

Phenomic Network Analysis for Depression Reveals Comorbidity Architecture, Genomic Relationships, and Pleiotropic Variants

Zhiyu Yang¹, Pritesh Jain¹, Petros Drineas^{*2}, and Peristera Paschou^{†1}

¹Department of Biological Sciences, Purdue University, West Lafayette, Indiana

²Department of Computer Science, Purdue University, West Lafayette, Indiana

Abstract

Depression is one of the most prevalent psychiatric disorders and is one of the leading causes of health ailment worldwide. It is known to be highly heritable and is frequently comorbid with other mental and physical traits. This observation motivated us to look deeper into the genetic and phenotypic connections between depression and other traits in order to identify correlations as well as potentially causal connections between them. In this study, we analyzed data from the UK biobank to systematically evaluate relationships between depression and other heritable traits both from a phenotypic and a genetic aspect. We compressed a total of 6,300 ICD codes into 412 heritable phecodes and we constructed a comorbidity network connecting depression and other disorders on over 300,000 participants of European ancestry. Additionally, we investigated the genetic correlation for each (phenotypic) connection in the resulting network. We also looked into potentially causal relationships using mendelian randomization for all pairs of significantly correlated disorders and uncovered horizontal pleiotropic genetic variants and genes contributing to disease etiologies. We found gastro-oesophageal reflux disease (GORD), body mass index, and osteoarthritis to be direct causes for depression, with GORD lying at the center of the causal network. Genes broadly expressed in various tissues, such as *NEGR1*, *TCF4*, and *BTN2A1* underlie the pathways that lead

*pdrineas@purdue.edu

†ppaschou@purdue.edu

not only to depression but also to other related disorders. Our work highlights the broad connections between depression and diverse traits, indicating a complex etiology and possible existence of subtypes for depression. Our findings highlight the value of cross-trait analysis in order to better understand the neurobiology of complex psychiatric disease.

1 Introduction

As one of the most prevalent psychiatric disorders, depression is a major burden on global health. Currently, it is affecting approximately 10% of the adult population in the United States (22) and shows an increasing trend in its incidence rate across the world (57; 69). According to the fifth edition of Diagnostic and Statistical Manual of Mental Disorders (DSM-V), a diagnosis of a major depressive episode requires at least one main symptom of depressed mood or *anhedonia* (defined as the loss of interest or pleasure), along with some secondary symptoms, including disturbance in appetite, sleep, or ability to engage in physical activity and cognition (1). However, despite the existence of precise diagnostic criteria, a high level of heterogeneity has been observed among depressive patients; this is common in many psychiatric traits that are nevertheless managed as dichotomous.

Individuals diagnosed with depression often also present with other mental and physical illness. On the psychiatric side, traits commonly co-occurring with depression include anxiety (97; 48), substance abuse (16), and post-traumatic stress disorder (PTSD) (28; 46). Meanwhile, depression is also reported to be highly comorbid with certain clinically distinct phenotypes, such as obesity (19; 4), cardiovascular diseases (41; 75), and even immune disorders including multiple sclerosis (MS) (30; 60) and inflammatory bowel disease (IBD) (62; 2). Given the wide range of phenotypes coexisting with depression, it is important to quantify and evaluate these comorbidities in depression patients systematically. However, perhaps due to lack of resources, such studies for depression are very limited. Most existing studies focusing on depression's comorbidity targeted only a small group of homogeneous diseases rather than a broader spectrum (36; 39; 99; 37; 40). Few studies aimed at recovering a more general comorbidity network of depression (84; 96; 63) but stopped at an epidemiological level, without further exploring underlying biological mechanisms. On the other hand, existing phenome-wide association studies (PheWAS) attempted to evaluate the association

between various phenotypes and the genetic risk of depression (72). However, using an individual's genetic risk score as a proxy may result in limited clinical interpretability.

The emergence of large comprehensive biobanks, as well as the abundance of publicly available genome-wide association studies (GWAS), allows us to systematically investigate depression and its comorbidities from both a phenotypic and a genetic perspective. Depression as well as most of the aforementioned mentioned co-occurring disorders have strong genetic foundations (89; 101; 64; 88; 38), with depression itself showing a heritability of approximately 30%-40% in family-based studies (32; 65; 91) and a SNP heritability of approximately 10% (55; 43). In this study, we investigate both the phenotypic and genetic architecture of depression and its comorbid disorders, leveraging the UK biobank datasets and largest available genomewide association studies (GWAS) to construct phenotypic and genetic networks. Our results shed light on the epidemiological and phenotypic relationships between depression and other disorders and help uncover biological insights that may underlie the clinically observed comorbidities.

2 Methods

2.1 Data

The UK Biobank dataset includes phenotypic and genetic data from approximately 500,000 individuals and was used to compute phenotypic correlations. As a quality control step, we removed samples with non-European British ancestry, based on their self-reported ancestry information. Next, we removed individuals with greater than third degree relatedness based on the kinship coefficient. Finally, we ran Principal Component Analysis (PCA) using TeraPCA (11) to remove individuals who do not overlap with European samples in the 1000 genomes dataset (mean principal component value on the top six principal components plus/minus three standard deviations). We further removed individuals with genotype missingness less than 0.98, thus retaining a dataset of 331,256 participants.

For the phenotypic data, we included 6,300 ICD-10 diagnoses (data field 41270) with a non-zero number of patients in our analyses. To reduce the dimensionality and increase the interpretability of our analyses, we further mapped the ICD-10 codes on PheCodes. Out of all 6,300 ICD10 diagnoses, 4,807 were mapped to at least one valid PheCode, for a total of 505 PheCodes and 1,434

child PheCodes. We removed PheCodes from categories that are dominated by non-genetic causes (infectious diseases, injuries, poisonings, pregnancy complications, etc.). This resulted in a total of 4,004 ICD-10 diagnoses mapped onto 411 parent PheCodes. We decided to separate depression from its parental PheCode 296 (mood disorders), which included bipolar and other mood disorders. We ended up with 412 phenotypic end points, including 410 parental PheCodes, plus one for depression and one for other mood disorders. (See Table S1 for the detailed list.)

2.2 Phenotypic network of depression and comorbid disorders

We computed phenotypic correlations between all pairs of PheCodes, a total of $412 \times 411 / 2 = 84,666$ pairs. Out of all these possible combinations, we observed 77,502 non-zero phenotypic correlations, where phenotypic correlation is defined as cosine of the angle formed between pairs of phenotype vectors. We focused our analyses on highly comorbid pairs of disorders that show phenotypic correlations in the top 1% quantile, i.e., we focused on the top 775 pairs. We extracted depression and its comorbid disorders from those top 775 pairs and we formed a phenotypic correlation network. For each trait that was deemed comorbid with depression, we used a hypergeometric test to analyze its enrichment in cases of depression patients.

2.3 Genetic network across depression and its comorbidities

To study the genetic relationship between depression and its comorbid disorders, we used linkage disequilibrium (LD) score regression (as implemented in LDSC (14)) to evaluate genetic correlations between all highly comorbid pairs in our network. For this analysis, we obtained publicly available summary statistics from GWAS of identical or similar traits (see Table 1). In order to retain independence between our phenotypic and genetic networks, we selected GWAS summary statistics that did not involve samples from the UK biobank. However, this was not possible for certain disorders (see Table 1). For all genetic correlation analyses, only SNPs that were matched with the HapMap3 SNP list were used. As the GWAS summary statistics are dominated by European samples, we used LD scores estimated from the European subgroup of the 1000 Genomes phase 3 project as both the independent variable and the regression weights. The significance threshold for this analysis was corrected by dividing by the number of tested pairs, which corresponds to the number of highly comorbid pairs within the phenotypic correlation network.

2.4 Phenome-wide causal inference

For all disorder pairs where a significant genetic correlation was observed, we further investigated their potential causative relationship using bidirectional, generalized, summary-statistic-based Mendelian randomization (GSMR) (111). For each analysis, we used SNPs that were associated with the exposure trait (p -value less than 5×10^{-6}) as instrument variables; we required a minimum of five independent SNPs ($r^2 < 0.05$) in this analysis. We excluded pleiotropic SNPs identified through the heterogeneity in the dependent instrument (HEIDI outlier approach, SNPs with $p_{HEIDI} < 0.01$) to guarantee that all instrument SNPs affect the outcome only via exposure. We used Bonferroni correction, correcting for the number of significant genetic correlations (multiplied by two for bidirectional tests) to determine the significance threshold of this analysis.

2.5 Evaluation of horizontal pleiotropy

We further looked into genetic variants that show horizontal pleiotropy effects between depression and disorders that appeared to be causally related to depression from the Mendelian randomization analysis. The analysis was carried out using Iterative Mendelian Randomization and Pleiotropy (IMRP) (110). Using causal effect sizes and standard errors estimated from the aforementioned Mendelian randomization analyses, we ran SNP-based pleiotropic test for depression and each causal exposure respectively. Similar to mendelian randomization, the test was carried out in a bi-directional manner. This is equivalent to running a GWAS for two traits: $\beta_{pleio1} = \beta_{depression} - \beta_{exp-dep} \times \beta_{exposure}$ and $\beta_{pleio2} = \beta_{exposure} - \beta_{dep-exp} \times \beta_{depression}$, where $\beta_{depression}$ and $\beta_{exposure}$ stand, respectively, for GWAS effect sizes for depression and the causal exposure; and $\beta_{exp-dep}$ and $\beta_{dep-exp}$ are the estimated causal effect sizes from the exposure trait to depression and from depression to the exposure trait, respectively. To identify pleiotropic SNPs between depression and the exposure trait, we first selected SNPs that were genome-wide significant ($p < 5 \times 10^{-8}$) in either direction. For a SNP to be considered pleiotropic, it has to satisfy two conditions: first, it must be genome-wise significant in one direction and, second, it must have a p -value below the Bonferroni-corrected threshold in the other direction. In more detail, suppose that we identified a total of n SNPs being genome-wide significant in either direction. Then, the pleiotropic SNPs will be the ones that have $p < 5 \times 10^{-8}$ in one direction and $p < 0.05/n$ in the other direction.

