Symptom network models in depression research

van Borkulo, Claudia Debora

CHAPTER 6

ASSOCIATION OF SYMPTOM NETWORK STRUCTURE WITH THE COURSE OF DEPRESSION

Adapted from:

Van Borkulo, C. D., Boschloo, L., Borsboom, D., Penninx, B.W.J.H., & Schoevers, R.A. (2015). Association of symptom network structure and the course of depression. *JAMA Psychiatry, 72* (12), 1219-1226.

Major depressive disorder (MDD) is a heterogeneous condition in terms of symptoms, course, and underlying disease mechanisms. Current classifications do not adequately address this complexity. In novel network approaches to psychopathology, psychiatric disorders are conceptualized as complex dynamic systems of mutually interacting symptoms. This perspective implies that a more densely connected network of symptoms is indicative of a poorer prognosis, but, to date, no previous study has examined whether network structure is indeed associated with the course of MDD. In this study, we examined whether the baseline network structure of MDD symptoms is associated with the course of MDD.

In this prospective study, in which remittent and persistent MDD was defined on the basis of a follow-up assessment after 2 years, 515 patients from the Netherlands Study of Depression and Anxiety with past-year MDD (established with the Composite International Diagnostic Interview) and at least moderate depressive symptoms (assessed with the Inventory of Depressive Symptomatology [IDS]) at baseline were studied. Baseline starting and ending dates were September 1, 2004, through February 28, 2007. Follow-up starting and ending dates were September 1, 2006, through February 28, 2009. Analysis was conducted August 2015. The MDD was considered persistent if patients had at least moderate depressive symptoms (IDS) at 2-year follow-up; otherwise, the MDD was considered remitted.

Sparse network structures of baseline MDD symptoms assessed via IDS were computed. Global and local connectivity of network structures were compared across persisters and remitters using a permutation test. Among the 515 patients, 335 (65.1%) were female, mean (*SD*) age was 40.9 (12.1) years, and 253 (49.1%) had persistent MDD at 2-year follow-up. Persisters ($n = 253$) had a higher baseline IDS sum score than remitters ($n = 262$) (mean [*SD*] score, 40.2 [8.9] vs 35.1 [7.1]; the test statistic for the difference in IDS sum score was 22 027; $P < .001$). The test statistic for the difference in network connectivity was 1.79 ($P = .01$) for the original data, 1.55 for data matched on IDS sum score ($P = .04$), and 1.65 for partialed out data ($P = .02$). At the symptom level, fatigue or loss of energy and feeling guilty had the largest difference in importance in persisters′ network compared with that of remitters (Cohen′s $d = 1.13$ and 1.18, respectively).

This study reports that symptom networks of patients with MDD are related to the course of MDD: persisters exhibited a more densely connected network at

baseline than remitters. More pronounced associations between symptoms may be an important determinant of persistence in MDD.

6.1 Introduction

Although major depressive disorder (MDD) has been intensively investigated in various scientific fields (eg, in genetic, biological, and clinical research), impairment has barely decreased for patients (Wichers, 2014). In addition, the large differences across patients with MDD in symptoms, disease origin, and treatment response are still not well understood. This limited extent of scientific progress may be related to the fundamental issue of what psychiatric disorders actually are (Borsboom, 2008; Insel et al., 2010; Kendler et al., 2011). Depressive symptoms have traditionally been assumed to arise from a common cause, analogous to classic physical disease models. However, psychometric assumptions underlying the common cause model may not be justified when studying psychopathology (Fried, 2015). This model, for example, implies that symptoms are psychometrically interchangeable (Borsboom, 2008; Cramer et al., 2010), and, consequently, summing symptoms to establish an MDD diagnosis, as in current classification systems, would be an efficient way of reducing measurement error (Lord, Novick, & Birnbaum, 1968). Rather than measurement error, the overt heterogeneity in symptom patterns for MDD appears to be a very real phenomenon (L. Chen, Eaton, Gallo, & Nestadt, 2000; Fried, Nesse, Zivin, Guille, & Sen, 2014; Fried & Nesse, 2014; Holtzheimer & Mayberg, 2011; Østergaard, Jensen, & Bech, 2011): MDD symptoms are associated with different risk factors (Fried, Nesse, et al., 2014), different patterns of comorbidity (Lux & Kendler, 2010), and different levels of impairment (Fried & Nesse, 2014). These findings suggest that the assumption of interchangeability of symptoms is violated; therefore, different perspectives have been pursued to explain the heterogeneity of MDD (Lamers et al., 2013; Van Loo, De Jonge, Romeijn, Kessler, & Schoevers, 2012; Vogelzangs et al., 2012).

