

## Original Article

**Cite this article:** Wainberg M, Zhukovsky P, Hill SL, Felsky D, Voineskos A, Kennedy S, Hawco C, Tripathy SJ (2023). Symptom dimensions of major depression in a large community-based cohort. *Psychological Medicine* 53, 438–445. <https://doi.org/10.1017/S0033291721001707>

Received: 15 January 2021

Revised: 9 April 2021

Accepted: 18 April 2021

First published online: 19 May 2021


**Key words:**

Major depression; symptom dimensions; comorbidity; polygenic risk scores; UK Biobank

**Author for correspondence:**

Shreejoy J. Tripathy,  
E-mail: [shreejoy.tripathy@camh.ca](mailto:shreejoy.tripathy@camh.ca)

## Symptom dimensions of major depression in a large community-based cohort

Michael Wainberg<sup>1</sup> , Peter Zhukovsky<sup>1</sup>, Sean L. Hill<sup>1,2,3,4</sup>, Daniel Felsky<sup>1,2,3</sup>, Aristotle Voineskos<sup>1,2,3</sup>, Sidney Kennedy<sup>2,3,5,6</sup>, Colin Hawco<sup>1,2,3</sup> and Shreejoy J. Tripathy<sup>1,2,3,4</sup>

<sup>1</sup>Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health, Toronto, ON, Canada; <sup>2</sup>Department of Psychiatry, University of Toronto, Toronto, Canada; <sup>3</sup>Institute of Medical Sciences, University of Toronto, Toronto, Canada; <sup>4</sup>Department of Physiology, University of Toronto, Toronto, Canada; <sup>5</sup>Krembil Research Institute, University Health Network, Toronto, Canada and <sup>6</sup>Li Ka Shing Knowledge Institute, Saint Michael's Hospital, Toronto, Canada

**Abstract**

**Background.** Our understanding of major depression is complicated by substantial heterogeneity in disease presentation, which can be disentangled by data-driven analyses of depressive symptom dimensions. We aimed to determine the clinical portrait of such symptom dimensions among individuals in the community.

**Methods.** This cross-sectional study consisted of 25 261 self-reported White UK Biobank participants with major depression. Nine questions from the UK Biobank Mental Health Questionnaire encompassing depressive symptoms were decomposed into underlying factors or ‘symptom dimensions’ via factor analysis, which were then tested for association with psychiatric diagnoses and polygenic risk scores for major depressive disorder (MDD), bipolar disorder and schizophrenia. Replication was performed among 655 self-reported non-White participants, across sexes, and among 7190 individuals with an ICD-10 code for MDD from linked inpatient or primary care records.

**Results.** Four broad symptom dimensions were identified, encompassing negative cognition, functional impairment, insomnia and atypical symptoms. These dimensions replicated across ancestries, sexes and individuals with inpatient or primary care MDD diagnoses, and were also consistent among 43 090 self-reported White participants with undiagnosed self-reported depression. Every dimension was associated with increased risk of nearly every psychiatric diagnosis and polygenic risk score. However, while certain psychiatric diagnoses were disproportionately associated with specific symptom dimensions, the three polygenic risk scores did not show the same specificity of associations.

**Conclusions.** An analysis of questionnaire data from a large community-based cohort reveals four replicable symptom dimensions of depression with distinct clinical, but not genetic, correlates.

**Introduction**

It has long been recognized that major depression is a heterogeneous disorder (Blumenthal, 1971). Indeed, it is increasingly appreciated that mental illnesses display both strong heterogeneity within disorders and strong pathophysiological and symptomatic overlap across disorders. Symptoms frequently transcend discrete, mutually exclusive diagnostic categories (Marshall, 2020): a Danish population-based study found that every mental illness is associated with an increased risk of every other mental illness (Plana-Ripoll et al., 2019). The same genetic variants often affect the risk of multiple mental illnesses (Anttila et al., 2018), to such an extent that all mental illnesses have been classified neurobiologically as variations along with a single ‘*p* factor’ (Caspi et al., 2014), albeit with some disorder-specific variation (Shanmugan et al., 2016). Depressive symptoms in particular do not only constitute an autonomous disorder, but may also arise reactively to the experience of environmental stressors or occur as a comorbidity in numerous other mental disorders, for instance in schizophrenia (Häfner et al., 2005).

These insights have driven a revisionary approach to psychiatric nosologies. The DSM-5 broadened the use of ‘specifiers’ (e.g. ‘with atypical features’, ‘with psychotic features’) in an attempt to refine clinical subtypes within a major depressive episode (American Psychiatric Association, 2013). The ICD-11 introduced an analogous notion of ‘qualifiers’, e.g. ‘with prominent anxiety symptoms’ (Stein et al., 2020). Clinical research, spurred partly by the US National Institute of Mental Health’s Research Domain Criteria (RDoC) initiative (Insel et al., 2010), has described several symptom-based (van Loo, de Jonge, Romeijn, Kessler, &

Schoevers, 2012) and biological (Beijers, Wardenaar, van Loo, & Schoevers, 2019; Drysdale et al., 2017) depressive subtypes.

The use of large-scale data is promising in its potential to identify latent dimensions of psychopathology (Weissman, 2020): the availability of large, deeply phenotyped cohort studies, perhaps most strongly exemplified by the UK Biobank (Bycroft et al., 2018), has the potential to enhance our understanding of the neuropsychiatric disease. What the UK Biobank lacks in depth of psychiatric measurement it makes up for in breadth, with genotypes and thousands of phenotypes assessed on up to half a million British participants. Analyses of this large-scale sample suggest that there are no detectable genetic subgroups of depressed patients (Howard et al., 2020) and that atypical depression (defined based on the presence of both hypersomnia and weight gain) is associated with more severe symptoms and more frequent psychiatric and non-psychiatric comorbidities (Brailean, Curtis, Davis, Dregan, & Hotopf, 2020).