On top of the SNP-based pleiotropic test, we also seek to interpret the results at a gene level. Therefore, to analyze depression-exposure pairs, we first carried out gene-based analyses using results from SNP-based pleiotropic tests for both directions using MAGMA (24). We subsequently performed gene-based pleiotropic screening on the MAGMA results. In this case, significant pleiotropic genes were defined as ones that are genome-wise significant ($p < 0.05$ divided by the total number of genes tested for the trait) in one direction *and* show p -value lower than a Bonferroni-corrected significance threshold (< 0.05 divided by the total number of genes satisfying the first condition) in the other direction.

3 Results

3.1 Phenotypic network of depression and its comorbidities

Among the 331,256 unrelated European samples that we retained from the UK biobank, 18,588 have been diagnosed with depression. Out of the 412 endpoint traits that we defined, we identified 12 frequent comorbidities with depression among the top 1% quantile of all pairwise phenotypic correlations. The most frequent one is hypertension (parent PheCode 401), with nearly half of all depression patients having been diagnosed with hypertension ($N_{Hypertension|Depression} = 9,196$). The next trait most frequently comorbid to depression, was osteoarthritis (parent PheCode 740), which was observed in more than one-third of all depression patients ($N_{Osteoarthritis|Depression} = 6,902$). Interestingly, out of the top 12 comorbidities for depression in the UK Biobank, only two are psychiatric or behavioral traits: anxiety disorders (parent PheCode 300, $N_{Anxiety|Depression} = 5,666$) and suicidal ideation or attempt (parent PheCode 297, $N_{Suicide|Depression} = 1,154$). Even though the number of comorbid cases with depression for either of these two psychiatric traits is not as high as hypertension or osteoarthritis, one has to consider anxiety disorders and suicidal ideation are not as prevalent in the UK biobank. Anxiety disorders and suicidal ideation are highly enriched within depression patients ($p_{Anxiety} = 5.84 \times 10^{-3711}$ and $p_{Suicide} = 2.56 \times 10^{-1120}$). The remaining eight traits span conditions that affect various organs, including gut, lungs, and processes of the metabolic system (see Table 2). Moreover, just as expected, all these comorbidities of depression are also tightly connected to each other (see Figure 1 and Table S2).

3.2 Genetic network of depression and its comorbidities

We observed a total of 56 connections in the phenotypic network of depression and its 12 comorbid traits discussed above. We further proceeded to examine the potential genetic relationships between disorder pairs. Focusing on the 13 traits in the phenotypic network, we were able to find publicly available GWAS summary statistics for all of them (see Table 1). For lipid metabolism disorders (parent PheCode 277), we used two sets of summary statistics, one corresponding to low-density lipoprotein levels and one corresponding to high-density lipoprotein levels. Similarly, for fluid and electrolyte disorders (parent PheCode 276), we also used two sets of summary statistics as proxies, one corresponding to sodium level measurements and one corresponding to potassium level measurements. These resulted in a total of 76 independent tests for genetic correlations and a Bonferroni-corrected significance threshold less than 6.58×10^{-4} . Using this threshold, we identified 18 genetically correlated pairs of disorders, resulting to a fully connected genetic network (see Table 3 and Figure 2). We found that even though at the phenotypic level all 12 traits are highly comorbid with depression, only six out of the 12 actually share a genetic basis with depression (see Table 3). More interestingly, these six traits are not always the most comorbid or enriched ones with depression. For example, surprisingly, hypertension (enrichment $p_{Hypertension} = 7.48 \times 10^{-833}$) was not found genetically correlated with depression. In the resulting genetic network, we noticed that gastro-oesophageal reflux disease (GORD) had the highest number of connections with all other traits. More precisely, GORD was genetically correlated with eight other traits in the genetic network, thus forming a “hub” in our network (Figure 2).

3.3 Phenome-wide causal inference

For the 18 genetically correlated pairs of disorders discussed in the previous section, we used Mendelian randomization as implemented in the bidirectional GSMR to test for the existence of putative causal association for each pair. Using a Bonferroni-corrected significance threshold equal to 1.39×10^{-3} , we identified 18 traits as being potential causal exposures for an outcome, while using (nearly independent) associated SNPs as instrumental variables. Among the six traits genetically correlated with depression, we identified three causal risk factors for depression: GORD (169 independently associated SNPs with $b_{xy} = 0.25$ and $p = 1.34 \times 10^{-30}$); body mass index (BMI)

(541 independently associated SNPs with $b_{xy} = 0.17$ and $p = 4.08 \times 10^{-14}$); and osteoarthritis (16 independently associated SNPs, $b_{xy} = 0.10$ and $p = 7.98 \times 10^{-4}$). Depression itself also appeared to have causal effects on both anxiety ($b_{xy} = 0.62$, $p = 2.99 \times 10^{-09}$, and $N_{SNP} = 50$) and GORD ($b_{xy} = 0.11$, $p = 2.66 \times 10^{-11}$, and $N_{SNP} = 64$). Recall that GORD was a hub in our genetic network and GORD is again a hub in the causality network (see Figure 3 and Table 4).

3.4 Pleiotropic variants underlying depression and potential causal exposures

For the three disorders that were found as potential immediate causal exposures of depression based on our Mendelian randomization analysis (GORD, BMI, and osteoarthritis), we further looked for evidence of horizontal pleiotropy (see Table S3 for SNP-based pleiotropic results). We identified 522 SNPs showing genome-wide significant pleiotropic effects in either of the tested directions between depression and GORD. Only one of them also surpassed the Bonferroni-corrected threshold ($p < 0.05/522 = 9.58 \times 10^{-5}$) in the other direction and was therefore identified as pleiotropic. SNP rs11040813 ($p_{pleio1} = 3.70 \times 10^{-05}$, $p_{pleio2} = 3.80 \times 10^{-08}$) is located in region 11p15.4 close to gene *CNGA4*. For BMI, we identified a total of 10,667 SNPs showing genome-wide significant pleiotropic effects with depression in at least one of the tested directions. 103 of these SNPs were identified as pleiotropic ($p_{pleio1} < 5.00 \times 10^{-08}$ and $p_{pleio2} < 4.69 \times 10^{-06}$, or $p_{pleio2} < 5.00 \times 10^{-08}$ and $p_{pleio1} < 4.69 \times 10^{-06}$). The SNPs span regions 1p31.1, 10q25.1 and 16q12.2. 13 of the 103 SNPs showed genome-wide significant pleiotropic effects in both directions and all of them are located in region 1p31.1 (gene *LOC105378797*). For depression and osteoarthritis, 472 SNPs were picked up as genome-wide significant for one of the tested directions. However, none of them satisfied the necessary requirements to be a bidirectional pleiotropic SNP.

To seek higher power and better interpretability, we also carried out pleiotropic screening at a gene level by performing gene-based analyses in addition to SNP-based analyses (see Table 5 for gene-based pleiotropic analyses). For depression and GORD, a total of 18,172 protein coding genes were tested, making the genome-wide significance threshold less than 2.75×10^{-06} . 58 genes were found significant at this threshold in at least one direction and eight were identified as pleiotropic ($p < 0.05/58 = 8.62 \times 10^{-4}$ in the other tested direction). Among these eight genes, the one that was most strongly associated with both directions was *TCF4* ($p_{pleio1} = 4.18 \times 10^{-6}$ and $p_{pleio2} = 1.67 \times 10^{-9}$), followed by *HIST1H2BN* ($p_{pleio1} = 7.34 \times 10^{-7}$ and $p_{pleio2} = 2.73 \times 10^{-6}$). For the trait pair

depression and BMI, we tested a total of 18,292 genes, resulting in a genome-wide gene significance threshold less than 2.73×10^{-06} . Using this threshold, 391 genes showed pleiotropic effect in either direction. Five of them also showed a p -value smaller than $0.05/391 = 1.28 \times 10^{-4}$ in the other direction and were identified as pleiotropic genes for the two traits. The list included *NEGR1* ($p_{pleio1} = 1.29 \times 10^{-20}$ and $p_{pleio2} = 6.30 \times 10^{-6}$) and *RAB27B* ($p_{pleio1} = 7.44 \times 10^{-7}$ and $p_{pleio2} = 9.47 \times 10^{-8}$). Finally, for the trait pair depression and osteoarthritis, 18,230 protein coding genes were tested and 31 were found significant (p value below 2.74×10^{-06} in one of the tested directions). However, similar to the SNP-based analysis, none of the genes satisfied the necessary requirements to be classified as bidirectional pleiotropic genes.

4 Discussion

Motivated by the highly heterogeneous nature of the phenotypic manifestations and genetic basis of depression and its many co-occurring disorders we embarked on a systematic evaluation of their phenotypic and genetic relationships seeking to gain insights into the biological mechanisms that may underlie comorbidities in depression. We uncovered intriguing phenotypic and genetic relationships between depression and GORD and BMI, as well as genes that could play a central role in governing the links between mental disorders and disorders that are not immediately thought of as relating to the nervous system.

Using ICD-10 records of the European samples in the UK Biobank, we began by constructing a (phenotypic) comorbidity network of depression. We mapped over six thousands ICD codes onto approximately four hundreds PheCodes of heritable traits. We discovered 12 traits that are often comorbid with depression, including mental disorders (for instance suicidal ideation and anxiety) we well as traits affecting the digestive system (abdominal hernia, gastritis and duodenitis, functional digestive disorders, diseases of esophagus), metabolic disorders (fluid and electrolyte disorders, lipid metabolism, obesity, etc.), asthma, hypertension, and osteoarthritis. Many of these comorbidities have been previously observed in epidemiological studies, i.e., such as GORD (20), obesity (17), asthma (47), hypertension condition (34), and the more well-studied anxiety disorders (98) and suicidal ideation (29; 73).