One recently proposed alternative is based on network models, in which disorders are conceptualized as complex dynamic systems of interacting symptoms (Boschloo et al., 2015; Cramer, van der Sluis, et al., 2012; Cramer et al., 2010; Kendler et al., 2011). This implies, for instance, that a person may experience sadness after a causal chain of feelings and emotions triggered by a stressful life event: insomnia leads to concentration problems to feeling worthless to feeling

sad to insomnia. Thus, in the network view, feedback loops may lead to circles of symptom coevolution, which can ultimately culminate in full-blown MDD. Support for this theoretical framework has come from, for example, intraindividual analyses revealing interactions among different mood states, in accordance with the idea that these form network structures (aan het Rot et al., 2012; Bringmann et al., 2013; Wichers, 2014). In addition, clinical experts view psychopathology as a system of causal relations where some symptoms play a more central role than others (Borsboom & Cramer, 2013; Kim & Ahn, 2002). An advantage of the network approach is that it, in contrast to the traditional common cause model, naturally accommodates the unique role of individual symptoms and their differences in risk factors and consequences (Cramer, Borsboom, et al., 2012; Fried, Nesse, et al., 2014; M. C. Keller, Neale, & Kendler, 2007; M. C. Keller & Nesse, 2005, 2006). This perspective accords well with recent advances in medicine and biology that indicate that physical diseases can be similarly analyzed as complex networks of factors that can contribute to the disease (Barabási, 2011; Schadt & Björkegren, 2012).

According to network approaches, more strongly connected networks will feature stronger feedback among their symptoms and may thus be related to a higher level of vulnerability to MDD and less positive prospects for recovery from MDD. If this is correct, we should expect symptoms to be more strongly connected in groups that have worse prognosis. This hypothesis may be investigated by examining patterns of symptom co-occurrence across cases, which can be used to construct an estimate of the (undirected) symptom network (the so-called Markov random field; Kindermann & Snell, 1980; Lauritzen, 1996). Assuming that individuals′ response patterns are realizations of a relatively homogeneous network structure, a stronger connection between 2 symptoms in the Markov random field indicates that symptoms tend to align their states more strongly while controlling for the value of the other variables in the network. This alignment may arise from a variety of causal and homeostatic mechanisms, which may be directional or bidirectional, so that connections in the Markov random field network can be viewed as a causal skeleton that encodes the existence but not the direction of putative causal relations in the population.

This study is the first, to our knowledge, to examine group-level differences in baseline network connectivity between patients with persistent vs remitted MDD at 2-year follow-up. Overall network connectivity is compared using the recently

developed Network Comparison Test (NCT; Van Borkulo, Epskamp, & Milner, 2016). In addition, local connectivity of individual symptoms in the networks is compared using 4 centrality measures (node strength, closeness, betweenness, and eigenvector centrality; Barrat et al., 2004; Boccaletti et al., 2006; Bonacich, 1987; Opsahl et al., 2010). Because centrality measures reveal how well connected each symptom is, they may identify symptoms that play an important role in the prognosis of MDD and thus suggest valuable targets for treatment.

6.2 Methods

6.2.1 Study Sample

Participants were selected from the Netherlands Study of Depression and Anxiety, an ongoing longitudinal cohort study designed to examine the long-term course and consequences of depressive and anxiety disorders in the adult population (aged 18-65 years; Penninx et al., 2008). Participants were recruited from the community (564 [18.9%]), general practice (1610 [54.0%]), and secondary mental health care (807 [27.1%]). Baseline starting and ending dates were September 1, 2004, through February 28, 2007. Follow-up starting and ending dates were September 1, 2006, through February 28, 2009. Baseline assessment included 2981 participants, consisting of people with current or a history of depressive and/or anxiety disorders, and a healthy control group. The medical ethics boards of the participating centers approved the study, and all participants signed written informed consent.