Rather than relying on established nosologies, we embarked upon an agnostic, data-driven approach. Specifically, we performed a factor analysis on nine questions from the UK Biobank Mental Health Questionnaire (Davis et al., 2020) pertaining to an individual's worst reported lifetime episode of depression. Factor analysis and other techniques for elucidating the underlying symptom structure of multi-dimensional data have a rich history in depression research, dating back to the development of early depression rating scales such as the Hamilton Depression Rating Scale (Hamilton, 1960), which found four symptom dimensions roughly corresponding to negative cognition/psychomotor retardation, gastrointestinal symptoms/initial insomnia/weight loss/anhedonia, anxiety/agitation and somatic anxiety/nighttime awakening. Other studies have since explored a richer set of symptoms by pooling questions across multiple depression rating scales (Ballard et al., 2018; Fried, 2017). For instance, one study found eight factors encompassing depressed mood, tension, negative cognition, suicidal thoughts, impaired sleep, reduced appetite, anhedonia and amotivation (Ballard et al., 2018).

However, the unique characteristics of the UK Biobank cohort allow us to go beyond previous studies of depressive symptoms in two key ways. First, its clinical heterogeneity enables a direct comparison between individuals with major depression and demographically similar individuals in the community with undiagnosed self-reported depression. Second, the UK Biobank's extraordinary breadth of phenotyping facilitates the painting of a rich clinical portrait of individual depressive symptom dimensions.

## Methods

### Participants

Participants were included from the UK Biobank (Fig. 1), a community-based cohort study with genetics and deep phenotyping on approximately half a million individuals from across the UK, aged 40–69 years at recruitment (Bycroft et al., 2018). A total of 157 338 participants completed an online Mental Health Questionnaire (Davis et al., 2020), of whom 33 414 (21%) reported ever being diagnosed with depression by a health professional, a case definition we call 'major depression' following the terminology of the Psychiatric Genomics Consortium (McIntosh, Sullivan, & Lewis, 2019).

A total of 85 943 participants of the 157 338 (55%) answered yes to the question 'Have you ever had a time in your life when you felt sad, blue, or depressed for two weeks or more in a row?'. Note that this 55% is likely larger than the percentage of the general population who would endorse this question, because of selection bias in who responded to the emailed questionnaire invitation (Davis et al., 2020). This question is analogous to one of the two questions on the Patient Health Questionnaire-2 (PHQ-2) (Kroenke, Spitzer, & Williams, 2003), a clinically validated screening tool for major depressive disorder (MDD) (Levis et al., 2020), as well as to one of the two questions on the Composite International Diagnostic Interview short-form (Kessler, Andrews, Mroczek, Ustun, & Wittchen, 1998). This question was a prerequisite for being asked the nine questions included in the factor analysis.

Of the 33 414 participants reporting a diagnosis of major depression, almost all, 31 675 (95%), also reported ever feeling 'sad, blue or depressed' for 2 weeks or more. These 31 675 participants were further subsetted to lists of self-reported White ( $N = 25\,261$ ) and non-White ( $N = 655$ ) participants.

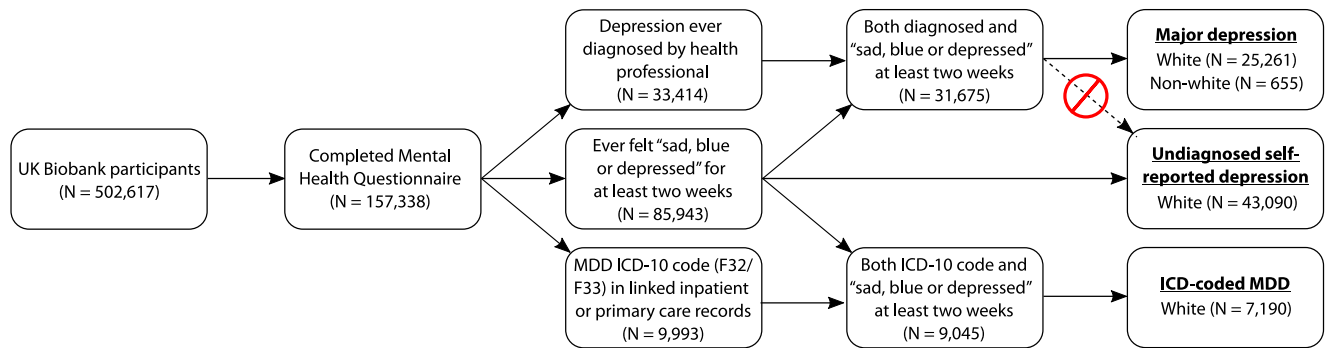
### Exploratory factor analysis

An exploratory factor analysis (maximum likelihood with oblimin rotation) was performed for the largest ancestry group, self-reported White participants, across nine questions from the Mental Health Questionnaire pertaining to an individual's worst reported lifetime episode of depression (Table 1), with the aim of identifying a small number of latent factors that could explain the majority of variance in responses across the nine questions. We note the possibility that these questions might pertain to a depressive episode even worse (from the patient's perspective) than the episode during which they were formally diagnosed, but we consider it unlikely that this even worse episode would not also merit the same formal depression diagnosis.

The exploratory factor analysis was conducted using version 1.9.12.31 of the *psych* package in version 3.5.3 of the R programming language. Specifically, the *polychoric* function was used to compute polychoric correlations among all pairs of the nine questions; then, a maximum likelihood factor analysis was run on the resulting correlation matrix using the *fa* function with oblimin rotation (Lawley, 1940). We selected the minimum number of factors with a high goodness of fit, defined as a Tucker–Lewis index (Tucker & Lewis, 1973) above 0.95 and root mean square error of approximation below 0.05. Correlation-preserving 'ten Berge' factor scores (ten Berge, Krijnen, Wansbeek, and Shapiro, 1999) were computed using the *factor.scores* function.

