

Research paper

Robust symptom networks in recurrent major depression across different levels of genetic and environmental risk



H.M. van Loo^{a,b,*}, C.D. Van Borkulo^c, R.E. Peterson^{a,d}, E.I. Fried^c, S.H. Aggen^a, D. Borsboom^c, K.S. Kendler^{a,d,e}

^a Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

^b Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen The Netherlands

^c Department of Psychology, University of Amsterdam, The Netherlands

^d Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA

^e Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA, USA

ARTICLE INFO

Keywords:

Major depressive disorder
Depressive symptoms
Complex networks
Genetic risk
Stressful life events

ABSTRACT

Background: Genetic risk and environmental adversity—both important risk factors for major depression (MD)—are thought to differentially impact on depressive symptom types and associations. Does heterogeneity in these risk factors result in different depressive symptom networks in patients with MD?

Methods: A clinical sample of 5784 Han Chinese women with recurrent MD were interviewed about their depressive symptoms during their lifetime worst episode of MD. The cases were classified into subgroups based on their genetic risk for MD (family history, polygenic risk score, early age at onset) and severe adversity (childhood sexual abuse, stressful life events). Differences in MD symptom network structure were statistically examined for these subgroups using permutation-based network comparison tests.

Results: Although significant differences in symptom endorsement rates were seen in 18.8% of group comparisons, associations between depressive symptoms were similar across the different subgroups of genetic and environmental risk. Network comparison tests showed no significant differences in network strength, structure, or specific edges (P -value > 0.05) and correlations between edges were strong (0.60–0.71).

Limitations: This study analyzed depressive symptoms retrospectively reported by severely depressed women using novel statistical methods. Future studies are warranted to investigate whether similar findings hold in prospective longitudinal data, less severely depressed patients, and men.

Conclusions: Similar depressive symptom networks for MD patients with a higher or lower genetic or environmental risk suggest that differences in these etiological influences may produce similar symptom networks downstream for severely depressed women.

1. Introduction

Major depression (MD) is a heterogeneous disorder in terms of presentation, course of illness and treatment response (Cuijpers et al., 2012; Eaton et al., 2008; Hardeveld et al., 2013; Holma et al., 2008; Oquendo et al., 2004; Simon and Perlis, 2010). This heterogeneity may also be reflected in its etiology. That is, it is likely that many different pathways can lead to the syndrome of MD as opposed to MD being the result of one common complex pathway (Kendler, 2013; Krishnan and Nestler, 2008; Milaneschi et al., 2016; Wichers, 2014).

Two of the most prominent risk factors for MD are genetic risk (Kendler et al., 2006; Sullivan et al., 2000) and environmental adversity (Chen et al., 2014; Kendler et al., 1999; Tao et al., 2011). These broad

distinct risk factors might be associated with differential pathways to MD, including specific (neuro)biological differences (Krishnan and Nestler, 2008; Pittenger and Duman, 2008). However, risk factors could also impact on aspects of the disorder further downstream, such as how symptoms interact within a complex network.

When viewing symptoms within a network theory perspective, MD develops and persists through direct interactions among its defining symptoms, such as depressed mood, insomnia, fatigue, and concentration difficulties (Borsboom, 2017; Cramer et al., 2010; Fried et al., 2017). Different risk factors related to genes or environmental adversity are hypothesized to influence the strength of these connections between symptoms, which may produce individual differences in liability to develop the disorder (Borsboom, 2017; Cramer et al., 2011; Wichers,

* Correspondence to: Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen The Netherlands
E-mail address: h.van.loo@umcg.nl (H.M. van Loo).

2014). There is some evidence that network connectivity may differ across persisting and remitting cases of MD (Van Borkulo et al., 2015). It is unclear, however, whether heterogeneity in these risk factors leads to clinical variation in the presentation of MD once a syndrome has occurred. In particular, it is unknown whether differences in symptom network structure can be detected across subgroups that differ on risk factors.

Finding potential differences in group-level symptom networks would be important to improve our understanding of MD's heterogeneity by clarifying potential self-sustaining mechanisms and tailoring treatments for specific subtypes of MD (Borsboom, 2017; Fried et al., 2017). For example, if insomnia would trigger many other symptoms specifically in patients with adversity-related depression, treatments targeting sleep problems would be expected to be more effective in this subgroup than in patients unexposed to adversity (Fried et al., 2017). On the other hand, if depressive symptom networks are robust across different levels of genetic and environmental risk, this would suggest that these distinct etiological pathways upstream might eventuate in a final common pathway downstream.

This study thus aims to examine differences in depressive symptom networks during episodes of MD across different levels of genetic risk and environmental adversity in a large sample of Han Chinese women with recurrent MD, using recently developed statistical methods to quantify network differences.

2. Methods

2.1. Data

Data are from the China, Oxford, and VCU Experimental Research on Genetic Epidemiology (CONVERGE) project, a study of Han Chinese women with recurrent major depression aimed at identifying genetic risk factors for MD in a rigorously ascertained cohort (CONVERGE Consortium, 2015). Major depression cases ($n = 5864$) were recruited from 58 provincial mental health centers and psychiatric departments of general medical hospitals in 45 cities and 23 provinces of China. Cases were between 30 and 60 years of age and had two or more episodes of MD meeting DSM-IV criteria with the first episode between ages 14 and 50, had not abused drugs or alcohol before their first depressive episode, and reported no history of schizophrenia or mania.

A diagnosis of MD was based on a face-to-face interview with a trained clinician using the Composite International Diagnostic Interview (CIDI; World Health Organization lifetime version 2.1; Chinese version), which classifies MD according to DSM-IV criteria (American Psychiatric Association, 1994). A previous study using latent variable measurement invariance testing showed that the 9 aggregated and 14 disaggregated DSM-criteria for MD perform similarly in the CONVERGE sample when compared to other matched samples from the United States and Europe (Kendler et al., 2015). The study protocol was approved by the Ethical Review Board of Oxford University and the ethics committees of all participating hospitals. All participants provided written informed consent. A more detailed description of the sample and study design can be found in (CONVERGE Consortium, 2015).

2.2. Measurements

2.2.1. Depressive symptoms

Trained postgraduate medical students, junior psychiatrists or senior nurses interviewed the ascertained MD patients on a broad set of depressive and anxiety symptoms during their most severe lifetime episode of MD using a structured computerized questionnaire. Assessment of depressive symptoms was based on the CIDI, supplemented by questions about symptoms of melancholia, anxiety, loss of self-confidence and self-esteem, hopelessness and helplessness, crying and decreased libido. For this study, the core MD criteria of depressed

mood and interest loss were dropped from the analyses due to the extremely high positive endorsement rates (99.6% and 98.9%, respectively). In total, we analyzed 24 dichotomously coded symptoms including a) the 12 disaggregated DSM-criteria for major depression and b) 12 additional symptoms commonly experienced by depressed patients derived from a range of sources including Beck's work (Beck et al., 1980) and the DSM-IV criteria for melancholia (American Psychiatric Association, 1994) in order to investigate the network structure of a broad set of DSM and non-DSM symptoms (Fried, 2017; Fried et al., 2016a; Guloksuz et al., 2017). Cases with missing data on one or more of the 24 symptoms were excluded for all statistical analyses. We excluded 70 cases with missing data on any of the 24 symptoms, resulting in a sample of 5784 cases included in the statistical analyses.

2.2.2. Genetic risk measures

To test for network differences related to genetic effects, we compared depressive symptom networks for patients with higher vs. lower genetic risk of MD. Three indexes of genetic risk were used: a positive family history of MD, a high polygenic risk score (PRS), and an earlier age at onset.

First, family history of lifetime MD was coded positive if the subject reported a lifetime episode of MD in at least one of her biological parents and siblings according to the Research Diagnostic Criteria Family History (Endicott et al., 1978). Family history was coded negative if the subject reported no family members with MD, and had reported on the MD status of at least two family members (otherwise family history was set to missing).