6.2.2 Persistence of MDD at Follow-up

We selected 585 participants with past-year MDD and at least moderate depressive symptoms at baseline. An MDD diagnosis (*DSM-IV-TR*) was assessed using the Composite Interview Diagnostic Instrument (CIDI; Wittchen, 1994). Severity of depressive symptoms in the week before baseline was measured with the 30-item, self-report Inventory of Depressive Symptomatology (IDS Rush et al., 1996) and was considered moderate for scores exceeding 25 (standard cut-off point; Rush et al., 1996). Persistence of MDD was defined as having at least moderate depressive symptoms (IDS score >25) at 2-year follow-up. The number of patients

with a past 6-month diagnosis at baseline (241 [95.3%] vs 247 [94.3%], $\chi_1 = 0.091$, $P = .77$) or a past month diagnosis at baseline (204 [80.6%] vs 195 [74.4%], $\chi_1 = 2.495$, $P = .11$) was comparable in persisters vs remitters. Seventy patients (12.0%) had missing data at follow-up and were excluded for further analyses. Included patients ($n = 515$) had lower IDS sum scores at baseline than excluded patients (mean [*SD*], 37.6 [8.4] vs 39.8 [S8.2]; $W = 21109$; $P = .02$), whereas sex (335 [65.1%] female vs 180 [71.4%], $\chi_1 = 0.85$, $P = .36$) and age (mean [*SD*], 40.9 [12.1] vs 43.6 [11.5]; $t_{90.98} = 1.85$; $P = .07$) were not related to inclusion.

6.2.3 Baseline *DSM-IV* Symptoms of MDD

Nine *DSM-IV-TR* criteria of MDD (American Psychiatric Association, 2000) were assessed at baseline with separate items of the IDS (Table 6.1) scored from 0 to 3. We disaggregated criteria where possible. As such, the criteria change in sleep and change in activity were disaggregated into an increase or a decrease. Criterion change in weight/appetite was retained as an aggregated symptom; participants were instructed to report either decreased or increased appetite, leading to perfectly negatively correlated variables. Because these associations are inherently different in nature (logical) than other associations (potentially causal), we did not include them in the network. The criteria change in weight/appetite and insomnia were therefore composed from multiple items by computing the mean score.

6.2.4 Statistical Analysis

6.2.4.1 General Differences

A Wilcoxon rank sum test for ordinal data was performed to test differences in baseline IDS sum scores and item scores of persisters and remitters. The significance level for all analyses was $\alpha = .05$.

6.2.4.2 Network Estimation

Network structures of baseline MDD symptoms were estimated separately for persisters and remitters using ℓ_1-regularized partial correlations among symptoms (Friedman et al., 2008; Tibshirani, 1996). Partial rather than zero-order correlations are used because, assuming that depressive symptoms arise from a limited set of

TABLE 6.1. Mapping of Items of the IDS to DSM-IV Criteria. Abbreviation: IDS, Inventory of Depressive Symptomatology.

DSM-IV Criterion		IDS Item	
Item	Description	Item	Description
A1	Depressed mood	5	Feeling sad
A2	Loss of interest/pleasure	19	General interest
A3	Weight/appetite change	11	Decreased appetite
		12	Increased appetite
		13	Decreased weight
		14	Increased weight
A4-a	Insomnia	1	Falling asleep
		2	Sleep during the night
		3	Waking up too early
A4-b	Hypersomnia	4	Sleeping too much
A5-a	Psychomotor retardation	23	Feeling slowed down
A5-b	Psychomotor agitation	24	Feeling restless
A6	Fatigue or loss of energy	20	Energy level
A7	Guilt/worthlessness	16	View of myself
A8	Concentration	15	Concentration/decision making
A9	Suicidality	18	Thoughts of death or suicide