Next, using GWAS summary statistics, we explored the genetic correlation between all highly

comorbid pairs of disorders in the phenotypic network. This analysis resulted in a, somewhat less well-connected network (compared to the phenotype-based one), where we observed that traits that are phenotypically connected with depression do not necessarily share a common genetic basis with it. Interestingly, hypertension, one of the most enriched conditions among depression patients, did not show significant genetic correlation with depression. This implies that high comorbidity rates between depression and hypertension (or even other cardiovascular traits), may not be due to a shared heritable etiology. Instead, such phenotypic commorbidities may be mediated by other factors (18), such as drug effects of antidepressants (95; 45). This lack of genetic connections between depression and hypertension is concordant with the observation that these two traits show unique, although interactive, effects on cognitive functions and brain volumes as has been previously described (87; 67).

Significant genetic correlations have been validated by our results between depression and psychiatric traits (anxiety and suicidal ideation) (70; 27). We also identified links connecting depression with GORD, BMI, chronic gastritis, and osteoarthritis, reinforcing the complexity of depression's etiology. Strong clinical relationships between the digestive system and psychiatric conditions have been known for over 35 years (31; 100). Recent studies also provided further evidence supporting genetic commonalities between depression and various digestive traits (107), including GORD (79) and chronic gastritis (3). GORD was a "hub" in our genetic network and appears significantly correlated with BMI and osteoarthritis, two traits that share a direct genetic basis with depression. GORD also creates connections between depression and hypertension, asthma, and IBD.

For all pairs of disorders that we found to be (significantly) genetically correlated we looked into potential causal relationships using Mendelian randomization. All the traits in the connected component of the genetic network (except suicidal ideation) remained connected in the causality network. We found three immediate risk factors for depression: GORD, BMI, and osteoarthritis. Additionally, depression arose as a significant risk factor for anxiety and GORD. The other traits, although not directly linked to depression, appeared to be connected via intermediate nodes. GORD appeared as a hub and, along with BMI, could be acting to connect depression to all other digestive traits. Interestingly, the asthma node also pointed towards depression via GORD. Phenotypic comorbidity between asthma and GORD has been widely reported (12; 7), whereas, to the best of our knowledge, evidence for a causal link between these two disorders has not been suggested

until now. We used SNP effects as instruments and we provided insights into possible relationships between asthma and GORD; biological explanations explaining such causal relationships might be due to the epithelial dysfunction in asthma, since such patients are more vulnerable to acid reflux (81). Our results also point towards a uni-directional causality from BMI to depression. This observation is supported by longitudinal studies, where obesity was found to increase the risk of depression (59). Many studies found that the distribution of depression incidence as a function of BMI is a U-shaped curve (61; 25; 109), indicating that not only obesity, but also any BMI abnormalities could result in a higher risk of depressive conditions. On the other hand, the relationship between depression and osteoarthritis has not received as much attention as the other trait pairs, perhaps due to the fact that the latter is strongly age-dependent; a relatively recent PheWAS discovered an association between the genetic risk score for depression and osteoarthritis in the UK biobank samples (72). On the clinical side, links between osteoarthritis/arthritis and depressive traits are more frequently explained by the pain caused by such conditions (58; 104; 94), which is also in concordance with our finding that osteoarthritis indicated a causal effect on depression, but not the other way around.

In order to further explore the biological mechanisms underlying connections between the aforementioned traits and depression, we investigated genetic horizontal pleiotropic effects at the SNP and gene levels for depression and its three most closely correlated traits: GORD, BMI, and osteoarthritis. We identified several multi-functional genetic regions and genes that play a role in the etiology of depression and its comorbidities. Our gene-based pleiotropy analysis of depression and GORD highlighted the role of transcription factor 4 (*TCF4*), which regulates the differentiation of various cell types and underlies the etiology of both diseases. This gene has also been reported to have an impact in depression and insomnia (15). Beyond depression and GORD, associations between *TCF4* and many other mental and physical conditions have been widely studied; such conditions include cognitive ability, autism, the development of certain cancer types, as well as rare diseases such as Pitt-Hopkins syndrome (33; 74; 51; 56; 5). Moreover, *TCF4* has been identified as one of the strongest genetic biomarkers for schizophrenia susceptibility in multiple genetic studies (44; 90; 52). Animal experiments suggest that the over-expression of this gene in the brain results in functional impairments that constitute schizophrenic symptoms (13) and that such effects further interact with different exposure conditions (8). Interestingly, in patients with depression,

a decreased *TCF4* expression at the mRNA and protein level was observed when compared to healthy controls (71). Our results indicate that, beyond schizophrenia, it may also be worthwhile to further investigate the pathology of *TCF4* in depressive conditions.

One of the top genes that was uncovered from the pleiotropic analysis between depression and BMI was *NEGR1* (Neuronal Growth Regulator 1). This is a known obesity-associated gene, whose effect has been verified in multiple populations (66; 76; 83; 85). *NEGR1* regulates neuronal control of food intake (9; 53) and along with *FTO*, which was also picked up in the SNP-based pleiotropic analysis, have been reported as having an impact on eating disorders (68; 35). Given its high expression level in the brain, its connections with psychiatric traits has been recently investigated and *NEGR1* has been associated with a wide range of mental traits. Apart from eating related disorders, it also plays a role in depression (43), cognitive performance (54), and Alzheimer's disease (77). As a functional validation, increased depressive, anxiety- and autistic-like behaviors were observed in *NEGR1*-deficient mice, along with reduced hippocampal neurogenesis and disruption in cortical development (78; 92). Finally, as a putative tumor suppressor, *NEGR1* has also been identified to be a down-regulated gene in various cancer cells (50; 93), including ovarian cancer (86; 21) and breast cancer (10; 82).

Regarding limitations for our study, first, our results are based on analysis of the UK Biobank and sample inclusion in the UK biobank is volunteer-based so might not represent the true underlying population distribution. However, most of the connections observed in our work have also been previously reported in epidemiological studies which provides further support for our findings. Second, our results used GWAS summary statistics, which exhibit some variance in their respective power. Additionally, while most of the GWAS that we used were performed on European populations, the GWAS for electrolyte disorders (GWAS of sodium and potassium Levels) was based on the Qatar population.

5 Conclusion

In this paper we systematically evaluated phenotypic and genetic connections between depression and other disorders and looked into their causality relationships. We also sought to uncover genes with potentially pleiotropic effects across depression and highly correlated disorders. Our results

suggest that beyond phenotypic comorbidities, depression shows genetic connections with multiple psychiatric and non-psychiatric traits and we identify multiple genetic variants that could be linking mental health disorders to disorders of other systems. Such common genetic background may help explain the heterogeneity of the symptoms of depressive patients and help further sub-typing of depression. Our work highlights the value of trans-diagnostic approaches towards uncovering insights into the the etiology of psychiatric disease.

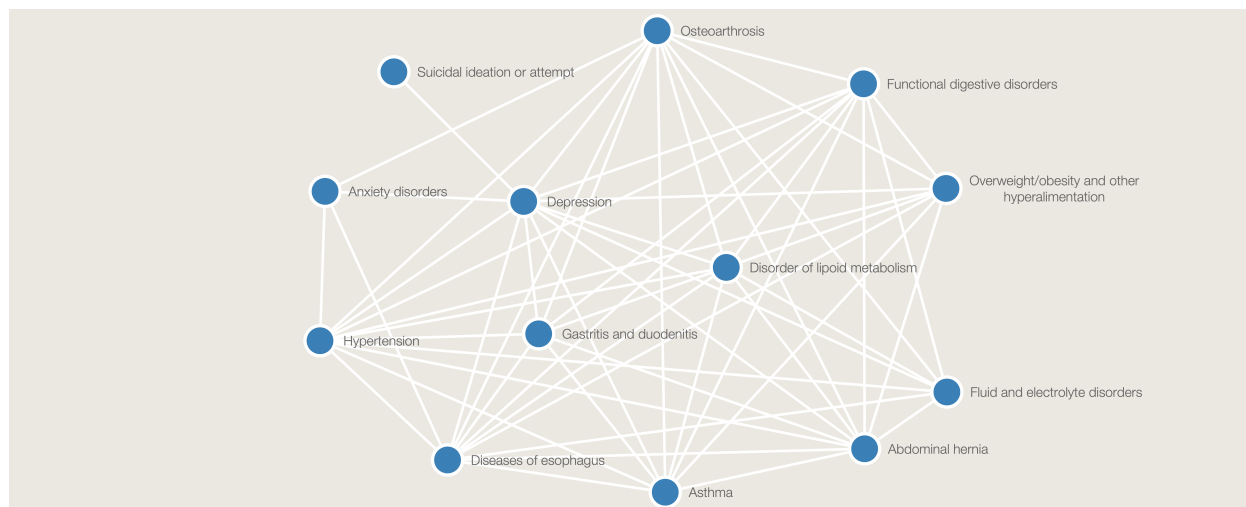


Figure 1: **Phenotypic disease network of depression.** The disease network of depression was obtained from the fully connected phenotypic network using phenotypic correlations as edge weights, by reducing the dense graph to a sparse network by keeping only the top 1% of the connections and extracting depression and its immediate neighbour nodes.