### Confirmatory factor analyses across diverse ancestries and depression case definitions

We performed several confirmatory factor analyses (CFAs) to replicate the symptom structure derived from the exploratory factor analysis of our primary cohort. First, to confirm generalizability to individuals of diverse ancestries, including those underrepresented in medical research (Smart & Harrison, 2017), we performed a CFA on 655 self-reported non-White participants. Second, we performed a CFA across sexes, in male and female White participants. Third, we performed a CFA on 7190 White participants with an ICD-10 code for MDD (F32 or F33) from linked inpatient or primary care records, which we call 'ICD-coded MDD'



**Fig. 1.** Flowchart of study inclusion criteria. Note that ICD-coded MDD is based on linked inpatient and primary care records, while other definitions are based exclusively on self-reporting of symptoms and/or diagnoses.

(Fig. 1) following the terminology of a recent genome-wide association study from the PGC (Howard et al., 2018). Finally, we performed CFA on 43 090 White participants who reported ever feeling ‘sad, blue, or depressed for two weeks or more in a row’ but not ever receiving a depression diagnosis, which we call ‘undiagnosed self-reported depression’ (Fig. 1). CFAs were conducted using the *cfa* function from version 0.6-6 of the *lavaan* R package (Rosseel, 2012), with default parameter settings.

### Polygenic risk scores

Polygenic risk scores (PRSs) were derived from public genome-wide association study (GWAS) results for MDD (Wray et al., 2018; [https://pgcdata.med.unc.edu/major\\_depressive\\_disorders/daner\\_pgc\\_mdd\\_meta\\_w2\\_no23andMe\\_rmUKBB.gz](https://pgcdata.med.unc.edu/major_depressive_disorders/daner_pgc_mdd_meta_w2_no23andMe_rmUKBB.gz)), bipolar disorder (Stahl et al., 2019; [https://pgcdata.med.unc.edu/bipolar\\_disorder/daner\\_PGC\\_BIP32b\\_mds7a\\_0416a.gz](https://pgcdata.med.unc.edu/bipolar_disorder/daner_PGC_BIP32b_mds7a_0416a.gz)) and schizophrenia (Pardiñas et al., 2018; [https://pgcdata.med.unc.edu/schizophrenia/SCZ\\_wave3/PGC3\\_SCZ\\_wave3\\_public.v2.tsv.gz](https://pgcdata.med.unc.edu/schizophrenia/SCZ_wave3/PGC3_SCZ_wave3_public.v2.tsv.gz)) across self-reported White participants.

The UK Biobank’s imputed genotypes were filtered using version 2.00 of the *plink* GWAS analysis software (Chang et al., 2015). Non-autosomal variants, duplicates, indels and variants with imputation INFO score <0.8 were removed, as were variants with Hardy–Weinberg equilibrium  $p$ -value < $10^{-10}$ , over 5% missingness, or minor allele frequency below 0.1% across self-reported White participants. Summary statistics were harmonized with the UK Biobank imputed genotypes with respect to reference/alternate allele and strand, using the allele harmonization framework from *munge\_sumstats.py* in the *ldsc* software package (Bulik-Sullivan et al., 2015). Ambiguous variants (A/T, C/G, G/C, T/A) and variants missing from the UK Biobank were excluded. Summary statistics were then subset to  $p < 0.05$ , a threshold found to be most predictive across all self-reported White participants in the UK Biobank (Table 2). Frequency-informed linkage disequilibrium (LD) pruning to  $r^2 > 0.2$  across the self-reported White participants was then performed using a 500 kb sliding window. The remaining variants constituted the trait’s PRS, with the variants’ effect sizes (beta coefficients for educational attainment, log odds ratios for the other three case-control studies) constituting the weights of the PRS. Finally, PRSs were scored on each individual in the study cohort by summing, across the variants in the PRS, the variant’s weight times the individual’s number of effect alleles of that variant; missing genotypes were mean-imputed.

### Associations with phenotypic variables and PRSs

The factors or ‘symptom dimensions’ resulting from factor analysis were associated with 19 mental illness-related fields, including 15 self-reported diagnosed mental illnesses, family history of severe depression, and the 3 PRSs described above. Statistical associations were performed using linear regression for continuous traits (i.e. PRSs), with results reported as effect sizes ( $\beta$  coefficients); or logistic regression for binary traits (i.e. diagnoses and family history), with results reported as odds ratios. Effect sizes and odds ratios were adjusted for age and sex, except for PRS associations, which were also adjusted for the top 10 genotype principal components.

Associations were performed using version 0.11.0 of the *statsmodels* Python package. For each phenotype and factor, a separate logistic (*statsmodels.Logit*; for binary phenotypes) or linear (*statsmodels.OLS*; for non-binary phenotypes) regression was conducted with the phenotype as the output variable and the factor scores as the input variable, with age and sex as covariates and the regression taking place across all individuals with non-missing values for the phenotype. Factor scores and non-binary phenotypes were both standardized to zero mean and unit variance, and can thus be interpreted as being per standard deviation increase in the factor score. To avoid convergence issues due to the presence of sex as a covariate, rare binary phenotypes exhibited by fewer than five males or five females were excluded.

Two-tailed  $p$ -values were calculated from the factor’s regression coefficient in the usual way, by dividing the coefficient by its standard error and then converting this  $z$ -score to a  $p$ -value by inverse-normal transformation; statistical significance was set at a false discovery rate of 5%. Since many phenotypes were correlated with most or all factors, but to varying degrees, we applied a difference-of-effect sizes test in order to compare the effect sizes with each other.  $z$ -scores for the difference of two regression coefficients  $\beta_1$  and  $\beta_2$  with standard errors  $\sigma_1$  and  $\sigma_2$  were calculated using the formula  $z_{\text{diff}} = (\beta_1 - \beta_2) / \sqrt{(\sigma_1^2 + \sigma_2^2)}$ .  $p$ -values were then computed by inverse-normal transforming these  $z$ -scores.

