Psychiatry Research 271 (2019) 640–648

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Capturing the risk of persisting depressive symptoms: A dynamic network investigation of patients' daily symptom experiences

Robin N. Groen[a,*], Evelien Snippe[a], Laura F. Bringmann[b], Claudia J.P. Simons[c,d], Jessica A. Hartmann[e], Elisabeth H. Bos[f], Marieke Wichers[a]

[a] Interdisciplinary Center for Psychopathology and Emotion regulation (ICPE), Dept. of Psychiatry (UCP), University of Groningen, University Medical Centre Groningen (UMCG), Groningen, The Netherlands
[b] Heymans Institute, Department of Psychometrics and Statistics, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands
[c] School for Mental Health and Neuroscience, Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands
[d] GGzE, Institute for Mental Health Care Eindhoven and De Kempen, Eindhoven, The Netherlands
[e] Orygen, the National Centre of Excellence in Youth Mental Health, University of Melbourne, Australia
[f] Heymans Institute, Department of developmental Psychology, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands

ARTICLE INFO

Keywords:
Network approach
Network connectivity
Major depressive disorder
Psychiatry
Multilevel vector autoregressive (VAR) modelling

ABSTRACT

What drives the large differences across patients in terms of treatment efficacy of major depressive disorder (MDD) is unclear. A network approach to psychopathology may help to reveal underlying mechanisms determining patients' capacity for recovery. We used daily diary MDD symptom data and six-month follow-up data on depression to examine how dynamic associations between symptoms relate to the future course of MDD. Daily experiences of depressive symptoms of 69 participants were assessed by means of the SCL-90-R depression subscale, three days a week for a period of six weeks, as part of a larger intervention study. Multilevel vector autoregressive modelling was used to estimate networks of dynamic symptom connections. Long-term outcome was determined by the percentage change in Hamilton Depression Rating Scale (HDRS) score between pre-intervention and six-month follow-up. For patients with more persisting symptoms, the symptom 'feeling everything is an effort' most strongly predicted other symptoms. The networks of the two groups did not significantly differ in overall connectivity. Findings suggest that future research should not solely focus on the presence or intensity of individual symptoms when predicting long-term outcomes, but should also examine the role of a specific symptom in the larger network of dynamic symptom-to-symptom interactions.

1. Introduction

While a variety of pharmacological and psychological treatment options are available for depression, response rates are fairly low. About 50 percent of all patients do not respond sufficiently to treatment (Fava et al., 2003; Papakostas, 2009). Currently, it appears difficult to determine who will respond. One approach that could potentially further our understanding of patients' capacity for recovery is to focus on processes occurring at the symptom level.

Specifically, symptom dynamics within individuals may be informative in understanding symptom persistence. It has been theorized that mental disorders can be understood as the result of a continuous interplay between symptoms (Cramer et al., 2010; Teasdale, 1988). By activating and deactivating each other over time, symptoms may form a complex system (i.e., a network) of mutually reinforcing dependencies (Schmittmann et al., 2013). Within this network, symptoms function as nodes and the dynamic associations between symptoms can be modelled as edges (Bringmann et al., 2013; Borsboom and Cramer, 2013). Individual differences in network structure are likely to exist, for example, differences concerning which symptoms are connected, how strong the associations are, and which symptoms are most important in the network (Booij et al., 2016; Stavrakakis et al., 2013; van Gils et al., 2014; Wichers, 2014). These inter-individual differences may help to understand variability in patients' capacity for recovery.

One network characteristic that may be clinically relevant is the overall connectivity of the network. In psychopathology research, overall connectivity is frequently calculated as the mean strength of the absolute connection weights between nodes (Bringmann et al., 2016; Pe et al., 2015). It has been hypothesized that networks with stronger overall connectivity are more vulnerable networks (Cramer et al.,

* Corresponding author at: University of Groningen, University Medical Center Groningen, Dept. of Psychiatry, Interdisciplinary Center Psychopathology and Emotion regulation, PO Box 30.001, 9700 RB Groningen, The Netherlands
E-mail address: r.n.groen@umcg.nl (R.N. Groen).

https://doi.org/10.1016/j.psychres.2018.12.054
Received 17 January 2018; Received in revised form 7 December 2018; Accepted 7 December 2018
Available online 09 December 2018
0165-1781/

2016): when symptoms are likely to trigger each other more easily and strongly, it becomes increasingly difficult to break the cycle of activation. Indeed, findings derived from simulated data (Cramer et al., 2016) and empirical momentary assessment data on mental states (Pe et al., 2015; Bringmann et al., 2016; Wigman et al., 2015) indicate that stronger overall connectivity may signal vulnerability to psychopathology. Cramer et al., (2016) found that simulated networks with strong connectivity induced a depressed state sooner than other networks and did not show spontaneous recovery. Moreover, healthy individuals with higher levels of neuroticism (Bringmann et al., 2016) and individuals with MDD appeared to have dynamic networks with higher overall connectivity compared to healthy individuals (Pe et al., 2015; Wigman et al., 2015). However, these findings need to be interpreted with care as a recent study suggests that group differences in connectivity may be confounded by group differences in severity (Terluin et al., 2016).

Another network characteristic with potential clinical relevance concerns symptom centrality; i.e. which specific symptoms have more and stronger connections with other symptoms. This may specifically hold for MDD, as there is a large variability in symptom profiles among patients with MDD (Fried and Nesse, 2015; Zimmerman et al., 2015). For instance, specific MDD symptoms have been differentially associated with genetic factors (Kendler et al., 2013), environmental risk factors (Keller et al., 2007), efficacy of antidepressants (Uher et al., 2012), and psychosocial impairment (Fried and Nesse, 2014). Concerning the course of depression, Conradi et al., (2011) observed that cognitive problems, lack of energy, and sleeping problems were more likely to linger during remission than other symptoms. Together, these findings illustrate the potential relevance of studying individual symptoms within a network.

Two studies examined whether network connectivity and symptom centrality are associated with persistence of MDD (van Borkulo et al., 2015; Schweren et al., 2018). Van Borkulo et al. (2015) reported that a baseline depressive symptom network obtained in patients with persistent MDD at two-year follow-up had higher overall connectivity as compared to the baseline network from patients with remitted symptoms at two-year follow-up. Schweren et al. (2018), however, did not find support for an association between increased overall network connectivity and poorer outcomes. In terms of symptom centrality, the symptoms 'concentration problems' (Schweren et al., 2018) and 'fatigue' and 'feeling guilty' (van Borkulo et al., 2015) were most central in the network of participants with persisting symptoms; 'loss of interest' was found to be most central in the network of the remitted MDD group (van Borkulo et al., 2015). As these studies investigated symptom networks based on cross-sectional data (symptoms assessed at a single time point), they indicate that the co-occurrence of symptoms may differ in groups of individuals with a poor versus a good clinical outcome.