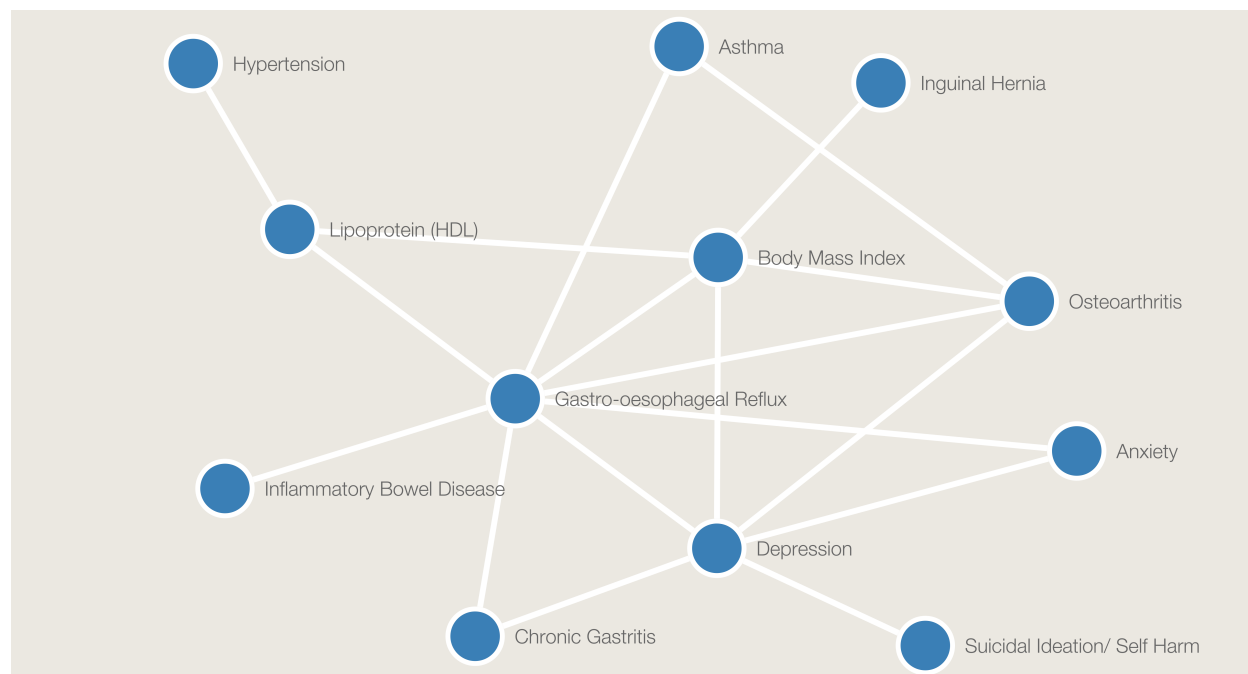


Figure 2: **Genetic disease network of depression.** For all connections observed in the phenotypic network of depression (figure 1), we evaluated the pairwise genetic correlation using LDSC(14). Edges in the graph show all significant genetic correlations between traits after correcting for multiple hypothesis testing.

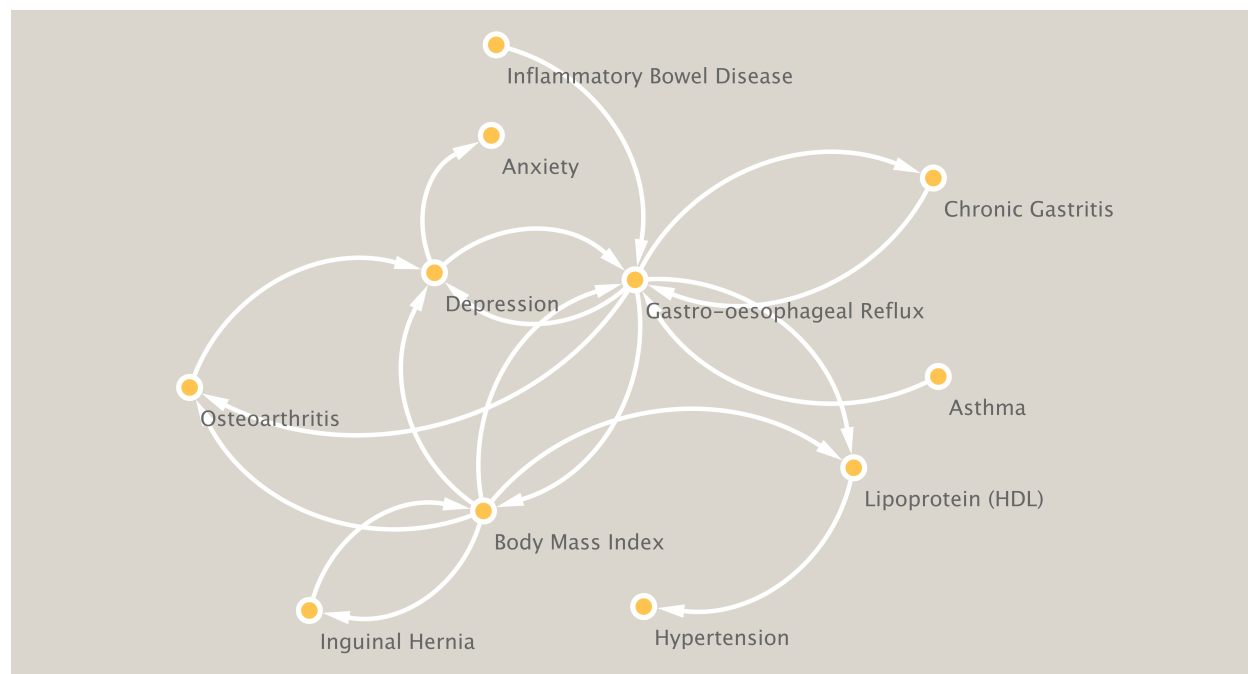


Figure 3: **Causal relationships between traits that are comorbid with depression.** We tested causal relationships between traits and depression using bidirectional Mendelian randomization for all genetically correlated pairs of disorders. Using SNP effects as instrumental variables, the resulting directed graph shows significant causalities between traits after correcting for multiple hypothesis testing.

Trait	Reference	Sample Size	comment
asthma	(26)	23,948 cases, 118,538 controls	
body mass index	(42)	100,418 samples	
gastro-oesophageal reflux	(6)	80,265 cases, 305,011 controls	
hypertension	(105)	27,123 cases, 22,018 controls	diverse population
osteoarthritis	(108)	30,727 cases and 297,191 controls	
depression	(106)	59,851 cases and 113,154 controls	
hernia	(102)	14,278 cases, 286,513 controls	
anxiety	(80)	7,016 cases, 14,745 controls	
low-density lipoprotein		94,595 samples	
high-density lipoprotein	(103)	94,595 samples	
sodium levels		6,018 samples	
potassium levels	(49)	6,017 samples	Qatar population
inflammatory bowel disease	(23)	25,042 cases, 34,915 controls	
suicide attempt	(102)	125,844 samples	
chronic gastritis	(102)	15,492 cases, 229,398 controls	

Table 1: **Sources for GWAS summary statistics.** Summary statistics are dominated by samples of European ancestry, with two exceptions (hypertension and sodium/potassium levels).

Trait	total N_{sample}	N_{sample} in depression	Enrichment P
All Sample	331,256	18,588	-
Depression	18,588	18,588	-
Disorder of lipid metabolism	46,738	5,394	1.84×10^{-644}
Osteoarthritis	59,572	6,902	1.99×10^{-887}
Diseases of esophagus	47,027	5,664	2.09×10^{-764}
Suicidal ideation or attempt	1,495	1,154	2.56×10^{-1120}
Asthma	29,737	3,892	3.10×10^{-578}
Abdominal hernia	54,605	5,297	3.34×10^{-391}
Functional digestive disorders (IBD)	31,818	4,342	4.09×10^{-711}
Gastritis and duodenitis (ChronicGastritis)	34,101	4,188	4.16×10^{-550}
Fluid and electrolyte disorders	16,418	2,879	4.25×10^{-685}
Overweight obesity and other hyperalimantation	21,266	3,529	4.26×10^{-787}
Anxiety disorders	13,186	5,666	5.84×10^{-3711}
Hypertension	94,738	9,196	7.48×10^{-833}

Table 2: **Enrichment of comorbidities in depression patients.** For all traits phenotypically correlated with depression (Figure 1), we tested the enrichment of cases in depression patients using a hyper-geometric test. Enrichment P denotes the hyper-geometric test p -value; total N_{sample} denotes the total number of cases of the trait in the UK biobank samples; N_{sample} in depression denotes the number of the cases having the respective trait *and* depression.

Trait 1	Trait 2	R_g	SE	P
Anxiety	Depression	0.8739	0.1678	1.90×10^{-07}
Anxiety	Gastro-oesophageal Reflux	0.4876	0.1190	4.17×10^{-05}
Asthma	Gastro-oesophageal Reflux	0.2047	0.0358	1.12×10^{-08}
Asthma	Osteoarthritis	0.2643	0.0668	7.65×10^{-05}
Body Mass Index	Depression	0.1140	0.0234	1.11×10^{-06}
Body Mass Index	Gastro-oesophageal Reflux	0.3007	0.0201	2.18×10^{-50}
Body Mass Index	Inguinal Hernia	-0.2361	0.0302	5.35×10^{-15}
Body Mass Index	Lipoprotein HDL	-0.2898	0.0300	4.71×10^{-22}
Body Mass Index	Osteoarthritis	0.4385	0.0437	1.17×10^{-23}
Chronic Gastritis	Depression	0.3735	0.0790	2.28×10^{-06}
Chronic Gastritis	Gastro-oesophageal Reflux	0.8434	0.0975	5.18×10^{-18}
Depression	Gastro-oesophageal Reflux	0.5221	0.0271	1.48×10^{-82}
Depression	Osteoarthritis	0.3556	0.0519	7.19×10^{-12}
Depression	Suicidal Ideation Self Harm	0.9687	0.1904	3.65×10^{-07}
Gastro-oesophageal Reflux	Inflammatory Bowel Disease	-0.1065	0.0293	2.84×10^{-04}
Gastro-oesophageal Reflux	Lipoprotein HDL	-0.1954	0.0246	1.72×10^{-15}
Gastro-oesophageal Reflux	Osteoarthritis	0.5348	0.0446	3.50×10^{-33}
Hypertension	Lipoprotein HDL	-0.2728	0.0712	10^{-04}

Table 3: **Pairwise significant genetic correlations.** R_g denotes the genetic correlation coefficient; SE and P denote the standard error and test p -value for R_g respectively. The genetic correlation was computed using LDSC (14). See also Figure 2.

