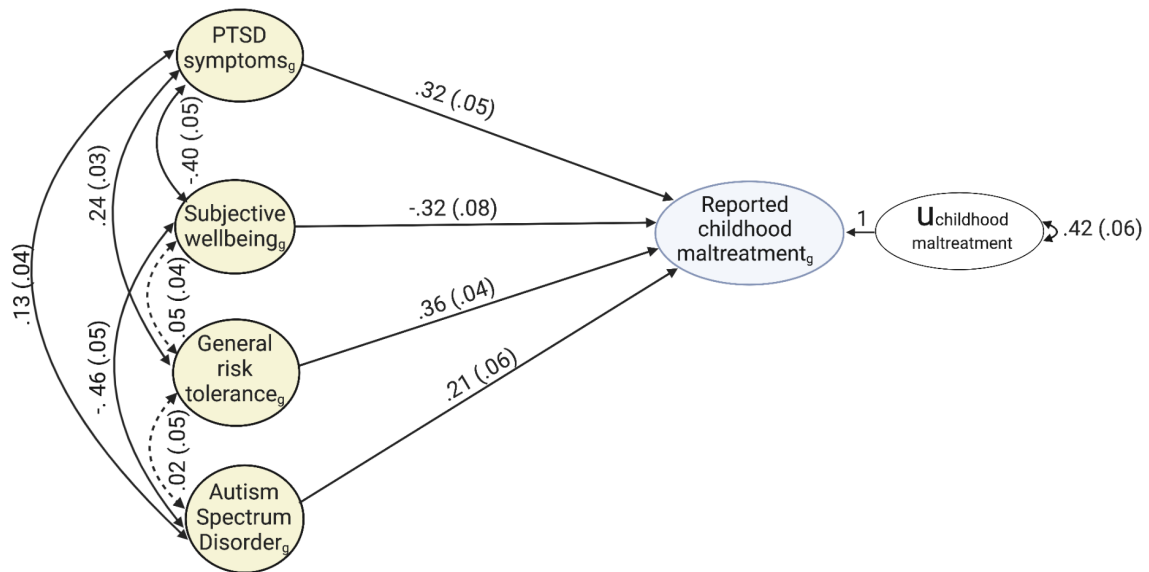
Supplementary results

Sensitivity analyses

Model of independently contributing health and psychiatric traits

All four traits had significant conditional associations with reported childhood maltreatment independent of the other traits; subjective well-being ($b_g = -0.32 \pm 0.08$; $P = 1.22 \times 10^{-4}$), general risk tolerance ($b_g = 0.36 \pm 0.04$; $P = 9.15 \times 10^{-21}$), PTSD symptoms ($b_g = 0.32 \pm 0.05$; $P = 7.98 \times 10^{-10}$) and ASD ($b_g = 0.21 \pm 0.06$; $P = 5.80 \times 10^{-4}$). The residual genetic variance of reported childhood maltreatment was 0.42 ± 0.06 ($P = 1.24 \times 10^{-12}$; **Supplementary Figure 1**). Direct estimation of the overlap in the residual genetic variance of childhood maltreatment that remained in the psychiatric disorder model and the health trait model was not possible. To address this, we specified a genomic multiple regression model that included the four independently contributing traits and two additional traits. These two additional traits broadly capture the shared genetic associations observed across the psychiatric disorders (ADHD) and health and behavioural traits (self-reported tiredness). In this sensitivity model, the residual genetic variance of childhood maltreatment was 0.40 ± 0.06 ($P = 2.11 \times 10^{-12}$), explaining 60% of the h^2_{SNP} of childhood maltreatment (**Supplementary Table 6**). The addition of these traits made only a ~2% difference in the residual genetic variance of childhood maltreatment. Thus, if all psychiatric, health, and behavioural traits could be accounted for in one model, it is unlikely this would greatly reduce the h^2_{SNP} of childhood maltreatment. This

means that 40–42% of the h^2_{SNP} of childhood maltreatment remains unexplained by the genetic components of traits included in our models.



Supplementary Figure 1. Path diagram representing results from genomic multiple regression analyses of independently contributing traits to the genetic component of childhood maltreatment.

Autism spectrum disorder (ASD), post-traumatic stress disorder (PTSD) symptoms, general risk tolerance, and subjective well-being were selected based on results shown in Figures 2A and 2B. A genomic multiple regression was specified in Genomic SEM. We regressed the genetic component of childhood maltreatment on the genetic components of four independently genetically associated traits. We used a weighted least-squares estimator. Single-headed arrows are conditional genetic associations ($bg \pm SE$) between the explanatory variables and childhood maltreatment independent of the genetic influences of the other explanatory variables. A solid line indicates that the conditional genetic association is significant. Double-headed arrows connecting explanatory variables represent genetic correlations ($rg \pm SE$). Double-headed arrows connecting the genetic component of childhood maltreatment to itself is the residual genetic variance of childhood maltreatment ($umaltreatment \pm SE$) that is unexplained by the genetic influence of the explanatory variables.

Self-reported lifetime trauma

The genetic correlation between reported childhood maltreatment and retrospectively reported lifetime trauma was high ($r_g = 0.884 \pm 0.04$; $P = 1.00 \times 10^{-122}$), but significantly different from 1 ($P = 0.0211$). Results for lifetime trauma are in **Supplementary Tables 7–11**. We identified 18 traits with significant genetic correlations ($r_g > \pm 0.25$) with retrospective lifetime trauma after Bonferroni correction for multiple testing. Most of these traits were also top genetic correlations with childhood maltreatment and categorised into the same two groups of psychiatric disorders and health and behavioural traits. In genomic multiple regression analyses, the residual genetic variance of lifetime trauma was 0.25 ± 0.13 ($P = 0.059$) and 0.47 ± 0.07 ($P = 8.00 \times 10^{-11}$) after taking into account the genetic associations with health and behavioural

traits and psychiatric disorders, respectively. In these models, independent genetic associations with lifetime trauma were identified for ASD ($b_g = 0.36 \pm 0.10$, $P = 2.28 \times 10^{-4}$), PTSD symptoms ($b_g = 0.28 \pm 0.09$; $P = 1.13 \times 10^{-3}$), subjective well-being ($b_g = -0.67 \pm 0.33$; $P = 0.048$) and general risk tolerance ($b_g = 0.42 \pm 0.10$; $P = 4.29 \times 10^{-5}$). When these four traits were included in one genomic multiple regression model, the residual genetic variance of lifetime trauma was 0.33 ± 0.07 ($P = 1.46 \times 10^{-6}$). Given the consistency in findings between retrospectively reported lifetime trauma and reported childhood maltreatment, this suggests that the inclusion of a small prospective sample in the reported childhood maltreatment meta-analysis did not impact our results. Our findings likely reflect associations with retrospectively reported and not prospectively reported childhood maltreatment, given the lack of power of the prospective GWAS ($h^2_{\text{SNP}} Z \text{ score} = 1.39$)⁵.

Supplementary references

1. Coleman, J. R. I. *et al.* Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Mol. Psychiatry* (2020) doi:10.1038/s41380-019-0546-6.
2. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
3. Davis, K. A. S. *et al.* Mental health in UK Biobank - development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open* **6**, e18 (2020).
4. Team, R. C. R Core Team (2019) R Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. *References-Scientific Research Publishing*.
5. Warrier, V. *et al.* Gene–environment correlations and causal effects of childhood maltreatment on physical and mental health: a genetically informed approach. *The Lancet Psychiatry* **8**, 373–386 (2021).

Supplementary Tables

Supplementary tables for reported childhood maltreatment

Supplementary Table 1: LDscore genetic correlation (r_g) results for all phenotypes with reported childhood maltreatment.

Traits with a r_g |Z| statistic ≥ 5 are in bold. For brevity, results are shown here for traits with r_g |Z| statistic ≥ 2 . Full results for all traits can be accessed via the following link: <https://tinyurl.com/Chapter3AppendixBTables>.

GWAS		r_g with childhood maltreatment				h^2_{SNP} (Observed scale)		
Phenotype	PMID	r_g	SE	Z	P	h^2_{SNP}	SE	Z
number of sexual partners (risky behaviour)	30643258	0.48	0.02	20.04	2.44E-89	0.10	0.00	29.24
Major depressive disorder (PGC2 + 23andme)	29700475	0.56	0.03	19.98	8.50E-89	0.06	0.00	23.08
Major depressive disorder (PGC2 + 23andme, no UKBiobank)	29700475	0.55	0.03	19.14	1.13E-81	0.06	0.00	22.28
PTSD Avoidance symptoms	33510476	0.58	0.03	18.72	3.51E-78	0.06	0.00	16.74
Total PTSD symptoms (MVP)	33510476	0.56	0.03	18.30	7.65E-75	0.07	0.00	18.31
Major depressive disorder (PGC2, no 23andme, no UKBiobank)	29700475	0.63	0.04	16.91	3.82E-64	0.07	0.00	16.11
PTSD Hyperarousal symptoms	33510476	0.54	0.03	16.64	3.42E-62	0.07	0.00	18.21
PTSD Re-experiencing symptoms	33510476	0.52	0.03	16.62	5.08E-62	0.07	0.00	17.57
General risk tolerance (UKB + 10 replication cohorts)	30643258	0.42	0.03	15.56	1.46E-54	0.05	0.00	22.52
Depressive symptoms	27089181	0.66	0.04	15.15	7.95E-52	0.05	0.00	13.06
ADHD	30478444	0.56	0.04	15.09	1.96E-51	0.24	0.01	16.38
Schizophrenia PGC 2018 Mixed ancestry	31740837	0.35	0.02	15.07	2.68E-51	0.34	0.01	30.45
ADHD (Mixed ancestry)	30478444	0.56	0.04	14.82	1.08E-49	0.36	0.02	16.51
Schizophrenia PGC 2014	25056061	0.36	0.03	13.64	2.35E-42	0.45	0.02	28.52
Schizophrenia PGC 2014 European	25056061	0.35	0.03	13.46	2.65E-41	0.47	0.02	27.87
ADHD male	29325848	0.58	0.05	12.88	5.79E-38	0.25	0.02	11.90
Schizophrenia PGC2 European + CLOZUK	29483656	0.36	0.03	12.20	2.95E-34	0.42	0.02	27.27
Self-rated health	27864402	-0.48	0.04	-12.09	1.22E-33	0.10	0.01	16.74
ever smoker (UKB + TAG)	30643258	0.32	0.03	11.95	6.59E-33	0.09	0.00	28.53
Neuroticism 300K meta-analysis female	30593698	0.34	0.03	11.63	2.86E-31	0.11	0.01	18.77
Reproductive behavior: Age First Birth	27798627	-0.40	0.03	-11.56	6.76E-31	0.06	0.00	17.81
Lifetime probable anxiety diagnosis	31748690	0.45	0.04	11.52	1.01E-30	0.10	0.01	15.10
Heavy versus never smoker in UKBiobank	28166213	0.45	0.04	11.36	6.69E-30	0.26	0.02	15.27
Self-reported tiredness	28322280	0.52	0.05	11.06	2.04E-28	0.07	0.01	11.57
Subjective well-being: survey questions on life satisfaction, positive affect, or happiness	27089181	-0.52	0.05	-10.73	7.10E-27	0.02	0.00	12.45
Reproductive behavior: Age First Birth Female	27798627	-0.41	0.04	-10.49	9.27E-26	0.05	0.00	14.53

GWAS		r_g with childhood maltreatment				h^2_{SNP} (Observed scale)		
Phenotype	PMID	r_g	SE	Z	P	h^2_{SNP}	SE	Z
Lifetime probable anxiety diagnosis female	31748690	0.50	0.05	10.41	2.33E-25	0.14	0.01	11.63
Cross-psychiatric disorders (ADHD, autism, bipolar disorder, and schizophrenia)	23453885	0.40	0.04	10.10	5.56E-24	0.18	0.01	16.37
Neuroticism 300K female.	30593698	0.34	0.04	9.81	9.81E-23	0.12	0.01	16.71
Insomnia (2019)	30804565	0.32	0.03	9.64	5.56E-22	0.05	0.00	21.86
PTSD PGC2 case control	31594949	0.76	0.08	9.55	1.28E-21	0.02	0.00	5.72
Autism spectrum disorder (PGC + iPSYCH)	30804558	0.41	0.04	9.48	2.51E-21	0.20	0.02	13.19
BMI meta-analysis of males and females	31852892	0.19	0.02	9.46	3.15E-21	0.23	0.01	28.64
BMI	31852892	0.19	0.02	9.35	8.71E-21	0.23	0.01	28.73
Neuroticism.	27089181	0.39	0.04	9.26	2.08E-20	0.09	0.01	12.00
Social deprivation, as measured using Townsend scores	27818178	0.50	0.06	8.92	4.83E-19	0.04	0.00	8.42
Household income, one person per household	27818178	-0.43	0.05	-8.81	1.30E-18	0.07	0.01	10.38
ADHD female	29325848	0.65	0.07	8.63	6.36E-18	0.14	0.03	4.87
Neuroticism 300K meta-analysis male	30593698	0.33	0.04	8.54	1.32E-17	0.11	0.01	12.26
BMI male	31852892	0.20	0.02	8.47	2.37E-17	0.25	0.01	25.06
PTSD MVP case control	33510476	0.42	0.05	8.46	2.68E-17	0.03	0.00	11.47
Household income	27818178	-0.41	0.05	-8.43	3.41E-17	0.06	0.01	10.54
Schizophrenia+bipolar	24280982	0.30	0.04	8.31	9.33E-17	0.38	0.02	16.73
Body mass index healthy meta-analysis (not alcohol/smoking-adjust)	30593698	0.20	0.02	8.18	2.85E-16	0.22	0.01	24.46
BMI healthy meta-analysis of males and females	31852892	0.21	0.03	8.15	3.78E-16	0.22	0.01	24.10
BMI healthy	31852892	0.20	0.03	8.13	4.23E-16	0.21	0.01	23.92
Neuroticism 150K healthy meta-analysis.	30593698	0.31	0.04	8.09	5.84E-16	0.10	0.01	12.32
Educational Attainment: EduYears	30038396	-0.19	0.02	-8.07	7.04E-16	0.11	0.00	36.26
BMI anorexia-free meta-analysis of males and females	31852892	0.20	0.03	8.05	7.96E-16	0.22	0.01	24.22
BMI anorexia-free	31852892	0.20	0.03	8.01	1.17E-15	0.21	0.01	23.79
BMI Locke & UKB (updated)	30124842	0.17	0.02	7.96	1.73E-15	0.17	0.01	28.70
Fat mass meta-analysis of males and females	31852892	0.16	0.02	7.88	3.39E-15	0.22	0.01	30.00
Fat mass	31852892	0.16	0.02	7.81	5.68E-15	0.22	0.01	29.96
Bipolar disorder	31043756	0.23	0.03	7.63	2.39E-14	0.35	0.02	20.75
Educational Attainment: EduYears	27225129	-0.21	0.03	-7.44	9.94E-14	0.17	0.01	27.56
BMI female	31852892	0.18	0.02	7.39	1.50E-13	0.23	0.01	26.20
Intelligence (Cognitive ability)	29942086	-0.20	0.03	-7.33	2.36E-13	0.19	0.01	28.06
Reproductive behavior: Age First Birth Male	27798627	-0.36	0.05	-7.19	6.48E-13	0.11	0.01	10.75
BMI anorexia-free male	31852892	0.21	0.03	7.13	1.00E-12	0.24	0.01	20.48
BMI healthy male	31852892	0.21	0.03	7.12	1.07E-12	0.24	0.01	20.50
Major depressive disorder (PGC1)	22472876	0.55	0.08	6.98	2.93E-12	0.16	0.03	5.95
Tobacco: smoking initiation (ever versus never been a regular smoker)	20418890	0.34	0.05	6.95	3.74E-12	0.08	0.01	11.80
Body fat percentage meta-analysis of males and females	30593698	0.16	0.02	6.94	3.80E-12	0.22	0.01	31.63
Body fat percentage	30593698	0.16	0.02	6.94	3.85E-12	0.21	0.01	31.96