To investigate whether symptom persistence is associated with how symptoms follow each other within individuals, a dynamic network approach may be more suitable than a cross-sectional network approach (Bos and Wanders, 2016). Recent findings also suggest that the two approaches may lead to different conclusions (Bos et al., 2017). Multilevel vector autoregressive (VAR) models are suited to estimate dynamic associations between multiple variables, as they allow to test temporal (lagged) associations as well as bidirectional associations (Bolger and Laurenceau, 2013; Rovine and Walls, 2006), which gives insight in how symptoms follow each other over time (Bringmann et al., 2013). Another advantage of multilevel VAR models is that between-group differences in within-person associations can be tested (Bringmann et al., 2013). One study has used repeated assessments (weekly) of depressive symptoms to estimate a dynamic symptom network (Bringmann et al., 2015). The authors found that specifically the symptoms 'loss of pleasure' and 'past failure' were most central. Yet, a study that relates the characteristics of a dynamic symptom network with long-term depressive symptom persistence is currently lacking.

The present study is the first to investigate whether the characteristics of a dynamic symptom network provide insight regarding long-term symptom persistence. Due to the recurrent nature of MDD, we considered long-term persistence most clinically relevant. We used temporal data, consisting of daily assessments of depressive symptoms to derive networks based on symptom dynamics within individuals. Data were collected during a randomized controlled trial (RCT) for patients with MDD (Kramer et al., 2014). Long-term changes in depressive symptoms were therefore expected to occur, enabling comparisons between patients who experienced reduced or persisting symptoms. Our aim was to examine whether overall connectivity (including inter-node and intra-node connectivity) can differentiate between individuals who differ with respect to improvement of depressive symptoms at six-month follow-up. We expected to observe stronger network connectivity in patients with persisting MDD symptoms as compared to patients experiencing symptom reductions at follow-up. Exploratory, we examined group-differences in centrality of individual symptoms and differences in dynamic associations between specific symptoms

2. Method

2.1. Sample

Data were derived from the intervention period of a RCT investigating the effectiveness of an Experience Sampling Method (ESM) intervention as adjunct to treatment-as-usual (TAU) for depression (Kramer et al., 2014).

Recruitment took place between January 2010 and February 2012 through health care professionals working in catchment areas of mental health facilities located around the Dutch cities of Eindhoven and Maastricht. Inclusion criteria were (i) aged between 18 and 65 years and (ii) a DSM-IV diagnosis of major depressive disorder with current or residual symptoms (17-item Hamilton Depression Rating Scale (HDRS) > 7 (Hamilton, 1960)), and (iii) current use of anti-depressants or mood stabilizers. Exclusion criteria included a diagnosis of non-affective psychosis (e.g. schizophrenia), schizoaffective disorder, or a (hypo) manic or mixed episode within the past month. The standing Ethics Committee approved the project and all participants gave written informed consent before enrolment. One hundred and two patients were included. For the present study, only the patients enrolled in the experimental (n = 33) and pseudo-experimental (n = 36) treatment arm were included.

2.2. Design

A detailed description of the ESM intervention can be found elsewhere (Hartmann et al., 2015; Kramer et al., 2014). In short, individuals in the treatment arms completed 10 ESM assessments a day for three consecutive days per week, for six weeks (i.e., 18 days in total). Individuals either received weekly feedback based on the ESM assessments (experimental group) or completed a weekly HDRS assessment (pseudo-experimental group). At the last ESM assessment of each day, individuals also reported levels of depressive symptoms, which we used to investigate symptom dynamics. Additionally, participants completed five follow-up assessments (1, 4, 8, 12, and 24 weeks after the end of the intervention), including assessments of depressive symptoms. The follow-up at 24 weeks (six months) was used to evaluate long-term symptom improvement and persistence.

2.3. Measures

2.3.1. Daily depressive symptoms

Daily levels of depressive symptoms were measured with the 13-item Symptom Checklist-90-Revised (SCL-90-R) depression subscale (Derogatis, 1992). Participants were asked: "How much were you bothered today by:" (see Table 2 for the exact wording of all included

items), and rated each item on a five-point Likert scale ranging from 1 = 'not at all' to 5 = 'extremely' . Items were included if the daily responses to these items displayed sufficient within-person variation. Two items, 'crying easily' and 'thoughts of ending life', were excluded because the majority of participants entered identical responses for more than 80% of the time (Lebo and Nesselroade, 1978).

2.3.2. Long-term change in depressive symptomatology

Depressive symptomatology was assessed with the 17-item HDRS at baseline and five follow-up assessments. Sensitivity and internal, interrater, and retest reliability of the 17-item HDRS are good (Bagby et al., 2004). The percentage change in HDRS score between baseline and six-month follow-up served to measure symptom improvement or persistence. If the HDRS score was missing at six-month follow-up, the HDRS score at three-month follow-up was used ($n = 7$). To compare participants with more persistent symptoms to those with reduced symptoms, we performed a median split on the percentage HDRS change scores as this maximized power for estimating the symptom associations in each group. Median HDRS percentage change was 21.1% (min: −950%, max: 100%), participants above the median formed the reduced symptoms group, and those below the median the more persistent symptoms group.

2.4. Statistical analyses

In the following section, we first describe how dynamic networks were estimated. Next, we discuss the permutation procedure to test for group differences in overall connectivity (including inter-node and intra-node connectivity), centrality of individual symptoms, and dynamic associations between specific symptoms.

2.4.1. Dynamic network estimation

Associations between symptoms over time were estimated with multilevel vector autoregressive modelling (Bringmann et al., 2013; Rovine and Walls, 2006). Multilevel VAR models assume stationarity of the modelled process, as trends in the data may lead to spurious correlations between lagged variables (Rovine and Walls, 2006). Hence, to account for changes in SCL-90-R depression symptoms over the course of the intervention, we performed a detrending method that subtracts both the person mean and linear trend line from the data (Curran and Bauer, 2011). For each participant, we regressed each symptom separately on a linear function of time (1,2,3…n), using ordinary least squares (OLS). The difference (residual) between the predicted and the participant's actual response on that symptom was used as predictor variable. Another advantage of this 'detrending and person-mean centering' procedure is that variance in the data now represented within-person variance only, instead of both within-person and between-person variance (Curran and Bauer, 2011). This is important, as we are interested in within-person effects. However, person-mean centering may lead to a slight underestimation of the autoregressive effects (Hamaker and Grasman, 2015). The newly computed variables were included as fixed effects in the multilevel model.