Exposure	Outcome	β_{xy}	SE	P	N_{SNPs}
Depression	Anxiety	0.6175	0.1041	2.99×10^{-09}	50
Asthma	Gastro-oesophageal Reflux	-0.0384	0.0099	1.07×10^{-04}	66
Body Mass Index	Depression	0.1678	0.0222	4.08×10^{-14}	541
Gastro-oesophageal Reflux	Body Mass Index	0.0815	0.0076	1.12×10^{-26}	166
Body Mass Index	Gastro-oesophageal Reflux	0.2204	0.0163	1.35×10^{-41}	526
Inguinal Hernia	Body Mass Index	-0.0147	0.0045	1.10×10^{-03}	80
Body Mass Index	Inguinal Hernia	-0.3566	0.0345	5.61×10^{-25}	525
Body Mass Index	Lipoprotein HDL	-0.2668	0.0115	3.67×10^{-119}	473
Body Mass Index	Osteoarthritis	0.5141	0.0420	1.99×10^{-34}	549
Gastro-oesophageal Reflux	Chronic Gastritis	0.4523	0.0331	1.79×10^{-42}	157
Chronic Gastritis	Gastro-oesophageal Reflux	0.1604	0.0353	5.60×10^{-06}	10
Gastro-oesophageal Reflux	Depression	0.2524	0.0220	1.34×10^{-30}	169
Depression	Gastro-oesophageal Reflux	0.1098	0.0165	2.66×10^{-11}	64
Osteoarthritis	Depression	0.1008	0.0301	7.98×10^{-04}	16
Inflammatory Bowel Disease	Gastro-oesophageal Reflux	-0.0145	0.0036	4.80×10^{-05}	443
Gastro-oesophageal Reflux	Lipoprotein HDL	-0.0669	0.0148	6.49×10^{-06}	119
Gastro-oesophageal Reflux	Osteoarthritis	0.2969	0.0408	3.46×10^{-13}	176
Lipoprotein HDL	Hypertension	-0.1298	0.0310	2.90×10^{-05}	216

Table 4: **Causalities between exposures and outcomes from the Mendelian randomization analysis.** β_{xy} , SE , and P are effect size, corresponding standard error, and test p -value, respectively. N_{SNPs} is the number of independent SNPs used as instruments in the analyses. Mendelian randomization was carried out using the GSMR tool implemented in GCTA (111). See Figure 3 as well.

GeneName	CHR	START	STOP	N_{SNPs}	Z_{pleio1}	P_{pleio1}	Z_{pleio2}	P_{pleio2}
Depression - GORD								
<i>CCKBR</i>	11	6280841	6293363	67	4.10	2.09×10^{-05}	5.18	1.14×10^{-07}
<i>PGBD1</i>	6	28249314	28270326	59	4.85	6.30×10^{-07}	4.06	2.41×10^{-05}
<i>HIST1H2BN</i>	6	27805544	27821533	32	4.82	7.34×10^{-07}	4.55	2.73×10^{-06}
<i>ZSCAN26</i>	6	28234788	28246001	18	4.87	5.56×10^{-07}	3.40	3.32×10^{-04}
<i>TCF4</i>	18	52889562	53303252	677	4.46	4.18×10^{-06}	5.91	1.67×10^{-09}
<i>ZKSCAN4</i>	6	28212404	28227030	33	4.99	2.94×10^{-07}	4.07	2.34×10^{-05}
<i>BTN2A1</i>	6	26458132	26476849	79	4.56	2.58×10^{-06}	4.67	1.54×10^{-06}
<i>BTN3A2</i>	6	26365387	26378548	97	5.03	2.49×10^{-07}	3.85	5.93×10^{-05}
Depression - BMI								
<i>IP6K1</i>	3	49761728	49823973	137	5.73	5.07×10^{-09}	3.70	1.07×10^{-04}
<i>RAB27B</i>	18	52385097	52562747	681	4.81	7.44×10^{-07}	5.21	9.47×10^{-08}
<i>DENND1A</i>	9	126141933	126692423	1825	8.32	4.51×10^{-17}	3.70	1.08×10^{-04}
<i>DNAJC11</i>	1	6694228	6761966	201	4.74	1.07×10^{-06}	3.75	8.84×10^{-05}
<i>NEGR1</i>	1	71868625	72748533	2550	9.24	1.29×10^{-20}	4.37	6.30×10^{-06}

Table 5: **Pleiotropic genes between depression and its immediate risk factors.** For all three immediate risk factors of depression, we looked into horizontal pleiotropic genetic variants shared between depression and the risk factor using IMRP (110). The analysis was done at both the SNP and the gene level. (See table S3 for SNP-based results.) GeneName, CHR, START, and STOP denote the gene and its genomic location; N_{SNPs} is number of SNPs analyzed for the respective gene; Z and P denote standardized pleiotropic effect sizes and test p -values; $pleio1$ stands for the SNP’s effect on depression, excluding the causality effect of the risk factor, and $pleio2$ stands for the SNP effect on the risk factor taking into account any causal effects from depression to the risk factor trait.

6 Supplementary table legends

Table S1. Parent phecodes and number of cases in our analysis. After pre-processing, we mapped 6,300 ICD-10 codes onto 412 parent phenodes, including depression. The table is a detailed list of final phecodes used in the analyses and the number of cases after quality control.

Table S2. Edges in the phenotypic network. We computed pairwise phenotypic correlation between all phecodes included in our study. We extracted the top 1% highly-correlated pairs. We further filtered out all immediate neighbour nodes of depression and the network spanning these nodes. This network includes a total of 56 edges (listed in the table). Edge weights are the phenotypic correlations between two nodes, defined as the cosine of the angle formed by the two phenotype vectors.

Table S3. SNP-based pleiotropic results for depression and GORD, and depression and BMI. Pleiotropic SNPs identified between depression and its immediate risk factors (GORD and BMI) were identified by the Mendelian randomization analysis.

References

- [1] *Diagnostic and statistical manual of mental disorders : DSM-5™*. American Psychiatric Publishing, a division of American Psychiatric Association, Washington, DC ;, 5th edition. edition, 2013 - 2013.
- [2] Áine Abautret-Daly, Elaine Dempsey, Adolfo Parra-Blanco, Carlos Medina, and Andrew Harkin. Gut–brain actions underlying comorbid anxiety and depression associated with inflammatory bowel disease. *Acta neuropsychiatrica*, 30(5):275–296, 2018.
- [3] Emmanuel O Adewuyi, Divya Mehta, Yadav Sapkota, Asa Auta, Kosuke Yoshihara, Mette Nyegaard, Lyn R Griffiths, Grant W Montgomery, Daniel I Chasman, and Dale R Nyholt. Genetic analysis of endometriosis and depression identifies shared loci and implicates causal links with gastric mucosa abnormality. *Human genetics*, 140(3):529–552, 2021.
- [4] Gajender Aleti, Jordan N Kohn, Emily A Troyer, Kelly Weldon, Shi Huang, Anupriya Tripathi, Pieter C Dorrestein, Austin D Swafford, Rob Knight, and Suzi Hong. Salivary bacterial signatures in depression-obesity comorbidity are associated with neurotransmitters and neuroactive dipeptides. *medRxiv*, 2021.
- [5] Jeanne Amiel, Marlène Rio, Loïc De Pontual, Richard Redon, Valérie Malan, Nathalie Boddaert, Perrine Plouin, Nigel P Carter, Stanislas Lyonnet, Arnold Munnich, et al. Mutations in *tcf4*, encoding a class i basic helix-loop-helix transcription factor, are responsible for pitt-hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *The American Journal of Human Genetics*, 80(5):988–993, 2007.
- [6] Jiyuan An, Puya Gharahkhani, Matthew H Law, Jue-Sheng Ong, Xikun Han, Catherine M Olsen, Rachel E Neale, John Lai, Tom L Vaughan, Ines Gockel, et al. Gastroesophageal reflux gwas identifies risk loci that also associate with subsequent severe esophageal diseases. *Nature communications*, 10(1):1–10, 2019.
- [7] Fehmi Ates and Michael F Vaezi. Insight into the relationship between gastroesophageal reflux disease and asthma. *Gastroenterology & hepatology*, 10(11):729, 2014.