## Results

### Identification of depressive symptom dimensions

Four factors (‘symptom dimensions’) were identified by exploratory factor analysis of nine questions from the UK Biobank Mental Health Questionnaire pertaining to an individual’s worst

**Table 1.** The nine questions from the UK Biobank Mental Health Questionnaire included in the factor analysis

Short name	Field ID	Question phrasing
Tiredness	20449	Did you feel more tired out or low on energy than is usual for you?
Worthlessness	20450	People sometimes feel down on themselves, no good, worthless. Did you feel this way?
Thoughts of death	20437	Did you think a lot about death – either your own, someone else's or death in general?
Difficulty concentrating	20435	Did you have a lot more trouble concentrating than usual?
Impact on normal roles	20440	Think about your roles at the time of this episode, including study/employment, childcare and housework, leisure pursuits. How much did these problems interfere with your life or activities? → coded as: not at all (0), a little (1), somewhat (2), a lot (3)
Weight change	20536	Did you gain or lose weight without trying, or did you stay about the same weight? → coded as: gained weight (+1), lost weight (-1), stayed about the same or was on a diet (0)
Sleeping too much	20532 and 20534	20532: Did your sleep change? 20534: (if yes to 20532) Was that: [re sleep change] Sleeping too much?
Trouble falling asleep	20532 and 20533	20532: Did your sleep change? 20533: (if yes to 20532) Was that: [re sleep change] Trouble falling asleep?
Waking too early	20532 and 20535	20532: Did your sleep change? 20535: (if yes to 20532) Was that: [re sleep change] Waking too early?

All questions pertain to an individual's worst lifetime episode of depression, and all were coded as binary variables unless otherwise indicated.

**Table 2.** Predictive accuracy of PRSs in the UK Biobank at various  $p$ -value thresholds

Polygenic risk score	ICD code(s) benchmarked on (inpatient or primary care)	Area under the curve (AUC)			
		$p < 5 \times 10^{-5}$	$p < 0.0005$	$p < 0.005$	$p < 0.05$
MDD	F32 + F33	0.513	0.524	0.530	<b>0.540</b>
Bipolar disorder	F31	0.566	0.579	0.603	<b>0.609</b>
Schizophrenia	F20	0.605	0.629	0.663	<b>0.688</b>

The area under the curve (AUC), also known as the area under the receiver operating characteristic curve (AUROC) or concordance statistic (C statistic), is the fraction of the time that the polygenic risk score would rank a randomly chosen case higher than a randomly chosen control.

reported lifetime episode of depression (Table 3) across 25 261 individuals with major depression. Factors were labeled Factor A through Factor D, in order of decreasing variance explained, and roughly corresponded to the categories of atypical depressive symptoms (Factor A), functional impairment (Factor B), insomnia (Factor C) and negative cognition (Factor D). Four was the minimum number of factors required for high goodness of fit to the underlying questionnaire data (Methods), with a Tucker–Lewis index of 0.963 and root mean squared error of approximation of 0.047. As expected, these four factors do not fully represent the cohort's symptom structure; there is marked heterogeneity in which of the questions associated with each factor were endorsed individually or in combination. The internal structure of each factor is shown in Fig. 2.

#### Symptom dimensions are consistent across ancestries, sexes and depression case definitions

The identified factor structure replicated across ancestries, in 655 non-White participants with major depression ( $\chi^2_{26} = 58.6$ ,  $p = 0.0002$ ). It also replicated across sexes, in 7960 male ( $\chi^2_{26} = 470.1$ ,  $p = 5 \times 10^{-83}$ ) and 17 301 female ( $\chi^2_{26} = 1302.9$ ,  $p = 1 \times 10^{-258}$ ) White participants with major depression.

The factor structure also replicated across depression case definitions. It replicated in 7190 White participants with ICD-coded MDD ( $\chi^2_{26} = 600.06$ ,  $p = 6 \times 10^{-110}$ ). It also replicated in 43 090

White participants with undiagnosed self-reported depression, who reported ever feeling 'sad, blue, or depressed for two weeks or more in a row' but never receiving a depression diagnosis ( $\chi^2_{26} = 2565.27$ ,  $p = 3 \times 10^{-529}$ ), with a similar pattern of correlations among symptoms as in the full cohort (Fig. 3). This suggests that at least some individuals with undiagnosed self-reported depression would have met the criteria to be diagnosed with major depression, if they had only sought help at the time.