For each symptom, and each outcome group we performed a separate multilevel analysis with the Stata mixed command (Stata version 14). In each model, a symptom at each time point (t) was modelled as a function of the lagged response on that same symptom (to control for possible autoregressive effects) and lagged values of all other symptoms (t-1). Furthermore, mean severity of depressive symptoms was added as a fixed effect to the model to control for differences in severity. The intervention's structure, with assessments during three consecutive days per week resulted in two lagged associations per week to be available. We estimated no lag between the last day of the previous week and first day of the following week, as associations with a lag of 4 days were expected to be different from associations with a lag of 1 day. A random intercept was included in each model. With 11 items, it was computationally not possible to include all 11 random slopes in the multilevel autoregressive model. Moreover, including all random effects in the model would have resulted in overfitting given the number of available observations and the number of estimated parameters. Therefore, we excluded random slopes from the model. Excluding random slopes did not affect the estimation of our p-values since we used a permutation testing procedure (Good, 2005), which does not make parametric assumptions about the null distribution of test statistics to generate p-values. The used permutation procedure developed by Wolfgang Viechtbauer (code attached in the supplementary materials, see Klippel et al., 2017, Snippe et al., 2017 for previous use of this permutation code) compares the results based on the observed data with a distribution of results derived from repeated permutation of the data. Networks were visualized with the qgraph package (Epskamp et al., 2012) in the statistical program R (R Core Team, 2017).

2.4.2. Network connectivity

Connectivity estimates were based on all network edges (including non-significant edges), and were calculated by averaging over the absolute values of all edges (overall connectivity), only cross-regressive edges (inter-node connectivity), and only autoregressive effects (intra-node connectivity). We used permutation tests investigate group differences in connectivity. Group differences were calculated by subtracting the connectivity estimates of the reduced group from the connectivity estimates of the more persistent group. To create the permutation distribution, first connectivity differences based upon the regression coefficients derived from fitting a model with the actual data were saved. Then, the group variable was randomly assigned to each individual, and connectivity differences were estimated based on regression coefficients derived from the model with the reshuffled data. This process was repeated 100,000 times. Subsequently, statistical significance was determined by comparing the size of the connectivity differences based on the actual data to the permutation distribution of connectivity differences. P-values were computed based on the proportion of times that the connectivity difference under the permutation distribution was as extreme or more extreme than the value of the observed connectivity difference (i.e. whether the difference fell into the area of the 2.5% most extreme values in the permutation distribution). By doubling the proportions, the two-sided p-values were obtained. Coefficients with a (two-sided) p-value below 0.05 were considered statistically significant.

2.4.3. Centrality indices

We used the centrality indices in-strength and out-strength- to examine symptoms' roles in the networks of the two groups. Centrality indices were based on all coefficients estimated in the multilevel autoregressive models (including non-significant coefficients) and quantified the importance of each symptom (Opsahl et al., 2010). *In-strength* and *out-strength* reflect the sum of ingoing edge weights and the sum of outgoing edge weights to the specific node, respectively. Due to the presence of negative associations in the network, we used a different approach than for instance Wigman et al. (2015) to estimate centrality indices. Because the sign of the association provides information on whether a symptom may have an activating or inhibiting role within the network, we estimated centrality indices separately for positive and negative associations. Significant differences in positive and negative, in-strength and out-strength between groups were determined with permutation testing, and permutation-based p-values were generated by permuting the group variable. With these separate measures, we evaluated whether the symptom mostly influenced other symptoms within the network or was more likely to receive input from other symptoms. In-strength and out-strength were visualized with the qgraph (Epskamp et al., 2012).

2.4.4. Edge differences

Statistical significance for group differences in network edges was determined by comparing the size of the edge-differences based on the

actual data to a permutation distribution of the edge-differences. To create a permutation distribution, the group label was randomly assigned to participants. This was repeated 100,000 times and each time a model was fitted with the reshuffled data. P-values were computed according to the same procedure as described before (see Section 2.4.2. for details).

3. Results

3.1. Descriptive statistics and baseline group differences

Nine participants were excluded from further analyses, because they did not start the intervention ($n = 2$) or had missing follow-up data ($n = 7$). Of the remaining 60 participants, 30 participants fell into the more persistent symptom group and 30 into the reduced symptom group. Groups did not significantly differ ($t(53) = 1.7$; $p = 0.1$) in HDRS sum scores post daily diary assessments, but did significantly differ ($t(58) = 7.2$; $p < 0.001$) in HDRS sum score at six-month follow-up (reduced group $M = 6.5$, $SD = 4.5$; more persistent group $M = 15.4$, $SD = 5.0$). Average change for the reduced group was $M = 7.2$ ($SD = 4.0$) and for the more persistent group $M = 1.3$ ($SD = 4.6$) which was significantly different ($t(58) = -7.6$; $p < 0.001$). Table 1 shows the demographic and baseline characteristics of each group. Groups did not differ in baseline HDRS sum scores. Groups did also not differ in average level, variances and mean squared successive differences (MSSDs; von Neumann et al., 1941) of any of the daily depressive symptoms (see Table 2 and Table S1 for MSSDs). Some participants completed more daily assessments than planned, which resulted in a higher mean number of observations than expected. However, groups did not differ on the number of available lagged observations. On average 11.6 lagged observations (median: 11, range 3–19) were available for participants in the persistent symptom group, compared to 11.0 lagged observations (median: 12, range 0–14) for participants in the reduced symptoms group. More participants in the reduced group had switched antidepressant medication or had started antidepressant medication in the 8 weeks prior to study entry.

3.2. Overall network connectivity

Permutation testing revealed no statistically significant difference between the two groups in overall network connectivity (difference = −0.0145, $p = 0.43$; more persistent group $B = 0.087$; reduced group $B = 0.102$). Neither did the two groups differ significantly in terms of intra-node connectivity (difference = −0.043, $p = 0.23$; more persistent group $B = 0.095$; reduced group $B = 0.138$) or inter-node connectivity (difference = −0.012, $p = 0.51$; more persistent group $B = 0.087$; reduced group $B = 0.098$).

3.3. Symptom centrality

The importance of specific symptoms within the networks was evaluated based on the centrality indices (see Fig. 1; unstandardized in Figure S1) that were derived from all (incl. non-significant) edges in the network. For the more persistent group, 'feeling everything is an effort' (effort) had the highest positive out-strength, suggesting that if intensity levels in effort would change, this would be followed by larger activation of other nodes in the network than if levels in other symptoms would change. In the reduced group's network 'worrying too much about things' (worrying) appeared to have the highest positive out-strength. Permutation testing revealed a difference between groups in symptom centrality for the symptom 'effort', which at the uncorrected α level had a higher positive out-strength in the 'more persistent' group than the 'reduced group' (see table 3). Groups did not significantly differ on any of the other centrality indices (see table S2 for permutation results).