- [8] DM Badowska, MM Brzózka, N Kannaiyan, C Thomas, P Dibaj, A Chowdhury, H Steffens, CW Turck, P Falkai, A Schmitt, et al. Modulation of cognition and neuronal plasticity in gain-and loss-of-function mouse models of the schizophrenia risk gene *tcf4*. *Translational psychiatry*, 10(1):1–12, 2020.
- [9] Florianne Bauer, Clara C Elbers, Roger AH Adan, Ruth JF Loos, N Charlotte Onland-Moret, Diederick E Grobbee, Jana V van Vliet-Ostaptchouk, Cisca Wijmenga, and Yvonne T van der Schouw. Obesity genes identified in genome-wide association studies are associated with adiposity measures and potentially with nutrient-specific food preference. *The American journal of clinical nutrition*, 90(4):951–959, 2009.
- [10] Julia Beretov, Valerie C Wasinger, Ewan KA Millar, Peter Schwartz, Peter H Graham, and Yong Li. Proteomic analysis of urine to identify breast cancer biomarker candidates using a label-free lc-ms/ms approach. *PloS one*, 10(11):e0141876, 2015.
- [11] Aritra Bose, Vassilis Kalantzis, Eugenia-Maria Kontopoulou, Mai Elkady, Peristera Paschou, and Petros Drineas. Terapca: a fast and scalable software package to study genetic variation in tera-scale genotypes. *Bioinformatics*, 35(19):3679–3683, 2019.
- [12] C Broers, J Tack, and A Pauwels. gastro-oesophageal reflux disease in asthma and chronic obstructive pulmonary disease. *Alimentary pharmacology & therapeutics*, 47(2):176–191, 2018.
- [13] Magdalena M Brzózka, Konstantin Radyushkin, Sven P Wichert, Hannelore Ehrenreich, and Moritz J Rossner. Cognitive and sensorimotor gating impairments in transgenic mice overexpressing the schizophrenia susceptibility gene *tcf4* in the brain. *Biological psychiatry*, 68(1):33–40, 2010.
- [14] Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J Daly, Alkes L Price, and Benjamin M Neale. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*, 47(3):291–295, 2015.
- [15] Lei Cai, Yiran Bao, Xiaoqian Fu, Hongbao Cao, Ancha Baranova, Xiangrong Zhang, Jing

- Sun, and Fuquan Zhang. Causal links between major depressive disorder and insomnia: A mendelian randomisation study. *Gene*, 768:145271, 2021.
- [16] Cali A Calarco and Mary Kay Lobo. Depression and substance use disorders: Clinical comorbidity and shared neurobiology. In *International Review of Neurobiology*, volume 157, pages 245–309. Elsevier, 2021.
- [17] Mariko Carey, Hannah Small, Sze Lin Yoong, Allison Boyes, Alessandra Bisquera, and Rob Sanson-Fisher. Prevalence of comorbid depression and obesity in general practice: a cross-sectional survey. *British Journal of General Practice*, 64(620):e122–e127, 2014.
- [18] Robert M Carney, Kenneth E Freedland, Gregory E Miller, and Allan S Jaffe. Depression as a risk factor for cardiac mortality and morbidity: a review of potential mechanisms. *Journal of psychosomatic research*, 53(4):897–902, 2002.
- [19] Woo Ri Chae, Anja Schienkiewitz, Yong Du, Ulfert Hapke, Christian Otte, and Niels Michalski. Comorbid depression and obesity among adults in germany: Effects of age, sex, and socioeconomic status. *Journal of affective disorders*, 299:383–392, 2022.
- [20] Po-Han Chou, Chih-Chien Lin, Ching-Heng Lin, Chia-Jui Tsai, Chin Cheng, Yi-Ping Chuo, Chin-Hong Chan, and Tsuo-Hung Lan. Prevalence of gastroesophageal reflux disease in major depressive disorder: a population-based study. *Psychosomatics*, 55(2):155–162, 2014.
- [21] Sarah Christen. The iglon family in ovarian cancer: Functional characterisation of lsamp and negr1 in comparison with the tumour suppressor opcml. 2013.
- [22] Michael Daly, Angelina R Sutin, and Eric Robinson. Depression reported by us adults in 2017–2018 and march and april 2020. *Journal of affective disorders*, 278:131–135, 2021.
- [23] Katrina M De Lange, Loukas Moutsianas, James C Lee, Christopher A Lamb, Yang Luo, Nicholas A Kennedy, Luke Jostins, Daniel L Rice, Javier Gutierrez-Achury, Sun-Gou Ji, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature genetics*, 49(2):256–261, 2017.
- [24] Christiaan A de Leeuw, Joris M Mooij, Tom Heskes, and Danielle Posthuma. Magma: generalized gene-set analysis of gwas data. *PLoS computational biology*, 11(4):e1004219, 2015.

- [25] Leonore M De Wit, Annemieke Van Straten, Marieke Van Hertem, Brenda WJH Penninx, and Pim Cuijpers. Depression and body mass index, a u-shaped association. *BMC public health*, 9(1):1–6, 2009.
- [26] Florence Demenais, Patricia Margaritte-Jeannin, Kathleen C Barnes, William OC Cookson, Janine Altmüller, Wei Ang, R Graham Barr, Terri H Beaty, Allan B Becker, John Beilby, et al. Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nature genetics*, 50(1):42–53, 2018.
- [27] Emily DiBlasi, Joeeun Kang, and Anna R Docherty. Genetic contributions to suicidal thoughts and behaviors. *Psychological medicine*, 51(13):2148–2155, 2021.
- [28] Ana Teresa D D’Elia, Mario F Juruena, Bruno M Coimbra, Marcelo F Mello, and Andrea F Mello. Posttraumatic stress disorder (ptsd) and depression severity in sexually assaulted women: hypothalamic-pituitary-adrenal (hpa) axis alterations. *BMC psychiatry*, 21(1):1–12, 2021.
- [29] Amy Farabaugh, Stella Bitran, Maren Nyer, Daphne J Holt, Paola Pedrelli, Irene Shyu, Steven D Hollon, Sidney Zisook, Lee Baer, Wilma Busse, et al. Depression and suicidal ideation in college students. *Psychopathology*, 45(4):228–234, 2012.
- [30] Anthony Feinstein, Sandra Magalhaes, Jean-Francois Richard, Blair Audet, and Craig Moore. The link between multiple sclerosis and depression. *Nature Reviews Neurology*, 10(9):507–517, 2014.
- [31] Mark Feldman, Pamela Walker, Janet L Green, and Kathy Weingarden. Life events stress and psychosocial factors in men with peptic ulcer disease: a multidimensional case-controlled study. *Gastroenterology*, 91(6):1370–1379, 1986.
- [32] Ana Maria Fernandez-Pujals, Mark James Adams, Pippa Thomson, Andrew G McKechnie, Douglas HR Blackwood, Blair H Smith, Anna F Dominiczak, Andrew D Morris, Keith Matthews, Archie Campbell, et al. Epidemiology and heritability of major depressive disorder, stratified by age of onset, sex, and illness course in generation scotland: Scottish family health study (gs: Sfhs). *PloS one*, 10(11):e0142197, 2015.

- [33] Marc P Forrest, Matthew J Hill, Andrew J Quantock, Enca Martin-Rendon, and Derek J Blake. The emerging roles of tcf4 in disease and development. *Trends in molecular medicine*, 20(6):322–331, 2014.
- [34] Gernot Fugger, Markus Dold, Lucie Bartova, Alexander Kautzky, Daniel Souery, Julien Mendlewicz, Alessandro Serretti, Joseph Zohar, Stuart Montgomery, Richard Frey, et al. Comorbid hypertension in patients with major depressive disorder—results from a european multicenter study. *European Neuropsychopharmacology*, 29(6):777–785, 2019.
- [35] C Gamero-Villaruel, L María González, I Gordillo, J Antonio Carrillo, A García-Herráiz, I Flores, R Rodríguez-López, and G Gervasini. Impact of *negr1* genetic variability on psychological traits of patients with eating disorders. *The pharmacogenomics journal*, 15(3):278–283, 2015.
- [36] Kylie P Glanville, Jonathan RI Coleman, Paul F O’Reilly, James Galloway, and Cathryn M Lewis. Investigating pleiotropy between depression and autoimmune diseases using the uk biobank. *Biological psychiatry global open science*, 1(1):48–58, 2021.
- [37] Alexander H Glassman. Depression and cardiovascular comorbidity. *Dialogues in clinical neuroscience*, 9(1):9, 2007.
- [38] Thelma B González-Castro, Carlos A Tovilla-Zárate, Alma D Genis-Mendoza, Isela E Juárez-Rojop, Humberto Nicolini, María L López-Narváez, and José J Martínez-Magaña. Identification of gene ontology and pathways implicated in suicide behavior: Systematic review and enrichment analysis of gwas studies. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 180(5):320–329, 2019.
- [39] Claudia Gagnoli. Hypothesis of the neuroendocrine cortisol pathway gene role in the comorbidity of depression, type 2 diabetes, and metabolic syndrome. *The application of clinical genetics*, 7:43, 2014.
- [40] A Halaris. Comorbidity between depression and cardiovascular disease. *International Angiology*, 28(2):92, 2009.

- [41] David L Hare, Samia R Toukhsati, Peter Johansson, and Tiny Jaarsma. Depression and cardiovascular disease: a clinical review. *European heart journal*, 35(21):1365–1372, 2014.
- [42] Thomas J Hoffmann, H el ene Choquet, Jie Yin, Yambazi Banda, Mark N Kvale, Maria Glymour, Catherine Schaefer, Neil Risch, and Eric Jorgenson. A large multiethnic genome-wide association study of adult body mass index identifies novel loci. *Genetics*, 210(2):499–515, 2018.
- [43] David M Howard, Mark J Adams, Toni-Kim Clarke, Jonathan D Hafferty, Jude Gibson, Masoud Shirali, Jonathan RI Coleman, Saskia P Hagenaars, Joey Ward, Eleanor M Wigmore, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience*, 22(3):343–352, 2019.
- [44] Masashi Ikeda, Atsushi Takahashi, Yoichiro Kamatani, Yukihide Momozawa, Takeo Saito, Kenji Kondo, Ayu Shimasaki, Kohei Kawase, Takaya Sakusabe, Yoshimi Iwayama, et al. Genome-wide association study detected novel susceptibility genes for schizophrenia and shared trans-populations/diseases genetic effect. *Schizophrenia bulletin*, 45(4):824–834, 2019.
- [45] Sinem Ilgin, Volkan Kilic, Merve Baysal, Gozde Aydogan-Kilic, Seyda Ucarcan, Begum Dermenci, and Ozlem Atli. Evidence for cardiotoxicity associated with sertraline in rats. *Toxicology research*, 7(5):817–825, 2018.
- [46] Hyojin Im, Laura ET Swan, Abdilkadir H Warsame, and Maimuna M Isse. Risk and protective factors for comorbidity of ptsd, depression, and anxiety among somali refugees in kenya. *International journal of social psychiatry*, 68(1):134–146, 2022.
- [47] Mingdi Jiang, Ping Qin, and Xu Yang. Comorbidity between depression and asthma via immune-inflammatory pathways: a meta-analysis. *Journal of affective disorders*, 166:22–29, 2014.
- [48] Tim Kaiser, Philipp Herzog, Ulrich Voderholzer, and Eva-Lotta Brakemeier. Unraveling the comorbidity of depression and anxiety in a large inpatient sample: Network analysis to examine bridge symptoms. *Depression and anxiety*, 38(3):307–317, 2021.