#### Associations with mental illness diagnoses, PRSs and family history

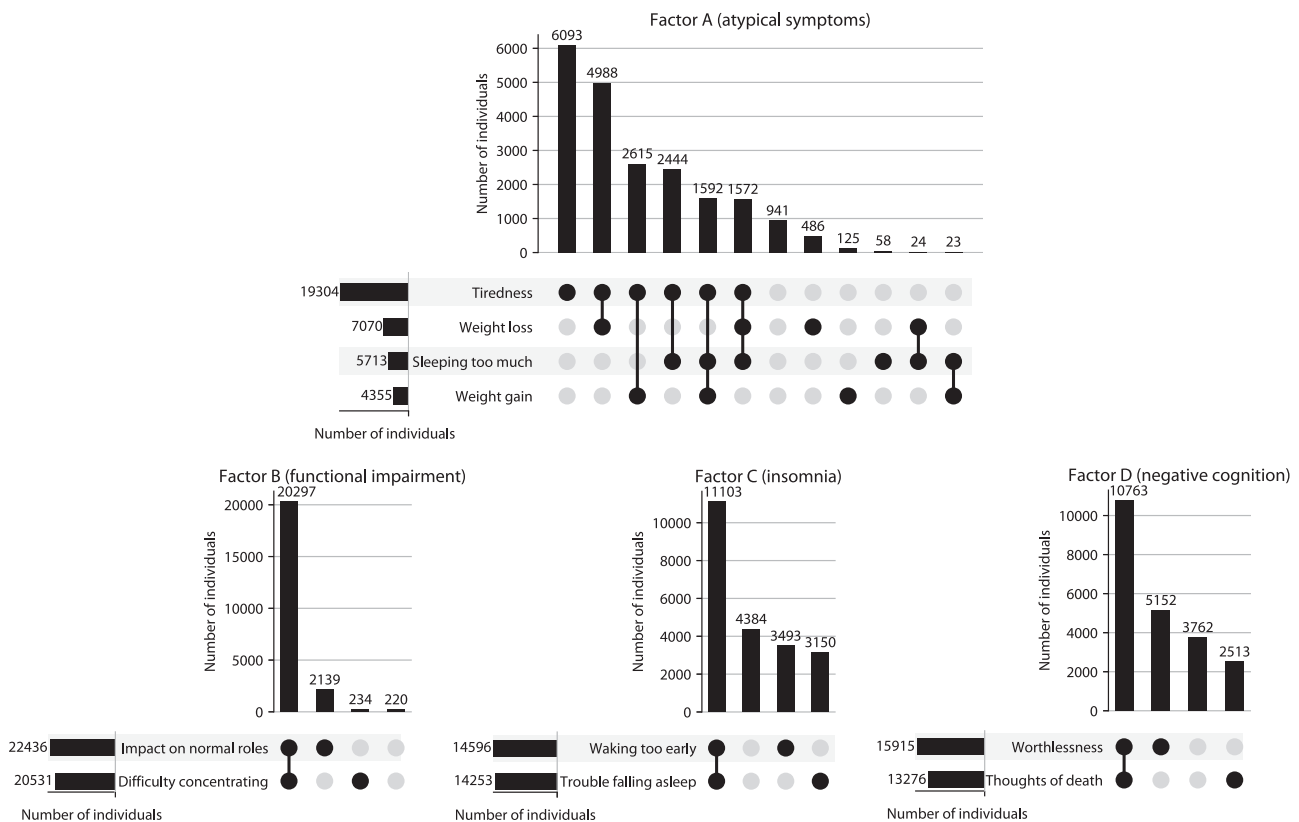
Strikingly, every symptom dimension was associated with increased risk of nearly every mental illness (according to self-report of professional diagnoses in the Mental Health Questionnaire) and PRS or family history thereof (Table 4). 44 of 60 factor-illness associations were significant after multiple testing correction. All but one of these 44 associations were between higher factor scores and increased risk of mental illness. Similarly, 10 of 12 factor-PRS associations and 4 of 4 factor-family history associations were between higher factor scores and increased risk of mental illness. This is reminiscent of how every mental illness has been associated with an increased risk of every other mental illness (Plana-Ripoll et al., 2019).

However, certain illnesses were particularly associated with specific symptom dimensions. For instance, anorexia nervosa

**Table 3.** Factor loadings

	Factor A: atypical symptoms	Factor B: functional impairment	Factor C: insomnia	Factor D: negative cognition
Sleeping too much	<b>0.95</b>	-0.02	-0.06	0.01
Tiredness	<b>0.46</b>	<u>0.26</u>	<u>0.31</u>	0.05
Weight change	<b>0.20</b>	<u>-0.14</u>	-0.03	0.07
Difficulty concentrating	-0.02	<b>0.96</b>	-0.02	0.00
Impact on normal roles	<u>0.14</u>	<b>0.41</b>	0.03	0.10
Waking too early	<u>-0.17</u>	0.00	<b>0.75</b>	0.03
Trouble falling asleep	0.09	0.00	<b>0.70</b>	-0.03
Worthlessness	0.00	0.00	-0.01	<b>1.00</b>
Thoughts of death	0.07	0.02	<u>0.14</u>	<b>0.36</b>

The largest factor loading for each symptom is bolded; loadings with a magnitude >0.1 are underlined. Symptoms refer to specific questions from the UK Biobank Mental Health Questionnaire (Table 1).

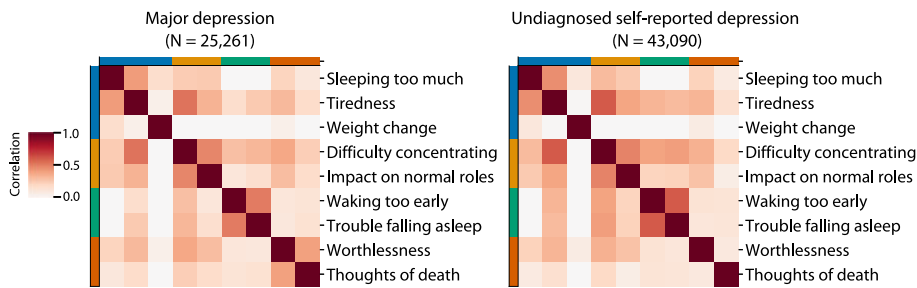


**Fig. 2.** Internal symptom structure of each of the four symptom dimensions. Each panel corresponds to a factor, showing how many people endorsed each of the factor's questions (left bar graph), and each combination thereof, including no questions (top bar graph). Note that for the purposes of this visualization, weight change is split into weight gain and weight loss, while the impact on normal roles is binarized as {a little, somewhat, a lot} v. {not at all}.

(AOR = 2.00 [1.53, 2.61]) and bulimia nervosa (AOR = 2.30 [1.60, 3.30]) were exclusively associated with factor D (negative cognition). Factor D was also significantly more associated than other factors with social anxiety/phobia (AOR = 2.79 [2.33, 3.35]), reflective of transdiagnostic contributions of negative cognition to multiple psychiatric illnesses (Ehring & Watkins, 2008). Despite being associated with nearly every symptom dimension, PRSs and family history did not display significant differential associations between symptom dimensions.

### Discussion

In this study, we analyzed the latent symptom structure of major depression in a population-based cohort of 25 261 self-reported White participants with a lifetime depression diagnosis. The identified symptom structure replicated across ancestries and case definitions. Each symptom dimension had a unique comorbidity profile, being associated with a specific combination of mental illnesses, though not showing any obvious genetic signatures



**Fig. 3.** Concordant symptom structure between major depression and undiagnosed self-reported depression. Matrices of polychoric correlations across 25 261 individuals with major depression (left) and 43 090 individuals with undiagnosed self-reported depression (right). Color bars at the top and left indicate membership in one of the four symptom dimensions (blue = atypical symptoms, orange = functional impairment, green = insomnia, red = negative cognition).