direct (pairwise) interactions among symptoms, observed correlations might have been indirect (spurious). In such cases, a partial correlation network is known to recover the causal skeleton of the network whereas a correlation network does not. ℓ_1-regularization is used to find the optimal balance between parsimony and goodness of fit of the network and to circumvent multiple testing problems that arise in conventional significance testing because a network of 11 variables would require 55 ($11 \times 10/2$) significance tests. If the data are indeed a realization of a sparse network of pairwise interactions, this procedure converges to the true network (Foygel & Drton, 2011). For completeness, however, networks based on Pearson correlations and partial correlations were also estimated. Model selection with ℓ_1-regularization is performed with the extended Bayesian information criterion (J. Chen & Chen, 2008). This procedure yields accurate network estimations (Foygel & Drton, 2010; Van Borkulo et al., 2014) and is implemented in R-package qgraph (Epskamp et al., 2012). The extension of the Bayesian information criterion encompasses a hyperparameter γ, which is assigned the number zero (see Section B.1 and Figure B.1 in Appendix B for the influence of γ on network estimation).

6.2.4.3 Differences in Overall Connectivity

The overall connectivity (or global strength) of the networks, defined as the weighted sum of the absolute connections (Barrat et al., 2004), is determined for persisters and remitters. Statistical assessment of the difference in overall connectivity between networks of both groups was performed using the NCT, which is implemented in the R-package NCT (Van Borkulo, Epskamp, & Milner, 2016). The NCT is a 2-tailed permutation test in which the difference between 2 groups (persisters and remitters) is calculated repeatedly (100 000 times) for randomly regrouped individuals. This results in a distribution under the null hypothesis (assuming that both groups are equal), which can be used to test the observed difference between the empirical groups. The observed difference is considered significant at the threshold of .05.

6.2.4.4 Controlling for Baseline Severity

Two additional analyses were performed to control for baseline differences in severity. First, groups were matched on IDS sum score. Second, groups were matched by regressing (or partialing) out general level of functioning as an indicator of severity — measured by the World Health Organization Disability Assessment Schedule II (WHODAS II; Chwastiak & Von Korff, 2003). For more detailed information on these analyses and a more general discussion on severity as a confounder, see Section B.2, Figure B.2 and B.3 (Appendix B).

6.2.4.5 Differences in Local Connectivity

To reveal which symptoms play an important role in activating (or being activated by) other symptoms, those that occupy critical positions in the network have to be identified. Differences in importance of specific symptoms may be quantified by computing the 4 best-known local (ie, node specific) centrality measures: node strength, closeness, betweenness, and eigenvector centrality (Barrat et al., 2004; Boccaletti et al., 2006; Bonacich, 1987; Opsahl et al., 2010). Node strength measures the weighted number of connections of a focal node and thereby the degree to which that node is involved in the network (Barrat et al., 2004). This measure, however, only considers the local structure of the focal node (Opsahl et al., 2010). Closeness also takes the global structure of the network into account because it measures how close the focal node is to other nodes; it is inversely

proportional to the mean shortest distance to all other nodes (Boccaletti et al., 2006). Betweenness measures the degree to which the central node acts as a bridge that connects different parts of the network and may reflect the degree to which the node can assert control over information flow through the network (Boccaletti et al., 2006; Opsahl et al., 2010). Eigenvector centrality measures the degree to which a node is connected to other central nodes; it is proportional to the sum of centralities of nodes connected to the focal node (Bonacich, 1987).

6.2.4.6 Centrality Analyses

Networks were analyzed with $\gamma = 0$. Stability analyses were performed to investigate the influence of the value of γ on local centrality measures. Centralities were most stable and networks were similar with $\gamma = 0$ and 0.1, confirming that $\gamma = 0$ is the optimal choice (Figure B.4 in Appendix B). Statistical analyses were performed using R package, version 3.0.2 (R Development Core Team, 2011). To determine which symptoms differentiate most among the networks, effect sizes for differences in mean centrality were calculated (see Section B.4 in the Appendix for an explanation on how effect sizes were calculated).