GWAS		r_g with childhood maltreatment				h^2_{SNP} (Observed scale)		
Phenotype	PMID	r_g	SE	Z	P	h^2_{SNP}	SE	Z
Educational Attainment: EduYears Men	27225129	-0.24	0.03	-6.90	5.12E-12	0.20	0.01	21.86
Fat mass male	31852892	0.17	0.02	6.72	1.79E-11	0.23	0.01	25.71
Insomnia	28604731	0.33	0.05	6.72	1.85E-11	0.05	0.00	9.82
Fat mass healthy meta-analysis of males and females	31852892	0.16	0.03	6.47	9.86E-11	0.21	0.01	25.49
Fat mass healthy	31852892	0.16	0.03	6.45	1.13E-10	0.21	0.01	25.53
Fat mass anorexia-free meta-analysis of males and females	31852892	0.16	0.03	6.40	1.58E-10	0.21	0.01	25.60
Father's age at death	27015805	-0.29	0.05	-6.36	2.00E-10	0.02	0.00	11.50
Fat mass anorexia-free	31852892	0.16	0.03	6.36	2.05E-10	0.21	0.01	25.34
Fat mass healthy meta-analysis (not alcohol/smoking-adjust)	31852892	0.16	0.03	6.35	2.09E-10	0.21	0.01	25.86
BMI (EUR Locke)	25673413	0.15	0.02	6.34	2.25E-10	0.14	0.01	18.13
BMI (MIX Locke)	25673413	0.15	0.02	6.31	2.79E-10	0.14	0.01	18.41
Mother attained age	27015805	0.30	0.05	6.28	3.47E-10	0.02	0.00	10.75
Parents attained age	27015805	0.25	0.04	6.28	3.49E-10	0.03	0.00	13.74
Fat mass female	31852892	0.15	0.02	6.24	4.33E-10	0.23	0.01	27.89
Alcohol dependence all ancestries	30482948	0.54	0.09	6.23	4.72E-10	0.05	0.01	5.42
Parents' age at death	27015805	-0.28	0.05	-6.22	5.10E-10	0.02	0.00	11.24
Neuroticism 150K healthy male.	30593698	0.32	0.05	6.17	6.79E-10	0.10	0.01	7.51
Body fat percentage female	30593698	0.16	0.03	6.05	1.44E-09	0.23	0.01	29.56
Neuroticism 150K healthy female.	30593698	0.29	0.05	6.04	1.58E-09	0.10	0.01	10.37
Mother's age at death	27015805	-0.30	0.05	-5.96	2.45E-09	0.01	0.00	9.13
BMI healthy female	31852892	0.19	0.03	5.95	2.61E-09	0.20	0.01	17.57
Alcohol dependence European	30482948	0.53	0.09	5.93	3.10E-09	0.05	0.01	5.41
Binary variable for college completion	23722424	-0.24	0.04	-5.78	7.34E-09	0.10	0.01	13.58
BMI anorexia-free female	31852892	0.19	0.03	5.77	7.81E-09	0.20	0.01	17.42
Intelligence	28530673	-0.20	0.04	-5.76	8.49E-09	0.19	0.01	18.25
Body fat percentage male	30593698	0.15	0.03	5.75	8.89E-09	0.24	0.01	26.74
Father attained age	27015805	0.25	0.04	5.74	9.50E-09	0.03	0.00	12.83
Attainment of a college or a university degree	27046643	-0.19	0.03	-5.65	1.60E-08	0.15	0.01	19.42
Reproductive behavior: Number of Children Ever Born	27798627	0.25	0.04	5.58	2.46E-08	0.02	0.00	14.00
Obesity class 1	23563607	0.17	0.03	5.53	3.14E-08	0.17	0.01	19.53
Body fat percentage healthy	30593698	0.15	0.03	5.48	4.14E-08	0.20	0.01	26.08
Fat mass anorexia-free male	31852892	0.16	0.03	5.48	4.29E-08	0.22	0.01	20.98
Fat mass healthy male	31852892	0.16	0.03	5.48	4.38E-08	0.22	0.01	21.01
Body fat percentage healthy meta-analysis of males and females	30593698	0.15	0.03	5.45	5.12E-08	0.21	0.01	26.15
Waist circumference all ancestries	25673412	0.14	0.03	5.44	5.34E-08	0.14	0.01	16.26
Waist circumference European	25673412	0.14	0.03	5.42	5.88E-08	0.14	0.01	16.29
Body fat percentage anorexia-free	30593698	0.15	0.03	5.38	7.64E-08	0.20	0.01	25.90
Body fat percentage anorexia-free meta-analysis of males and females	30593698	0.15	0.03	5.36	8.28E-08	0.21	0.01	25.96

GWAS		r_g with childhood maltreatment				h^2_{SNP} (Observed scale)		
Phenotype	PMID	r_g	SE	Z	P	h^2_{SNP}	SE	Z
BMI (Locke) female	25673413	0.15	0.03	5.36	8.36E-08	0.17	0.01	18.51
Educational Attainment: EduYears Women	27225129	-0.17	0.03	-5.35	8.67E-08	0.18	0.01	23.06
Body fat percentage healthy meta-analysis (not alcohol/smoking-adjust)	30593698	0.15	0.03	5.34	9.11E-08	0.21	0.01	26.25
BMI (Locke) male	25673413	0.15	0.03	5.29	1.25E-07	0.17	0.01	15.19
Body fat percentage (Lu. et al)	26833246	0.20	0.04	5.24	1.62E-07	0.11	0.01	13.20
Lifetime probable anxiety diagnosis male	31748690	0.35	0.07	5.20	1.97E-07	0.07	0.01	6.02
Undersleeper	27494321	0.29	0.06	5.08	3.87E-07	0.04	0.00	10.48
Coronary Artery Disease	28714975	0.15	0.03	5.02	5.10E-07	0.04	0.00	14.67
Obesity class 2	23563607	0.19	0.04	5.02	5.10E-07	0.16	0.01	15.05
Verbal-numerical reasoning	27046643	-0.23	0.05	-4.99	5.92E-07	0.20	0.02	11.23
Longest 10% survival	27015805	-0.28	0.06	-4.95	7.62E-07	0.01	0.00	8.88
Insomnia female	28604731	0.32	0.07	4.87	1.12E-06	0.05	0.01	5.81
Problematic consequences of drinking (AUDIT-P)	30336701	0.21	0.04	4.81	1.54E-06	0.06	0.00	12.20
Fat mass healthy female	31852892	0.15	0.03	4.80	1.57E-06	0.21	0.01	18.55
Tobacco: smoking cessation (former versus current smokers)	20418890	-0.35	0.07	-4.70	2.55E-06	0.06	0.01	6.05
Fat mass anorexia-free female	31852892	0.15	0.03	4.68	2.89E-06	0.21	0.01	18.44
Lean mass (fat free mass)	30593698	0.09	0.02	4.54	5.56E-06	0.29	0.01	24.77
Overweight	23563607	0.15	0.03	4.53	5.96E-06	0.09	0.01	16.57
Lifetime cannabis use	27023175	0.30	0.07	4.47	7.68E-06	0.09	0.02	5.59
Waist circumference European female	25673412	0.14	0.03	4.46	8.08E-06	0.16	0.01	16.56
Reproductive behavior: Number of Children Ever Born Female	27798627	0.24	0.05	4.45	8.63E-06	0.02	0.00	10.48
Severe early onset obesity vs controls	30677029	0.26	0.06	4.44	9.00E-06	0.32	0.04	7.36
Body fat percentage anorexia-free male	30593698	0.14	0.03	4.44	9.14E-06	0.23	0.01	23.15
Body fat percentage healthy female	30593698	0.15	0.03	4.43	9.42E-06	0.20	0.01	19.58
Body fat percentage healthy male	30593698	0.14	0.03	4.43	9.62E-06	0.23	0.01	22.97
Oversleeper	27494321	0.39	0.09	4.38	1.17E-05	0.02	0.00	4.61
Lean mass (fat free mass) meta-analysis of males and females	30593698	0.09	0.02	4.36	1.29E-05	0.28	0.01	25.32
ADHD symptoms	27663945	0.58	0.13	4.29	1.79E-05	0.07	0.03	2.51
Insomnia male	28604731	0.29	0.07	4.25	2.14E-05	0.06	0.01	6.77
Body fat percentage anorexia-free female	30593698	0.14	0.03	4.24	2.20E-05	0.20	0.01	19.51
Waist-hip ratio all ancestries	25673412	0.14	0.03	4.22	2.44E-05	0.10	0.01	16.32
Waist-hip ratio EUR	25673412	0.14	0.03	4.20	2.62E-05	0.10	0.01	16.16
Reproductive behavior: Number of Children Ever Born Male	27798627	0.25	0.06	4.05	5.06E-05	0.03	0.01	6.94
Lean mass (fat free mass) mass healthy	30593698	0.10	0.02	4.05	5.16E-05	0.30	0.01	20.76
Lean mass (fat free mass) anorexia-free	30593698	0.10	0.02	4.03	5.48E-05	0.30	0.01	20.81
Primary sclerosing cholangitis	27992413	-0.23	0.06	-3.97	7.06E-05	0.15	0.03	4.42
Autism spectrum disorder worldwide	28540026	0.21	0.05	3.96	7.39E-05	0.32	0.04	9.07
Extreme BMI	23563607	0.17	0.04	3.94	8.17E-05	0.59	0.05	12.54

GWAS		r_g with childhood maltreatment				h^2_{SNP} (Observed scale)		
Phenotype	PMID	r_g	SE	Z	P	h^2_{SNP}	SE	Z
Lean mass (fat free mass) male	30593698	0.09	0.02	3.93	8.48E-05	0.33	0.02	21.37
Age at Menarche (marker of puberty in females)	25231870	-0.12	0.03	-3.93	8.51E-05	0.16	0.01	20.16
PTSD European female	28439101	0.34	0.09	3.92	8.76E-05	0.46	0.12	3.76
Type-2 Diabetes 2017	28566273	0.14	0.04	3.89	1.00E-04	0.08	0.01	13.72
Hip circumference all ancestries	25673412	0.11	0.03	3.87	1.00E-04	0.15	0.01	16.62
Lean mass (fat free mass) anorexia-free meta-analysis of males and females	30593698	0.09	0.02	3.84	1.00E-04	0.29	0.01	21.51
Hip circumference European	25673412	0.11	0.03	3.84	1.00E-04	0.15	0.01	16.66
Lean mass (fat free mass) healthy meta	30593698	0.09	0.02	3.84	1.00E-04	0.29	0.01	21.47
Weight EU female	23754948	0.12	0.03	3.78	2.00E-04	0.19	0.01	17.43
Fat-free mass/lean mass healthy meta-analysis (not alcohol/smoking-adjust)	31852892	0.09	0.02	3.74	2.00E-04	0.29	0.01	21.73
Waist circumference European male	25673412	0.14	0.04	3.71	2.00E-04	0.18	0.02	10.59
Lean mass (fat free mass) female	30593698	0.08	0.02	3.64	3.00E-04	0.30	0.01	23.99
Lean mass (fat free mass) healthy male	30593698	0.09	0.03	3.61	3.00E-04	0.34	0.02	17.04
Coronary Artery Disease (European, South Asian and East Asian) additive model	26343387	0.13	0.03	3.60	3.00E-04	0.07	0.01	12.94
Lean mass (fat free mass) anorexia-free male	30593698	0.09	0.03	3.60	3.00E-04	0.34	0.02	17.03
Childhood Intelligence	23358156	-0.27	0.07	-3.59	3.00E-04	0.28	0.04	6.26
Citrate	27005778	-0.34	0.10	-3.57	4.00E-04	0.07	0.02	3.44
Weight EU male	23754948	0.12	0.03	3.55	4.00E-04	0.19	0.01	14.92
PTSD European	28439101	0.53	0.15	3.52	4.00E-04	0.12	0.05	2.28
Aggression	26087016	0.57	0.16	3.51	4.00E-04	0.05	0.02	2.08
Bipolar disorder	21926972	0.17	0.05	3.48	5.00E-04	0.43	0.04	11.48
Obesity class 3	23563607	0.17	0.05	3.48	5.00E-04	0.12	0.01	8.98
Anxiety Disorders meta-analysis: case-control	26754954	0.48	0.14	3.46	5.00E-04	0.08	0.03	2.53
Hip circumference all ancestries female	25673412	0.11	0.03	3.36	8.00E-04	0.18	0.01	16.88
Hip circumference European female	25673412	0.11	0.03	3.34	8.00E-04	0.18	0.01	17.04
Autism spectrum disorder European	28540026	0.19	0.06	3.29	0.001	0.33	0.04	7.73
Age at Menarche (marker of puberty in females)	28436984	-0.08	0.02	-3.26	0.0011	0.16	0.01	19.78
High FEV1 versus average FEV1 never smokers	28166213	0.21	0.06	3.20	0.0014	0.21	0.03	6.32
Schizophrenia-vs-bipolar	24280982	0.19	0.06	3.17	0.0015	0.31	0.04	7.66
Total cholesterol in HDL	20686565	-0.11	0.04	-3.07	0.0022	0.11	0.02	5.45
Lean mass (fat free mass) anorexia-free female	30593698	0.09	0.03	3.01	0.0026	0.30	0.02	18.23
Lean mass (fat free mass) healthy female	30593698	0.09	0.03	3.00	0.0027	0.30	0.02	18.16
Consumption (AUDIT-C)	30336701	-0.11	0.04	-2.95	0.0031	0.08	0.01	14.96
Alzheimer's Disease 2019 (clinically diagnosed AD and AD-by-proxy)	30617256	0.25	0.08	2.93	0.0034	0.01	0.00	1.90
Persistent thin individuals vs severe early onset obesity	30677029	0.16	0.05	2.91	0.0037	1.32	0.15	8.69
Childhood obesity	22484627	0.14	0.05	2.87	0.0041	0.41	0.05	8.99
Tobacco: CPD (cigaretts per day)	20418890	0.23	0.08	2.84	0.0046	0.06	0.02	3.67
ADHD	20732625	0.29	0.11	2.75	0.006	0.25	0.09	2.85