**Table 4.** Associations of factors with mental illness diagnoses, polygenic risk scores and family history.

	Factor A: atypical symptoms	Factor B: functional impairment	Factor C: insomnia	Factor D: negative cognition
Mental health problems ever diagnosed by a professional (age- and sex-adjusted odds ratio)				
Other psychotic illness (N = 328)	<b>2.01 [1.60, 2.54]</b>	<b>↑ 4.19 [2.59, 6.80]</b>	<b>1.60 [1.24, 2.06]</b>	<b>2.26 [1.69, 3.03]</b>
(Hypo)mania/bipolar/manic depression (N = 384)	<b>2.47 [2.00, 3.05]</b>	<b>3.36 [2.24, 5.05]</b>	<b>↓ 1.31 [1.04, 1.64]</b>	<b>3.09 [2.29, 4.17]</b>
Any personality disorder (N = 221)	<b>1.95 [1.47, 2.59]</b>	<b>1.79 [1.16, 2.75]</b>	<i>1.13 [0.84, 1.52]</i>	<b>3.23 [2.14, 4.86]</b>
Social anxiety/phobia (N = 977)	<b>1.59 [1.38, 1.83]</b>	<b>1.89 [1.54, 2.33]</b>	<b>↓ 1.27 [1.10, 1.47]</b>	<b>↑ 2.79 [2.33, 3.35]</b>
Over- or binge-eating (N = 386)	<b>1.86 [1.50, 2.31]</b>	<b>2.03 [1.45, 2.85]</b>	<b>↓ 1.17 [0.93, 1.46]</b>	<b>2.54 [1.92, 3.38]</b>
Bulimia nervosa (N = 244)	<i>0.93 [0.69, 1.25]</i>	<i>1.07 [0.75, 1.52]</i>	<i>1.05 [0.79, 1.39]</i>	<b>↑ 2.30 [1.60, 3.30]</b>
Obsessive-compulsive disorder (N = 465)	<b>1.66 [1.36, 2.03]</b>	<b>1.37 [1.05, 1.80]</b>	<i>1.14 [0.93, 1.40]</i>	<b>2.25 [1.75, 2.89]</b>
Autism spectrum disorder (N = 100)	<b>2.22 [1.47, 3.35]</b>	<i>1.45 [0.79, 2.65]</i>	<i>0.86 [0.56, 1.32]</i>	<b>1.88 [1.12, 3.16]</b>
Agoraphobia (N = 322)	<b>1.93 [1.52, 2.45]</b>	<b>2.20 [1.52, 3.16]</b>	<i>1.30 [1.01, 1.66]</i>	<b>1.64 [1.26, 2.13]</b>
Schizophrenia (N = 61)	<b>2.20 [1.29, 3.75]</b>	<i>1.49 [0.71, 3.14]</i>	<b>↓ 0.46 [0.27, 0.81]</b>	<i>1.80 [0.96, 3.35]</i>
ADD/ADHD (N = 62)	<b>1.93 [1.13, 3.28]</b>	<i>0.83 [0.45, 1.56]</i>	<i>0.85 [0.49, 1.47]</i>	<b>2.15 [1.11, 4.16]</b>
Anxiety/nerves/GAD (N = 8170)	<b>↓ 1.40 [1.32, 1.49]</b>	<b>↑ 2.09 [1.94, 2.26]</b>	<b>1.71 [1.61, 1.81]</b>	<b>1.76 [1.65, 1.87]</b>
Panic attacks (N = 3516)	<b>1.52 [1.40, 1.65]</b>	<b>↑ 2.03 [1.81, 2.27]</b>	<b>1.59 [1.46, 1.72]</b>	<b>1.52 [1.40, 1.65]</b>
Anorexia nervosa (N = 381)	<b>↓ 0.81 [0.63, 1.04]</b>	<i>1.31 [0.98, 1.76]</i>	<i>1.18 [0.94, 1.48]</i>	<b>↑ 2.00 [1.53, 2.61]</b>
Other phobia (N = 601)	<b>1.40 [1.16, 1.68]</b>	<b>1.59 [1.24, 2.02]</b>	<b>1.28 [1.07, 1.54]</b>	<b>1.32 [1.09, 1.58]</b>
Polygenic risk scores (age- and sex-adjusted effect size, β)				
Schizophrenia	<b>0.06 [0.03, 0.09]</b>	<b>0.07 [0.03, 0.10]</b>	<b>0.04 [0.01, 0.06]</b>	<b>0.08 [0.05, 0.10]</b>
Major depressive disorder	<b>0.07 [0.04, 0.10]</b>	<b>0.03 [0.00, 0.07]</b>	<i>0.02 [-0.01, 0.05]</i>	<b>0.03 [0.00, 0.06]</b>
Bipolar disorder	<i>0.02 [-0.01, 0.05]</i>	<b>0.04 [0.00, 0.07]</b>	<b>0.04 [0.02, 0.07]</b>	<b>0.04 [0.02, 0.07]</b>
Family history (age-, sex- and genotype principal component-adjusted odds ratio)				
Severe depression (N = 6564)	<b>1.21 [1.13, 1.29]</b>	<b>1.42 [1.31, 1.53]</b>	<b>1.18 [1.11, 1.25]</b>	<b>1.38 [1.29, 1.47]</b>

Mental illnesses and polygenic risk scores are ordered in descending order by the largest odds ratio/effect size across the four factors. Significant associations (FDR < 0.1) are **bolded** while non-significant ones are *italicized*; associations significantly **larger** (↑) or **smaller** (↓) than for all other factors (*p* < 0.05, difference-of-effect-sizes test) are denoted in red and blue, respectively. For binary traits, *N* denotes the number of people with the trait. ADD/ADHD = attention-deficit (and hyperactivity) disorder, GAD = generalized anxiety disorder.

relative to other dimensions. To our knowledge, this study represents the largest-ever analysis of the structure of depressive symptoms.