- [49] Masahiro Kanai, Masato Akiyama, Atsushi Takahashi, Nana Matoba, Yukihide Momozawa, Masashi Ikeda, Nakao Iwata, Shiro Ikegawa, Makoto Hirata, Koichi Matsuda, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nature genetics*, 50(3):390–400, 2018.
- [50] Hyejin Kim, Ji-Sook Hwang, Jinpyo Hong, Bogman Lee, and Soojin Lee. Newly identified cancer-associated role of human neuronal growth regulator 1 (*negr1*). *Journal of Cancer*, 5(7):598, 2014.
- [51] Lydia Kriegl, David Horst, Jana A Reiche, Jutta Engel, Thomas Kirchner, and Andreas Jung. *Lef-1* and *tcf4* expression correlate inversely with survival in colorectal cancer. *Journal of translational medicine*, 8(1):1–8, 2010.
- [52] E Kwon, W Wang, and LH Tsai. Validation of schizophrenia-associated genes *csmd1*, *c10orf26*, *cacna1c* and *tcf4* as *mir-137* targets. *Molecular psychiatry*, 18(1):11–12, 2013.
- [53] AWS Lee, H Hengstler, K Schwald, M Berriel-Diaz, D Loreth, et al. Functional inactivation of the genome-wide association study obesity gene. 2012.
- [54] James J Lee, Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, Tuan Anh Nguyen-Viet, Peter Bowers, Julia Sidorenko, Richard Karlsson Linnér, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature genetics*, 50(8):1112–1121, 2018.
- [55] Daniel F Levey, Murray B Stein, Frank R Wendt, Gita A Pathak, Hang Zhou, Mihaela Aslan, Rachel Quaden, Kelly M Harrington, Yaira Z Nuñez, Cassie Overstreet, et al. Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in 1.2 million individuals highlight new therapeutic directions. *Nature neuroscience*, 24(7):954–963, 2021.
- [56] Hong Li, Ying Zhu, Yury M Morozov, Xiaoli Chen, Stephanie Cerceo Page, Matthew D Rannals, Brady J Maher, and Pasko Rakic. Disruption of *tcf4* regulatory networks leads to abnormal cortical development and mental disabilities. *Molecular psychiatry*, 24(8):1235–1246, 2019.

- [57] Qingqing Liu, Hairong He, Jin Yang, Xiaojie Feng, Fanfan Zhao, and Jun Lyu. Changes in the global burden of depression from 1990 to 2017: Findings from the global burden of disease study. *Journal of psychiatric research*, 126:134–140, 2020.
- [58] Santiago A Lozano-Calderon, J Sebastiaan Souer, Jesse B Jupiter, and David Ring. Psychological differences between patients that elect operative or nonoperative treatment for trapeziometacarpal joint arthrosis. *Hand*, 3(3):271–275, 2008.
- [59] Floriana S Luppino, Leonore M de Wit, Paul F Bouvy, Theo Stijnen, Pim Cuijpers, Brenda WJH Penninx, and Frans G Zitman. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Archives of general psychiatry*, 67(3):220–229, 2010.
- [60] Ruth Ann Marrie. Comorbidity in multiple sclerosis: implications for patient care. *Nature Reviews Neurology*, 13(6):375–382, 2017.
- [61] E Martin-Rodriguez, F Guillen-Grima, E Aubá, A Martí, and A Brugos-Larumbe. Relationship between body mass index and depression in women: a 7-year prospective cohort study. the apna study. *European psychiatry*, 32:55–60, 2016.
- [62] Marta Martin-Subero, George Anderson, Buranee Kanchanatawan, Michael Berk, and Michael Maes. Comorbidity between depression and inflammatory bowel disease explained by immune-inflammatory, oxidative, and nitrosative stress; tryptophan catabolite; and gut–brain pathways. *CNS spectrums*, 21(2):184–198, 2016.
- [63] Peter Marx, Peter Antal, Bence Bolgar, Gyorgy Bagdy, Bill Deakin, and Gabriella Juhasz. Comorbidities in the diseasome are more apparent than real: what bayesian filtering reveals about the comorbidities of depression. *PLoS computational biology*, 13(6):e1005487, 2017.
- [64] Dermot PB McGovern, Subra Kugathasan, and Judy H Cho. Genetics of inflammatory bowel diseases. *Gastroenterology*, 149(5):1163–1176, 2015.
- [65] Matt McGue and Kaare Christensen. The heritability of depression symptoms in elderly danish twins: occasion-specific versus general effects. *Behavior genetics*, 33(2):83–93, 2003.

- [66] Aurora Mejía-Benítez, Miguel Klünder-Klünder, Loic Yengo, David Meyre, Celia Aradillas, Esperanza Cruz, Elva Pérez-Luque, Juan Manuel Malacara, Maria Eugenia Garay, Jesús Peralta-Romero, et al. Analysis of the contribution of *fto*, *npc1*, *enpp1*, *negr1*, *gnpda2* and *mc4rgenes* to obesity in mexican children. *BMC medical genetics*, 14(1):1–6, 2013.
- [67] Maaïke Meurs, Nynke A Groenewold, Annelieke M Roest, Nic JA van der Wee, Dick J Veltman, Marie-José van Tol, and Peter de Jonge. The associations of depression and hypertension with brain volumes: Independent or interactive? *NeuroImage: Clinical*, 8:79–86, 2015.
- [68] Nadia Micali, Alison E Field, Janet L Treasure, and David M Evans. Are obesity risk genes associated with binge eating in adolescence? *Obesity*, 23(8):1729–1736, 2015.
- [69] Darío Moreno-Agostino, Yu-Tzu Wu, Christina Daskalopoulou, M Tasdik Hasan, Martijn Huisman, and Matthew Prina. Global trends in the prevalence and incidence of depression: a systematic review and meta-analysis. *Journal of Affective Disorders*, 281:235–243, 2021.
- [70] Genevieve Morneau-Vaillancourt, Jonathan RI Coleman, Kirstin L Purves, Rosa Cheesman, Christopher Rayner, Gerome Breen, and Thalia C Eley. The genetic and environmental hierarchical structure of anxiety and depression in the uk biobank. *Depression and anxiety*, 37(6):512–520, 2020.
- [71] Joanna Mossakowska-Wójcik, Agata Orzechowska, Monika Talarowska, Janusz Szemraj, and Piotr Gałeccki. The importance of *tcf4* gene in the etiology of recurrent depressive disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 80:304–308, 2018.
- [72] Anwar Mulugeta, Ang Zhou, Catherine King, and Elina Hyppönen. Association between major depressive disorder and multiple disease outcomes: a phenome-wide mendelian randomisation study in the uk biobank. *Molecular psychiatry*, 25(7):1469–1476, 2020.
- [73] Sharif Mustaffa, Rashid Aziz, Mohd Nasir Mahmood, and Shukri Shuib. Depression and suicidal ideation among university students. *Procedia-Social and Behavioral Sciences*, 116:4205–4208, 2014.
- [74] Abdolrahman S Nateri, Bradley Spencer-Dene, and Axel Behrens. Interaction of phospho-

- rylated c-jun with tcf4 regulates intestinal cancer development. *Nature*, 437(7056):281–285, 2005.
- [75] Charles B Nemeroff and Pascal J Goldschmidt-Clermont. Heartache and heartbreak—the link between depression and cardiovascular disease. *Nature Reviews Cardiology*, 9(9):526–539, 2012.
- [76] Maggie CY Ng, Claudia HT Tam, Wing Yee So, Janice SK Ho, Alfred W Chan, Heung Man Lee, Ying Wang, Vincent KL Lam, Juliana CN Chan, and Ronald CW Ma. Implication of genetic variants near *negr1*, *sec16b*, *tmem18*, *etv5/dgkg*, *gnpda2*, *lin7c/bdnf*, *mtch2*, *bcdin3d/faim2*, *sh2b1*, *fto*, *mc4r*, and *kctd15* with obesity and type 2 diabetes in 7705 chinese. *The Journal of Clinical Endocrinology & Metabolism*, 95(5):2418–2425, 2010.
- [77] Hua Ni, Min Xu, Gui-Lai Zhan, YU Fan, Hejiang Zhou, Hong-Yan Jiang, Wei-Hong Lu, Liwen Tan, Deng-Feng Zhang, Yong-Gang Yao, et al. The gwas risk genes for depression may be actively involved in alzheimer’s disease. *Journal of Alzheimer’s Disease*, 64(4):1149–1161, 2018.
- [78] Kyungchul Noh, Hyunkyung Lee, Tae-Yong Choi, Yeonhee Joo, Soo-Joeng Kim, Hyejin Kim, Jin Young Kim, Jeong Won Jahng, Soojin Lee, Se-Young Choi, et al. *Negr1* controls adult hippocampal neurogenesis and affective behaviors. *Molecular psychiatry*, 24(8):1189–1205, 2019.
- [79] Jue-Sheng Ong, Jiyuan An, Xikun Han, Matthew H Law, Priyanka Nandakumar, Johannes Schumacher, Ines Gockel, Anne Bohmer, Janusz Jankowski, Claire Palles, et al. Multitrait genetic association analysis identifies 50 new risk loci for gastro-oesophageal reflux, seven new loci for barrett’s oesophagus and provides insights into clinical heterogeneity in reflux diagnosis. *Gut*, 2021.
- [80] Takeshi Otowa, Karin Hek, Minyoung Lee, Enda M Byrne, Saira S Mirza, Michel G Nivard, Timothy Bigdeli, Steven H Aggen, Daniel Adkins, Aaron Wolen, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular psychiatry*, 21(10):1391–1399, 2016.