GWAS		r_g with childhood maltreatment				h^2_{SNP} (Observed scale)		
Phenotype	PMID	r_g	SE	Z	P	h^2_{SNP}	SE	Z
Cross-disorder: ADHD subset	23453885	0.26	0.10	2.74	0.0061	0.26	0.08	3.05
Coronary Artery Disease (European, South Asian and East Asian) recessive model	26343387	0.11	0.04	2.71	0.0066	0.02	0.00	8.90
Glycine male	31043756	0.16	0.06	2.64	0.0084	0.19	0.06	3.44
Fasting glucose (age- & sex-adjusted)	20081858	0.14	0.06	2.62	0.0089	0.09	0.02	4.61
Mean corpuscular hemoglobin MCH	27863252	0.06	0.02	2.60	0.0093	0.22	0.03	8.72
Loneliness: Case-control	27629369	0.38	0.15	2.58	0.01	0.12	0.06	1.96
Loneliness: Linear	27629369	0.48	0.19	2.57	0.0102	0.10	0.06	1.74
BMI childhood: 3-10 years	26604143	0.11	0.04	2.53	0.0115	0.25	0.02	10.06
Visceral adipose tissue volume	27918534	0.17	0.07	2.51	0.012	0.14	0.03	5.52
Height (2018)	30124842	-0.05	0.02	-2.51	0.0122	0.50	0.02	22.01
Anxiety Disorders meta-analysis: factor scores	26754954	0.31	0.13	2.49	0.0127	0.07	0.03	2.46
PTSD Mixed Ancestry	28439101	0.60	0.25	2.44	0.0148	0.06	0.04	1.48
Hip circumference all ancestries male	25673412	0.09	0.04	2.43	0.015	0.19	0.02	11.16
Hip circumference European male	25673412	0.09	0.04	2.36	0.0184	0.19	0.02	11.43
Pericardial adipose tissue volume female	27918534	0.24	0.10	2.35	0.019	0.19	0.07	2.79
eGFRcys (serum cystatin)	26831199	-0.11	0.05	-2.33	0.0198	0.17	0.07	2.51
Concentration of very small VLDL particles	27005778	0.17	0.07	2.29	0.0219	0.12	0.04	3.25
Mean corpuscular volume	27863252	0.06	0.02	2.29	0.022	0.24	0.03	9.29
Concentration of small LDL particles	27005778	0.21	0.09	2.29	0.0223	0.10	0.04	2.28
Mono-unsaturated fatty acids	27005778	0.24	0.10	2.28	0.0227	0.11	0.04	2.47
Physical activity (objectively-measured) female	34753499	-0.11	0.05	-2.26	0.0238	0.16	0.01	11.18
Subcutaneous adipose tissue volume	27918534	0.14	0.06	2.26	0.0238	0.16	0.03	6.37
Visceral adipose tissue volume female	27918534	0.19	0.09	2.26	0.0239	0.15	0.04	3.70
Total lipids in very small VLDL	27005778	0.18	0.08	2.17	0.0301	0.11	0.04	3.01
Chronotype	27494321	-0.09	0.04	-2.15	0.0319	0.10	0.01	17.72
Hip circumference European male BMI-adjusted	25673412	-0.07	0.03	-2.14	0.032	0.20	0.02	12.31
Total lipids in small LDL	27005778	0.21	0.10	2.14	0.0321	0.09	0.04	2.02
Obsessive compulsive disorder	28761083	0.13	0.06	2.13	0.0331	0.34	0.05	6.86
Anorexia nervosa (PGC2)	31308545	0.08	0.04	2.10	0.0354	0.18	0.01	14.79
Serum total cholesterol	27005778	0.27	0.13	2.10	0.0354	0.07	0.04	1.76
Asthma	20860503	0.12	0.06	2.10	0.0355	0.12	0.02	5.31
Omega-9 and saturated fatty acids	27005778	0.27	0.13	2.10	0.0359	0.09	0.04	1.95
Concentration of IDL particles	27005778	0.23	0.11	2.10	0.036	0.08	0.04	1.87
Triglycerides in IDL	27005778	0.17	0.08	2.09	0.0363	0.11	0.03	3.09
UACR (urinary albumin-to-creatinine ratio)	26631737	0.15	0.07	2.07	0.0388	0.05	0.01	5.15
Fasting insulin (BMI-adjusted)	22581228	0.13	0.06	2.07	0.0389	0.07	0.01	6.97
Triglycerides	20686565	0.08	0.04	2.06	0.0393	0.15	0.03	5.75
PTSD Mixed Ancestry	28439101	0.73	0.35	2.06	0.0395	0.03	0.02	1.15

GWAS		r_g with childhood maltreatment				h^2_{SNP} (Observed scale)		
Phenotype	PMID	r_g	SE	Z	P	h^2_{SNP}	SE	Z
TWEAK	28240269	-0.33	0.16	-2.05	0.0408	1.06	0.64	1.67
Apolipoprotein B	27005778	0.21	0.10	2.05	0.0408	0.08	0.04	2.13
Hip circumference all ancestries male BMI-adjusted	25673412	-0.07	0.03	-2.04	0.0412	0.20	0.02	12.33
Hip circumference European BMI-adjusted	25673412	-0.06	0.03	-2.04	0.0413	0.15	0.01	16.46
Hip circumference all ancestries BMI-adjusted	25673412	-0.06	0.03	-2.02	0.0431	0.15	0.01	16.61
Total lipids in IDL	27005778	0.23	0.11	2.01	0.0443	0.08	0.04	1.78
HOMA-IR: insulin resistance (age- & sex-adjusted)	20081858	0.14	0.07	2.00	0.0454	0.07	0.01	5.20
Glycine	27005778	0.12	0.06	2.00	0.0459	0.21	0.12	1.86

Supplementary Table 2: LDscore genetic correlation (r_g) results for genetically correlated phenotypes with reported childhood maltreatment $|r_g| > 0.25$ and $r_g |Z|$ statistic ≥ 5 , a heritability Z statistics ≥ 5 , mean chi-square > 1.02 .

Traits that met criteria and were selected for multiple regression analyses in table 2A. Traits with $|r_g| > 0.25$ and $r_g |Z|$ statistic ≥ 5 but were not included in downstream analyses in 2B. These traits were excluded for the following reasons: the sample was mixed ancestry, or a very similar trait was the preferred phenotype for analyses due to power, sample characteristics or phenotyping.

2A) 22 top r_g traits included in analyses on childhood maltreatment													
GWAS				r_g with reported childhood maltreatment						h^2_{SNP} (Observed scale)			Assigned category for analyses
Phenotype	PMID	Year	mean chi2	r_g	se	z	p	Gcov int	Gcov int se	h^2_{SNP}	SE	Z	
ADHD	30478444	2017	1.29	0.56	0.04	15.09	1.96E-51	0.00	0.01	0.24	0.01	16.38	Psychiatric Disorder
Total PTSD symptoms (MVP)	33510476	2021	1.35	0.56	0.03	18.30	7.65E-75	0.01	0.01	0.07	0.00	18.31	Psychiatric Disorder
MDD (PGC2 + 23andme + UKB)	29700475	2018	1.53	0.56	0.03	19.98	8.50E-89	0.01	0.01	0.06	0.00	23.08	Psychiatric Disorder
Alcohol dependence European	30482948	2017	1.07	0.53	0.09	5.93	3.10E-09	0.00	0.01	0.05	0.01	5.41	Psychiatric Disorder
Self-reported tiredness	28322280	2017	1.14	0.52	0.05	11.06	2.04E-28	0.02	0.01	0.07	0.01	11.57	Health & behavioural
Subjective well-being: survey questions on life satisfaction, positive affect, or happiness	27089181	2016	1.15	-0.52	0.05	-10.73	7.10E-27	-0.02	0.01	0.02	0.00	12.45	Psychiatric Disorder
Social deprivation, as measured using Townsend scores	27818178	2016	1.09	0.50	0.06	8.92	4.83E-19	0.02	0.00	0.04	0.00	8.42	Health & behavioural
number of sexual partners (risky behaviour)	30643258	2019	1.72	0.48	0.02	20.04	2.44E-89	0.06	0.01	0.10	0.00	29.24	Health & behavioural
Self-rated health	27864402	2016	1.23	-0.48	0.04	-12.09	1.22E-33	-0.03	0.01	0.10	0.01	16.74	Health & behavioural
Lifetime probable anxiety diagnosis	31748690	2020	1.16	0.45	0.04	11.52	1.01E-30	0.15	0.01	0.10	0.01	15.10	Psychiatric Disorder
General risk tolerance (UKB + 10 replication cohorts)	30643258	2019	1.48	0.42	0.03	15.56	1.46E-54	0.05	0.01	0.05	0.00	22.52	Health & behavioural
Autism spectrum disorder (PGC + iPSYCH)	30804558	2019	1.19	0.41	0.04	9.48	2.51E-21	0.01	0.01	0.20	0.02	13.19	Psychiatric Disorder
Reproductive behavior: Age First Birth	27798627	2016	1.22	-0.40	0.03	-11.56	6.76E-31	-0.01	0.01	0.06	0.00	17.81	Health & behavioural
Schizophrenia PGC 2014 European	25056061	2014	1.74	0.35	0.03	13.46	2.65E-41	0.01	0.01	0.47	0.02	27.87	Psychiatric Disorder

Neuroticism 300K meta-analysis. (UKB)	30593698	2017	1.67	0.34	0.03	11.63	2.86E-31	0.07	0.01	0.11	0.01	18.77	Health & behavioural
Insomnia (2019)	30804565	2019	1.36	0.32	0.03	9.64	5.56E-22	0.03	0.01	0.05	0.00	21.86	Health & behavioural
Tobacco: ever smoker (UKB + TAG)	30643258	2019	1.92	0.32	0.03	11.95	6.59E-33	0.05	0.01	0.09	0.00	28.53	Health & behavioural
Parents attained age	27015805	2017	1.34	0.25	0.04	6.28	3.49E-10	0.01	0.01	0.03	0.00	13.74	Health & behavioural

B) Top rg traits excluded from analyses of childhood maltreatment

GWAS				rg with reported childhood maltreatment						h ² _{SNP} (Observed scale)			Reason for exclusion
Phenotype	PMID	Year	mean chi2	rg	se	z	p	Gcov int	Gcov int se	h ² _{SNP}	SE	Z	
PTSD PGC2 case control	31594949	2019	1.08	0.76	0.08	9.55	1.28E-21	0.10	0.01	0.02	0.00	5.72	> powerful MVP PTSD total symptoms included
Depressive symptoms	27089181	2016	1.15	0.66	0.04	15.15	7.95E-52	0.02	0.01	0.05	0.00	13.06	> powerful depression GWAS included
MDD (PGC2, no 23andme, no UKB)	29700475	2018	1.20	0.63	0.04	16.91	3.82E-64	0.00	0.01	0.07	0.00	16.11	> powerful MDD GWAS included
PTSD Avoidance symptoms	33510476	2021	1.31	0.58	0.03	18.72	3.51E-78	0.01	0.01	0.06	0.00	16.74	> powerful MVP PTSD tot. symp. included (rg across these traits is not signif. diff. from 1)
ADHD male	29325848	2017	1.18	0.58	0.05	12.88	5.79E-38	0.00	0.01	0.25	0.02	11.90	> powerful ADHD GWAS included
ADHD (Mixed Ancestry)	30478444	2017	1.29	0.56	0.04	14.82	1.08E-49	0.00	0.01	0.36	0.02	16.51	Mixed Ancestry
MDD (PGC1)	22472876	2013	1.07	0.55	0.08	6.98	2.93E-12	0.00	0.01	0.16	0.03	5.95	> powerful MDD GWAS included
MDD (PGC2 + 23andme, no UK)	29700475	2018	1.48	0.55	0.03	19.14	1.13E-81	0.00	0.01	0.06	0.00	22.28	> powerful MDD GWAS included
Alcohol dependence all ancestries	30482948	2017	1.07	0.54	0.09	6.23	4.72E-10	-0.01	0.01	0.05	0.01	5.42	Mixed Ancestry
PTSD Hyperarousal symptoms	33510476	2021	1.35	0.54	0.03	16.64	3.42E-62	0.00	0.01	0.07	0.00	18.21	> powerful MVP PTSD tot. symp. included (rg across these traits is not signif. diff. from 1)
PTSD Re-experiencing symptoms	33510476	2021	1.32	0.52	0.03	16.62	5.08E-62	0.00	0.01	0.07	0.00	17.57	> powerful MVP PTSD tot. symp. included (rg across these traits is not signif. diff. from 1)
Lifetime probable anxiety diagnosis female	31748690	2020	1.10	0.50	0.05	10.41	2.33E-25	0.13	0.00	0.14	0.01	11.63	> powerful Anxiety GWAS included
Heavy versus never smoker in UKBileve	28166213	2017	1.29	0.45	0.04	11.36	6.69E-30	0.01	0.01	0.26	0.02	15.27	> powerful tobacco smoking GWAS included
Household income, 1 per household	27818178	2016	1.15	-0.43	0.05	-8.81	1.30E-18	-0.02	0.00	0.07	0.01	10.38	> powerful household income GWAS included
PTSD MVP case control	33510476	2021	1.17	0.42	0.05	8.46	2.68E-17	0.00	0.01	0.03	0.00	11.47	> powerful MVP PTSD total symptoms included
Reproductive behavior: Age First Birth Female	27798627	2016	1.16	-0.41	0.04	-10.49	9.27E-26	-0.01	0.01	0.05	0.00	14.53	> powerful Age First Birth GWAS included
Household income	27818178	2016	1.16	-0.41	0.05	-8.43	3.41E-17	-0.03	0.01	0.06	0.01	10.54	Social depriv. GWAS included (rg across these traits is not signif. diff. from 1)
Cross-psychiatric disorders (ADHD, ASD, bipolar disorder, & SCZ)	23453885	2013	1.22	0.40	0.04	10.10	5.56E-24	0.00	0.01	0.18	0.01	16.37	ADHD GWAS and ASD GWAS included
Neuroticism.	27089181	2016	1.31	0.39	0.04	9.26	2.08E-20	0.04	0.01	0.09	0.01	12.00	> powerful Neuroticism GWAS included
SCZ PGC2 European + CLOZUK	29483656	2018	2.01	0.36	0.03	12.20	2.95E-34	0.01	0.01	0.42	0.02	27.27	SGZ PGC 2014 GWAS included
Reproductive behavior: Age First Birth Male	27798627	2016	1.05	-0.36	0.05	-7.19	6.48E-13	0.00	0.00	0.11	0.01	10.75	> powerful Age First Birth GWAS included
SCZ PGC 2014 (EUR and East Asian)	25056061	2014	1.76	0.36	0.03	13.64	2.35E-42	0.00	0.01	0.45	0.02	28.52	Mixed Ancestry
Lifetime probable anxiety diagnosis male	31748690	2020	1.06	0.35	0.07	5.20	1.97E-07	0.08	0.01	0.07	0.01	6.02	> powerful Anxiety GWAS included
SCZ PGC Mixed ancestry	31740837	2019	2.13	0.35	0.02	15.07	2.68E-51	0.01	0.01	0.34	0.01	30.45	Mixed Ancestry
Neuroticism 300K female. (UKB)	30593698	2017	1.35	0.34	0.04	9.81	9.81E-23	0.05	0.01	0.12	0.01	16.71	> powerful Anxiety GWAS included

Tobacco: smoking initiation (ever versus never been a regular smoker)	20418890	2010	1.10	0.34	0.05	6.95	3.74E-12	0.00	0.01	0.08	0.01	11.80	> powerful tobacco smoking GWAS included
Insomnia	28604731	2017	1.11	0.33	0.05	6.72	1.85E-11	0.02	0.00	0.05	0.00	9.82	> powerful Insomnia 2019 GWAS included
Neuroticism meta-analysis. (UKB)	30593698	2017	1.34	0.33	0.04	8.54	1.32E-17	0.05	0.01	0.11	0.01	12.26	> powerful neuroticism GWAS included
Neuroticism 150K healthy meta-analysis.	30593698	2017	1.33	0.31	0.04	8.09	5.84E-16	0.05	0.01	0.10	0.01	12.32	> powerful neuroticism GWAS
Mother's age at death	27015805	2016	1.17	-0.30	0.05	-5.96	2.45E-09	-0.02	0.01	0.01	0.00	9.13	Parents age GWAS included
Schizophrenia+bipolar	24280982	2014	1.31	0.30	0.04	8.31	9.33E-17	0.00	0.01	0.38	0.02	16.73	SCZ GWAS included
Undersleeper	27494321	2016	1.11	0.29	0.06	5.08	3.87E-07	0.02	0.01	0.04	0.00	10.48	> powerful insomnia 2019 GWAS included
Father's age at death	27015805	2016	1.19	-0.29	0.05	-6.36	2.00E-10	-0.01	0.01	0.02	0.00	11.50	Parents age GWAS included
Neuroticism 150K healthy female. (UKB)	30593698	2017	1.15	0.29	0.05	6.04	1.58E-09	0.04	0.01	0.10	0.01	10.37	> powerful neuroticism GWAS included
Parents' age at death	27015805	2017	1.21	-0.28	0.05	-6.22	5.10E-10	-0.01	0.01	0.02	0.00	11.24	> powerful parents attained age GWAS included

Supplementary table 3: A-K Genomic SEM multiple regression models where each psychological, social and health trait is introduced in iterative stages, in order of the most highly genetically correlated trait with reported childhood maltreatment.