Every symptom dimension was associated with an increased risk of nearly every mental illness. In part, this reflects shared diagnostic criteria between depression and comorbid disorders. For instance, functional impairment (‘clinically significant distress or impairment in social, occupational, or other important areas of functioning’) forms part of the DSM-5 diagnostic criteria for both MDD and generalized anxiety disorder (American Psychiatric Association, 2013), which might help explain why anxiety is

significantly more associated with the functional impairment symptom dimension than with any other dimension. It is also consistent with the notion of transdiagnostic subtypes – a key focus of RDoC (Insel et al., 2010) – in which neurobiologically similar subtypes cut across existing diagnostic categories (Grisanzio et al., 2018) and genetic variants have pleiotropic effects across multiple dimensions of psychopathology (Anttila et al., 2018).

Similarly, every symptom dimension was positively associated with every or nearly every PRS. However, unlike for diagnosed mental illnesses, symptom dimensions were not *differentially*

associated with polygenic risk. This is concordant with a recent study finding no evidence of genetically defined depressive subtypes (Howard et al., 2020).

Dimensions showed highly distinct patterns of association with comorbid mental illnesses. Clinicians should be aware of these associations between specific types of depressive symptoms and specific comorbidities, as patients presenting with one may also have the other. Moreover, treating certain of these comorbidities may lead to concomitant improvement in depressive symptoms.

Despite the UK Biobank's substantial size, breadth and diversity of phenotyping, it has at least three major disadvantages for this application. First, the Mental Health Questionnaire does not fully correspond to established rating scales and DSM-5 specifiers. For instance, it lacks questions on rejection sensitivity and leaden paralysis that would improve ascertainment of atypical depression, and it lacks questions on psychomotor agitation that would improve ascertainment of depression with mixed features. While the Mental Health Questionnaire does ask about psychotic experiences, anxiety and mania, they are not asked with reference to a particular depressive episode, so we chose not to use them within our factor structure here. Second, the Mental Health Questionnaire's temporal ascertainment is limited: an individual's worst reported lifetime episode of depression may be only a small fraction of what is often a prolonged course of illness, with many relapses and remissions. Third, being specific to a single developed country, the dataset lacks a broad representation of the world's population, despite our trans-ancestral replication.

The use of self-report data is considered controversial (Abbasi, 2017), but has at least two related advantages for this application. First, it enables a direct comparison between major depression and undiagnosed self-reported depression among demographically similar individuals, which we find largely shares the same symptom structure. Thus, this enables ascertainment of a potentially broad segment of the population who may well have experienced an episode of *bona fide* major depressive disorder, but not sought help at the time (Boerema et al., 2016), and consequently go undiagnosed and untreated. For instance, a meta-analysis by the World Health Organization found that across 24 countries, 56.3% of individuals with depression and 56.0% with dysthymia did not receive any treatment for their illness (Kohn, Saxena, Levav, & Saraceno, 2004). Second, as the cohort is composed of a broad spectrum of individuals in the community, rather than merely those seeking treatment at a psychiatric research hospital, it is arguably more representative of the general population than the patients typically recruited into psychiatric research protocols. The high prevalence of undiagnosed self-reported depression with a concordant symptom structure to major depression is consistent with the notion of a large burden of untreated patients (Kohn et al., 2004) who might benefit from psychiatric care.

On the whole, this study provides perhaps the highest-resolution view to date of depressive symptom dimensions in the community. Additional research is needed to further elucidate the underlying neurobiological correlates of these dimensions in ways that can inform treatment decisions.

**Acknowledgements.** MW and SJT were funded by the Kavli Foundation, Krembil Foundation, CAMH Discovery Fund, the McLaughlin Foundation, NSERC (RGPIN-2020-05834 and DGEGR-2020-00048) and CIHR (NGN-171423). PZ was funded by the Labatt Family Postdoctoral Fellowship in Depression Biology. DF is supported by the Koerner Family Foundation New Scientist Program. CH was funded by the CAMH Foundation, the Brain and Behavior Research Foundation, and the NIMH. This work was conducted under the auspices of UK Biobank application 61530, 'Multimodal subtyping

of mental illness across the adult lifespan through integration of multi-scale whole-person phenotypes'. The authors declare no conflicts of interest.