Second, a binary variable indicating high PRS for MD was created based on a PRS within the highest quartile of all cases (Peterson et al., 2016). Details on DNA sequencing and imputation have been previously reported (Cai et al., 2017; CONVERGE Consortium, 2015; Peterson et al., 2016). In brief, CONVERGE obtained sequence data of 11,670 samples and imputed genotypes using methods previously described in Cai et al. (2017). After quality control, 5303 cases of MD and 5337 controls remained included for genome wide association analysis (CONVERGE Consortium, 2015). In this sample, we derived PRS for MD by randomly dividing all cases and controls into independent training and testing sets (50–50% split). We estimated SNP effects for risk of MD by the best linear unbiased prediction (BLUP) method implemented in GCTA (Yang et al., 2011) using the training set and tested these aggregate genome-wide scores in the testing set. PRS scores were constructed using the profile option in PLINK (Chang et al., 2015; Purcell et al., 2007) and SNPs were weighted by their BLUP parameter estimates. To reduce the effect of population stratification (Patterson et al., 2006; Price et al., 2006), the PRS was adjusted for the first two principal components reflecting ancestry, which distinguished north-south regional differences. This PRS was significantly associated with case-control status ($P < 4.6 \times 10^{-5}$), accounting for 1.1% of the variability in MD risk (Peterson et al., 2016).

Third, early age of onset was defined as cases who reported a first episode of MD before the age of 28 (i.e. lowest quartile); normal or older age at onset was defined as a first episode at age 28 years or older. A younger age at first onset is a modest index of familial liability, particularly before age 25, up to age 35 (Docherty et al., 2017; Kendler et al., 2005). Recent research found differences in genetic architecture for age at onset of MD, with cases with early age at onset showing greater genetic similarities with schizophrenia and bipolar (Power et al., 2017; Verduijn et al., 2017), although a recent study in CONVERGE did not find an association between age at onset and common variant PRS for MD (Docherty et al., 2017). To keep the sample size of the early age at onset subsample sufficiently large, we selected cases within the lowest quartile of age at onset, instead of studying a subsample with a lower age at onset.

2.2.3. Adversity measure

To test for network differences related to severe environmental adversity, we compared depressive symptom networks for patients with higher vs. lower environmental risk of MD. A binary index of adversity was created from self-reported stressful life events (SLE) and childhood sexual abuse (CSA) in order to identify individuals exposed to severe environmental adversities. The SLE questionnaire was adapted from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (Kendler and Prescott, 2006) and assessed 16 traumatic lifetime events and the age of their occurrence (Tao et al., 2011) (Additional Table 1). Because there is evidence that sensitive subjects like CSA are more accurately reported on using more confidential methods of assessment (Laumann et al., 1994), participants were asked to complete a paper questionnaire about CSA. The CSA questionnaire was a shortened version of a scale developed by Martin et al. (1993) and queried whether, before the subject was 16, had any adult, or any other older person, initiated any unwanted sexual advances like kissing or hugging in a sexual way, touching or fondling private parts, showing their sex organs, making them touch the person in a sexual way, or attempting or having sexual intercourse (Chen et al., 2014). The possible responses were “never,” “once,” and “more than once.” While it is known that the patterns of association with MD differ between some forms of environmental stressors (Kessler, 1997), this heterogeneity is expected to be relatively small and so for these analyses we grouped them together to maximize statistical power and reduce multiple testing. Subjects were considered an “adversity case” if they i) had non-missing data on SLE and CSA scales and ii) endorsed any CSA and/or had high aggregate SLE scores (+3 SD). Since SLE vary in severity, the SLE score was constructed by weighting each item by their effect size on MD and summed across each of the 16 items. SLE for MD cases were only included if they preceded MD onset. Using this adversity exposure status, we grouped the sample into those “exposed” and those “unexposed”.

2.3. Statistical methods

2.3.1. Assessing differences in symptom endorsements

Four pairwise comparisons were carried out assessing both symptom endorsement and network structure for all the defined subsamples: 1) family history positive vs. negative; 2) high vs. low PRS; 3) earlier vs. later age at onset; 4) adversity exposed vs. unexposed. Differences in endorsement rates for all 24 symptoms were assessed by chi-squared tests using Bonferroni-corrected alphas ($\alpha = 0.0005$). Similarities in endorsement rates were examined by Spearman rank-order correlations.

2.3.2. Network estimation

We estimated the appropriate network model for dichotomous data – the Ising Model – via the IsingFit package ($\gamma = 0.25$; Van Borkulo et al., 2014). In these networks, “edges” (i.e. relationships among symptoms) are estimated using regularized logistic regressions (Van Borkulo et al., 2014). Each symptom is regressed on all others, and an $L1$ -penalty is imposed on the regression coefficients for an optimal balance between parsimony and goodness of fit. The strength of the edge is the average of the regression coefficients from the two regularized logistic regressions of symptom 1 on symptom 2 (adjusted for all other symptoms) and symptom 2 on symptom 1 (adjusted for all other symptoms) (note that due to applying the AND-rule, both regression coefficients need to be non-zero). This estimation method has a high specificity (i.e. connections that are included in the estimated network almost certainly reflect nonzero parameters at population level) and acceptable sensitivity (Van Borkulo et al., 2014). For an extensive explanation of the method we refer to Van Borkulo et al. (2014). The accuracy and stability of the network estimation was further examined using non-parametric bootstrap methods (R-package *bootnet*, 1000 bootstrapped samples) (Epskamp et al., 2017).

2.3.3. Network comparisons

First, the recently developed Network Comparison Test (NCT) was used to statistically compare two networks on three aspects: (1) network structure, (i.e., are the distributions of edge weights similar), (2) global strength (i.e., are the sums of all absolute edge weights similar), and (3) individual edge strength (i.e., are individual edge weights similar) (Van Borkulo et al., 2017). NCT uses two-tailed permutation tests in which the original group members are repeatedly randomly reassigned to new subsamples that maintain the original sample sizes (1000 times), after which their network structures were compared on the three aspects described above. The resulting distribution under the null hypothesis (under the assumption that groups are equal) can be used to test whether the observed differences between the original groups are statistically significant ($\alpha = 0.05$). The test on individual edge weights uses a Holm-Bonferroni correction for multiple testing. See Van Borkulo et al. (2017) for a more detailed description of NCT.

Second, the degree of similarity among individual edge weights across networks were assessed with Spearman rank-order correlations.

Third, to get an impression of the independent tendency of symptoms to be non-zero, we compared the intercepts for each symptom (also referred to as “thresholds”) across the subsamples. The thresholds represent the intercept for each symptom in the regularized logistic regression analyses used to estimate the networks. All analyses were performed in R (R Core Team, 2017). We used R-package *qgraph* to visualize networks (Epskamp et al., 2012). For further details, we refer to the R-script included in the Additional material.

3. Results

3.1. Sample

The 5784 ascertained women with recurrent MD and complete data on all 24 depressive symptoms had a mean age at interview of 44.4 years (SD 8.9) and reported on average 5.2 previous depressive episodes (SD 8.9). Their mean age at onset was 34.8 years (SD 10.0) and their mean age at the worst episode was 40.3 years (SD 9.7), resulting in a mean recall period regarding their symptoms during the worst depressive episode of 4.0 years (SD 5.7). Of these cases, 27.7% reported a positive family history, 75.6% reported at least one lifetime SLE, and 10.3% experienced any CSA (8.7% nongenital, 6.5% genital, 2.8% intercourse). In total, 34.0% classified as adversity exposed (i.e. reported CSA and/or high aggregate SLE scores (+3 SD)). Additional Table 1 shows endorsement rates of each SLE among MD-cases.

The three indexes of genetic risk and the index of environmental risk were at most weakly positively correlated. Family history, early age at onset, and adversity were weakly positively correlated, whereas a high PRS was not correlated with any of the other indicators of genetic risk or adversity (Table 1).

3.2. Symptom endorsement rates

As reported elsewhere (Kendler et al., 2015; Li et al., 2014), endorsement rates in the MD case sample were high (> 90%) for concentration difficulties, fatigue, distinct quality of depressed mood,

Table 1
Correlations between measures of genetic and environmental risk.

	FH positive	AAO < 28	Adversity
AAO < 28	0.14 (0.09–0.19)		
Adversity	0.20 (0.15–0.25)	0.14 (0.10–0.18)	
PRS high	0.03 (−0.02–0.08)	0.02 (−0.03–0.07)	−0.03 (−0.08–0.02)

AAO, age at onset; FH, family history; PRS, polygenic risk score.

Numbers represent bivariate tetrachoric correlations and 95% confidence intervals in parentheses.

Table 2

Symptom endorsement rates (%) by comparison groups.