6.3 Results

6.3.1 General Differences

In our sample of 515 patients, 335 (65.1%) were female, and mean (*SD*) age was 40.9 (12.1) years. In total, 253 patients (49.1%) had persistent MDD at 2-year follow-up. Persisters had a higher baseline IDS sum score than remitters (mean [*SD*], 40.2 [8.9] vs 35.1 [7.1]; the test statistic for the difference in IDS sum score was 22 027; $P < .001$). Persisters had higher scores than remitters on depressed mood, loss of interest, insomnia, psychomotor retardation, fatigue or loss of energy, concentration/decision making, and suicidality (Table 6.2). After matching on severity was performed, only hypersomnia and weight/appetite change differed significantly (see Table B.3 in Appendix B).

TABLE 6.2. Analysis of Item Scores of Persisters and Remitters. [a]The test statistic from the Wilcoxon rank sum test.

Symptom (Abbreviation)	Mean (SD)			
	Persisters (n = 253)	Remitters (n = 262)	Statistic[a]	*P* Value
Depressed mood (dep)	1.85 (0.75)	1.53 (0.72)	25 446	<.001
Loss of interest or pleasure (int)	1.38 (0.71)	1.12 (0.61)	26 493	<.001
Weight/appetite change (wap)	1.16 (0.79)	1.24 (0.79)	34 990	.27
Insomnia (ins)	1.39 (0.81)	1.15 (0.71)	27 506	.001
Hypersomnia (hyp)	0.68 (0.87)	0.79 (0.88)	35 646	.11
Psychomotor agitation (agi)	1.30 (0.85)	1.23 (0.90)	31 683	.36
Psychomotor retardation (ret)	1.26 (0.94)	0.89 (0.90)	25 864	<.001
Fatigue or loss of energy (ene)	1.89 (0.76)	1.62 (0.70)	26 568	<.001
Feeling guilty (gui)	1.89 (1.12)	1.78 (1.15)	31 448	.28
Concentration/decision making (con)	1.73 (0.77)	1.47 (0.76)	27 039	<.001
Suicidality (sui)	0.99 (0.82)	0.82 (0.85)	29 236	.01

6.3.2 Differences in Overall Connectivity

The network of persisters was more strongly connected than that of the remitters (Figure 6.3.2, left panels). Additional analyses to control for differences in baseline severity revealed that differences in connectivity were still present after matching on depression severity (IDS sum score) and after partialing out general functioning (WHODAS II; Figure 6.3.2, middle and right panels). The NCT confirmed that differences in connectivity were statistically significant for all analyses. The test statistic for the difference in network connectivity was 1.79 ($P = .01$) for the original data, 1.55 for data matched on IDS sum score ($P = .04$), and 1.65 for WHODAS II partialed out data ($P = .02$). For results of NCT across the entire range of γ, see Table B.2 of the accompanying Appendix. Networks based on ordinary Pearson correlations and nonregularized partial correlations also yielded qualitatively similar results (see Figure B.5 in Appendix B) and other global connectivity measures (Table B.4 in Appendix B).

6.3.3 Differences in Local Connectivity

To investigate differences in local connectivity, we compared the networks of persisters and remitters on 4 centrality measures (Figure 6.3.3). Considering node strength (Figure 6.3.3), similar patterns were found. However, depressed mood,