3.4. Edge differences between groups

Fig. 2 displays the 'full' dynamic symptom networks for the reduced and more persistent symptom group, which include all non-significant edges with coefficients > 0.10. We report all dynamic associations in Table S3 in the supplementary materials. In the network, each node represents a symptom and each edge is the lagged association of that symptom at *t-1* (yesterday) with another symptom at *t* (today).

Table 1
Demographic and baseline clinical characteristics of major depression course subgroups.

	Reduced symptom group ($n = 30$)	More persistent symptom group ($n = 30$)	df	Test parameter	*p*
Age (mean ± SD)	49.5 ± 9.7	46.5 ± 10.0	58	$t = -1.15$	0. 25
Percentage female, *N* (%)	14 (47%)	19 (63%)	1	$\chi^2 = 1.67$	0.19
Education, *N* (%)			2	$\chi^2 = 0.90$	0.64
Low	5 (17%)	8 (27%)			
Middle	12 (40%)	11 (37%)			
High	13 (43%)	11(37%)			
Full- or part-time work, *N* (%)	7 (23%)	12 (40%)	1	$\chi^2 = 1.93$	0.17
Married or living together (%)	15 (50%)	15(50%)	1	$\chi^2 = 0.00$	1.00
Treatment arm (%)			1	$\chi^2 = 0.07$	0.80
Experimental	14 (43%)	13 (47%)			
Pseudo-experimental	16 (57%)	17 (53%)			
HDRS total score (mean ± SD)	13.8 ± 5.3	14.1 ± 6.9	58	$t = 0.21$	0.83
Number of depressive episodes (mean ± SD)	8.9 ± 17.2	*9.9* ± 18.4	53	$t = 0.21$	0.84
Comorbid diagnosis			1	$\chi^2 = 1.1$	0.29
Present	10	14			
Absent	20	16			
Duration antidepressant pharmacotherapy prior to study entrance (%)				*Fisher's exact test*	**0.01***
New / switch (< 8 weeks)	9 (30%)	1 (3%)			
Maintenance (> 8 weeks)	21 (70%)	29 (97%)			
Number of daily diary assessments completed (mean ± SD)	15.3 ± 4.0	14.8 ± 3.8	58	$t = -0.43$	0.67
Number of available lagged observations (mean ± SD)	11.03 ± 3.1	11.6 ± 3.5	58	$t = 0.70$	0.49

Note: Statistics are *t*-test for continuous variables (means), and chi-square or Fisher's exact test for categorical (%) variables.
Bold values are significant *p*-values.
* Significant at $\alpha = 0.05$.

Table 2
Mean and standard deviations for the SCL-90-R depression subscale items for the major depression course subgroups.

SCL-90-R items depression subscale (mean ± SD)	Reduced symptom group (n = 30)	More persistent symptom group (n = 30)	p
Feeling low in energy or slowed down	2.39 ± 0.75	2.54 ± 0.84	0.30
Feeling of being trapped or caught	1.83 ± 0.63	2.08 ± 0.61	0.83
Loss of sexual interest or pleasure	1.98 ± 0.62	2.41 ± 0.62	0.99
Feeling lonely	1.82 ± 0.54	2.14 ± 0.64	0.26
Blaming yourself for things	1.93 ± 0.51	1.96 ± 0.59	0.30
Feeling blue	2.07 ± 0.62	2.48 ± 0.69	0.47
Worrying too much about things	2.57 ± 0.68	2.84 ± 0.67	0.90
Feeling no interest in things	2.09 ± 0.55	2.15 ± 0.58	0.72
Feeling hopeless about the future	2.07 ± 0.61	2.47 ± 0.75	0.10
Feeling everything is an effort	2.46 ± 0.64	2.75 ± 0.71	0.44
Feelings of worthlessness	1.88 ± 0.57	1.9 ± 0.56	0.91
SCL-90-R depression subscale score	2.01 ± 0.36	2.19 ± 0.39	0.56

Note: SD for each variable reflects the average SD per person within the respective group.