Model K standardised results is the final model included in the main analyses. Each model takes into account the genetic correlations across Psychological and Social traits, but for simplicity, these genetic correlations are only shown in Table L.

A) Childhood maltreatment ~ Tiredness								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.05	0.00	0.73	0.06	0.73	0.00
Childhood maltreatment	~	Tiredness	0.54	0.06	0.52	0.06	0.52	0.00
B) Childhood maltreatment ~ Tiredness + Subjective well-being								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.05	0.00	0.65	0.06	0.65	0.00
Childhood maltreatment	~	Tiredness	0.34	0.08	0.33	0.08	0.33	0.00
Childhood maltreatment	~	Subj. wellbeing	-0.57	0.14	-0.33	0.08	-0.33	0.00
C) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.04	0.00	0.58	0.07	0.58	0.00
Childhood maltreatment	~	Tiredness	0.19	0.10	0.18	0.09	0.18	0.05
Childhood maltreatment	~	Subj. wellbeing	-0.55	0.14	-0.32	0.08	-0.32	0.00
Childhood maltreatment	~	Social dep	0.42	0.12	0.31	0.08	0.31	0.00
D) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.04	0.00	0.51	0.06	0.51	0.00
Childhood maltreatment	~	Tiredness	0.20	0.09	0.20	0.09	0.20	0.03
Childhood maltreatment	~	Subj. wellbeing	-0.52	0.13	-0.31	0.08	-0.31	0.00
Childhood maltreatment	~	Sexual partners	0.26	0.04	0.30	0.05	0.30	0.00

Childhood maltreatment	~	Social dep	0.22	0.13	0.16	0.09	0.16	0.08
E) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners + self rated health								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.04	0.00	0.51	0.06	0.51	0.00
Childhood maltreatment	~	Tiredness	0.16	0.12	0.15	0.12	0.15	0.20
Childhood maltreatment	~	Subj. wellbeing	-0.52	0.13	-0.31	0.08	-0.31	0.00
Childhood maltreatment	~	Sexual partners	0.27	0.05	0.31	0.05	0.31	0.00
Childhood maltreatment	~	Social dep	0.18	0.16	0.13	0.11	0.13	0.25
Childhood maltreatment	~	Health	-0.07	0.11	-0.08	0.12	-0.08	0.54
F) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners + self-rated health + general risk tolerance								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.00	0.44	0.06	0.44	0.00
Childhood maltreatment	~	Tiredness	0.13	0.13	0.12	0.13	0.12	0.32
Childhood maltreatment	~	Subj. wellbeing	-0.64	0.15	-0.38	0.09	-0.38	0.00
Childhood maltreatment	~	Sexual partners	0.08	0.07	0.10	0.08	0.10	0.21
Childhood maltreatment	~	Risk	0.38	0.07	0.32	0.06	0.32	0.00
Childhood maltreatment	~	Social dep	0.26	0.17	0.19	0.13	0.19	0.14
Childhood maltreatment	~	Health	-0.03	0.11	-0.03	0.13	-0.03	0.82
G) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners + self-rated health + general risk tolerance + Age first birth								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.00	0.42	0.06	0.42	0.00
Childhood maltreatment	~	Tiredness	0.18	0.14	0.17	0.13	0.17	0.19
Childhood maltreatment	~	Subj. wellbeing	-0.71	0.16	-0.42	0.09	-0.42	0.00
Childhood maltreatment	~	Sexual partners	0.09	0.07	0.11	0.08	0.11	0.17
Childhood maltreatment	~	Risk	0.35	0.07	0.30	0.06	0.30	0.00
Childhood maltreatment	~	Social dep	0.16	0.19	0.12	0.14	0.12	0.39
Childhood maltreatment	~	Health	0.09	0.13	0.11	0.16	0.11	0.50
Childhood maltreatment	~	Age first birth	-0.24	0.10	-0.21	0.09	-0.21	0.02
H) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners + self-rated health + general risk tolerance + Age first birth + Neuroticism								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.00	0.42	0.07	0.42	0.00
Childhood maltreatment	~	Tiredness	0.22	0.20	0.22	0.19	0.22	0.26
Childhood maltreatment	~	Subj. wellbeing	-0.78	0.27	-0.46	0.16	-0.46	0.00
Childhood maltreatment	~	Sexual partners	0.08	0.07	0.10	0.08	0.10	0.25
Childhood maltreatment	~	Risk	0.35	0.07	0.30	0.06	0.30	0.00
Childhood maltreatment	~	Social dep	0.14	0.20	0.10	0.15	0.10	0.48
Childhood maltreatment	~	Health	0.12	0.18	0.14	0.21	0.14	0.50

Childhood maltreatment	~	Age first birth	-0.27	0.14	-0.23	0.13	-0.23	0.07
Childhood maltreatment	~	Neuroticism	-0.06	0.12	-0.07	0.14	-0.07	0.63

I) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners + self-rated health + general risk tolerance + Age first birth + Neuroticism + Insomnia

Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.01	0.42	0.07	0.42	0.00
Childhood maltreatment	~	Insomnia	-0.07	0.06	-0.07	0.06	-0.07	0.25
Childhood maltreatment	~	Tiredness	0.25	0.21	0.24	0.21	0.24	0.25
Childhood maltreatment	~	Subj. wellbeing	-0.80	0.28	-0.47	0.17	-0.47	0.00
Childhood maltreatment	~	Sexual partners	0.08	0.07	0.10	0.09	0.10	0.25
Childhood maltreatment	~	Risk	0.35	0.07	0.30	0.06	0.30	0.00
Childhood maltreatment	~	Social dep	0.14	0.20	0.10	0.15	0.10	0.49
Childhood maltreatment	~	Health	0.12	0.18	0.14	0.21	0.14	0.52
Childhood maltreatment	~	Age first birth	-0.28	0.15	-0.25	0.13	-0.25	0.06
Childhood maltreatment	~	Neuroticism	-0.05	0.12	-0.07	0.15	-0.07	6.439661e-

J) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners + self-rated health + general risk tolerance + Age first birth + Neuroticism + Insomnia + Ever smoker

Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.00	0.41	0.07	0.41	0.00
Childhood maltreatment	~	Neuroticism	-0.03	0.12	-0.04	0.15	-0.04	0.78
Childhood maltreatment	~	Insomnia	-0.07	0.06	-0.07	0.06	-0.07	0.26
Childhood maltreatment	~	Tiredness	0.21	0.22	0.21	0.21	0.21	0.34
Childhood maltreatment	~	Subj. wellbeing	-0.78	0.28	-0.46	0.17	-0.46	0.01
Childhood maltreatment	~	Sexual partners	0.11	0.07	0.13	0.09	0.13	0.12
Childhood maltreatment	~	Risk	0.35	0.07	0.30	0.06	0.30	0.00
Childhood maltreatment	~	Social dep	0.19	0.23	0.14	0.17	0.14	0.40
Childhood maltreatment	~	Health	0.09	0.19	0.11	0.22	0.11	0.63
Childhood maltreatment	~	Age first birth	-0.28	0.15	-0.24	0.13	-0.24	0.07
Childhood maltreatment	~	Ever smoker	-0.09	0.06	-0.10	0.07	-0.10	0.16

K) Childhood maltreatment ~ Tiredness + Subjective well-being + social deprivation + number of sexual partners + self-rated health + general risk tolerance + Age first birth + Neuroticism + Insomnia + Ever smoker + Parents attained age

Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.01	0.41	0.07	0.41	7.76E-09
Childhood maltreatment	~	Neuroticism	-0.04	0.13	-0.05	0.16	-0.05	7.60E-01
Childhood maltreatment	~	Insomnia	-0.07	0.06	-0.07	0.06	-0.07	2.49E-01
Childhood maltreatment	~	Parents attained age	0.10	0.18	0.06	0.12	0.06	5.85E-01
Childhood maltreatment	~	Tiredness	0.24	0.25	0.23	0.24	0.23	3.45E-01
Childhood maltreatment	~	Subj. wellbeing	-0.80	0.31	-0.47	0.18	-0.47	1.00E-02
Childhood maltreatment	~	Sexual partners	0.11	0.07	0.14	0.09	0.14	1.25E-01

Childhood maltreatment	~	Risk	0.36	0.07	0.31	0.06	0.31	9.96E-07
Childhood maltreatment	~	Social dep	0.17	0.25	0.12	0.18	0.12	5.02E-01
Childhood maltreatment	~	Health	0.14	0.25	0.16	0.29	0.16	5.84E-01
Childhood maltreatment	~	Age first birth	-0.29	0.16	-0.25	0.14	-0.25	8.05E-02
Childhood maltreatment	~	Ever smoker	-0.09	0.06	-0.10	0.07	-0.10	1.21E-01

L) Cross Psychological & Social trait genetic correlations									
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value_diff_0	p_value_diff_1
Neuroticism	~~	Insomnia	0.04	0.00	0.37	0.03	0.37	2.85E-38	1.8E-106
Neuroticism	~~	Parents attained age	0.00	0.00	0.04	0.04	0.04	2.86E-01	1.2E-124
Insomnia	~~	Parents attained age	0.01	0.00	0.26	0.04	0.26	5.34E-12	5.92E-87
Tiredness	~~	Neuroticism	0.05	0.00	0.57	0.04	0.57	1.39E-40	2.33E-23
Tiredness	~~	Insomnia	0.04	0.00	0.53	0.04	0.53	1.71E-38	1.05E-31
Tiredness	~~	Parents attained age	0.01	0.00	0.31	0.05	0.31	1.33E-10	9.43E-46
Tiredness	~~	Subj. wellbeing	-0.02	0.00	-0.57	0.06	-0.57	4.07E-24	1.06E-14
Tiredness	~~	Sexual partners	0.02	0.00	0.23	0.04	0.23	9.80E-11	8.4E-101
Tiredness	~~	Risk	0.01	0.00	0.13	0.04	0.13	4.64E-04	4.8E-118
Tiredness	~~	Social dep	0.03	0.00	0.51	0.07	0.51	2.45E-13	8.47E-13
Tiredness	~~	Health	-0.06	0.00	-0.77	0.06	-0.77	2.17E-37	9.82E-05
Tiredness	~~	Age first birth	-0.02	0.00	-0.34	0.04	-0.34	2.76E-15	2.1E-51
Tiredness	~~	Ever smoker	0.02	0.00	0.26	0.03	0.26	2.35E-14	2.8E-109
Subj. wellbeing	~~	Neuroticism	-0.04	0.00	-0.66	0.05	-0.66	2.29E-45	7.02E-13
Subj. wellbeing	~~	Insomnia	-0.02	0.00	-0.43	0.04	-0.43	1.78E-22	1.2E-39
Subj. wellbeing	~~	Parents attained age	0.00	0.00	-0.16	0.05	-0.16	3.29E-03	4.16E-57
Subj. wellbeing	~~	Sexual partners	-0.01	0.00	-0.17	0.04	-0.17	7.81E-07	1.3E-122
Subj. wellbeing	~~	Risk	0.00	0.00	0.05	0.04	0.05	1.58E-01	5.5E-148
Subj. wellbeing	~~	Social dep	-0.01	0.00	-0.31	0.07	-0.31	2.00E-05	3.99E-21
Subj. wellbeing	~~	Health	0.02	0.00	0.48	0.05	0.48	1.67E-19	5.54E-23
Subj. wellbeing	~~	Age first birth	0.00	0.00	0.09	0.05	0.09	4.71E-02	4.5E-88
Subj. wellbeing	~~	Ever smoker	-0.01	0.00	-0.16	0.03	-0.16	1.51E-06	2.3E-143
Sexual partners	~~	Neuroticism	0.00	0.00	0.01	0.03	0.01	7.70E-01	8.1E-279
Sexual partners	~~	Insomnia	0.01	0.00	0.15	0.03	0.15	3.76E-08	5E-204
Sexual partners	~~	Parents attained age	0.01	0.00	0.13	0.03	0.13	4.14E-05	2.5E-162
Sexual partners	~~	Risk	0.04	0.00	0.59	0.03	0.59	1.31E-81	2.06E-39
Sexual partners	~~	Social dep	0.03	0.00	0.48	0.05	0.48	3.37E-25	3.01E-29
Sexual partners	~~	Health	-0.02	0.00	-0.19	0.03	-0.19	1.51E-09	1E-152
Sexual partners	~~	Age first birth	-0.02	0.00	-0.23	0.03	-0.23	1.03E-17	1.4E-181
Sexual partners	~~	Ever smoker	0.05	0.00	0.54	0.03	0.54	1.75E-99	6.61E-75
Risk	~~	Neuroticism	0.00	0.00	-0.05	0.03	-0.05	6.98E-02	7.2E-226
Risk	~~	Insomnia	0.01	0.00	0.09	0.03	0.09	1.93E-03	4.1E-221

Risk	~~	Parents attained age	0.00	0.00	0.04	0.03	0.04	2.57E-01	3.7E-179
Risk	~~	Social dep	0.01	0.00	0.21	0.05	0.21	3.65E-05	1.29E-55
Risk	~~	Health	-0.01	0.00	-0.11	0.04	-0.11	1.60E-03	3.6E-132
Risk	~~	Age first birth	-0.01	0.00	-0.22	0.03	-0.22	1.67E-11	9.8E-123
Risk	~~	Ever smoker	0.02	0.00	0.32	0.03	0.32	8.67E-35	1.4E-147
Social dep	~~	Neuroticism	0.01	0.00	0.11	0.04	0.11	8.12E-03	2.38E-94
Social dep	~~	Insomnia	0.02	0.00	0.34	0.05	0.34	1.62E-10	1.05E-36
Social dep	~~	Parents attained age	0.02	0.00	0.50	0.06	0.50	3.00E-16	4.64E-16
Social dep	~~	Health	-0.04	0.00	-0.62	0.06	-0.62	4.32E-23	1.7E-09
Social dep	~~	Age first birth	-0.03	0.00	-0.57	0.06	-0.57	5.02E-20	5.67E-12
Social dep	~~	Ever smoker	0.03	0.00	0.58	0.04	0.58	1.05E-41	1.53E-22
Health	~~	Neuroticism	-0.03	0.00	-0.31	0.05	-0.31	5.83E-11	5.55E-50
Health	~~	Insomnia	-0.04	0.00	-0.49	0.04	-0.49	2.11E-38	5.2E-42
Health	~~	Parents attained age	-0.03	0.00	-0.60	0.04	-0.60	5.81E-44	3.67E-21
Health	~~	Age first birth	0.04	0.00	0.60	0.04	0.60	2.10E-49	6.08E-23
Health	~~	Ever smoker	-0.04	0.00	-0.38	0.03	-0.38	5.26E-40	9.2E-106
Age first birth	~~	Neuroticism	-0.01	0.00	-0.13	0.03	-0.13	1.31E-05	1.9E-176
Age first birth	~~	Insomnia	-0.02	0.00	-0.32	0.04	-0.32	4.14E-20	4.43E-81
Age first birth	~~	Ever smoker	-0.03	0.00	-0.40	0.03	-0.40	3.96E-39	2.06E-88
Ever smoker	~~	Neuroticism	0.01	0.00	0.10	0.02	0.10	1.44E-05	0.00
Ever smoker	~~	Insomnia	0.02	0.00	0.21	0.03	0.21	2.15E-16	5.5E-199
Ever smoker	~~	Parents attained age	0.02	0.00	0.36	0.03	0.36	1.84E-39	3.8E-119

Supplementary table 4: A-G Genomic SEM multiple regression models where each Psychiatric disorder is introduced in iterative stages, in order of the most highly genetically correlated trait with reported childhood maltreatment.