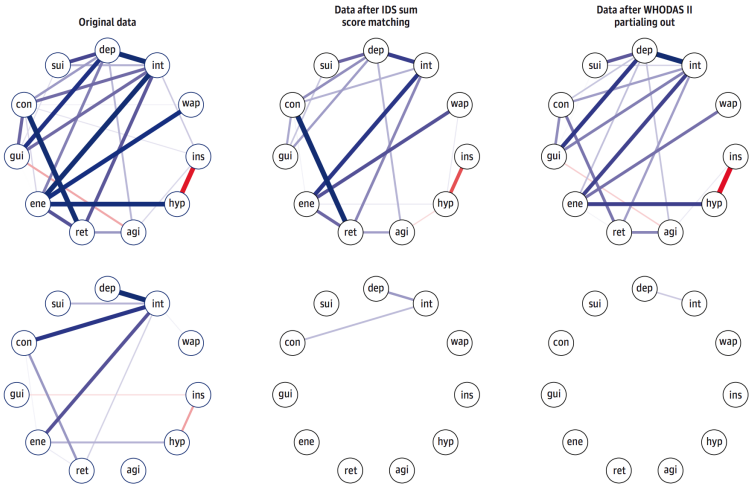

FIGURE 6.1. Network Structures of Persisters and Remitters Before and After Controlling for Severity. Network structures of persisters (n = 253) and remitters (n = 262) based on original data, data after matching on Inventory of Depressive Symptomatology (IDS) sum scores (n = 172 for both groups), and data after World Health Organization Disability Assessment Schedule II (WHODAS II) partialing out. Blue connections represent positive associations, whereas red connections represent negative associations. Thicker edges represent stronger associations (positive or negative). agi indicates psychomotor agitation; con, concentration/decision making; dep, depressed mood; ene, fatigue or loss of energy; gui, feeling guilty; hyp, hypersomnia; ins, insomnia; int, loss of interest or pleasure; ret, psychomotor retardation; sui, suicidality; wap, weight/appetite change.

fatigue or loss of energy, and feeling guilty had relatively higher values in the persisters′ network than in the remitters′ network. The pattern of closeness is also similar in both networks, but persisters had relatively higher values on feeling guilty, psychomotor retardation, and weight and/or appetite change compared with remitters (Figure 6.3.3). Regarding betweenness, fatigue or loss of energy had the highest value in the persisters′ network, whereas loss of interest had the highest value in the remitters′ network (Figure 6.3.3). The eigenvector centrality also follows a similar pattern in both networks. Symptom loss of interest features the highest value in both networks. The largest difference lies in the role of feeling

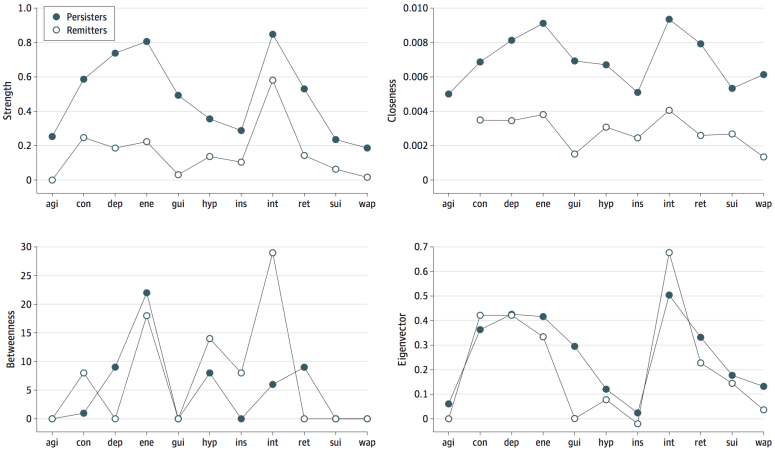

FIGURE 6.2. Centrality Measures. Four node centrality measures of persisters and remitters: strength, closeness, betweenness, and eigenvector centrality. agi indicates psychomotor agitation; con, concentration/decision making; dep, depressed mood; ene, fatigue or loss of energy; gui, feeling guilty; hyp, hypersomnia; ins, insomnia; int, loss of interest or pleasure; ret, psychomotor retardation; sui, suicidality; wap, weight/appetite change.

guilty; this symptom has a relatively high value in persisters′ network but has one of the lowest in remitters′ network.

Symptoms that have the largest difference in importance in persisters compared with remitters across all 4 centrality measures are fatigue or loss of energy and feeling guilty (Cohen′s d = 1.13 and 1.18, respectively; see Table B.3 for all effect sizes).