	All	FH neg.	FH pos.	χ^2	P	PRS low	PRS high	χ^2	P	AAO \geq 28	AAO < 28	χ^2	P	No adv.	Adv.	χ^2	P
N	5784	3907	1494			3909	1301			4149	1477			3511	1812		
Appetite loss	83.4	83.5	83.7	0.1	0.778	82.8	84.9	3.1	0.080	85.0	79.7	21.2	0.000	84.5	81.1	9.8	0.002
Weight loss	58.8	57.8	61.6	6.4	0.011	58.9	58.3	0.1	0.745	60.5	54.4	15.5	0.000	58.6	58.8	0.0	0.934
Appetite gain	9.5	9.1	10.2	1.2	0.266	10.2	7.1	10.0	0.002	8.4	12.3	18.7	0.000	7.8	13.0	36.3	0.000
Weight gain	6.7	6.2	7.3	1.9	0.163	6.9	5.5	2.4	0.124	6.1	8.5	9.5	0.002	5.5	9.1	24.7	0.000
Insomnia	92.3	92.5	92.2	0.0	0.877	92.5	91.6	0.5	0.469	93.4	90.0	17.5	0.000	93.3	91.1	8.8	0.003
Hypersomnia	13.1	11.7	16.4	20.8	0.000	12.7	13.5	0.4	0.541	11.6	17.3	30.0	0.000	11.7	16.2	19.8	0.000
Psychomotor retardation	76.0	76.0	77.2	0.9	0.346	75.3	77.8	3.1	0.077	75.9	76.5	0.2	0.642	76.2	77.0	0.2	0.627
Psychomotor agitation	72.5	72.4	74.0	1.5	0.221	73.2	71.6	1.0	0.321	73.4	70.3	5.0	0.025	73.2	71.6	1.6	0.201
Fatigue	93.4	92.9	95.2	10.8	0.001	93.2	94.0	0.9	0.344	94.0	91.8	7.5	0.006	93.5	93.1	0.5	0.489
Worthlessness	90.1	89.5	92.1	8.6	0.003	90.0	89.4	0.4	0.552	89.3	92.3	11.0	0.001	89.6	92.1	8.2	0.004
Reduced concentration	97.5	97.4	98.2	3.8	0.051	97.7	97.6	0.0	0.995	97.7	97.6	0.0	0.950	97.7	98.1	0.3	0.558
Suicidal ideation	76.2	75.9	77.3	1.3	0.253	76.1	75.1	0.5	0.500	75.9	77.5	1.6	0.203	74.1	80.0	21.1	0.000
Lost interest in (nearly) all activities	78.4	78.9	78.8	0.0	0.855	78.4	80.0	0.7	0.394	79.8	76.1	8.2	0.004	80.0	76.8	7.4	0.006
Decreased self esteem	84.2	83.9	85.3	1.6	0.208	84.8	82.8	2.8	0.093	83.5	86.5	7.6	0.006	83.5	85.5	3.4	0.063
Decreased self confidence	86.1	86.3	86.9	0.5	0.489	86.2	85.2	0.9	0.351	85.4	88.1	7.0	0.008	86.0	86.4	0.1	0.754
Depression different from grief	93.0	92.9	93.6	0.6	0.425	93.0	93.3	0.1	0.756	92.6	94.0	2.6	0.105	93.3	92.4	1.2	0.280
Diurnal variation	62.0	60.6	65.9	13.1	0.000	61.8	61.6	0.0	0.842	62.2	61.7	0.1	0.773	63.4	60.1	5.4	0.020
Decreased libido	88.8	88.3	91.0	8.0	0.005	88.4	88.9	0.2	0.694	89.8	86.1	14.2	0.000	89.5	88.2	1.7	0.192
Loss of ability to enjoy good things	86.2	86.1	87.2	1.1	0.304	86.5	85.1	1.5	0.221	87.0	83.6	10.3	0.001	86.7	84.5	4.4	0.035
Irritability or anger	74.7	73.4	77.1	7.9	0.005	75.2	71.9	5.5	0.019	73.9	76.5	3.7	0.054	72.6	79.0	25.7	0.000
Hopeless	80.6	79.5	83.4	10.3	0.001	80.0	81.6	1.3	0.263	78.9	85.1	26.2	0.000	78.6	84.2	23.2	0.000
Cry a lot	67.4	66.6	69.4	3.9	0.048	67.8	67.0	0.3	0.597	66.2	71.2	11.9	0.001	65.7	71.2	16.2	0.000
Helpless	89.2	88.2	92.2	18.0	0.000	89.1	89.4	0.0	0.841	88.5	91.6	11.0	0.001	87.8	92.6	28.3	0.000
Nervous, jittery, anxious	89.2	87.7	93.4	36.3	0.000	89.5	88.3	1.4	0.245	88.8	90.6	3.5	0.062	88.6	90.5	4.3	0.039
Average symptom endorsement ^a	72.9	72.4	74.6	-6.2	0.000	72.9	72.6	0.9	0.348	72.8	73.3	-1.3	0.192	72.6	73.9	-3.8	0.000

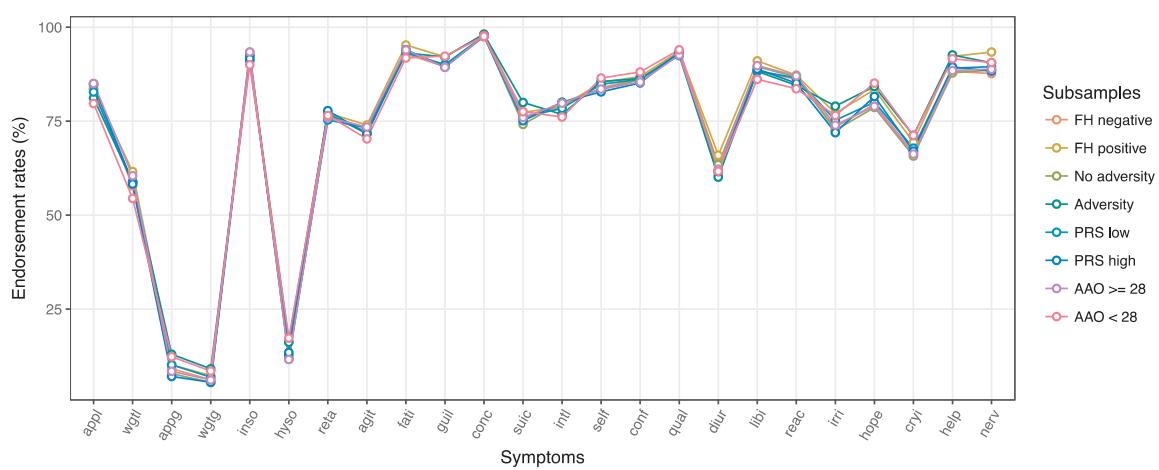
AAO, age at onset; Adv., adversity; FH, family history; PRS, polygenic risk score; χ^2 , chi-squared; P, P-value. Endorsement rates are expressed in percentages.In bold: P-values < 0.0005 (Bonferroni-corrected).^a Statistic for average symptom endorsement is t-test.

Fig. 1. Average symptom endorsement rates per subsample. AAO, age at onset; FH, family history; PRS, polygenic risk score. For abbreviation of symptoms see legend Fig. 2. Mean endorsement rates for each subsample for all 24 symptoms during the worst depressive lifetime episode.

insomnia, and feelings of guilt or worthlessness. In total, 12 out of 24 symptoms were endorsed by $> 80\%$ of the sample. Less frequently experienced symptoms included those described in the DSM-specifier for atypical depression: hypersomnia (13.1%), increased appetite (9.5%),

and weight gain (6.7%) (Table 2, Fig. 1).

Symptom endorsement rates were very similar for the different subsamples: 78 out of 96 chi-squared tests (81.3%) comparing symptom endorsement rates for each of the four subsample comparisons were not

significantly different and correlations of endorsement rates were very high (Spearman's r ranging from 0.95 to 0.99). If there were differences in endorsement rates, absolute differences between the subsamples were modest (maximum of 6.2%). The maximum statistical differences in endorsement rates (in terms of χ^2) concerned the symptoms 1) feeling nervous, jittery, and anxious, more often observed in cases with a positive instead of a negative family history of MD, 2) hypersomnia in cases with an earlier instead of a later age at onset, and 3) appetite gain, in adversity-exposed instead of adversity-unexposed cases (Table 2, Additional Fig. 1).