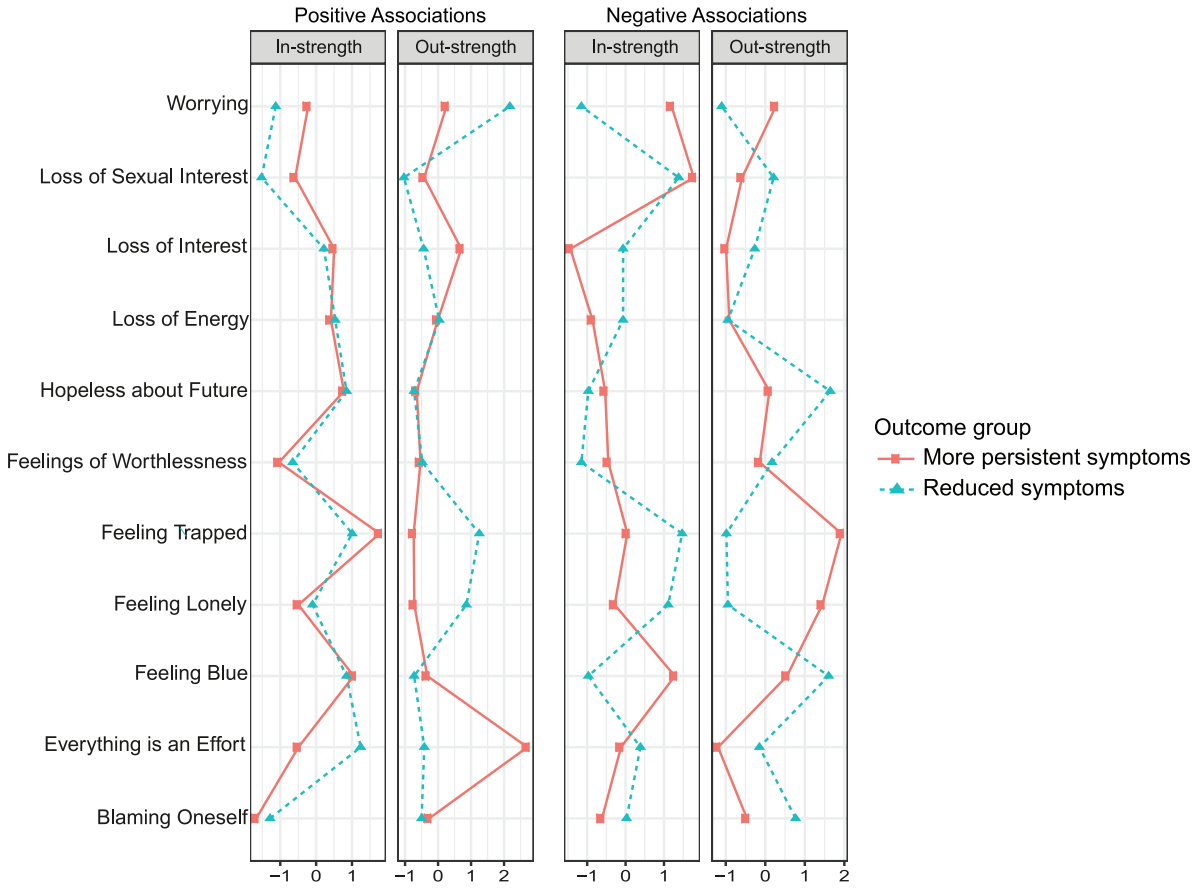

Fig. 1. Standardized centrality indices for each symptom, separately for positive associations (left) and negative associations (right), within the networks of the more persistent (solid red) and reduced group (dotted blue).

Autoregressive effects (the curved arrows) show the influence of the symptom on itself over time. After permutation testing, 11 of these edges were significantly different between the two groups at the uncorrected α level (indicated with an asterisk in Fig. 2).

4. Discussion

This is the first study that used intensive longitudinal assessments of daily depressive symptom intensity to investigate whether symptom dynamics differentiate between participants who experienced symptom reduction and participants with more persistent symptoms at six-month follow-up. This study does not support the hypothesis that vulnerability for more persistent psychopathology is associated with higher overall connectivity of dynamic symptom networks. However, exploratory findings of the current study suggest that patients with more persistent symptoms may differ in specific network characteristics: which symptoms are significantly associated with one another and which symptom has highest out-strength. This may suggest that identifying symptom dynamics is informative of individuals' capacity for recovery in addition to looking at the overall intensity of symptoms.

In the networks of individuals with different clinical outcomes, different symptoms seemed particularly important in predicting other symptoms the next day. Within the network of participants who experienced more persistent depressive symptoms six months after the daily diary assessments, 'feeling everything is an effort' (effort) had the highest out-strength, which may indicate that it is the most influential symptom. In the network of the reduced symptoms group, 'worrying too much about things' (worrying) had the highest out-strength, suggesting

Table 3
Significant positive out-strength differences major depression course subgroups.

SCL-90-R items depression subscale	Differences More persistent group vs Reduced group		Coefficients More persistent group	Reduced group
	Difference	*p*	*B*	*B*
Feeling low in energy or slowed down	−0.52	0.20	0.39	0.91
Feeling of being trapped or caught	−0.82	0.13	0.13	0.95
Loss of sexual interest or pleasure	0.33	0.38	0.47	0.14
Feeling lonely	−1.05	0.13	0.05	1.10
Blaming yourself for things	0.17	0.79	0.42	0.25
Feeling blue	0.64	0.09	0.77	0.12
Worrying too much about things	−1.19	0.33	0.85	2.04
Feeling no interest in things	0.53	0.44	1.08	0.58
Feeling hopeless about the future	−0.08	0.16	0.18	0.19
Feeling everything is an effort	1.65	0.02*	1.68	0.03
Feelings of worthlessness	−0.12	0.77	0.31	0.42

* significant at uncorrected α level

that this symptom played a key role in activating other symptoms in the network. However, when directly tested whether these symptoms played a different role in both groups, only out-strength of effort significantly differed between the two groups, indicating that this symptom played a more central role in the persistent than in the reduced symptoms group. This finding is not in line with the results of previous cross-sectional network studies, which showed a central role for loss of energy/fatigue and feeling guilty (van Borkulo et al., 2015) and concentration problems (Schweren et al., 2018) associated with MDD persistence. As this is a first attempt relating average dynamic within-person associations among depressive symptoms to clinical outcomes, current interpretations based on our results remain speculative. Nonetheless, if the results regarding effort are replicated, it may suggest that depressive symptoms may be less likely to change when 'effort' has an activating role in the network.

To what extent the difference in symptom reduction between the two groups is caused by the observed difference in network structure is unclear. This difference could be due to more individuals in the reduced group having switched or started antidepressant treatment in the eight weeks prior to study entry. However, we do not think that this confounds the results, since there is currently no evidence that antidepressant medication directly affects network structures of mental states (Snippe et al., 2017). Moreover, the networks may mainly reflect that one system is more likely to change than the other, independently of what caused symptom reduction or persistence. Whether networks explain why reductions occur, or mainly signal in whom they are more likely to occur needs further investigation.