Model (G) standardised results is the final model included in the main analyses. Each model takes into account the genetic correlations across Psychiatric disorders, but for simplicity, these genetic correlations are only shown in Table H.

A) Childhood maltreatment ~ ADHD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.05	0.00	0.69	0.06	0.69	8.35E-27
Childhood maltreatment	~	ADHD	0.32	0.02	0.56	0.04	0.56	1.07E-36
B) Childhood maltreatment ~ ADHD + PTSD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.04	0.00	0.58	0.06	0.58	5.49E-24
Childhood maltreatment	~	ADHD	0.21	0.03	0.37	0.05	0.37	1.99E-11
Childhood maltreatment	~	PTSD	0.38	0.05	0.38	0.05	0.38	7.10E-14
C) Childhood maltreatment ~ ADHD + PTSD + MDD								

Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.04	0.00	0.53	0.05	0.53	9.84E-24
Childhood maltreatment	~	ADHD	0.17	0.03	0.31	0.06	0.31	6.43E-08
Childhood maltreatment	~	PTSD	0.24	0.06	0.24	0.06	0.24	7.30E-05
Childhood maltreatment	~	MDD	0.26	0.05	0.28	0.05	0.28	1.38E-07

D) Childhood maltreatment ~ ADHD + PTSD + MDD + alcohol dependence								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.04	0.00	0.51	0.06	0.51	5.57E-17
Childhood maltreatment	~	ADHD	0.15	0.04	0.26	0.07	0.26	0.00
Childhood maltreatment	~	PTSD	0.20	0.07	0.20	0.07	0.20	0.00
Childhood maltreatment	~	MDD	0.21	0.06	0.23	0.06	0.23	0.00
Childhood maltreatment	~	ALCH	0.18	0.11	0.20	0.13	0.20	0.11

E) Childhood maltreatment ~ ADHD + PTSD + MDD + alcohol dependence + anxiety (ANX)								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.04	0.00	0.49	0.07	0.49	8.88E-13
Childhood maltreatment	~	ADHD	0.13	0.05	0.23	0.09	0.23	0.01
Childhood maltreatment	~	PTSD	0.23	0.08	0.22	0.08	0.22	0.00
Childhood maltreatment	~	MDD	0.38	0.17	0.42	0.18	0.42	0.02
Childhood maltreatment	~	ALCH	0.21	0.13	0.24	0.15	0.24	0.11
Childhood maltreatment	~	ANX	-0.16	0.15	-0.23	0.22	-0.23	0.29

F) Childhood maltreatment ~ ADHD + PTSD + MDD + alcohol dependence + anxiety (ANX) + autism spectrum disorder								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.01	0.45	0.08	0.45	5.12E-08
Childhood maltreatment	~	ADHD	0.07	0.07	0.13	0.12	0.13	0.28
Childhood maltreatment	~	PTSD	0.30	0.09	0.30	0.09	0.30	0.00
Childhood maltreatment	~	MDD	0.25	0.18	0.27	0.20	0.27	0.17
Childhood maltreatment	~	ALCH	0.28	0.16	0.32	0.18	0.32	0.08
Childhood maltreatment	~	ANX	-0.17	0.17	-0.25	0.24	-0.25	0.30
Childhood maltreatment	~	ASD	0.21	0.08	0.27	0.10	0.27	0.01

G) Childhood maltreatment ~ ADHD + PTSD + MDD + alcohol dependence + anxiety (ANX) + autism spectrum disorder + Schizophrenia								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.01	0.44	0.08	0.44	9.23E-09
Childhood maltreatment	~	ADHD	0.09	0.07	0.15	0.12	0.15	0.21
Childhood maltreatment	~	PTSD	0.29	0.09	0.29	0.09	0.29	0.00
Childhood maltreatment	~	MDD	0.24	0.18	0.27	0.19	0.27	0.17
Childhood maltreatment	~	ALCH	0.24	0.17	0.28	0.19	0.28	0.15
Childhood maltreatment	~	ANX	-0.18	0.16	-0.26	0.23	-0.26	0.26
Childhood maltreatment	~	ASD	0.19	0.08	0.25	0.11	0.25	0.02

Childhood maltreatment	~	SCZ	0.06	0.03	0.11	0.06	0.11	0.06
------------------------	---	-----	------	------	------	------	------	------

H) Cross psychiatric disorder genetic correlations									
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value_diff_0	abs_p_value_diff_1
MDD	~~	ANX	0.0984	0.0049	0.8643	0.0427	0.8643	3.04E-91	1.48E-03
ANX	~~	PTSD	0.0584	0.0044	0.5614	0.0426	0.5614	9.77E-40	6.57E-25
MDD	~~	PTSD	0.0449	0.0024	0.5753	0.0303	0.5753	1.63E-80	1.06E-44
ANX	~~	ASD	0.0491	0.0076	0.3686	0.0571	0.3686	1.1E-10	2.10E-28
MDD	~~	ASD	0.0473	0.0035	0.4737	0.0346	0.4737	1.11E-42	2.78E-52
PTSD	~~	ASD	0.0118	0.0040	0.1295	0.0438	0.1295	0.0031	4.86E-88
MDD	~~	ALCH	0.0422	0.0052	0.4738	0.0585	0.4738	5.78E-16	2.47E-19
ANX	~~	ALCH	0.0602	0.0106	0.5057	0.0891	0.5057	1.37E-08	2.87E-08
PTSD	~~	ALCH	0.0387	0.0066	0.4746	0.0814	0.4746	5.46E-09	1.06E-10
ASD	~~	ALCH	0.0038	0.0085	0.0366	0.0817	0.0366	0.6542	4.15E-32
MDD	~~	ADHD	0.0608	0.0047	0.4419	0.0340	0.4419	1.4E-38	1.67E-60
ANX	~~	ADHD	0.0616	0.0077	0.3356	0.0418	0.3356	9.29E-16	5.52E-57
PTSD	~~	ADHD	0.0637	0.0048	0.5064	0.0385	0.5064	1.6E-39	1.19E-37
ASD	~~	ADHD	0.0582	0.0084	0.3614	0.0524	0.3614	5.45E-12	3.89E-34
ALCH	~~	ADHD	0.0658	0.0115	0.4576	0.0801	0.4576	0.0000	1.25E-11
PTSD	~~	SCZ	0.0329	0.0037	0.2414	0.0268	0.2414	2.35E-19	7.10E-176
ADHD	~~	SCZ	0.0300	0.0076	0.1249	0.0316	0.1249	0.0001	6.30E-169
MDD	~~	SCZ	0.0502	0.0040	0.3371	0.0269	0.3371	3.84E-36	1.75E-134
ALCH	~~	SCZ	0.0567	0.0082	0.3637	0.0529	0.3637	6.28E-12	2.71E-33
ANX	~~	SCZ	0.0689	0.0071	0.3467	0.0357	0.3467	3.01E-22	1.15E-74
ASD	~~	SCZ	0.0388	0.0069	0.2221	0.0396	0.2221	2.02E-08	6.40E-86

Supplementary table 5: Genomic SEM genomic multiple regression analyses of independently contributing traits selected based on results shown in figure 2A/ supplementary table 3K and figure 2B/ supplementary table 4G.

The model takes into account the genetic correlations across traits.

Final independent traits model: Childhood maltreatment ~ ASD + Risk + Subjective wellbeing									
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value	
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.00	0.42	0.06	0.42	1.24E-12	
Childhood maltreatment	~	ASD	0.16	0.05	0.21	0.06	0.21	5.8E-04	
Childhood maltreatment	~	Risk	0.42	0.04	0.36	0.04	0.36	9.15E-21	
Childhood maltreatment	~	Subjective_wellbeing	-0.54	0.14	-0.32	0.08	-0.32	1.2E-04	
Childhood maltreatment	~	PTSD	0.32	0.05	0.32	0.05	0.32	7.98E-10	
ASD	~~	PTSD	0.01	0.00	0.13	0.04	0.13	3.1E-03	

Risk	~~	PTSD	0.01	0.00	0.24	0.03	0.24	6.66E-18
Subjective_wellbeing	~~	PTSD	-0.02	0.00	-0.40	0.04	-0.40	3.39E-19
PTSD	~~	PTSD	0.07	0.00	1.00	0.05	1.00	1.80E-75
ASD	~~	ASD	0.12	0.01	1.00	0.08	1.00	5.0E-39
ASD	~~	Risk	0.00	0.00	0.02	0.05	0.02	7.3E-01
ASD	~~	Subjective wellbeing	-0.02	0.00	-0.46	0.05	-0.46	4.3E-18
Risk	~~	Risk	0.05	0.00	1.00	0.04	1.00	9.6E-115
Risk	~~	Subjective wellbeing	0.00	0.00	0.05	0.04	0.05	1.6E-01
Subjective wellbeing	~~	Subjective wellbeing	0.02	0.00	1.00	0.08	1.00	5.9E-35

Supplementary table 6 of sensitivity analysis: Genomic SEM genomic multiple regression analyses of independently contributing traits (in figure 3/ supplementary table 5) and two additional traits that broadly capture the shared genetic associations observed across the psychiatric disorders (as captured by ADHD) and health traits (as captured by self-reported tiredness).

The model takes into account the genetic correlations across traits.

Sensitivity analysis model: Childhood maltreatment ~ ASD + General risk-tolerance + Subjective wellbeing + PTSD + ADHD + Self-repored tiredness								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Childhood maltreatment	~~	Childhood maltreatment	0.03	0.00	0.40	0.06	0.40	2.11E-12
Childhood maltreatment	~	ASD	0.12	0.05	0.16	0.07	0.16	1.69E-02
Childhood maltreatment	~	Risk	0.37	0.05	0.31	0.04	0.31	1.20E-12
Childhood maltreatment	~	Subjective wellbeing	-0.54	0.15	-0.32	0.09	-0.32	4.23E-04
Childhood maltreatment	~	PTSD	0.26	0.09	0.26	0.09	0.26	6.52E-03
Childhood maltreatment	~	Tiredness	-0.01	0.13	-0.01	0.12	-0.01	9.56E-01
Childhood maltreatment	~	ADHD	0.09	0.04	0.17	0.07	0.17	1.34E-02
ASD	~~	ASD	0.12	0.01	1.00	0.08	1.00	4.99E-39
ASD	~~	Risk	0.00	0.00	0.02	0.05	0.02	7.29E-01
ASD	~~	Subjective wellbeing	-0.02	0.00	-0.46	0.05	-0.46	4.33E-18
ASD	~~	PTSD Avoidance	0.01	0.00	0.15	0.05	0.15	1.22E-03
ASD	~~	Tiredness	0.03	0.00	0.32	0.05	0.32	1.25E-09
Risk	~~	Risk	0.05	0.00	1.00	0.04	1.00	9.56E-115
Risk	~~	Subjective wellbeing	0.00	0.00	0.05	0.04	0.05	1.57E-01
Risk	~~	PTSD Avoidance	0.01	0.00	0.24	0.03	0.24	4.58E-16
Risk	~~	Tiredness	0.01	0.00	0.13	0.04	0.13	4.64E-04
Subjective wellbeing	~~	Subjective wellbeing	0.02	0.00	1.00	0.08	1.00	5.91E-35
Subjective wellbeing	~~	PTSD Avoidance	-0.02	0.00	-0.44	0.05	-0.44	2.06E-19
Subjective wellbeing	~~	Tiredness	-0.02	0.00	-0.57	0.06	-0.57	4.07E-24
PTSD Avoidance	~~	PTSD Avoidance	0.06	0.00	1.00	0.06	1.00	5.00E-63
PTSD Avoidance	~~	Tiredness	0.05	0.00	0.71	0.05	0.71	9.76E-48
Tiredness	~~	Tiredness	0.07	0.01	1.00	0.09	1.00	1.99E-30

Supplementary tables for self-reported lifetime trauma

Link to access to Supplementary Tables 7-8: LD Score genetic correlation (rg) results between self-reported trauma and 576 phenotypes. Genetic correlation results between self-reported lifetime trauma and 576 complex traits were highly similar to genetic correlation results observed with reported childhood maltreatment. For brevity, Supplementary Table 7 and 8 are not presented here, and can be accessed via the link <https://tinyurl.com/Chapter3AppendixBTables>. **Supplementary Table 7** shows the genetic correlation results for all phenotypes with self-reported trauma, and **Supplementary Table 8** shows significantly genetically correlated traits and those selected for genomic multiple regression.

Supplementary table 9: A-K Genomic SEM multivariable models where each psychological, social and health trait is introduced in iterative stages, in order of the most highly genetically correlated trait with self-reported lifetime trauma. Model K standardised results is the final model included in the main analyses. Each model takes into account the genetic correlations across Psychological and Social traits, but for simplicity, these genetic correlations are only shown in Table L.