On average, subjects with a family history of MD endorsed more symptoms than subjects without a family history of MD (74.6% vs 72.4%, $P < 0.0005$) and had significantly higher endorsement rates for hypersomnia, diurnal variation, helplessness, and feeling nervous. Also, adversity exposed subjects endorsed more symptoms than adversity unexposed subjects (73.9% vs 72.6%, $P < 0.0005$) with significantly higher scores for appetite and weight gain, hypersomnia, suicidal ideations, irritability, hopelessness, crying and helplessness (Table 1).

In contrast, average endorsement rates for subjects with an earlier age of onset compared to those with a later age of onset were very similar (73.3% vs. 72.8%, $P = 0.19$), but scores on specific individual symptoms differed across these groups. Women who were younger at onset reported significantly more appetite gain, hypersomnia, and hopelessness, whereas women who were older at onset reported significantly more appetite and weight loss, insomnia and decreased libido. There were no significant differences in average endorsement rates or on any of the 24 symptoms between cases with a high versus low PRS.

3.3. Network structure in the entire sample

The network estimated using all cases ($n = 5784$) including all 24 symptoms displayed characteristics that were also found in the subsequent network comparisons (Fig. 2, Additional Table 2 and Additional Fig. 2). These network results are consistent with previous findings in 3463 depressed outpatients (Fried et al., 2016a) and in a previous study in MD-cases in CONVERGE, which aimed to investigate the potential difference between DSM- and non-DSM depressive symptoms (Kendler et al., n.d.). Note that the current study and this previous study in CONVERGE used a largely overlapping sample and network estimation methods, but with a different aim and using a slightly different set of symptoms (the current study investigated 24 instead of 19 depressive symptoms, as it included all disaggregated

psychomotor symptoms and excluded interest loss). In all three studies, the networks included a mix of both DSM and non-DSM symptoms. The network findings did not support a clear separation between these two classes of symptoms.

The strongest positive edges were observed between increased appetite and weight gain (penalized OR 34.3), decreased appetite and weight loss (OR 5.7), hopelessness and suicidal ideation (OR 5.2), low self-esteem and low self-confidence (OR 4.18), hopelessness and helplessness (OR 3.6). As might be expected, the strongest negative edges were observed between insomnia and hypersomnia (OR 0.2), appetite gain and appetite loss (OR 0.3), weight gain and weight loss (OR 0.4).

The number of connecting edges per symptom varied. Psychomotor retardation was connected to more than 50% of the symptoms (15 out of 23), which was also true for psychomotor agitation, hopelessness (13 out of 23) and appetite loss (12 out of 23). The remaining symptoms were connected with at least three other symptoms, except for the symptom distinct quality of depressed mood, which was not connected to any other symptoms.

The intercepts – i.e. the independent tendency for symptoms to be non-zero as estimated by the average intercept for each symptom in the logistic regression analyses used to estimate the networks – were highest for distinct quality of depressed mood, insomnia, and fatigue and lowest for weight gain, psychomotor retardation, and hopelessness (Additional Table 2). Bootstrapping indicated good accuracy and stability with narrow confidence intervals for the edge weights and stable centrality estimates (Additional Figs. 3–6).

3.4. Network comparisons

First, network comparison tests showed no significant differences in network structure across subsamples with higher vs. lower genetic risk and higher vs. lower environmental risk (Fig. 3). P -values for network structure and global strength – representing the proportion of differences based on randomly drawn subsamples that were at least as extreme as the observed test statistic – were larger than 0.12. In addition, post-hoc analyses that were performed to assess the potential influence of differences in subsample size did not result in significant differences in network structure or global strength (Additional Table 3). With regard to individual edges, we found no differences in edge strength for the subsamples examined, with one exception. The edge between appetite loss and psychomotor agitation was significantly stronger in adversity unexposed cases (OR = 1.82) than in adversity exposed cases (OR = 1.00).

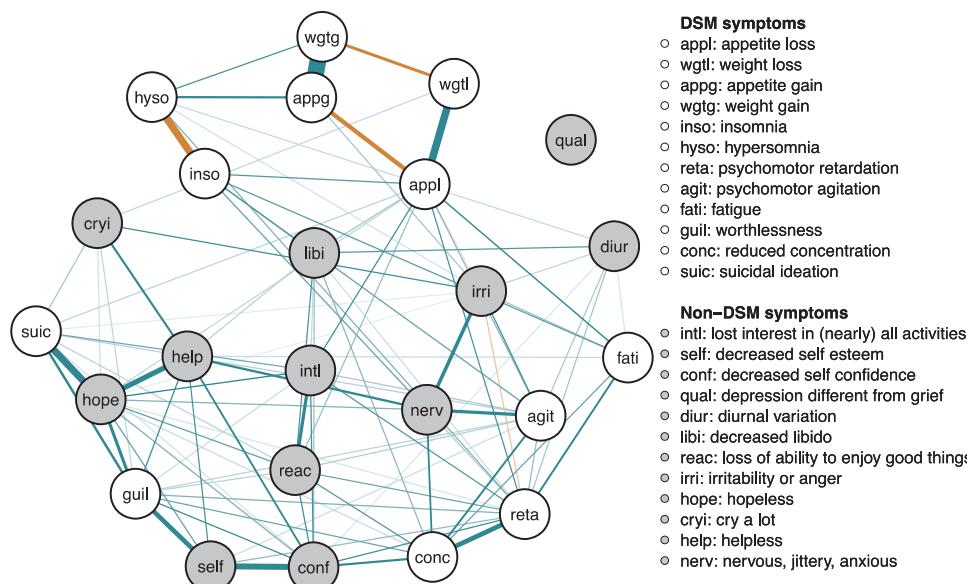


Fig. 2. Estimated regularized symptom network for the full sample ($n = 5784$). Green edges indicate positive (partial) associations; red edges indicate negative associations. Thicker/darker lines indicate stronger associations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