Further research is also needed with regard to the connectivity-

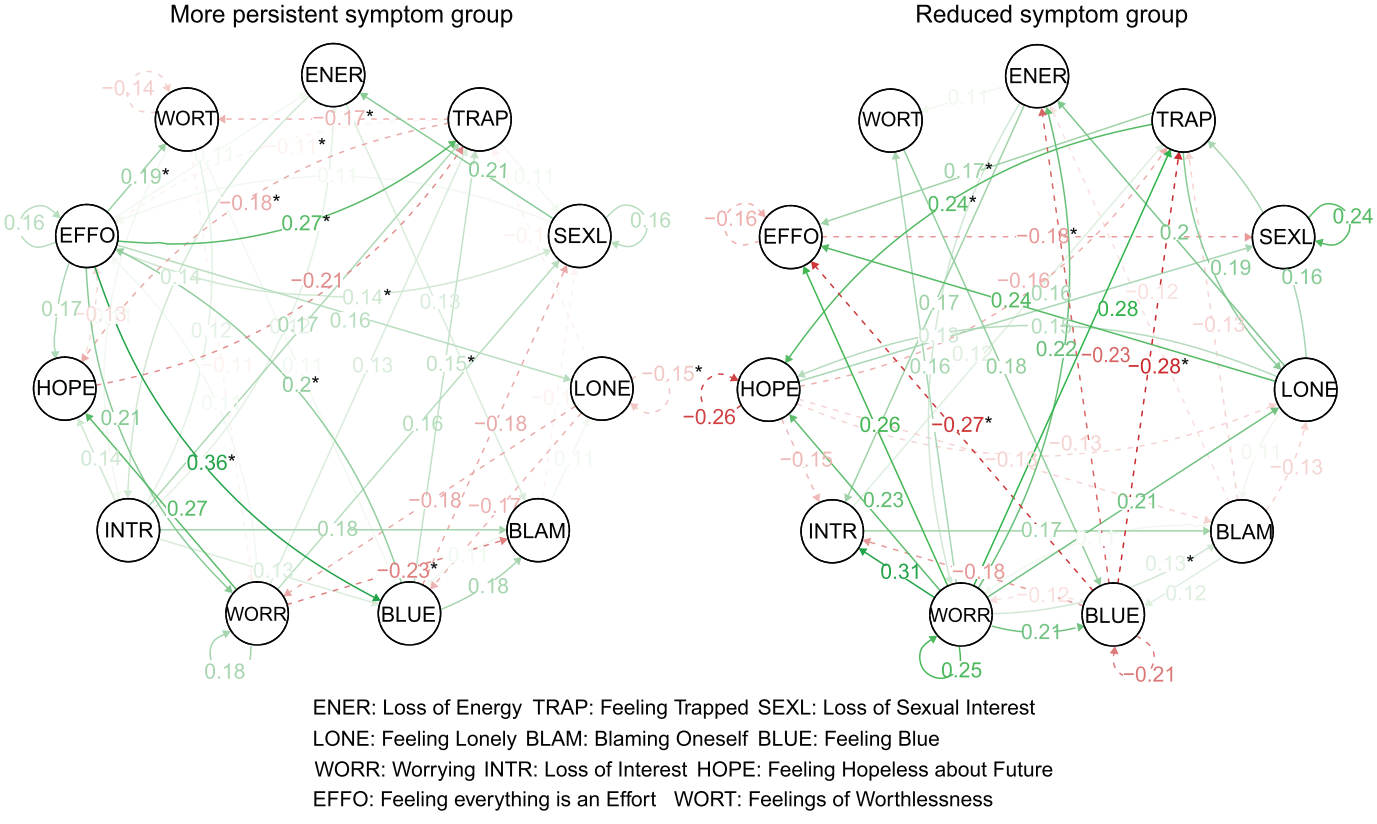

Fig. 2. Symptom networks for the more persistent- and reduced symptoms group. Networks include all edges with coefficients > 0.10. Strength of the edge coefficients were based on the estimates of the multilevel VAR models. Each edge represents the association between symptoms or the autoregressive effect over time (one lag is one day). Thicker/darker edges imply stronger associations. Green solid lines represent positive associations over time while red dotted lines show the negative associations. Edges that significantly differed between groups are marked with an asterisk (*) in each group's network.

vulnerability hypothesis in symptom networks. Networks with higher overall connectivity are considered to be more vulnerable networks (Cramer et al., 2016). We did not find support for this hypothesis in the context of day-to-day symptom dynamics; overall network connectivity did not differ between the more vulnerable group (more persistent symptoms) and the reduced symptom group. While our findings contrast with prior findings from simulations (Cramer et al., 2016) and empirical investigations (Pe et al., 2015; Wigman et al., 2015) that involved dynamic symptom assessments, some other recent studies likewise did not find an association between connectivity and increased vulnerability (Schweren et al., 2018; de Vos et al., 2017).

Several factors may have contributed to the observed differences between our study and previous research. A first difference between the current and previous studies is the timeframe under investigation. Symptom dynamics may render different findings than momentary affect dynamics investigated in experience sampling studies. Not all symptom-symptom associations do perhaps occur at the daily level, but may occur at a slower or faster rate. Hence, the specific timeframe in which various symptoms influence each other seems an important question that needs to be further addressed. Another contributing factor to the observed differences may be that linear trends from individuals' time-series were removed in the present study but not in these previous studies. This 'person-mean centering and detrending' method was needed in order to control for possible spurious correlations between lagged variables that may be induced by a change in symptoms over time (Rovine and Walls, 2006). This procedure does however not account for any non-linear changes in symptoms over time. Person-mean centering the predictor variables may lead to an underestimation of the autoregressive effects (Hamaker and Grasman, 2015), which may have contributed to the small number of significant autoregressive effects. However, it is unlikely that this would have differentially affected the two groups and thus would have led to consistent differences in connectivity. To account for potential changes in auto-correlations or cross-correlations over time (which is not possible in a regular vector autoregressive model), future studies (with sufficient time points) could employ novel models such as time-varying autoregressive (TV-AR, Bringmann et al., 2017) and vector-autoregressive (TV-VAR, Bringmann et al., 2018) models. With these models, various types of non-stationarity can be detected and accounted for, and change in the network structure itself could be considered (Bringmann et al., 2017).

As this is a first study with dynamic symptoms and long-term follow-up, we can draw no definite conclusions concerning the connectivity-vulnerability hypothesis. We believe that only through replication of varying designs we will be able to establish whether consistent differences exist in findings concerning the connectivity-vulnerability hypothesis, and whether these are related to type of the design (contemporaneous, cross-sectional or varying of lag-length) or level (symptom level or daily experience level). Moreover, we think that alternative hypotheses in addition to the connectivity-vulnerability hypothesis should also be considered in future research. For instance, one relevant hypothesis based on complex dynamical systems theory poses that large-scale synchronization of components, visible in higher overall connectivity, may signal increased likelihood of impending critical transitions (Chen et al., 2012; Schiepek et al., 2011; van de Leemput et al., 2014). Hence, according to this theory, we would expect an increase in network connectivity over time in the group with the largest upcoming change in symptoms, even if this concerns a transition towards lower symptom levels (i.e., the reduced symptom group). However, this hypothesis would need to be investigated in personalized network models using sufficient data points to examine whether connectivity increases within individuals over time (Bos and De Jonge, 2014).

The main strength of our study is its unique design, with both within-individual dynamic measures of MDD symptom change and a follow-up six months after the daily symptom assessments providing insight on individuals' long-term course of depressive symptoms. Secondly, we controlled for severity differences in our models and observed no differences in symptom intensities or variance between groups. This strengthens our findings as this may influence group differences in dynamic associations (Terluin et al., 2016). However, there were also some limitations. Dynamic associations between symptoms were modelled based on the associations during three consecutive days within each week of the intervention period, instead of using continuous observations over time. With this design, two lags were available each week (on average 11 per person in total). While adherence was rather high, a continuous range of observations would have been more optimal, as more data points would have been available for analysis and no gap would have existed between observations. With more observations available for analysis and more individuals per group, we would have had more power, which is necessary to be able to detect smaller group differences. This is a limitation of our study. Likewise, with more observations available, we could have obtained more reliable estimates of the individual differences in the associations by including random slopes in the models. We did not model random slopes in the current study because it would have resulted in over-fitting, which is a limitation of our study. Moreover, the low number of assessments makes it difficult to apply certain cross-validation procedures (i.e., those discussed in Bulteel et al., 2018) to assess the stability of the estimates, and is thus a limitation of the current study. The development of multilevel VAR bootstrap methods to gauge stability would strengthen future research. Another potential limitation is that we used a median split to assess individuals' long-term course of depressive symptoms. However, we do believe that the median split has resulted in two meaningfully different groups. Moreover, it allowed us to maximize power in both groups, which was advantageous for estimating symptom associations. A comparison of the individuals at the extreme ends of the HDRS change score spectrum as a robustness check also revealed that this did not affect our main results. Yet, we do think it is important to keep in mind that the current findings may not generalize to groups of individuals who have completely recovered. Moreover, patients in the current sample participated in the study in addition to completing their treatment-as-usual (pharmacotherapy), and the majority had experienced previous depressive episodes. Findings may thus not generalize to other patient populations.