A) Lifetime trauma ~ Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.1268	0.0100	0.6986	0.0550	0.6986	5.64E-37
Lifetime trauma	~	Sexual Partners	0.7424	0.0399	0.5490	0.0295	0.5490	3.60E-77
B) Lifetime trauma ~ Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0925	0.0114	0.5097	0.0630	0.5097	6.12E-16
Lifetime trauma	~	Sexual Partners	0.6018	0.0436	0.4450	0.0322	0.4450	2.10E-43
Lifetime trauma	~	Tiredness	0.7347	0.0936	0.4469	0.0569	0.4469	4.07E-15
C) Lifetime trauma ~ Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0812	0.0111	0.4471	0.0612	0.4471	2.84E-13
Lifetime trauma	~	Sexual Partners	0.5838	0.0447	0.4317	0.0330	0.4317	4.64E-39
Lifetime trauma	~	Tiredness	0.4564	0.1320	0.2776	0.0803	0.2776	5.44E-04
Lifetime trauma	~	Subj.wellbeing	-0.8207	0.2134	-0.3041	0.0791	-0.3041	1.20E-04

D) Lifetime trauma ~ Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0661	0.0122	0.3640	0.0673	0.3640	6.36E-08
Lifetime trauma	~	Sexual Partners	0.2834	0.0664	0.2096	0.0491	0.2096	2.00E-05
Lifetime trauma	~	Tiredness	0.3854	0.1375	0.2344	0.0837	0.2344	5.07E-03
Lifetime trauma	~	Subj.wellbeing	-1.0421	0.2332	-0.3861	0.0864	-0.3861	7.84E-06
Lifetime trauma	~	Risk	0.6885	0.1015	0.3677	0.0542	0.3677	1.20E-11
E) Lifetime trauma ~ Social deprivation + Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0640	0.0121	0.3527	0.0669	0.3527	1.37E-07
Lifetime trauma	~	Sexual Partners	0.1989	0.0977	0.1471	0.0722	0.1471	4.17E-02
Lifetime trauma	~	Tiredness	0.2896	0.1708	0.1761	0.1039	0.1761	9.00E-02
Lifetime trauma	~	Subj.wellbeing	-1.0468	0.2328	-0.3879	0.0863	-0.3879	6.93E-06
Lifetime trauma	~	Risk	0.7186	0.1064	0.3838	0.0568	0.3838	1.44E-11
Lifetime trauma	~	Social dep	0.3011	0.2407	0.1378	0.1101	0.1378	2.11E-01
F) Lifetime trauma ~ Health + Social dep + Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0607	0.0134	0.3345	0.0739	0.3345	5.97E-06
Lifetime trauma	~	Sexual Partners	0.1379	0.1251	0.1020	0.0925	0.1020	2.70E-01
Lifetime trauma	~	Tiredness	0.5097	0.2210	0.3100	0.1344	0.3100	2.11E-02
Lifetime trauma	~	Subj.wellbeing	-1.0953	0.2537	-0.4058	0.0940	-0.4058	1.57E-05
Lifetime trauma	~	Risk	0.7523	0.1156	0.4017	0.0617	0.4017	7.49E-11
Lifetime trauma	~	Social dep	0.5046	0.3315	0.2309	0.1517	0.2309	1.28E-01
Lifetime trauma	~	Health	0.3257	0.2059	0.2388	0.1510	0.2388	1.14E-01
G) Lifetime trauma ~ Age first birth + Health + Social dep + Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0561	0.0144	0.3089	0.0792	0.3089	9.59E-05
Lifetime trauma	~	Sexual Partners	0.1504	0.1281	0.1112	0.0947	0.1112	2.40E-01
Lifetime trauma	~	Tiredness	0.5914	0.2316	0.3597	0.1408	0.3597	1.06E-02
Lifetime trauma	~	Subj.wellbeing	-1.2061	0.2859	-0.4468	0.1059	-0.4468	2.47E-05
Lifetime trauma	~	Risk	0.7001	0.1190	0.3739	0.0636	0.3739	4.04E-09
Lifetime trauma	~	Social dep	0.3391	0.3692	0.1552	0.1689	0.1552	3.58E-01
Lifetime trauma	~	Health	0.5231	0.2448	0.3836	0.1795	0.3836	3.26E-02
Lifetime trauma	~	Age first birth	-0.4110	0.2126	-0.2263	0.1171	-0.2263	5.32E-02

H) Lifetime trauma ~ Ever Smoker + Age first birth + Health + Social dep + Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0556	0.0144	0.3063	0.0792	0.3063	1.10E-04
Lifetime trauma	~	Sexual Partners	0.1822	0.1235	0.1347	0.0913	0.1347	1.40E-01
Lifetime trauma	~	Tiredness	0.5731	0.2353	0.3485	0.1431	0.3485	1.49E-02
Lifetime trauma	~	Subj.wellbeing	-1.2040	0.2856	-0.4461	0.1058	-0.4461	2.49E-05
Lifetime trauma	~	Risk	0.7045	0.1203	0.3762	0.0643	0.3762	4.76E-09
Lifetime trauma	~	Social dep	0.3907	0.3996	0.1788	0.1829	0.1788	3.28E-01
Lifetime trauma	~	Health	0.5051	0.2457	0.3704	0.1802	0.3704	3.98E-02
Lifetime trauma	~	Age first birth	-0.4178	0.2110	-0.2300	0.1162	-0.2300	4.77E-02
Lifetime trauma	~	Ever smoker	-0.0967	0.0926	-0.0687	0.0657	-0.0687	2.96E-01
I) Lifetime trauma ~ Neuroticism + Ever Smoker + Age first birth + Health + Social dep + Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0532	0.0175	0.2930	0.0962	0.2930	2.31E-03
Lifetime trauma	~	Neuroticism	-0.2462	0.2578	-0.1938	0.2029	-0.1938	3.40E-01
Lifetime trauma	~	Sexual Partners	0.1337	0.1470	0.0988	0.1087	0.0988	3.63E-01
Lifetime trauma	~	Tiredness	0.7916	0.4134	0.4815	0.2515	0.4815	5.55E-02
Lifetime trauma	~	Subj.wellbeing	-1.5155	0.5728	-0.5615	0.2122	-0.5615	8.15E-03
Lifetime trauma	~	Risk	0.7016	0.1315	0.3747	0.0702	0.3747	9.45E-08
Lifetime trauma	~	Social dep	0.2843	0.4785	0.1301	0.2190	0.1301	5.52E-01
Lifetime trauma	~	Health	0.6516	0.3785	0.4779	0.2776	0.4779	8.52E-02
Lifetime trauma	~	Age first birth	-0.5221	0.3189	-0.2875	0.1756	-0.2875	1.02E-01
Lifetime trauma	~	Ever smoker	-0.0504	0.1246	-0.0357	0.0885	-0.0357	6.86E-01
J) Lifetime trauma ~ Number of offspring + Neuroticism + Ever Smoker + Age first birth + Health + Social dep + Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0513	0.0206	0.2826	0.1132	0.2826	1.26E-02
Lifetime trauma	~	Neuroticism	-0.3229	0.3775	-0.2542	0.2972	-0.2542	3.92E-01
Lifetime trauma	~	Number offspring	-0.4613	0.6782	-0.1669	0.2454	-0.1669	4.96E-01
Lifetime trauma	~	Sexual partners	0.0839	0.1948	0.0620	0.1441	0.0620	6.67E-01
Lifetime trauma	~	Tiredness	0.9590	0.6509	0.5833	0.3959	0.5833	1.41E-01
Lifetime trauma	~	Subj.wellbeing	-1.6329	0.7584	-0.6050	0.2810	-0.6050	3.13E-02
Lifetime trauma	~	Risk	0.7707	0.1724	0.4116	0.0921	0.4116	7.84E-06
Lifetime trauma	~	Social dep	0.2354	0.5586	0.1077	0.2556	0.1077	6.73E-01
Lifetime trauma	~	Health	0.8305	0.6314	0.6090	0.4630	0.6090	1.88E-01
Lifetime trauma	~	Age first birth	-0.8258	0.7510	-0.4547	0.4135	-0.4547	2.72E-01
Lifetime trauma	~	Ever smoker	-0.0329	0.1514	-0.0234	0.1074	-0.0234	8.28E-01

K) Lifetime trauma ~ Insomnia + Number of offspring + Neuroticism + Ever Smoker + Age first birth + Health + Social dep + Risk + Subj.wellbeing + Tiredness + Sexual Partners								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0451	0.0239	0.2483	0.1316	0.2483	5.92E-02
Lifetime trauma	~	Neuroticism	-0.3443	0.4291	-0.2711	0.3378	-0.2711	4.22E-01
Lifetime trauma	~	Number offspring	-0.6231	0.8197	-0.2254	0.2965	-0.2254	4.47E-01
Lifetime trauma	~	Insomnia	-0.3507	0.2304	-0.2404	0.1579	-0.2404	1.28E-01
Lifetime trauma	~	Sexual partners	0.0791	0.2128	0.0585	0.1574	0.0585	7.10E-01
Lifetime trauma	~	Tiredness	1.1702	0.8056	0.7117	0.4900	0.7117	1.46E-01
Lifetime trauma	~	Subj.wellbeing	-1.7841	0.9023	-0.6610	0.3343	-0.6610	4.80E-02
Lifetime trauma	~	Risk	0.7841	0.1917	0.4187	0.1024	0.4187	4.29E-05
Lifetime trauma	~	Social dep	0.2508	0.6178	0.1148	0.2827	0.1148	6.85E-01
Lifetime trauma	~	Health	0.8974	0.7256	0.6581	0.5321	0.6581	2.16E-01
Lifetime trauma	~	Age first birth	-1.0058	0.9178	-0.5538	0.5054	-0.5538	2.73E-01
Lifetime trauma	~	Ever smoker	-0.0292	0.1683	-0.0207	0.1195	-0.0207	8.62E-01

L) Cross Psychological & Social trait genetic correlations									
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value_diff_0	p_value_diff_1
Neuroticism	~~	Number offspring	0.0017	0.0018	0.0322	0.0346	0.0322	3.52E-01	3.652E-172
Neuroticism	~~	Insomnia	0.0370	0.0043	0.3779	0.0444	0.3779	1.58E-17	1.3302E-44
Number offspring	~~	Insomnia	0.0054	0.0028	0.1199	0.0632	0.1199	5.79E-02	4.426E-44
Sexual partners	~~	Neuroticism	0.0009	0.0029	0.0081	0.0278	0.0081	7.70E-01	8.117E-279
Sexual partners	~~	Number offspring	0.0075	0.0017	0.1536	0.0342	0.1536	7.10E-06	3.215E-135
Sexual partners	~~	Insomnia	0.0190	0.0041	0.2067	0.0440	0.2067	2.66E-06	1.1423E-72
Sexual partners	~~	Tiredness	0.0190	0.0029	0.2327	0.0360	0.2327	9.8E-11	8.439E-101
Sexual partners	~~	Subj.wellbeing	-0.0086	0.0017	-0.1734	0.0351	-0.1734	7.83E-07	1.26E-122
Sexual partners	~~	Risk	0.0425	0.0022	0.5928	0.0310	0.5928	1.32E-81	2.0617E-39
Sexual partners	~~	Social dep	0.0295	0.0028	0.4803	0.0463	0.4803	3.32E-25	3.087E-29
Sexual partners	~~	Health	-0.0184	0.0030	-0.1866	0.0309	-0.1866	1.51E-09	1.029E-152
Sexual partners	~~	Age first birth	-0.0170	0.0020	-0.2299	0.0268	-0.2299	1.03E-17	1.392E-181
Sexual partners	~~	Ever smoker	0.0511	0.0024	0.5367	0.0253	0.5367	1.75E-99	6.6078E-75
Tiredness	~~	Neuroticism	0.0498	0.0037	0.5728	0.0429	0.5728	1.39E-40	2.326E-23
Tiredness	~~	Number offspring	0.0089	0.0023	0.2217	0.0573	0.2217	1.09E-04	5.0585E-42
Tiredness	~~	Insomnia	0.0436	0.0050	0.5758	0.0665	0.5758	4.74E-18	1.7831E-10
Tiredness	~~	Subj.wellbeing	-0.0232	0.0023	-0.5670	0.0560	-0.5670	4.07E-24	1.0575E-14
Tiredness	~~	Risk	0.0078	0.0022	0.1315	0.0376	0.1315	4.64E-04	4.804E-118
Tiredness	~~	Social dep	0.0256	0.0035	0.5059	0.0691	0.5059	2.45E-13	8.6458E-13
Tiredness	~~	Health	-0.0620	0.0049	-0.7663	0.0600	-0.7663	2.17E-37	9.8199E-05

Tiredness	~~	Age first birth	-0.0209	0.0026	-0.3439	0.0435	-0.3439	2.76E-15	2.103E-51
Tiredness	~~	Ever smoker	0.0201	0.0026	0.2560	0.0335	0.2560	2.35E-14	2.814E-109
Subj.wellbeing	~~	Neuroticism	-0.0351	0.0025	-0.6633	0.0469	-0.6633	2.28E-45	7.017E-13
Subj.wellbeing	~~	Number offspring	-0.0001	0.0013	-0.0021	0.0547	-0.0021	9.69E-01	2.3457E-74
Subj.wellbeing	~~	Insomnia	-0.0201	0.0028	-0.4356	0.0602	-0.4356	4.75E-13	6.8904E-21
Subj.wellbeing	~~	Risk	0.0019	0.0013	0.0518	0.0366	0.0518	1.57E-01	5.539E-148
Subj.wellbeing	~~	Social dep	-0.0096	0.0022	-0.3117	0.0730	-0.3117	1.98E-05	4.1494E-21
Subj.wellbeing	~~	Health	0.0236	0.0026	0.4779	0.0529	0.4779	1.66E-19	5.6409E-23
Subj.wellbeing	~~	Age first birth	0.0034	0.0017	0.0908	0.0457	0.0908	4.71E-02	4.499E-88
Subj.wellbeing	~~	Ever smoker	-0.0076	0.0016	-0.1587	0.0330	-0.1587	1.51E-06	2.301E-143
Risk	~~	Neuroticism	-0.0041	0.0023	-0.0535	0.0295	-0.0535	6.98E-02	7.197E-226
Risk	~~	Number offspring	0.0107	0.0015	0.3052	0.0433	0.3052	1.89E-12	6.0783E-58
Risk	~~	Insomnia	0.0058	0.0030	0.0869	0.0450	0.0869	5.33E-02	1.5407E-91
Risk	~~	Social dep	0.0092	0.0022	0.2083	0.0504	0.2083	3.62E-05	1.3262E-55
Risk	~~	Health	-0.0081	0.0026	-0.1144	0.0362	-0.1144	1.60E-03	3.564E-132
Risk	~~	Age first birth	-0.0119	0.0018	-0.2224	0.0330	-0.2224	1.67E-11	9.099E-123
Risk	~~	Ever smoker	0.0222	0.0018	0.3222	0.0262	0.3222	8.69E-35	1.441E-147
Social dep	~~	Neuroticism	0.0074	0.0028	0.1139	0.0430	0.1139	8.14E-03	2.3762E-94
Social dep	~~	Number offspring	0.0084	0.0021	0.2811	0.0715	0.2811	8.56E-05	8.7725E-24
Social dep	~~	Insomnia	0.0233	0.0044	0.4101	0.0773	0.4101	1.11E-07	2.3239E-14
Social dep	~~	Health	-0.0379	0.0038	-0.6219	0.0628	-0.6219	4.33E-23	1.7366E-09
Social dep	~~	Age first birth	-0.0261	0.0029	-0.5711	0.0623	-0.5711	4.98E-20	5.8019E-12
Social dep	~~	Ever smoker	0.0343	0.0025	0.5811	0.0429	0.5811	1.04E-41	1.5979E-22
Health	~~	Neuroticism	-0.0320	0.0049	-0.3058	0.0467	-0.3058	5.83E-11	5.5535E-50
Health	~~	Number offspring	-0.0108	0.0024	-0.2248	0.0508	-0.2248	9.55E-06	1.4159E-52
Health	~~	Insomnia	-0.0466	0.0053	-0.5109	0.0583	-0.5109	1.99E-18	4.8877E-17
Health	~~	Age first birth	0.0439	0.0030	0.5996	0.0406	0.5996	2.1E-49	6.0784E-23
Health	~~	Ever smoker	-0.0357	0.0027	-0.3775	0.0285	-0.3775	5.26E-40	9.24E-106
Age first birth	~~	Neuroticism	-0.0105	0.0024	-0.1334	0.0306	-0.1334	1.31E-05	1.946E-176
Age first birth	~~	Number offspring	-0.0241	0.0016	-0.6669	0.0456	-0.6669	1.62E-48	2.7763E-13
Age first birth	~~	Insomnia	-0.0238	0.0036	-0.3477	0.0530	-0.3477	5.49E-11	8.2492E-35
Age first birth	~~	Ever smoker	-0.0281	0.0021	-0.3960	0.0303	-0.3960	3.96E-39	2.0642E-88
Ever smoker	~~	Neuroticism	0.0101	0.0023	0.1000	0.0231	0.1000	1.44E-05	0.00E+00
Ever smoker	~~	Number offspring	0.0098	0.0016	0.2104	0.0345	0.2104	1.07E-09	6.267E-116
Ever smoker	~~	Insomnia	0.0205	0.0035	0.2317	0.0395	0.2317	4.35E-09	2.8786E-84

Supplementary table 10: A-G Genomic SEM multivariable models where each psychiatric disorder is introduced in iterative stages, in order of the most highly genetically correlated trait with self-reported lifetime trauma.