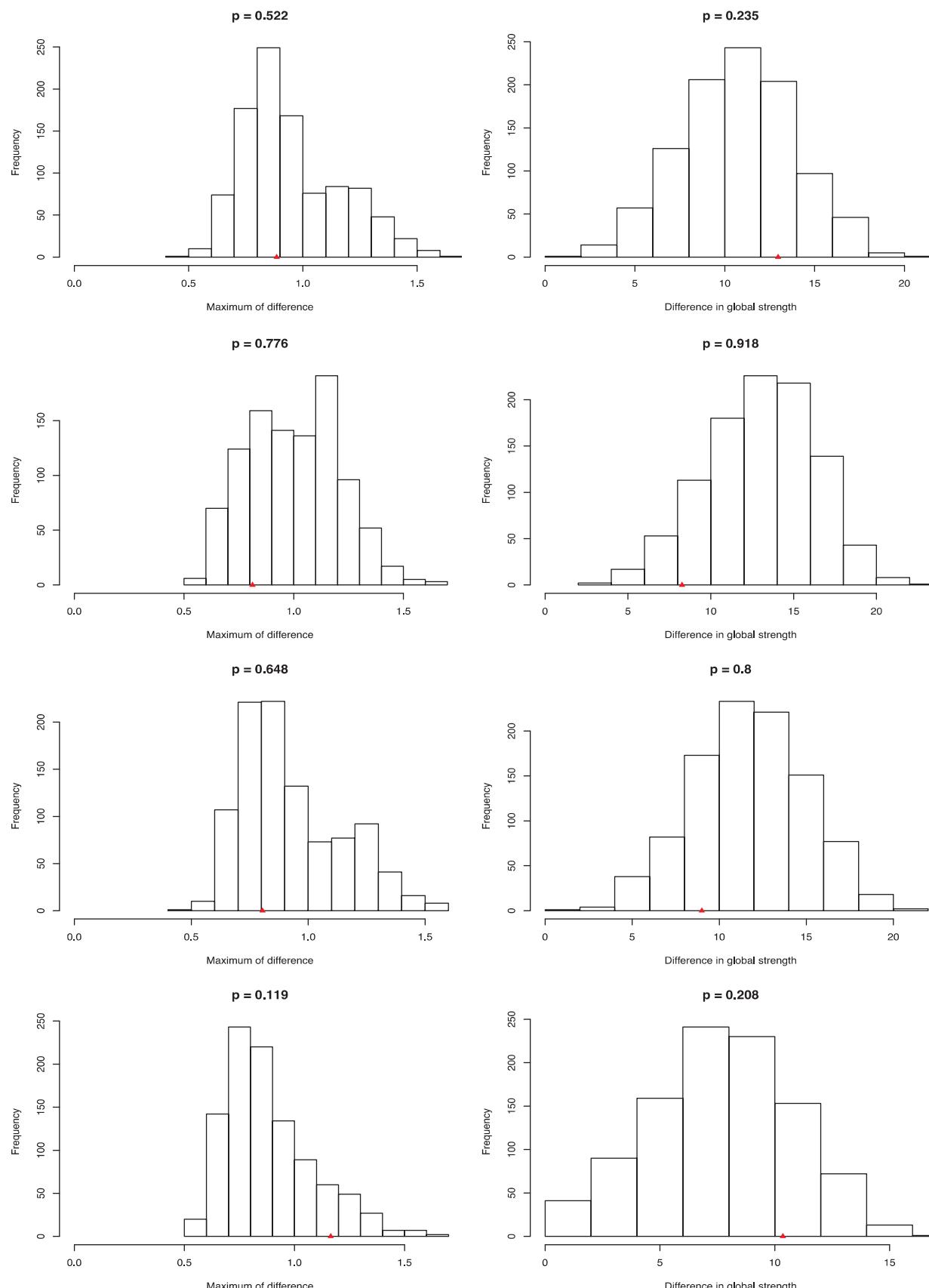


Fig. 3. Permutation testing results for network comparisons of network structure and global strength. From top to bottom: Each row represents the distribution of permutation tests for 1. Family history; 2. Polygenic risk score high vs. low; 3. AAO < 28 vs. ≥ 28 years; 4. Adversity exposed vs. unexposed. p = The P -value equals the proportion of network differences from the 1000 randomly regrouped subsamples at least as extreme as the network differences in the original subsamples. The triangle on the x-axis represents the estimated difference in the original data.

Second, to estimate the similarity of the networks, we correlated the edge weights with each other for each of the four subsample comparisons. Spearman rank-order correlations ranged from 0.60 (family history negative vs. positive) to 0.71 (PRS low vs. high), indicating strong similarities.

Third, on average, symptom intercepts were higher in women with higher genetic or environmental risk, indicating a tendency for the networks to be more activated. The difference in intercepts was most pronounced for cases with a positive family history as compared to cases with a negative family history (Additional Table 4).

3.5. Sensitivity analyses

Post-hoc sensitivity analyses were performed to assess the potential influence of recall bias on the results. We compared 2797 cases with an interval ≤ 1 year between the interview and the worst lifetime episode of MD with 2956 cases with an interval > 1 year between interview and the worst lifetime episode of MD. Symptom endorsement rates were largely similar in these two subsamples. Out of 24 symptoms, two symptoms were significantly more often endorsed in the subsample with a longer recall period, viz. weight loss and cry a lot ($P < 0.0005$). Spearman rank-order correlations between mean endorsement rates in the two subsamples were 0.99. Network comparison tests showed no significant differences in global strength ($P = 0.92$), network structure ($P = 0.822$) or individual edges ($P > 0.28$).

4. Discussion

To further our understanding of the heterogeneity of MD, we investigated symptom patterns and symptom networks across different levels of genetic and environmental risk in a large sample of Han Chinese women with recurrent MD. Apart from some moderate differences – for example, women with an earlier age at onset reported more often appetite gain, hypersomnia and hopelessness than women with an older age at onset – symptom endorsement rates were very similar for women with higher versus lower genetic risk and adversity exposed versus unexposed women. In addition, network comparison testing did not identify statistically significant differences in network structure, strength or associations between symptoms in these subgroups despite the high power due to large samples. Overall our findings suggest that types of symptoms and interconnectivity between symptoms for women experiencing a full MD syndrome are robust across different degrees of genetic and environmental risk.

4.1. Relation to previous studies

Although a range of studies have investigated differences in endorsement rates of depressive symptoms in subgroups of population-based samples with depressive symptoms (Fried et al., 2015b, 2014; Hybels et al., 2012; Keller et al., 2007; Keller and Nesse, 2006, 2005), no studies have investigated network structure in a clinical sample across different levels of genetic or environmental risk for MD like we did in this study.

The most comparable studies to date also investigated differences in associations between depressive symptoms after life events (Cramer et al., 2012; Fried et al., 2015a). Cramer et al. compared depressive symptom networks in a sample from the general population experiencing at least two depressive symptoms after four distinct stressful life events had occurred: stress, romantic loss, health problems, and interpersonal conflicts. Average tetrachoric correlations between depressive symptoms diverged across the subsamples who experienced one of the four different stressful life events, with for instance 'loss of interest' being more highly connected to other depressive symptoms after romantic loss than after health problems (Cramer et al., 2012). Fried et al. (2015a) compared depressive symptoms after partner loss in a prospective study of married elderly couples. This study included 'partner

loss' as a node in the depressive symptom network. Partner loss mainly affected loneliness, and loneliness in turn was associated with feeling sad and less happy, which could be an indication of network differences in widow(er)s versus controls.

The different findings in these two prior studies compared to our study could indicate that symptom networks diverge after specific types of stressful life events as opposed to after adversity in general. However, given the substantial differences between our study and the former studies in sample (population-based samples with depressive symptoms versus clinical sample with recurrent fully syndromal MD) and statistical methods (network estimation and comparison), follow-up studies are needed to clarify which of these aspects explain the different findings.

4.2. Study limitations

Our sample selection counts as one of the limitations of our study. We used data from a sample of Han Chinese women treated for recurrent, severe depression; endorsement rates for 12 of the 24 symptoms were higher than 80%. It might be that this high degree of severity affected the network structure, for example because of ceiling effects (Terluin et al., 2016). In that case, samples with lower degrees of MD-severity could have shown more differences in symptom networks, possibly related to higher variance and unidimensionality (Fried et al., 2016b). Furthermore, our sample included only female patients, whereas gender differences in depressive symptoms have been found in previous studies (Khan et al., 2002; Romans et al., 2007). Future studies are needed to investigate whether similar conclusions hold in male patients with MD.

Second, we used patients' retrospective reports of symptoms during their lifetime worst episode of MD, so we can only examine average undirected relations between symptoms in networks. Longitudinal data, on the other hand, would have given the opportunity to discover specific temporal or directed relationships between symptoms for specific subsamples (e.g., guilt predicting suicidality in one subgroup, but suicidality predicting guilt in another subgroup) or for specific individuals (De Vos et al., 2017). In addition, more fine-grained assessments of symptoms as opposed to dichotomous (present/absent) assessments of symptoms could have improved power to pick up differences in network structure.

Third, the three measures of genetic risk – family history of MD, earlier age of onset, and PRS for MD – all capture a limited degree of genetic risk for MD (Kendler et al., 2005; Peterson et al., 2016). Under the one child policy in China, sibship sizes in the families were small, possibly contributing to the limited power of this method. Early age of onset and a family history of MD might partly be attributed to early environmental factors, and the PRS for MD explains only 1.1% of variability for MD risk in cases versus controls (Peterson et al., 2016). This limitation has likely decreased power to detect differences across different levels of genetic risk. In addition, the PRS used in this study was designed to discriminate between MD-case and control status, as opposed to discriminate between more and less severely ill MD-cases. Future studies are needed to investigate whether different genetic risk factors are involved in MD-onset versus subsequent MD-severity, and eventually create PRS for MD-severity within cases.

Fourth, simulation studies testing the Network Comparison Test and network estimation methods for binary data resulted in very low rates of false-positive significant edges or network differences, but higher false negative rates (Van Borkulo et al., 2017, 2014). Despite the large sample size in this study, these methods could have failed to identify very subtle differences between networks, although the clinical relevance of finding such minor differences is disputable. Nevertheless, future studies should assess the replicability of our findings based on these new statistical methods.

Fifth, estimated intercepts for symptoms in the networks were on average higher in the subsamples with higher genetic risk or adversity,

which possibly indicates a tendency for the networks to be more activated. We were not able to test whether these potential important differences were statistically significant, as the required permutation tests are currently not available. Extensions of network comparison tests are warranted to investigate these potential differences between the networks.