- [81] Jeanne-Marie Perotin, Gabrielle Wheway, Kamran Tariq, Adnan Azim, Robert A Ridley, Jonathan A Ward, James PR Schofield, Clair Barber, Peter Howarth, Donna E Davies, et al. Vulnerability to acid reflux of the airway epithelium in severe asthma. *European Respiratory Journal*, 2021.
- [82] Victoria Kristina Perry. *Integration of copy number aberration and genomic expression data in preinvasive breast cancer*. PhD thesis, Boston University, 2012.
- [83] Alaitz Poveda, María Eugenia Ibáñez, and Esther Rebato. Common variants in *bdnf*, *faim2*, *fto*, *mc4r*, *negr1*, and *sh2b1* show association with obesity-related variables in a spanish roma population. *American Journal of Human Biology*, 26(5):660–669, 2014.
- [84] Hang Qiu, Liya Wang, Xianrong Zeng, and Jingping Pan. Comorbidity patterns in depression: A disease network analysis using regional hospital discharge records. *Journal of Affective Disorders*, 296:418–427, 2022.
- [85] Sobia Rana and Maha Mobin. Association of the *negr1* rs2815752 with obesity and related traits in pakistani females. *Upsala journal of medical sciences*, 125(3):226–234, 2020.
- [86] Lihong Ruan, Yuanyuan Xie, Fangmei Liu, and Xuehua Chen. Serum mir-1181 and mir-4314 associated with ovarian cancer: Mirna microarray data analysis for a pilot study. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 222:31–38, 2018.
- [87] Angelo Scuteri, Liana Spazzafumo, Luca Cipriani, Walter Gianni, Andrea Corsonello, Luca Cravello, Lazzaro Repetto, Silvia Bustacchini, Fabrizia Lattanzio, and Maurizio Sebastiani. Depression, hypertension, and comorbidity: disentangling their specific effect on disability and cognitive impairment in older subjects. *Archives of gerontology and geriatrics*, 52(3):253–257, 2011.
- [88] Mihoko Shimada-Sugimoto, Takeshi Otowa, and John M Hettema. Genetics of anxiety disorders: Genetic epidemiological and molecular studies in humans. *Psychiatry and clinical neurosciences*, 69(7):388–401, 2015.
- [89] JR Speakman, RJF Loos, Stephen O’Rahilly, JN Hirschhorn, and DB Allison. Gwas for

- bmi: a treasure trove of fundamental insights into the genetic basis of obesity. *International journal of obesity*, 42(8):1524–1531, 2018.
- [90] Stacy Steinberg, Simone de Jong, Ole A Andreassen, Thomas Werge, Anders D Børghlum, Ole Mors, Preben B Mortensen, Omar Gustafsson, Javier Costas, Olli PH Pietiläinen, et al. Common variants at *vrk2* and *tcf4* conferring risk of schizophrenia. *Human molecular genetics*, 20(20):4076–4081, 2011.
- [91] Patrick F Sullivan, Michael C Neale, and Kenneth S Kendler. Genetic epidemiology of major depression: review and meta-analysis. *American journal of psychiatry*, 157(10):1552–1562, 2000.
- [92] Joanna Szczurkowska, Francesca Pischedda, Bruno Pinto, Francesca Managò, Carola A Haas, Maria Summa, Rosalia Bertorelli, Francesco Papaleo, Michael K Schäfer, Giovanni Piccoli, et al. *Negr1* and *fgfr2* cooperatively regulate cortical development and core behaviours related to autism disorders in mice. *Brain*, 141(9):2772–2794, 2018.
- [93] Junko Takita, Yuyan Chen, Jun Okubo, Masashi Sanada, Masatoki Adachi, Kentaro Ohki, Riki Nishimura, Ryoji Hanada, Takashi Igarashi, Yasuhide Hayashi, et al. Aberrations of *negr1* on 1p31 and *myeov* on 11q13 in neuroblastoma. *Cancer science*, 102(9):1645–1650, 2011.
- [94] Monika Talarowska-Bogusz, Antoni Florkowski, Agata Radomska, Krzysztof Zboralski, and Piotr Gałeczki. Depression and pain in arthrosis of spine and hip in elderly patients. *Polski Merkurusz Lekarski: Organ Polskiego Towarzystwa Lekarskiego*, 21(126):566–569, 2006.
- [95] D Taylor. Antidepressant drugs and cardiovascular pathology: a clinical overview of effectiveness and safety. *Acta Psychiatrica Scandinavica*, 118(6):434–442, 2008.
- [96] Warren D Taylor, Douglas R McQuoid, and K Ranga Rama Krishnan. Medical comorbidity in late-life depression. *International Journal of Geriatric Psychiatry*, 19(10):935–943, 2004.
- [97] Wendela G Ter Meulen, Stasja Draisma, Albert M van Hemert, Robert A Schoevers, Ralph W Kupka, Aartjan TF Beekman, and Brenda WJH Penninx. Depressive and anxiety disorders

- in concert—a synthesis of findings on comorbidity in the nesda study. *Journal of affective disorders*, 284:85–97, 2021.
- [98] John WG Tiller. Depression and anxiety. *The Medical Journal of Australia*, 199(6):S28–S31, 2013.
- [99] Marion L Vetter, Thomas A Wadden, Jeffrey Lavenberg, Reneé H Moore, Sheri Volger, Joanna L Perez, David B Sarwer, and Adam Gilden Tsai. Relation of health-related quality of life to metabolic syndrome, obesity, depression and comorbid illnesses. *International journal of obesity*, 35(8):1087–1094, 2011.
- [100] Edward A Walker, Martin D Gelfand, Ann N Gelfand, Francis Creed, and Wayne J Katon. The relationship of current psychiatric disorder to functional disability and distress in patients with inflammatory bowel disease. *General hospital psychiatry*, 18(4):220–229, 1996.
- [101] Yan Wang and Ji-Guang Wang. Genome-wide association studies of hypertension and several other cardiovascular diseases. *Heart Genomics*, pages 1–29, 2018.
- [102] Kyoko Watanabe, Sven Stringer, Oleksandr Frei, Maša Umićević Mirkov, Christiaan de Leeuw, Tinca JC Polderman, Sophie van der Sluis, Ole A Andreassen, Benjamin M Neale, and Danielle Posthuma. A global overview of pleiotropy and genetic architecture in complex traits. *Nature genetics*, 51(9):1339–1348, 2019.
- [103] Cristen J Willer, Ellen M Schmidt, Sebanti Sengupta, Gina M Peloso, Stefan Gustafsson, Stavroula Kanoni, Andrea Ganna, Jin Chen, Martin L Buchkovich, Samia Mora, et al. Discovery and refinement of loci associated with lipid levels. *Nature genetics*, 45(11):1274, 2013.
- [104] Jolanta Wilmańska and Ewa Gułaj. Coincidence of arthrosis and depression in elderly patients-therapeutic implications. *Ortopedia, traumatologia, rehabilitacja*, 8(6):686–692, 2006.
- [105] Genevieve L Wojcik, Mariaelisa Graff, Katherine K Nishimura, Ran Tao, Jeffrey Haessler, Christopher R Gignoux, Heather M Highland, Yesha M Patel, Elena P Sorokin, Christy L Avery, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature*, 570(7762):514–518, 2019.

- [106] Naomi R Wray, Stephan Ripke, Manuel Mattheisen, Maciej Trzaskowski, Enda M Byrne, Abdel Abdellaoui, Mark J Adams, Esben Agerbo, Tracy M Air, Till MF Andlauer, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics*, 50(5):668–681, 2018.
- [107] Yeda Wu, Graham K Murray, Enda M Byrne, Julia Sidorenko, Peter M Visscher, and Naomi R Wray. Gwas of peptic ulcer disease implicates helicobacter pylori infection, other gastrointestinal disorders and depression. *Nature communications*, 12(1):1–17, 2021.
- [108] Eleni Zengini, Konstantinos Hatzikotoulas, Ioanna Tachmazidou, Julia Steinberg, Fernando P Hartwig, Lorraine Southam, Sophie Hackinger, Cindy G Boer, Unnur Styrkarsdottir, Arthur Gilly, et al. Genome-wide analyses using uk biobank data provide insights into the genetic architecture of osteoarthritis. *Nature genetics*, 50(4):549–558, 2018.
- [109] G Zhao, Earl S Ford, Satvinder Dhingra, Chaoyang Li, Tara W Strine, and AH Mokdad. Depression and anxiety among us adults: associations with body mass index. *International journal of obesity*, 33(2):257–266, 2009.
- [110] Xiaofeng Zhu, Xiaoyin Li, Rong Xu, and Tao Wang. An iterative approach to detect pleiotropy and perform mendelian randomization analysis using gwas summary statistics. *Bioinformatics*, 37(10):1390–1400, 2021.
- [111] Zhihong Zhu, Zhili Zheng, Futao Zhang, Yang Wu, Maciej Trzaskowski, Robert Maier, Matthew R Robinson, John J McGrath, Peter M Visscher, Naomi R Wray, et al. Causal associations between risk factors and common diseases inferred from gwas summary data. *Nature communications*, 9(1):1–12, 2018.