*Model (G) standardised results is the final model included in the main analyses. Each model takes into account the genetic correlations across Psychiatric disorders, but for simplicity, these genetic correlations are only shown in **Supplementary Table 4H**.*

A) Lifetime trauma ~ MDD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.1179	0.0111	0.6496	0.0612	0.6496	2.43E-26
Lifetime trauma	~	MDD	0.8637	0.0447	0.5920	0.0306	0.5920	2.57E-83
B) Lifetime trauma ~ Anxiety + MDD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.1166	0.0102	0.6426	0.0563	0.6426	3.21E-30
Lifetime trauma	~	MDD	0.6535	0.2292	0.4479	0.1571	0.4479	4.35E-03
Lifetime trauma	~	ANX	0.1822	0.1841	0.1667	0.1684	0.1667	3.22E-01
C) Lifetime trauma ~ PTSD symptoms + Anxiety + MDD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.1080	0.0106	0.5951	0.0586	0.5951	3.16E-24
Lifetime trauma	~	MDD	0.5134	0.2264	0.3518	0.1552	0.3518	2.34E-02
Lifetime trauma	~	ANX	0.1074	0.1897	0.0983	0.1736	0.0983	5.71E-01
Lifetime trauma	~	PTSD	0.4303	0.0921	0.2697	0.0577	0.2697	3.00E-06
D) Lifetime trauma ~ ASD + PTSD symptoms + Anxiety + MDD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0943	0.0115	0.5195	0.0633	0.5195	2.15E-16
Lifetime trauma	~	MDD	0.1941	0.2709	0.1331	0.1857	0.1331	4.74E-01
Lifetime trauma	~	ANX	0.1458	0.2064	0.1334	0.1888	0.1334	4.80E-01
Lifetime trauma	~	PTSD	0.5338	0.1017	0.3345	0.0638	0.3345	1.55E-07
Lifetime trauma	~	ASD	0.3976	0.0880	0.3191	0.0706	0.3191	6.18E-06
E) Lifetime trauma ~ Alcohol dependence + ASD + PTSD symptoms + Anxiety + MDD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0864	0.0135	0.4757	0.0743	0.4757	1.50E-10
Lifetime trauma	~	MDD	0.1290	0.2932	0.0884	0.2009	0.0884	6.60E-01
Lifetime trauma	~	ANX	0.0642	0.2428	0.0587	0.2221	0.0587	7.92E-01
Lifetime trauma	~	PTSD	0.4377	0.1151	0.2743	0.0722	0.2743	1.44E-04
Lifetime trauma	~	ASD	0.4563	0.0943	0.3662	0.0757	0.3662	1.30E-06
Lifetime trauma	~	ALCH	0.3579	0.1829	0.2565	0.1310	0.2565	5.03E-02

F) Lifetime trauma ~ ADHD + Alcohol dependence + ASD + PTSD symptoms + Anxiety + MDD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0862	0.0138	0.4750	0.0758	0.4750	3.64E-10
Lifetime trauma	~	MDD	0.1406	0.3096	0.0964	0.2122	0.0964	6.50E-01
Lifetime trauma	~	ANX	0.0496	0.2710	0.0454	0.2479	0.0454	8.55E-01
Lifetime trauma	~	PTSD	0.4602	0.1415	0.2884	0.0887	0.2884	1.15E-03
Lifetime trauma	~	ASD	0.4716	0.1157	0.3785	0.0929	0.3785	4.60E-05
Lifetime trauma	~	ALCH	0.3759	0.2174	0.2694	0.1558	0.2694	8.37E-02
Lifetime trauma	~	ADHD	-0.0338	0.1042	-0.0373	0.1152	-0.0373	7.46E-01
G) Lifetime trauma ~ SCZ + ADHD + Alcohol dependence + ASD + PTSD symptoms + Anxiety + MDD								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0847	0.0130	0.4667	0.0718	0.4667	8.00E-11
Lifetime trauma	~	MDD	0.1328	0.2980	0.0910	0.2042	0.0910	6.56E-01
Lifetime trauma	~	ANX	0.0426	0.2612	0.0390	0.2389	0.0390	8.70E-01
Lifetime trauma	~	PTSD	0.4466	0.1371	0.2799	0.0859	0.2799	1.13E-03
Lifetime trauma	~	ASD	0.4429	0.1202	0.3555	0.0964	0.3555	2.28E-04
Lifetime trauma	~	ALCH	0.3253	0.2256	0.2331	0.1617	0.2331	1.49E-01
Lifetime trauma	~	ADHD	-0.0149	0.1048	-0.0164	0.1159	-0.0164	8.87E-01
Lifetime trauma	~	SCZ	0.0857	0.0453	0.1026	0.0542	0.1026	5.84E-02

Supplementary table 11: Genomic SEM genomic multiple regression analyses of independently contributing traits to self-reported lifetime trauma, selected based on results shown in supplementary table 9K and supplementary table 10G.

The model takes into account the genetic correlations across traits.

Final independent traits model: Lifetime trauma ~ ASD + Risk + Subjective wellbeing + PTSD Symptoms								
Phenotype 1	op	Phenotype 2	Unstand_Est	Unstand_SE	STD_Genotype	STD_Genotype_SE	STD_All	p_value
Lifetime trauma	~~	Lifetime trauma	0.0598	0.0124	0.3292	0.0683	0.3292	1.46E-06
Lifetime trauma	~	Subjective wellbeing	-0.8918	0.2254	-0.3304	0.0835	-0.3304	7.60E-05
Lifetime trauma	~	Risk	0.8521	0.0833	0.4550	0.0445	0.4550	1.45E-24
Lifetime trauma	~	ASD	0.3528	0.0756	0.2832	0.0607	0.2832	3.06E-06
Lifetime trauma	~	PTSD	0.3947	0.0932	0.2474	0.0584	0.2474	2.27E-05
ASD	~~	PTSD	0.0118	0.0040	0.1295	0.0438	0.1295	3.1E-03
Subjective wellbeing	~~	ASD	-0.0248	0.0029	-0.4603	0.0531	-0.4603	4.33E-18
Risk	~~	ASD	0.0013	0.0036	0.0161	0.0464	0.0161	7.3E-01
Subjective wellbeing	~~	PTSD	-0.0169	0.0019	-0.4007	0.0447	-0.4007	3.39E-19
Risk	~~	PTSD	0.0148	0.0017	0.2436	0.0283	0.2436	6.66E-18
Subjective wellbeing	~~	Risk	0.0019	0.0013	0.0518	0.0366	0.0518	1.6E-01
Trauma	~~	Trauma	0.0598	0.0124	0.3292	0.0683	0.3292	1.46E-06
Subjective wellbeing	~~	Subjective wellbeing	0.0249	0.0020	1.0000	0.0811	1.0000	5.91E-35
Risk	~~	Risk	0.0518	0.0023	1.0000	0.0439	1.0000	9.56E-115
ASD	~~	ASD	0.1169	0.0089	1.0000	0.0765	1.0000	4.99E-39
PTSD	~~	PTSD	0.0713	0.0039	1.0000	0.0544	1.0000	1.80E-75

Link access to Supplementary Table 12: GWAS summary statistics data availability information.

Details on data availability and download link for each GWAS summary statistic used in Chapter 3 is available via the following link

<https://tinyurl.com/Chapter3AppendixBTables>.

Appendix C. Supplementary materials for Chapter 4

Supplementary Tables	242
<i>Supplementary Table 1: Proportion of participants with lifetime experience of a depressive or anxiety disorder diagnoses in the GLAD+ and AGDS datasets, total N = 15,131.....</i>	<i>242</i>
<i>Supplementary Table 2: Proportion of types of reported psychological treatments.</i>	<i>243</i>
<i>Supplementary Table 3: Genome-wide association study summary statistics used to derived polygenic scores in the GLAD+ dataset</i>	<i>244</i>
<i>Supplementary Table 4: Comparison of models predicting retrospectively self-reported psychological treatment outcomes using genetic, sociodemographic and clinical predictors, and all predictors in the GLAD+ dataset (N = 4,439).....</i>	<i>245</i>
<i>Supplementary Table 5: Standardised coefficients for the predictors retained in the final model based on the average optimal λ tuning parameter following 10-fold cross validation</i>	<i>245</i>
Supplementary Figures	247
<i>Supplementary Figure 1: Nested 10-fold cross validation.</i>	<i>247</i>
<i>Supplementary Figure 2: Q-Q plot (upper) and Manhattan plot (lower) of results from GWAS of retrospectively self-reported treatment outcomes in the GLAD+ dataset.</i>	<i>248</i>
<i>Supplementary Figure 3: Q-Q plot (upper) and Manhattan plot (lower) of results from GWAS of retrospectively self-reported treatment outcomes in the AGDS dataset.....</i>	<i>249</i>
References	250

Supplementary Tables

Supplementary Table 1: Proportion of participants with lifetime experience of a depressive or anxiety disorder diagnoses in the GLAD+ and AGDS datasets, total N = 15,131

	GLAD+		AGDS
	GLAD Study N = 5,667	COPING Study N = 1,869	N = 7,595
Depression	3102 (96.8%)	1692 (90.5%)	7595 (100.0%)
Generalised anxiety disorder	2762 (86.2%)	1037 (55.5%)	5274 (69.4%)
Panic disorder	1393 (43.4%)	241 (12.9%)	2496 (32.9%)
Agoraphobia	562 (17.5%)	83 (4.4%)	489 (6.4%)
Social anxiety	1430 (44.6%)	243 (13.0%)	1555 (20.5%)
Specific phobia	810 (25.3%)	195 (10.4%)	941 (12.3%)

Lifetime experience of a depressive or anxiety disorder is defined by participants either meeting diagnostic criteria based on reported symptoms or reporting receiving a diagnosis from a medical professional (any doctor, nurse or person with specialist training, such as a psychologist or psychiatrist). Further detail is given in Davies et al.¹.

Supplementary Table 2: Proportion of types of reported psychological treatments.

Treatment type	GLAD+		AGDS
	GLAD Study	COPING Study	
Self-reported benefit from therapy - adapted CIDI-SF modules*	n = 3206	n = 1869	
Cognitive behavioural therapy (CBT)	2075 (64.7%)	878 (47.0%)	-
Counselling	2579 (80.4%)	1569 (83.9%)	-
Group therapy	652 (20.3%)	244 (13.1%)	-
Guided self-help	897 (28.0%)	365 (19.5%)	-
Workshops	441 (13.8%)	156 (8.3%)	-
Relationship therapy	272 (8.5%)	203 (10.9%)	-
Family therapy	218 (6.8%)	97 (5.2%)	-
Online therapy	519 (16.2%)	161 (8.6%)	-
Other	279 (8.7%)	102 (5.5%)	-
Self-reported improvement in symptoms/functioning - optional lifetime treatment history questionnaire	N =2461		N=7595
Cognitive behavioural therapy (CBT)	839 (34.1%)	-	2462 (32.4%)
Counselling	1076 (43.7%)	-	4041 (53.2%)
Group therapy	241 (9.8%)	-	240 (3.2%)
Other	305 (12.4%)	-	852 (11.2%)

**For individuals from the GLAD+ dataset with self-reported benefit from therapy in relation to multiple disorders, responses were selected hierarchically. Responses from the major depression module were preferentially selected (94% of responses) to correspond as closely as possible with participants from AGDS, who were all ascertained on the basis of a lifetime depression diagnosis. When unavailable, responses were selected in relation to generalised anxiety disorder (4%), panic disorder (2%), agoraphobia (<1%), specific phobia (<1%), then social anxiety (<1%). Breakdown of treatment types is provided in relation to the selected disorder. Rates of different treatment types were comparable across disorder, but are not given for each disorder due to small sample sizes for some disorders resulting in potentially identifiable data. In the CIDI-SF modules, participants were able to report receiving more than one type of treatment, hence the total responses is greater than the sample size.*

Supplementary Table 3: Genome-wide association study summary statistics used to derived polygenic scores in the GLAD+ dataset

Trait	GWAS reference	N (N cases)	h^2_{SNP} (SE)
ADHD	Demontis et al. (2019)	53,293 (19,099)	0.240 (0.015)
Daily alcohol use	Schumann et al. (2016)	70,460	0.052 (0.008)
Anorexia nervosa	Watson et al. (2019)	72,517 (16,992)	0.174 (0.011)
Anxiety	Purves et al. (2018)	83,566 (25,453)	0.104 (0.07)
Autism	Grove et al. (2019)	46,350 (18,381)	0.203 (0.015)
Bipolar disorder	Stahl et al. (2019)	51,710 (20,352)	0.347 (0.016)
Body mass index	Locke et al. (2018)	693,529	0.158 (0.005)
Sleep duration	Jones et al. (2016)	128,266	0.055 (0.005)
Cross disorder	PGC (2013)	61,220 (33,332)	0.175 (0.011)
Depressive symptoms	Okbay et al. (2016)	161,460	0.047 (0.004)
Major depressive disorder (PGC2, no 23andMe)	Wray et al. (2018)	143,265 (45,591)	0.072 (0.001)
Broad depression	Howard et al. (2019) excl 23andme	500,199 (329,443)	0.089 (0.003)
Diabetes Type 2	Scott et al. (2017)	159,208 (26,676)	0.078 (0.005)
Educational attainment	Lee et al. (2018)	766,345	0.107 (0.003)
Health	Harris et al. (2017)	111,483	0.098 (0.006)
Household income	Hill et al. (2016)	112,151	0.060 (0.006)
Insomnia	Jansen et al. (2019)	386,533 (109,402)	0.0459 (0.002)
Intelligence	Savage et al. (2018)	269,867	0.183 (0.006)
Memory	Davies et al. (2016)	112,067	0.043 (0.005)
Neuroticism	Hübel et al. (2019)	287,535	0.110 (0.005)
Obsessive compulsive disorder	IOCDF (2018)	9,725 (2,688)	0.338 (0.048)
Physical activity	Unpublished - Hanscombe (2018)	66,224	0.143 (0.010)
Post-traumatic stress disorder	Nievergelt et al. (2019)	174,659 (23,212)	0.016 (0.003)
Age at first birth	Barban et al. (2016)	251,151	0.055 (0.003)
General risk tolerance	Linnér et al. (2019)	466,571	0.052 (0.002)
Schizophrenia	PGC (2014)	77,096 (33,640)	0.455 (0.016)
Ever smoker	Linnér et al. (2019)	518,663	0.109 (0.003)