## References

- Abbasi, J. (2017). 23andMe, Big data, and the genetics of depression. *JAMA: The Journal of the American Medical Association*, 317(1), 14–16.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5<sup>®</sup>)*. Arlington, VA: American Psychiatric Pub.
- Brainstorm Consortium, Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., ... Murray, R. (2018). Analysis of shared heritability in common disorders of the brain. *Science (New York, N.Y.)*, 360(6395). <https://doi.org/10.1126/science.aap8757>.
- Ballard, E. D., Yarrington, J. S., Farmer, C. A., Lener, M. S., Kadriu, B., Lally, N., ... Zarate, C. A., Jr. (2018). Parsing the heterogeneity of depression: An exploratory factor analysis across commonly used depression rating scales. *Journal of Affective Disorders*, 231, 51–57.
- Beijers, L., Wardenaar, K. J., van Loo, H. M., & Schoevers, R. A. (2019). Data-driven biological subtypes of depression: Systematic review of biological approaches to depression subtyping. *Molecular Psychiatry*, 24(6), 888–900.
- Blumenthal, M. D. (1971). Heterogeneity and research on depressive disorders. *Archives of General Psychiatry*, 24, 524. <https://doi.org/10.1001/archpsyc.1971.01750120040007>.
- Boerema, A. M., Kleiboer, A., Beekman, A. T. F., van Zoonen, K., Dijkshoorn, H., & Cuijpers, P. (2016). Determinants of help-seeking behavior in depression: A cross-sectional study. *BMC Psychiatry*, 16, 78.
- Brailean, A., Curtis, J., Davis, K., Dregan, A., & Hotopf, M. (2020). Characteristics, comorbidities, and correlates of atypical depression: Evidence from the UK biobank mental health survey. *Psychological Medicine*, 50(7), 1129–1138.
- Bulik-Sullivan, B. K., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., ... Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47, pp. 291–295. <https://doi.org/10.1038/ng.3211>.
- Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., ... Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726), 203–209.
- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... Moffitt, T. E. (2014). The *p* factor. *Clinical Psychological Science*, 2, 119–137. <https://doi.org/10.1177/2167702613497473>.
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7.
- Davis, K. A. S., Coleman, J. R. I., Adams, M., Allen, N., Breen, G., Cullen, B., ... Hotopf, M. (2020). Mental health in UK Biobank – development, implementation and results from an online questionnaire completed by 157 366 participants: A reanalysis. *BJPsych Open*, 6(2), E18. <https://doi.org/10.1192/bjo.2019.100>.
- Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., ... Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, 23(1), 28–38.
- Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy*, 1, 192–205. <https://doi.org/10.1521/ijct.2008.1.3.192>.
- Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191–197.
- Grisanzio, K. A., Goldstein-Piekarski, A. N., Wang, M. Y., Rashed Ahmed, A. P., Samara, Z., & Williams, L. M. (2018). Transdiagnostic symptom clusters and associations with brain, behavior, and daily function in mood, anxiety, and trauma disorders. *JAMA Psychiatry*, 75(2), 201–209.
- Häfner, H., Maurer, K., Trendler, G., W an der Heiden, Schmidt, M., & Könecke, R. (2005). Schizophrenia and depression: Challenging the paradigm of two separate diseases – a controlled study of schizophrenia, depression and healthy controls. *Schizophrenia Research*, 77(1). <https://doi.org/10.1016/j.schres.2005.01.004>.
- Hamilton, M. (1960). A RATING SCALE FOR DEPRESSION. *Journal of Neurology, Neurosurgery & Psychiatry*, 23, 56–62. <https://doi.org/10.1136/jnnp.23.1.56>.

- Howard, D. M., Adams, M. J., Shirali, M., Clarke, T.-K., Marioni, R. E., Davies, G., ... McIntosh, A. M. (2018). Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nature Communications*, 9(1), 1470.
- Howard, D. M., Folkersen, L., Coleman, J. R. I., Adams, M. J., Glanville, K., Werge, T., ... McIntosh, A. M. (2020). Genetic stratification of depression in UK Biobank. *Translational Psychiatry*, 10(1), 163.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167, pp. 748–751. <https://doi.org/10.1176/appi.ajp.2010.09091379>.
- Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H.-U. (1998). The world health organization composite international diagnostic interview short-form (CIDI-SF). *International Journal of Methods in Psychiatric Research*, 7, pp. 171–185. <https://doi.org/10.1002/mpr.47>.
- Kohn, R., Saxena, S., Levav, I., & Saraceno, B. (2004). The treatment gap in mental health care. *Bulletin of the World Health Organization*, 82(11), 858.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The patient health questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 41(11), 1284–1292.
- Lawley, D. N. (1940). VI.—The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Society of Edinburgh*, 60, pp. 64–82. <https://doi.org/10.1017/s037016460002006x>.
- Levis, B., Sun, Y., He, C., Wu, Y., Krishnan, A., & Bhandari, P. M., ... Depression Screening Data (DEPRESSD) PHQ Collaboration. (2020). Accuracy of the PHQ-2 alone and in combination With the PHQ-9 for screening to detect Major depression: Systematic review and meta-analysis. *JAMA: The Journal of the American Medical Association*, 323(22), 2290–2300.
- Marshall, M. (2020). The hidden links between mental disorders. *Nature*, 581(7806), 19–21.
- McIntosh, A. M., Sullivan, P. F., & Lewis, C. M. (2019). Uncovering the genetic architecture of major depression. *Neuron*, 102(1), 91–103.
- Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., ... Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, 50(3), 381–389.
- Plana-Ripoll, O., Pedersen, C. B., Holtz, Y., Benros, M. E., Dalsgaard, S., de Jonge, P., ... McGrath, J. J. (2019). Exploring comorbidity within mental disorders Among a danish national population. *JAMA Psychiatry*, 76(3), 259–270.
- Rosseel, Y. (2012). lavaan: An RPackage for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>.
- Shanmugan, S., Wolf, D. H., Calkins, M. E., Moore, T. M., Ruparel, K., Hopson, R. D., ... Satterthwaite, T. D. (2016). Common and dissociable mechanisms of executive system dysfunction across psychiatric disorders in youth. *The American Journal of Psychiatry*, 173(5), 517–526.
- Smart, A., & Harrison, E. (2017). The under-representation of minority ethnic groups in UK medical research. *Ethnicity & Health*, 22(1), 65–82.
- Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., & Trubetskoy, V., ... Bipolar Disorder Working Group of the Psychiatric Genomics Consortium. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*, 51(5), 793–803.
- Stein, D. J., Szatmari, P., Gaebel, W., Berk, M., Vieta, E., Maj, M., ... Reed, G. M. (2020). Mental, behavioral and neurodevelopmental disorders in the ICD-11: An international perspective on key changes and controversies. *BMC Medicine*, 18(1), 21.
- ten Berge, J. M. F., Krijnen, W. P., Wansbeek, T., & Shapiro, A. (1999). Some new results on correlation-preserving factor scores prediction methods. *Linear Algebra and its Applications*, 289, 311–318. [https://doi.org/10.1016/s0024-3795\(97\)10007-6](https://doi.org/10.1016/s0024-3795(97)10007-6).
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1–10. <https://doi.org/10.1007/bf02291170>.
- van Loo, H. M., de Jonge, P., Romeijn, J.-W., Kessler, R. C., & Schoevers, R. A. (2012). Data-driven subtypes of major depressive disorder: A systematic review. *BMC Medicine*, 10, 156.
- Weissman, M. M. (2020). Big data begin in psychiatry. *JAMA Psychiatry*, 77(9), 967–973. <https://doi.org/10.1001/jamapsychiatry.2020.0954>.
- Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., & Abdellaoui, A., ... Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), 668–681.