4.3. Implications of findings

Given the fact that this is the first study of its kind in terms of study sample and methods, and other studies have found indications that symptom patterns might differ across different types of life events (e.g., Cramer et al., 2012; Fried et al., 2015a, 2015b), we should be careful in drawing conclusions. If, however, the current findings prove to be robust in replication studies, how might the similar types and associations of symptoms in patients with MD across different levels of genetic and environmental risk be interpreted?

The findings could be compatible with the concept of 'equifinality', derived from general systems theory, which refers to the fact that the same end state may be reached from a variety of different initial conditions and through different processes (Cicchetti and Rogosch, 1996). In other words, different risk factors upstream might result in a similar symptom structure downstream. Genetic risk and environmental risk for MD might thus be associated with different pathways upstream, but eventuate in similar depressive symptom feedback loops and mutual reinforcements downstream. In this sense, the syndrome of MD could be compared to other medical conditions, such as that bacterial infections, auto-immune processes and cancer all of which can lead to a similar syndrome with symptoms of fever, malaise, and anorexia.

For clinical practice, the findings can be viewed both as reassuring and disappointing. On the one hand, the robust structure of MD could indicate that mechanisms on the symptom level (how symptoms are connected) might not be highly diverse in patients with MD – e.g., insomnia is associated with fatigue in most patients – although we should be cautious in translating group-level associations to individuals (Molenaar, 2008). On the other hand, within this sample of female patients with recurrent MD, this study did not identify potential specific mechanisms on the symptom level for MD subtypes with different levels of genetic or environmental risk, and therefore cannot provide much insight for developing personalized treatments based on these distinctions.

For research, these findings suggest that future studies should look beyond depressive symptoms in search of pathways leading to MD. If the syndrome of MD is very similar in patients with different proximal risk factors, and we aim to trace the differential pathways from the proximal risk factors to the full syndrome of MD, we should probably expand our search to different types of information instead of studying depressive symptoms only. Depressive symptoms might not convey enough information to trace differential pathways leading to the syndrome of MD. Similarly, the identification of specific course of illness patterns in patients with MD improved by including a broad range of clinical predictors in addition to depressive symptoms (van Loo et al., 2015; Wardenaar et al., 2014). Another approach would be to focus on pathways to individual depressive symptoms instead of the full syndrome of MD (Fried et al., 2015b, 2014).

Future studies are warranted to study 1) whether specific instead of aggregate genetic and environmental risk factors have a stronger impact on symptom networks, 2) whether different degrees of disease severity (e.g. in terms of number of symptoms, recurrences, types of symptoms) impact on symptom networks, and 3) how different types of risk factors upstream eventuate in a final common pathway and a robust structure of MD downstream. The aim of these studies would be to improve our understanding in the processes leading to MD and identify new possibilities to interfere.

Acknowledgements

The authors acknowledge the support of the CONVERGE consortium (China, Oxford and Virginia Commonwealth University Experimental Research on Genetic Epidemiology) for this work and gratefully acknowledge the support of all partners in hospitals across China. Special thanks to all the CONVERGE collaborators and patients who made this work possible.

Role of the funding source

The collection and analysis of the CONVERGE sample was funded by the Wellcome Trust (WT090532/Z/09/Z, WT083573/Z/07/Z and WT089269/Z/09/Z) and by NIMH grant MH100549.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.jad.2017.10.038>.

References

- American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV, 4th ed. American Psychiatric Publishing, Inc., Arlington, VA, US.
- Borsboom, D., 2017. A network theory of mental disorders. *World Psychiatry* 16, 5–13. <http://dx.doi.org/10.1002/wps.20375>.
- Cai, N., Bigdeli, T.B., Kretschmar, W.W., Li, Y., Liang, J., Hu, J., Peterson, R.E., Bacanu, S., Webb, B.T., Riley, B., Li, Q., Marchini, J., Mott, R., Kendler, K.S., Flint, J., 2017. 11,670 whole-genome sequences representative of the Han Chinese population from the CONVERGE project. *Sci. Data* 4, 170011. <http://dx.doi.org/10.1038/sdata.2017.11>.
- Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4, 7. <http://dx.doi.org/10.1186/s13742-015-0047-8>.
- Chen, J., Cai, Y., Cong, E., Liu, Y., Gao, J., Li, Y., Tao, M., Zhang, K., Wang, X., Gao, C., Yang, L., Li, K., Shi, J., Wang, G., Liu, L., Zhang, J., Du, B., Jiang, G., Shen, J., Zhang, Z., Liang, W., Sun, J., Hu, J., Liu, T., Wang, X., Miao, G., Meng, H., Li, Y., Hu, C., Huang, G., Li, G., Ha, B., Deng, H., Mei, Q., Zhong, H., Gao, S., Sang, H., Zhang, Y., Fang, X., Yu, F., Yang, D., Liu, T., Chen, Y., Hong, X., Wu, W., Chen, G., Cai, M., Song, Y., Pan, J., Dong, J., Pan, R., Zhang, W., Shen, Z., Liu, Z., Gu, D., Wang, X., Liu, X., Zhang, Q., Li, Y., Chen, Y., Kendler, K.S., Shi, S., Flint, J., 2014. Childhood sexual abuse and the development of recurrent major depression in Chinese women. *PLoS One* 9, e87569. <http://dx.doi.org/10.1371/journal.pone.0087569>.
- Cicchetti, D., Rogosch, F.A., 1996. Equifinality and multifinality in developmental psychopathology. *Dev. Psychopathol.* 8, 597. <http://dx.doi.org/10.1017/S0954579400007318>.
- CONVERGE Consortium, 2015. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* 523, 588–591. <http://dx.doi.org/10.1038/nature14659>.
- Cramer, A.O., Borsboom, D., Aggen, S.H., Kendler, K.S., 2012. The pathoplasticity of dysphoric episodes: differential impact of stressful life events on the pattern of depressive symptom inter-correlations. *Psychol. Med.* 42, 957–965. <http://dx.doi.org/10.1017/S003329171100211X>.
- Cramer, A.O.J., Kendler, K.S., Borsboom, D., 2011. Where are the genes? The implications of a network perspective on gene hunting in psychopathology. [A commentary on Johnson et al.]. *Eur. J. Pers.* 25, 270–271. <http://dx.doi.org/10.1002/per.834>.
- Cramer, A.O.J., Waldorp, L.J., van, der, M., Borsboom, D., 2010. Comorbidity: a network perspective. *Behav. Brain Sci.* 33, 137–150.
- Cuijpers, P., Reynolds 3rd, C.F., Donker, T., Li, J., Andersson, G., Beekman, A., 2012. Personalized treatment of adult depression: medication, psychotherapy, or both? A systematic review. *Depression Anxiety* 29, 855–864. <http://dx.doi.org/10.1002/da.21985>.
- De Vos, S., Wardenaar, K.J., Bos, E.H., Wit, E.C., Bouwmans, M.E.J., De Jonge, P., 2017. An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PLoS One*. <http://dx.doi.org/10.1371/journal.pone.0178586>.
- Docherty, A.R., Edwards, A.C., Yang, F., Peterson, R.E., Sawyers, C., Adkins, D.E., Moore, A.A., Webb, B.T., Bacanu, S.A., Flint, J., Kendler, K.S., 2017. Age of onset and family history as indicators of polygenic risk for major depression. *Depression Anxiety*. <http://dx.doi.org/10.1002/da.22607>.
- Eaton, W.W., Shao, H., Nestadt, G., Lee, B.H., Bienvenu, O.J., Zandi, P., 2008. Population-based study of first onset and chronicity in major depressive disorder. *Arch. Gen. Psychiatry* 65, 513–520. <http://dx.doi.org/10.1001/archpsyc.65.5.513>.
- Endicott, J., Andreasen, N., Spitzer, R.L., 1978. Family History Research Diagnostic Criteria. Biometrics Research, New York State Psychiatric Institute, New York, NY.
- Epskamp, S., Borsboom, D., Fried, E.I., 2017. Estimating psychological networks and their accuracy: a tutorial paper. *Behav. Res.* <http://dx.doi.org/10.3758/s13428-017-0862-1>.
- Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmitmann, V.D., Borsboom, D., 2012.

graph: network visualizations of relationships in psychometric data. *J. Stat. Softw.* 48, 1–18.