A more general challenge to consider when interpreting these results is that we cannot be sure that the symptoms within the network are truly dynamically associated with each other. There is also the possibility that an underlying factor is responsible for their associations over time (Wichers et al., 2017). Furthermore, we observed negative associations in the network, which we included in the connectivity estimates. To what extent negative associations should be expected in dynamic symptom networks (also observed by de Vos et al., 2017, Lutz et al., 2018) and whether they have implications for the connectivity-vulnerability hypothesis needs to be further investigated. Lastly, it is uncertain whether the observed differences in dynamics exist for all participants within each group or merely for a subgroup of participants. Future research may collect multiple intensive assessments for each individual (>100 time points per person) to be able to derive personalized models to address these questions.

In short, the exploratory findings from our study suggest that the symptom 'feeling everything is an effort' more strongly predicts other symptoms the next day in a group of individuals with more persistent symptoms at follow-up. Yet, we did not find evidence that connectivity is higher in the symptom network of individuals with more persistent symptoms. The current findings illustrate that examining the role of a specific symptom in dynamic symptom-to-symptom interactions may provide additional insight concerning symptom persistence, which should be explored further in future research.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Acknowledgements

The authors thank all patients for participating and all collaborating mental health centres for their support in recruiting patients.

Funding

This work was supported by the Netherlands Organization for Scientific Research (NWO) (R.N.G., grant number 406.16.507), (M.W., Aspasia grant), (J.A.H., grant number 825.15.015); Dutch Health Research Council (ZON-MW) (grant numbers 171001002, 91501003); and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovative programme (M. W., grant number 681466). The tool with which momentary assessments were performed (the PsyMate) is developed under the auspices of the Maastricht University technology transfer office, partially supported by unrestricted grants from Servier and Janssen-Cilag, and by funding from the European Community's Seventh Framework Program (grant number HEALTH-F2-2009-241909).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2018.12.054.