Supplementary Table 4: Comparison of models predicting retrospectively self-reported psychological treatment outcomes using genetic, sociodemographic and clinical predictors, and all predictors in the GLAD+ dataset (N = 4,439)

Model 1	Model 2	Pearson correlation between model predictions and outcome (R)		Difference in R	p-value of difference in R
		Model 1	Model 2		
	Sociodemographic and clinical predictors				
Genetic predictors		0.041	0.123	0.082	3.9×10^{-5}
Genetic predictors	All predictors	0.041	0.115	0.075	3.2×10^{-9}
Sociodemographic and clinical predictors	All predictors	0.123	0.115	-0.008	0.409

Supplementary Table 5: Standardised coefficients for the predictors retained in the final model based on the average optimal λ tuning parameter following 10-fold cross validation

Predictor	Beta
(Intercept)	1.392
<i>Sociodemographic</i>	
Sex	0.024
University degree	0.122
<i>Clinical</i>	
Number of episodes	-0.239
<i>Genetic</i>	
ADHD	-0.081
Daily alcohol use	0.003
Anorexia nervosa	.
Anxiety	0.026
Autism	.
Bipolar disorder	0.036
Body mass index	-0.031
Sleep duration	0.044
Cross disorder	-0.075
Depressive symptoms	-0.042
Major depressive disorder (PGC2, no 23andMe)	.
Broad depression	0.076
Diabetes Type 2	0.014
Educational attainment	0.001
Self-reported health	0.066
Household income	-0.020
Insomnia	-0.038
Intelligence	-0.008

Memory	-0.014
Neuroticism	0.043
Obsessive compulsive disorder	-0.029
Physical activity	0.001
Post-traumatic stress disorder	-0.016
Age at first birth	0.060
General risk tolerance	0.051
Schizophrenia	0.023
Ever smoker	.
Variance explained on the liability scale (R^2)	6.01%

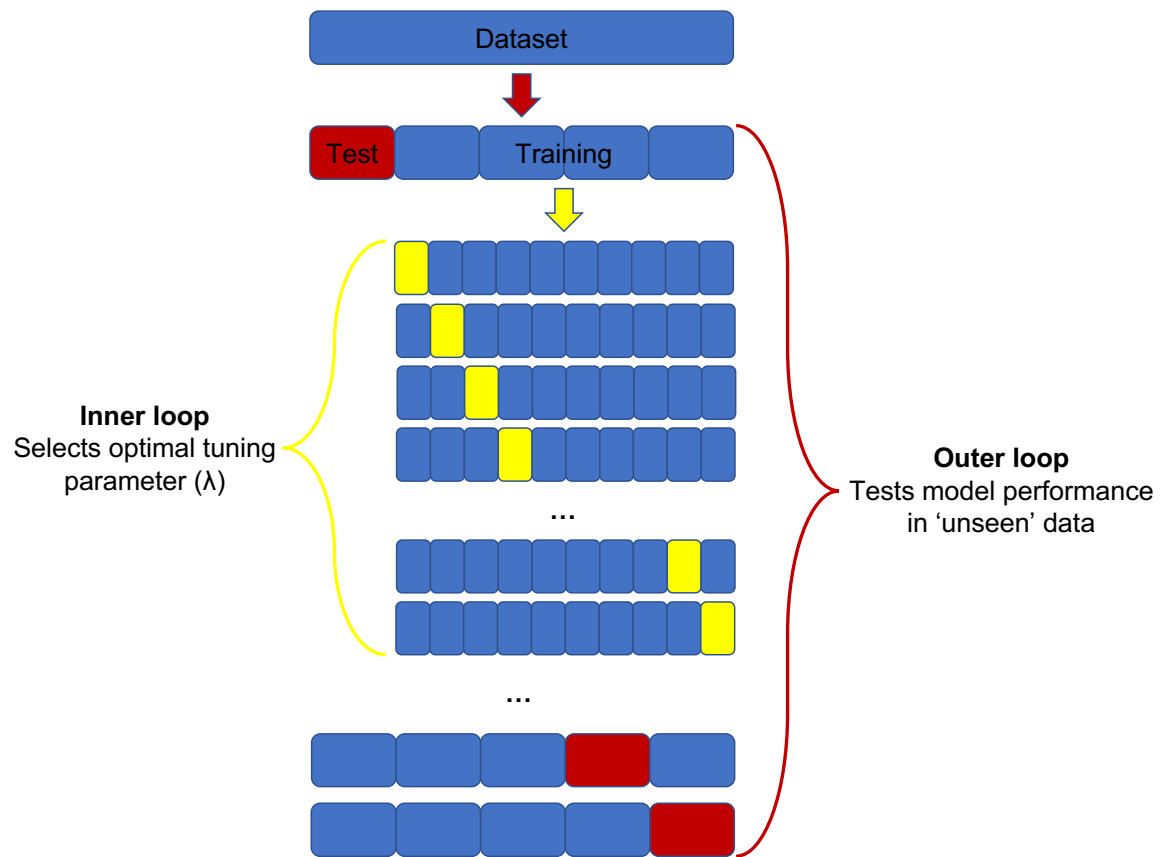
R^2 represents variance explained on the liability scale within the full dataset, determined using a prevalence of 41% for response to psychological therapy ².

Standardised coefficients for the predictors retained in the model following elastic net regularisation were obtained by re-estimating the cross-validated model in the full data set using the average optimal λ tuning parameter, identified from each inner model development loop (Supplementary Figure S1). Out of the 30 predictors, 26 were retained in the final model, based on an average λ tuning parameter of 0.009210859 (range: 0.008790644–0.009591377) and α of 0.1.

Supplementary Figures

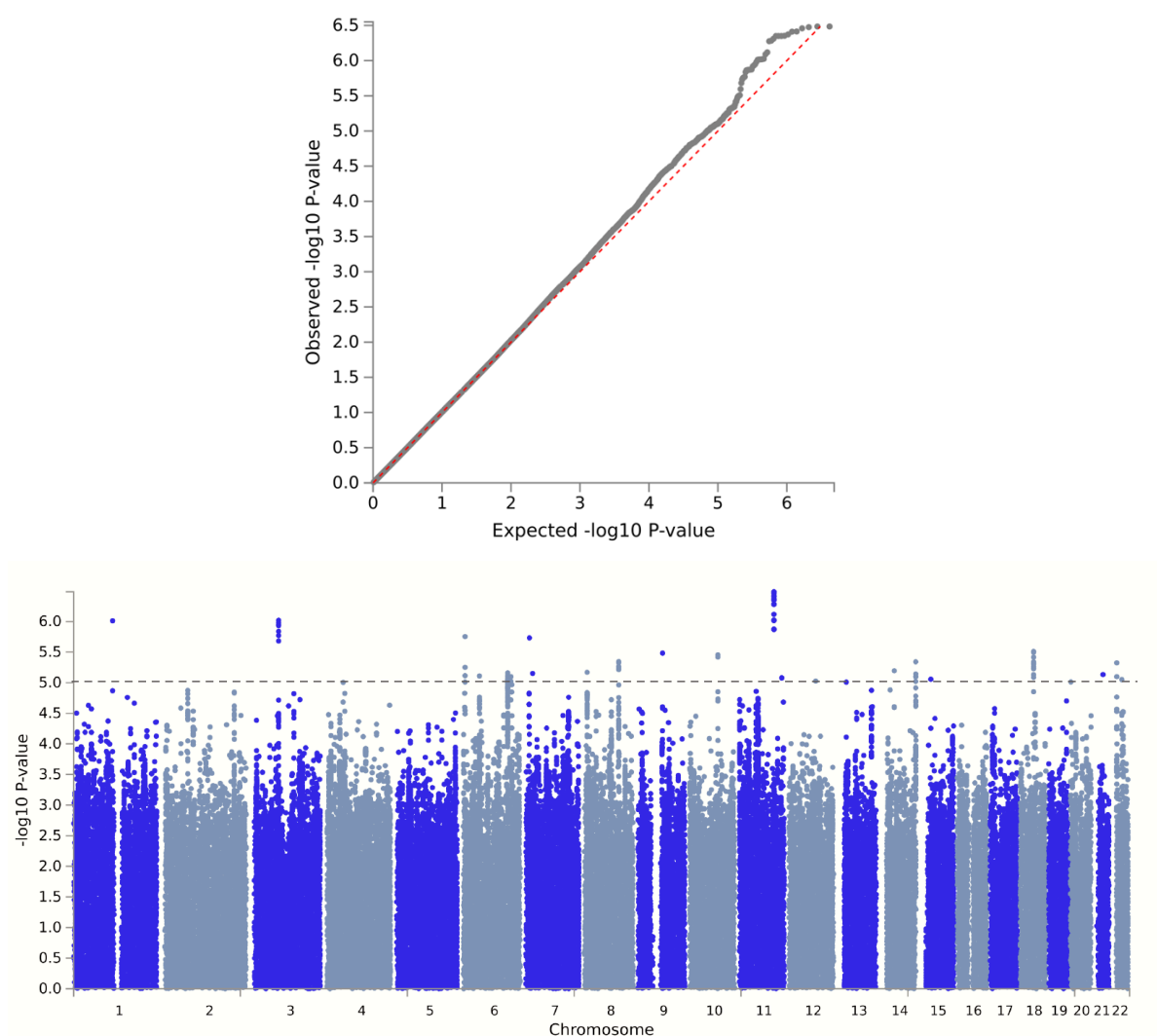
Supplementary Figure 1: Nested 10-fold cross validation.

Figure adapted from Meehan et al.³



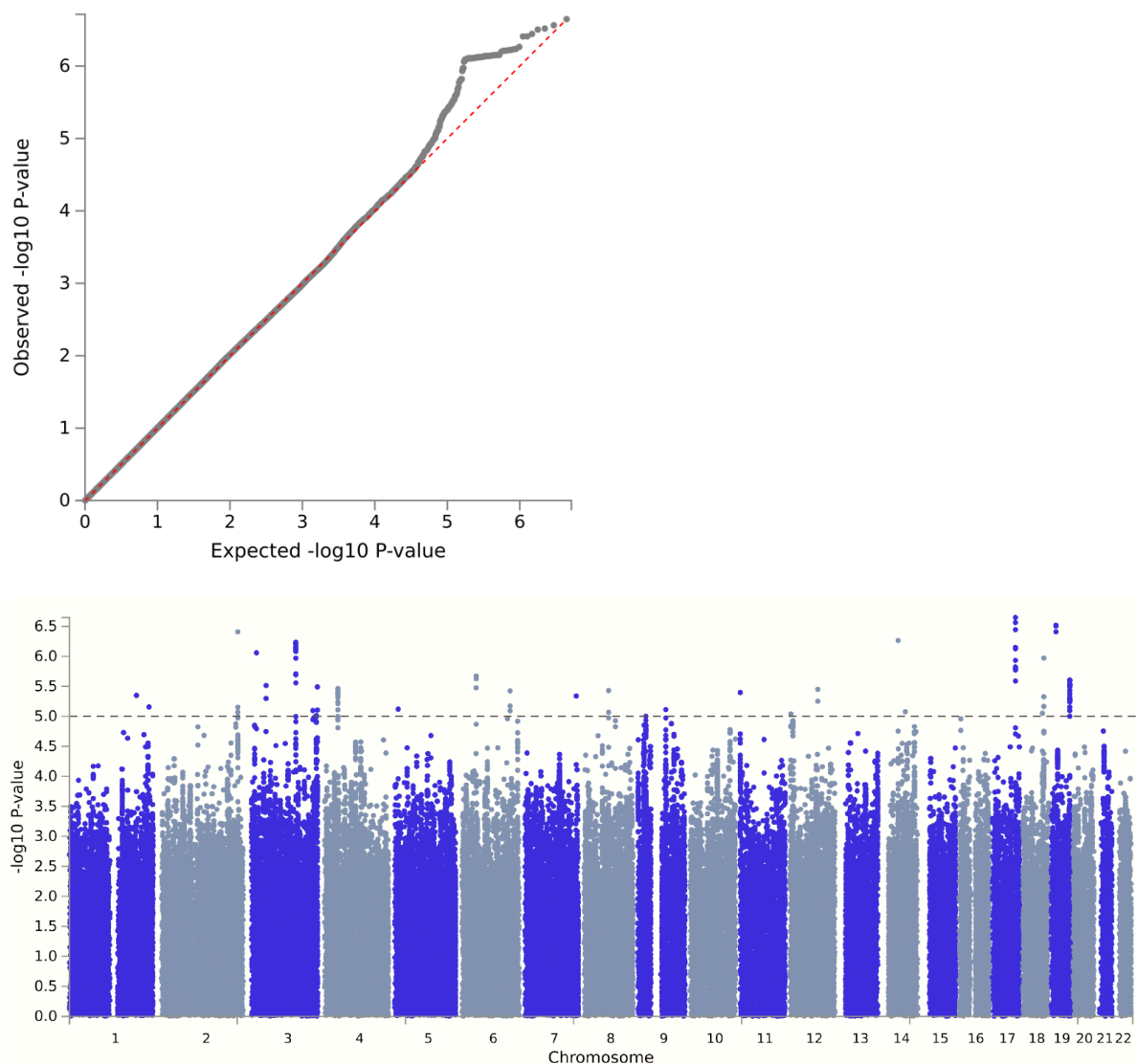
Supplementary Figure 2: Q-Q plot (upper) and Manhattan plot (lower) of results from GWAS of retrospectively self-reported treatment outcomes in the GLAD+ dataset.

In the Manhattan plot, dashed line indicates suggestive threshold ($p < 1 \times 10^{-5}$) for genetic variants associated with retrospectively self-reported treatment outcomes ($N = 5693$; 75% reported benefit/improvement).



Supplementary Figure 3: Q-Q plot (upper) and Manhattan plot (lower) of results from GWAS of retrospectively self-reported treatment outcomes in the AGDS dataset.

In the Manhattan plot, dashed line indicates suggestive threshold ($p < 1 \times 10^{-5}$) for genetic variants associated with retrospectively self-reported treatment outcomes ($N = 5715$; 75% reported improvement)).



References

1. Davies, M. R. et al. Comparison of symptom-based versus self-reported diagnostic measures of anxiety and depression disorders in the GLAD and COPING cohorts. *J. Anxiety Disord.* 85, 102491 (2022).
2. Cuijpers, P. et al. The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatr. Scand.* 144, 288–299 (2021).
3. Meehan, A. J. et al. Developing an individualized risk calculator for psychopathology among young people victimized during childhood: A population-representative cohort study. *J. Affect. Disord.* 262, 90–98 (2020).