6.4 Discussion

This study is the first, to our knowledge, to find that the baseline MDD symptom network of patients with persistent MDD at follow-up was more densely connected than that of patients who recovered. With a focus on individual symptoms and their connections, fatigue or loss of energy and feeling guilty featured the largest increase in connectivity in the persisters′ network compared with the remitters′ network. Although baseline severity differed between the groups,

controlling for severity affirmed the main results; hence, it is highly unlikely that severity was a confounder in this study.

Our results could be interpreted in the light of other research, such as the recent findings on uncomplicated and complicated MDD (Parker, Paterson, & Hadzi-Pavlovic, 2015; Wakefield & Schmitz, 2014). Uncomplicated MDD is primarily characterized by normal intense distress reactions (eg, sadness and insomnia) and has positive prospects. Complicated MDD is not just a more severe condition but also features pathogenic reactions (eg, feeling worthless or suicidal ideation) and has an unfavorable course. In addition, our findings could be interpreted using the clinical staging model. Following other domains of medicine, this model is gaining popularity in psychiatry because it postulates that psychiatric disorders develop in consecutive stages: from subthreshold symptoms to chronic, persistent MDD (Hetrick et al., 2008; McGorry, Hickie, Yung, Pantelis, & Jackson, 2006). Indeed, there is empirical evidence that progression of psychopathologic disease is related to stronger and more viable interactions of mental states over time in a general population sample (Wigman et al., 2013). This more refined form of diagnosis can distinguish patients who seem misleadingly similar because they share the same diagnosis (Cosci & Fava, 2012) and seeks to determine whether different interventions may apply according to disease stage (Boschloo et al., 2014; Hetrick et al., 2008; McGorry, 2007).

Information on local connectivity may guide clinical therapy because important symptoms, identified by local centrality measures, could be specifically targeted using microinterventions. Because fatigue or loss of energy, feeling guilty, and psychomotor retardation were identified as important symptoms in the persisters′ network, these targets are particularly plausible for intervention. However, additional research is warranted to confirm this hypothesis. For example, it has yet to be established which centrality measure is clinically most relevant in identifying the importance of symptoms. In addition, directionality of the networks may be established where relevant. Although a central symptom is likely to have an influence on other nodes, it may be a more efficient target for intervention if associations with other symptoms are directed outward or are at least bidirectional.

The few studies that investigated centrality measures found largely similar central symptoms (ie, loss of interest, depressed mood, and fatigue or loss of energy; Bringmann et al., 2015; Cramer et al., 2010). However, these results were

based on different questionnaires — CIDI and the Beck Depression Inventory (Bringmann et al., 2015; Cramer et al., 2010) — and network types (dynamic; Bringmann et al., 2015), so the question of how these results relate to each other should be considered open. However, the general pattern emerging from research in this area is that the variables that function as core criteria in current diagnostic systems (depressed mood and loss of interest) are more central in networks of MDD cases defined in current psychiatric studies.

Strengths of this study are 2-fold. First, data come from a high-quality longitudinal study with well-characterized patients from different levels of health care and low levels of loss to follow-up, strengthening ecologic validity. Second, in contrast to previous studies (Bringmann et al., 2013; Wigman et al., 2015) that relied solely on perceived differences in networks to compare network structures of different groups, we were able to perform statistical comparison based on a newly developed test for differences.

Limitations of this study are as follows. First, presented networks are based on a between-subjects design. These networks may be representative of individuals as long as the groups are homogenous. Although the distinction between persisters and remitters has made groups already more homogenous, research is warranted on whether presented network structures are indeed generalizable to individual patients. This requires longitudinal within-person studies (ecologic momentary assessment or experience sampling; aan het Rot et al., 2012; Bouwmans et al., 2015; Wichers, 2014). In such a full prospective design, comparison of the individual network structures of patients who remit within 2 years with those of patients who do not may then reveal whether differences in network connectivity are also found at the level of the individual patient. Second, this study focused on the persistence of MDD, defined as having at least moderate depressive symptoms in the week before 2-year follow-up. Consequently, it is possible that a patient marked as a persister had experienced remission and recurrence during follow-up. However, the median percentage of time with depressive symptom was 96.0% for persisters (in contrast to 27.0% in remitters), indicating that most patients did not experience remission.