References

Bagby, R.M., Ryder, A.G., Schuller, D.R., Marshall, M.B., 2004. The Hamilton depression rating scale: has the gold standard become a lead weight? Am. J. Psychiatry 161, 2163–2177.
Bolger, N., Laurenceau, J., 2013. Intensive Longitudinal Methods: an Introduction to Diary and Experience Sampling Research. Guilford Press, New York, NY.
Booij, S.H., Bos, E.H., de Jonge, P., Oldehinkel, A.J., 2016. The temporal dynamics of cortisol and affective states in depressed and non-depressed individuals. Psychoneuroendocrinology 69, 16–25.
Borsboom, D., Cramer, A.O.J., 2013. Network analysis: an integrative approach to the structure of psychopathology. Annu. Rev. Clin. Psychol. 9, 91–121.
Bos, E.H., De Jonge, P., 2014. "Critical slowing down in depression" is a great idea that still needs empirical proof. Proc. Natl. Acad. Sci. USA 111, E878.
Bos, E.H., Wanders, R.K., 2016. Group-level symptom networks in depression. JAMA Psychiatry 73, 411 411.
Bos, F.M., Snippe, E., de Vos, S., Hartmann, J.A., Simons, C.J.P., van der Krieke, L., et al., 2017. Can we jump from cross-sectional to dynamic interpretations of networks? Implications for the network perspective in psychiatry. Psychother. Psychosom. 86, 175–177.
Bringmann, L.F., Ferrer, E., Hamaker, E.L., Borsboom, D., Tuerlinckx, F., 2018. Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model. Multivar. Behav. Res. 53 (3), 293–314.
Bringmann, L.F., Hamaker, E.L., Vigo, D.E., Aubert, A., Borsboom, D., Tuerlinckx, F., 2017. Changing dynamics: Time-varying autoregressive models using generalized additive modeling. Psychol. Methods 22, 409–425.
Bringmann, L., Lemmens, L., Huibers, M., Borsboom, D., Tuerlinckx, F., 2015. Revealing the dynamic network structure of the Beck Depression Inventory-II. Psychol. Med. 45, 747–757.
Bringmann, L.F., Pe, M.L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., et al., 2016. Assessing temporal emotion dynamics using networks. Assessment 23, 425–435.
Bringmann, L.F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., et al., 2013. A network approach to psychopathology: new insights into clinical longitudinal data. PLoS one 8, e60188 e60188.
Bulteel, K., Mestdagh, M., Tuerlinckx, F., Ceulemans, E., 2018. VAR(1) based models do not always outpredict AR(1) models in typical psychological applications. Psychol. Methods 23, 740–756.
Chen, L., Liu, R., Liu, Z., Li, M., Aihara, K., 2012. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. Sci. Rep. 2, 342.
Conradi, H.J., Ormel, J., de Jonge, P., 2011. Presence of individual (residual) symptoms during depressive episodes and periods of remission: a 3-year prospective study. Psychol. Med. 41, 1165–1174.
Cramer, A.O.J., van Borkulo, C.D., Giltay, E.J., van, d.M., Kendler, K.S., Scheffer, M., Borsboom, D., 2016. Major depression as a complex dynamic system. PLoS One 11, e0167490.
Cramer, A.O.J., Waldorp, L.J., van der Maas, H.L.J., Borsboom, D., 2010. Comorbidity: a network perspective. Behav. Brain Sci. 33, 137–150.
Curran, P.J., Bauer, D.J., 2011. The disaggregation of within-person and between-person effects in longitudinal models of change. Annu. Rev. Psychol. 62, 583–619.
de Vos, S., Wardenaar, K.J., Bos, E.H., Wit, E.C., Bouwmans, M.E., de Jonge, P., 2017. An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. PLoS One 12, e0178586.
Derogatis, L.R., 1992. SCL-90-R: Administration, scoring & procedures manual-II for the (revised) version and other instruments of the psychopathology rating scale series. Clin. Psychom. Res. 1–16.
Epskamp, S., Cramer, A.O., Waldorp, L.J., Schmittmann, V.D., Borsboom, D., 2012. Qgraph: network visualizations of relationships in psychometric data. J. Stat. Softw. 48.
Fava, M., Rush, A.J., Trivedi, M.H., Nierenberg, A.A., Thase, M.E., Sackeim, H.A., et al., 2003. Background and rationale for the sequenced treatment alternatives to Relieve Depression (STAR* D) study. Psychiatr. Clin. North Am. 26, 457–494.
Fried, E.I., Nesse, R.M., 2015. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. BMC Med. 13, 72.
Fried, E.I., Nesse, R.M., 2014. The impact of individual depressive symptoms on impairment of psychosocial functioning. PLoS One 9, e90311.
Good, P.I., 2005. Permutation, Parametric and Bootstrap Tests of hypotheses: a Practical Guide to Resampling Methods For Testing Hypotheses, 3rd ed. Springer, New York, NY.
Hamaker, E.L., Grasman, R.P., 2015. To center or not to center? Investigating inertia with a multilevel autoregressive model. Front. Psychol. 5, 1492.
Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry. 23, 56–62.
Hartmann, J.A., Wichers, M., Menne-Lothmann, C., Kramer, I., Viechtbauer, W., Peeters, F., et al., 2015. Experience sampling-based personalized feedback and positive affect: a randomized controlled trial in depressed patients. PLoS One 10, e0128095.
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.
Kendler, K.S., Aggen, S.H., Neale, M.C., 2013. Evidence for multiple genetic factors underlying DSM-IV criteria for major depression. JAMA Psychiatry 70, 599–607.
Klippel, A., Viechtbauer, Reininghaus, W., Wigman, U., van Borkulo, J., MERGE, C., et al., 2017. The cascade of stress: a network approach to explore differential dynamics in populations varying in risk for psychosis. Schizophr. Bull sbx037.
Kramer, I., Simons, C.J., Hartmann, J.A., Menne-Lothmann, C., Viechtbauer, W., Peeters, F., et al., 2014. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. World Psychiatry 13, 68–77.
Lebo, M.A., Nesselroade, J.R., 1978. Intraindividual differences dimensions of mood change during pregnancy identified in five P-technique factor analyses. J. Res. Personal. 12, 205–224.
Lutz, W., Schwartz, B., Hofmann, S.G., Fisher, A.J., Husen, K., Rubel, J.A., 2018. Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: a methodological proof-of-concept study. Sci. Rep. 8, 7819.
Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks: generalizing degree and shortest paths. Soc. Netw. 32, 245–251.
Papakostas, G.I., 2009. Managing partial response or nonresponse: switching, augmentation, and combination strategies for major depressive disorder. J. Clin. Psychiatry 70 (Suppl 6), 16–25.
Pe, M.L., Kircanski, K., Thompson, R.J., Bringmann, L.F., Tuerlinckx, F., Mestdagh, M., et al., 2015. Emotion-network density in major depressive disorder. Clin. Psychol. Sci. 3, 292–300.
Core Team, R, 2017. R: A language and Environment For Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, pp. 2016.
Rovine, M.J., Walls, T.A., 2006. Multilevel autoregressive modeling of interindividual differences in the stability of a process. Eds. In: Walls, T.A., Schafer, J.L. (Eds.), Models For Intensive Longitudinal Data. Oxford University Press, New York, NY, US, pp. 124–147.
Schiepek, G., Fartacek, C., Sturm, J., Kralovec, K., Fartacek, R., Plöderl, M., 2011. Nonlinear dynamics: theoretical perspectives and application to suicidology. Suicide Life-Threat. Behav. 41, 661–675.
Schmittmann, V.D., Cramer, A.O., Waldorp, L.J., Epskamp, S., Kievit, R.A., Borsboom, D., 2013. Deconstructing the construct: a network perspective on psychological phenomena. New Ideas Psychol. 31, 43–53.
Schweren, L., van Borkulo, C.D., Fried, E., Goodyer, I.M., 2018. Assessment of symptom network density as a prognostic marker of treatment response in adolescent depression. JAMA Psychiatry 75, 98.
Snippe, E., Viechtbauer, W., Geschwind, N., Klippel, A., de Jonge, P., Wichers, M., 2017. The impact of treatments for depression on the dynamic network structure of mental states: Two randomized controlled trials. Sci. Rep. 7, 46523.
Stavrakakis, N., Roest, A.M., Verhulst, F., Ormel, J., de Jonge, P., Oldehinkel, A.J., 2013. Physical activity and onset of depression in adolescents: a prospective study in the general population cohort TRAILS. J. Psychiatr. Res. 47, 1304–1308.
Teasdale, J.D., 1988. Cognitive vulnerability to persistent depression. Cogn. Emot. 2, 247–274.
Terluin, B., de Boer, M.R., de Vet, Henrica, 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, e0155205.
Uher, R., Perlis, R.H., Henigsberg, N., Zobel, A., Rietschel, M., Mors, O., Hauser, J., et al., 2012. Depression symptom dimensions as predictors of antidepressant treatment

outcome: replicable evidence for interest-activity symptoms. Psychol. Med. 42, 967–980.
van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B.W.J.H., Waldorp, L.J., Schoevers, R.A., 2015. Association of symptom network structure with the course of longitudinal depression. JAMA Psychiatry 1219–1226.
van de Leemput, I.A., Wichers, M., Cramer, A.O.J., Borsboom, D., Tuerlinckx, F., Kuppens, P., et al., 2014. Critical slowing down as early warning for the onset and termination of depression. Procee. Natl. Acad. Sci. 111, 87–92.
van Gils, A., Burton, C., Bos, E.H., Janssens, K.A.M., Schoevers, R.A., Rosmalen, J.G.M., 2014. Individual variation in temporal relationships between stress and functional somatic symptoms. J. Psychosom. Res. 77, 34–39.
von Neumann, J., Kent, R.H., Bellinson, H.R., Hart, B.I., 1941. The mean square successive difference. Ann. Math. Stat. 12, 153–162.
Wichers, M., Wigman, J.T., Bringmann, L.F., de Jonge, P., 2017. Mental disorders as networks: some cautionary reflections on a promising approach. Soc. Psychiatry Psychiatr. Epidemiol. 1–3.
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.
Wigman, J., Van Os, J., Borsboom, D., Wardenaar, K., Epskamp, S., Klippel, A., et al., 2015. Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. Psychol. Med. 45, 2375–2387.
Zimmerman, M., Ellison, W., Young, D., Chelminski, I., Dalrymple, K., 2015. How many different ways do patients meet the diagnostic criteria for major depressive disorder? Compr. Psychiatry 56, 29–34.