Fried, E.I., 2017. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. *J. Affect. Disord.* 208, 191–197. <http://dx.doi.org/10.1016/j.jad.2016.10.019>.

Fried, E.I., Bockting, C., Arjadi, R., Borsboom, D., Amshoff, M., Cramer, A.O.J., Epskamp, S., Tuerlinckx, F., Carr, D., Stroebe, M., 2015a. From loss to loneliness: the relationship between bereavement and depressive symptoms. *J. Abnorm. Psychol.* 124, 256–265. <http://dx.doi.org/10.1037/abn0000028>.

Fried, E.I., Epskamp, S., Nesse, R.M., Tuerlinckx, F., Borsboom, D., 2016a. What are “good” depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *J. Affect. Disord.* 189, 314–320. <http://dx.doi.org/10.1016/j.jad.2015.09.005>.

Fried, E.I., Nesse, R.M., Guille, C., Sen, S., 2015b. The differential influence of life stress on individual symptoms of depression. *Acta Psychiatr. Scand.* 131, 465–471. <http://dx.doi.org/10.1111/acps.12395>. (doi).

Fried, E.I., Nesse, R.M., Zivin, K., Guille, C., Sen, S., 2014. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychol. Med.* 44, 2067–2076. <http://dx.doi.org/10.1017/S0033291713002900>.

Fried, E.I., Van Borkulo, C.D., Cramer, A.O.J., Boschloo, L., Schoevers, R.A., Borsboom, D., 2017. Mental disorders as networks of problems: a review of recent insights. *Soc. Psychiatry Psychiatr. Epidemiol.* 52, 1–10. <http://dx.doi.org/10.1007/s00127-016-1319-z>.

Fried, E.I., Van Borkulo, C.D., Epskamp, S., Schoevers, R.A., Tuerlinckx, F., Borsboom, D., 2016b. Measuring depression over time... Or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychol. Assess.* 28, 1354–1367. <http://dx.doi.org/10.1037/pas0000275>.

Guloksuz, S., Pries, L.-K., Van Os, J., 2017. Application of network methods for understanding mental disorders: pitfalls and promise. *Psychol. Med.* <http://dx.doi.org/10.1017/S0033291717001350>.

Hardevelde, F., Spijkerman, J., De Graaf, R., Hendriks, S.M., Licht, C.M., Nolen, W.A., Penninx, B.W., Beekman, A.T., 2013. Recurrence of major depressive disorder across different treatment settings: results from the NESDA study. *J. Affect. Disord.* 147, 225–231. <http://dx.doi.org/10.1016/j.jad.2012.11.008>.

Holma, K.M., Holma, I.A.K., Melartin, T.K., Ryttsälä, H.J., Isometsä, E.T., 2008. Long-term outcome of major depressive disorder in psychiatric patients is variable. *J. Clin. Psychiatry* 69, 196–205. <http://dx.doi.org/10.4088/JCP.v69n0205>.

Hybels, C.F., Landerman, L.R., Blazer, D.G., 2012. Age differences in symptom expression in patients with major depression. *Int. J. Geriatr. Psychiatry* 27, 601–611. <http://dx.doi.org/10.1002/gps.2759>.

Keller, M.C., Neale, M.C., Kendler, K.S., 2007. Association of different adverse life events with distinct patterns of depressive symptoms. *Am. J. Psychiatry* 164, 1521–1529. <http://dx.doi.org/10.1176/appi.ajp.2007.06091564>.

Keller, M.C., Nesse, R.M., 2006. The evolutionary significance of depressive symptoms: different adverse situations lead to different depressive symptom patterns. *J. Pers. Soc. Psychol.* 91, 316–330. <http://dx.doi.org/10.1037/0022-3514.91.2.316>.

Keller, M.C., Nesse, R.M., 2005. Is low mood an adaptation? Evidence for subtypes with symptoms that match precipitants. *J. Affect. Disord.* 86, 27–35. <http://dx.doi.org/10.1016/j.jad.2004.12.005>.

Kendler, K.S., 2013. What psychiatric genetics has taught us about the nature of psychiatric illness and what is left to learn. *Mol. Psychiatry* 18, 1058–1066. <http://dx.doi.org/10.1038/mp.2013.50>.

Kendler, K.S., Aggen, S.H., Flint, J., Borsboom, D., Fried, E.I., n.d. The centrality of DSM and non-DSM depressive symptoms in Han Chinese women with major depression. *J. Affect. Disord.*

Kendler, K.S., Aggen, S.H., Li, Y., Lewis, C.M., Breen, G., Boomsma, D.I., Bot, M., Penninx, B.W., Flint, J., 2015. The similarity of the structure of DSM-IV criteria for major depression in depressed women from China, the United States and Europe. *Psychol. Med.* 45, 1945–1954. <http://dx.doi.org/10.1017/S0033291714003067>.

Kendler, K.S., Gatz, M., Gardner, C.O., Pedersen, N.L., 2006. A Swedish national twin study of lifetime major depression. *Am. J. Psychiatry* 163, 109–114. <http://dx.doi.org/10.1176/appi.ajp.163.1.109>.

Kendler, K.S., Gatz, M., Gardner, C.O., Pedersen, N.L., 2005. Age at onset and familial risk for major depression in a Swedish national twin sample. *Psychol. Med.* 35, 1573–1579. <http://dx.doi.org/10.1017/S0033291705005714>.

Kendler, K.S., Karkowski, L.M., Prescott, C.A., 1999. Causal relationship between stressful life events and the onset of major depression. *Am. J. Psychiatry* 156, 837–841. <http://dx.doi.org/10.1176/ajp.156.6.837>.

Kendler, K.S., Prescott, C.A., 2006. *Genes, Environment and Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders*. Guilford Press, New York.

Kessler, R.C., 1997. The effects of stressful life events on depression. *Annu. Rev. Psychol.* 48, 191–214. <http://dx.doi.org/10.1146/annurev.psych.48.1.191>.

Khan, A.A., Gardner, C.O., Prescott, C.A., Kendler, K.S., 2002. Gender differences in the symptoms of major depression in opposite-sex dizygotic twin pairs. *Am. J. Psychiatry* 159, 1427–1429. <http://dx.doi.org/10.1176/appi.ajp.159.8.1427>.

Krishnan, V., Nestler, E.J., 2008. The molecular neurobiology of depression. *Nature* 455, 894–902. <http://dx.doi.org/10.1038/nature07455>.

Laumann, E.O., Gagnon, J.H., Michael, R.T., Micheals, S., Heiman, J., 1994. *The Social Organization of Sexuality: Sexual Practices in the United States, Social Organization of Sexuality: Sexual Practices in the United States*. University of Chicago Press, Chicago.

Li, Y., Aggen, S., Shi, S., Gao, J., Li, Y., Tao, M., Zhang, K., Wang, X., Gao, C., Yang, L., Liu, Y., Li, K., Shi, J., Wang, G., Liu, L., Zhang, J., Du, B., Jiang, G., Shen, J., Zhang, Z., Liang, W., Sun, J., Hu, J., Liu, T., Wang, X., Miao, G., Meng, H., Li, Y., Hu, C., Li, Y., Huang, G., Li, G., Ha, B., Deng, H., Mei, Q., Zhong, H., Gao, S., Sang, H., Zhang, Y.,

Fang, X., Yu, F., Yang, D., Liu, T., Chen, Y., Hong, X., Wu, W., Chen, G., Cai, M., Song, Y., Pan, J., Dong, J., Pan, R., Zhang, W., Shen, Z., Liu, Z., Gu, D., Wang, X., Liu, X., Zhang, Q., Flint, J., Kendler, K.S., 2014. The structure of the symptoms of major depression: exploratory and confirmatory factor analysis in depressed Han Chinese women. *Psychol. Med.* 44, 1391–1401. <http://dx.doi.org/10.1017/S03329171300192X>.

Martin, J., Anderson, J., Romans, S., Mullen, P., O’Shea, M., 1993. Asking about child sexual abuse: methodological implications of a two stage survey. *Child Abus. Negl.* 17, 383–392. [http://dx.doi.org/10.1016/0145-2134\(93\)90061-9](http://dx.doi.org/10.1016/0145-2134(93)90061-9).

Milaneschi, Y., Lamers, F., Peyrot, W.J., Abdellaoui, A., Willemsen, G., Hottenga, J.J., Jansen, R., Mbarek, H., Dehghan, A., Lu, C., group, C. inflammation working, Boomsma, D.I., Penninx, B.W., 2016. Polygenic dissection of major depression clinical heterogeneity. *Mol. Psychiatry* 21, 516–522. <http://dx.doi.org/10.1038/mp.2015.86>.

Molenaar, P.C., 2008. On the implications of the classical ergodic theorems: analysis of developmental processes has to focus on intra-individual variation. *Dev. Psychobiol.* 50, 60–69. <http://dx.doi.org/10.1002/dev.20262>.

Oquendo, M.A., Barrera, A., Ellis, S.P., Li, S., Burke, A.K., Grunbaum, M., Endicott, J., Mann, J.J., 2004. Instability of symptoms in recurrent major depression: a prospective study. *Am. J. Psychiatry* 161, 255–261. <http://dx.doi.org/10.1176/appi.ajp.161.2.255>.

Patterson, N., Price, A.L., Reich, D., 2006. Population structure and eigen analysis. *PLoS Genet.* 2, e190. <http://dx.doi.org/10.1371/journal.pgen.0020190>.

Peterson, R.E., Cai, N., Bigdeli, T.B., Li, Y., Reimers, M., Nikulova, A., Webb, B.T., Bacanu, S.A., Riley, B.P., Flint, J., Kendler, K.S., 2016. The genetic architecture of major depressive disorder in Han Chinese Women. *JAMA Psychiatry*. <http://dx.doi.org/10.1001/jamapsychiatry.2016.3578>.

Pittenger, C., Duman, R.S., 2008. Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology* 33, 88–109. <http://dx.doi.org/10.1038/sj.npp.1301574>.

Power, R.A., Tansey, K.E., Buttenschon, H.N., Cohen-Woods, S., Bigdeli, T., Hall, L.S., Kutzlik, Z., Lee, S.H., Ripke, S., Steinberg, S., Teumer, A., Viktorin, A., Wray, N.R., Arolt, V., Baune, B.T., Boomsma, D.I., Borglum, A.D., Byrne, E.M., Castelao, E., Craddock, N., Craig, I.W., Dannowski, U., Deary, I.J., Degenhardt, F., Forstner, A.J., Gordon, S.D., Grabe, H.J., Grove, J., Hamilton, S.P., Hayward, C., Heath, A.C., Hocking, L.J., Homuth, G., Hottenga, J.J., Kloiber, S., Krogh, J., Landen, M., Lang, M., Levinson, D.F., Lichtenstein, P., Lucae, S., MacIntyre, D.J., Madden, P., Magnusson, P.K., Martin, N.G., McIntosh, A.M., Middeldorp, C.M., Milaneschi, Y., Montgomery, G.W., Mors, O., Muller-Mylsok, B., Nyholt, D.R., Oskarsson, H., Owen, M.J., Padmanabhan, S., Penninx, B.W., Pergadia, M.L., Porteous, D.J., Potash, J.B., Preisig, M., Rivera, M., Shi, J., Shyn, S.I., Sigurdsson, E., Smit, J.H., Smith, B.H., Stefansson, H., Stefansson, K., Strohmaier, J., Sullivan, P.F., Thomson, P., Thorgeirsson, T.E., Van der Auwera, S., Weissman, M.M., CONVERGE Consortium GERAD1 Consortium, Cardi, C., Breen, G., Lewis, C.M., 2017. Genome-wide association for major depression through age at onset stratification: major depressive disorder working group of the psychiatric genomics consortium. *Biol. Psychiatry* 81, 325–335 (doi:10.1006/sj.bip.2016.3223)(16)(32386-1)(pii).

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909. <http://dx.doi.org/10.1038/ng.1847>.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. <http://dx.doi.org/10.1086/519795>.

R Core Team, 2017. R: A Language and Environment for Statistical Computing.

Romans, S.E., Tyas, J., Cohen, M.M., Silverstone, T., 2007. Gender differences in the symptoms of major depressive disorder. *J. Nerv. Ment. Dis.* 195, 905–911. <http://dx.doi.org/10.1097/nmd.0b013e3181594cb7>.

Simon, G.E., Perlis, R.H., 2010. Personalized medicine for depression: can we match patients with treatments? *Am. J. Psychiatry* 167, 1445–1455. <http://dx.doi.org/10.1176/appi.ajp.2010.09111680>.

Sullivan, P.F., Neale, M.C., Kendler, K.S., 2000. Genetic epidemiology of major depression: review and meta-analysis. *Am. J. Psychiatry* 157, 1552–1562. <http://dx.doi.org/10.1176/appi.ajp.157.10.1552>.

Tao, M., Xie, D., Wang, Z., Qiu, J., Wu, W., Sun, J., Tao, D., Zhao, H., Tian, T., Zhang, J., Gao, C., Niu, Q., Liu, S., Liu, J., Zhang, Y., He, Q., Rong, H., Gan, Z., Li, J., Chen, X., Pan, J., Cui, Y., Han, W., Ma, H., Xie, S., Jin, G., Li, L., Zhang, R., Tan, Q., Guan, J., Shi, S., Chen, Y., Kendler, K.S., Flint, J., Gao, J., Li, Y., Li, Q., 2011. Examining the relationship between lifetime stressful life events and the onset of major depression in Chinese women. *J. Affect. Disord.* 135, 95–99. <http://dx.doi.org/10.1016/j.jad.2011.06.054>.

Terluin, B., De Boer, M.R., De Vet, H.C.W., 2016. Differences in connection strength between mental symptoms might be explained by differences in variance: reanalysis of network data did not confirm staging. *PLoS One* 11. <http://dx.doi.org/10.1371/journal.pone.0155205>.

Van Borkulo, C.D., Boschloo, L., Lynn, Kossakowski, J., Tio, P., Schoevers, Robert A., Borsboom, D., Waldorp, L.J., 2017. Comparing network structures on three aspects: a permutation test. <http://doi.org/10.13140/RG.2.2.29455.38569>.

Van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B.W., Waldorp, L.J., Schoevers, R.A., 2015. Association of symptom network structure with the course of [corrected] depression. *JAMA Psychiatry* 72, 1219–1226. <http://dx.doi.org/10.1001/jamapsychiatry.2015.2079>.

Van Borkulo, C.D., Borsboom, D., Epskamp, S., Blanken, T.F., Boschloo, L., Schoevers, R.A., Waldorp, L.J., 2014. A new method for constructing networks from binary data. *Sci. Rep.* 4. <http://dx.doi.org/10.1038/srep05918>.

van Loo, H.M., Aggen, S.A., Chardner, C.O., Kendler, K.S., 2015. Multiple risk factors

predict recurrence of major depressive disorder in women. *J. Affect. Disord.* 180, 52–61. <http://dx.doi.org/10.1016/j.jad.2015.03.045>.

Verduijn, J., Milaneschi, Y., Peyrot, W.J., Hottenga, J.J., Abdellaoui, A., de Geus, E.J.C., Smit, J.H., Breen, G., Lewis, C.M., Boomsma, D.I., Beekman, A.T.F., Penninx, B.W.J.H., 2017. Using clinical characteristics to identify which patients with major depressive disorder have a higher genetic load for three psychiatric disorders. *Biol. Psychiatry* 81, 316–324. <http://dx.doi.org/10.1016/j.biopsych.2016.05.024>.

Wardenaar, K.J., van Loo, H.M., Cai, T., Fava, M., Gruber, M.J., Li, J., de Jonge, P., Nierenberg, A.A., Petukhova, M.V., Rose, S., Sampson, N.A., Schoevers, R.A., Wilcox, M.A., Alonso, J., Bromet, E.J., Bunting, B., Florescu, S.E., Fukao, A., Gureje, O., Hu, C., Huang, Y.Q., Karam, A.N., Levinson, D., Medina Mora, M.E., Posada-Villa, J., Scott, K.M., Taib, N.I., Viana, M.C., Xavier, M., Zarkov, Z., Kessler, R.C., 2014. The effects of co-morbidity in defining major depression subtypes associated with long-term course and severity. *Psychol. Med.* 44, 3289–3302. <http://dx.doi.org/10.1017/S0033291714000993>.

Wichers, M., 2014. The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychol. Med.* 44, 1349–1360.

Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. <http://dx.doi.org/10.1016/j.ajhg.2010.11.011>.