Supplementary Online Content

van Borkulo C, Boschloo Ly, Borsboom D, Penninx BWJH, Waldorp LJ, Schoevers RA. Association of symptom network structure with the course of longitudinal depression. *JAMA Psychiatry.* Published online November 11, 2015. doi:10.1001/jamapsychiatry.2015.2079.

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix 1. The Influence of γ on Network Estimation

In our network estimation procedure, hyperparameter γ was involved. The value of this hyperparameter can range from 0 to 1[e1]. When $\gamma = 0$, the network will be minimally sparse, resulting in a network with relatively many connections. Alternatively, when $\gamma = 1$, the network will be maximally sparse with relatively few (if any) connections. In the main analyses, we used $\gamma = 0$. The rationale of this choice is that with increasing values of γ, sensitivity of the network estimation procedure decreases; with increasing γ, networks tend to become similar (i.e., empty; see eFigure 1).

For completeness, we also performed NCT across the entire range of γ. Networks of original data differed significantly, whereas those of data controlled for severity did not (see eTable 2). While networks based on data controlled for severity were clearly different, NCT failed to confirm this. As can be seen in eFigure 1 (b and c), this was due to lack of sensitivity; with increasing γ, differences in connectivity were artificially (almost) absent.

eAppendix 2. Is Severity a Confound With Respect to Network Connectivity?

Although persisters and remitters were selected with the same baseline characteristics (past-year MDD and at least moderate depressive symptoms at baseline), persisters had significantly higher scores on most symptoms than remitters. This raises the question whether a difference in baseline severity is a confound with respect to network connectivity. In this section, we will discuss conceivable confounds.

If severity in itself were a confound with respect to network connectivity, then it should be the case that if a group has higher means on a set of variables, that group should also have a more connected network. However, the means of the variables do not play a role in the construction of the network; only the covariances do. In fact, one may standardize the variables without loss of generality: this will lead to exactly the same network, even though all variables in all groups would then have a mean of zero. Thus, mean level of the variables in itself cannot be a confound.

Although mean level of the variables in itself cannot be a confound, it is possible that something associated with severity, and which does influence network connectivity, plays the role of confound. An important candidate in this respect is variance. If, due to a methodological artifact, the variance in the individual item scores is lowered in the less severe group, so that it is associated with the mean levels of the variables in the network, then that could lead to a lower network connectivity due to restriction of range. A plausible mechanism that could produce this situation is the existence of floor and/or ceiling effects. If the group with low connectivity shows symptom score distributions with floor and/or ceiling effects while the group with high connectivity does not, the floor and/or ceiling effects might be a confound with respect to network connectivity.

Another possible mechanism that could lead to increased network connectivity in the more severe group is the presence of unmodelled latent variables. That is, if symptoms in the persister group were influenced more strongly by a latent variable (which would have to also be related to severity systematically), then the connectivity of persisters' network would be higher as a result. If, after controlling for such a latent variable, differences in connectivity disappear, the original difference was due to the latent variable. Conversely, if a difference in connectivity sustains, the latent variable cannot explain the difference.

eMethods 1. Analyses of Conceivable Confounds in Network Connectivity

Possible confounds in network connectivity are restriction of range and severity (discussed in eAppendix 1). These confounds were investigated with three analyses: (1) inspection of symptom score distributions and controlling for baseline severity by (2) matching on IDS score level and (3) regressing (or partialling) out an external measure of severity.

First, distributions of symptom scores were inspected with density plots. As shown in eFigure 4, symptom distributions are seemingly similar across groups; floor and/or ceiling effects, if

present, adhere to both groups. Hence, it is implausible that differences in restriction of range produced differences in network connectivity.

Second, groups were matched on IDS sum scores. Both groups are composed to contain the same number of patients with a sum score of 26, 27, and so on. This resulted in samples of 172 remitters and 172 persisters with the same mean IDS score (M = 36.6, SD = 7.1). Also, differences in individual symptom scores were reduced; only two symptoms (hypersomnia and weight/appetite change) differed significantly after matching (see eTable 1).

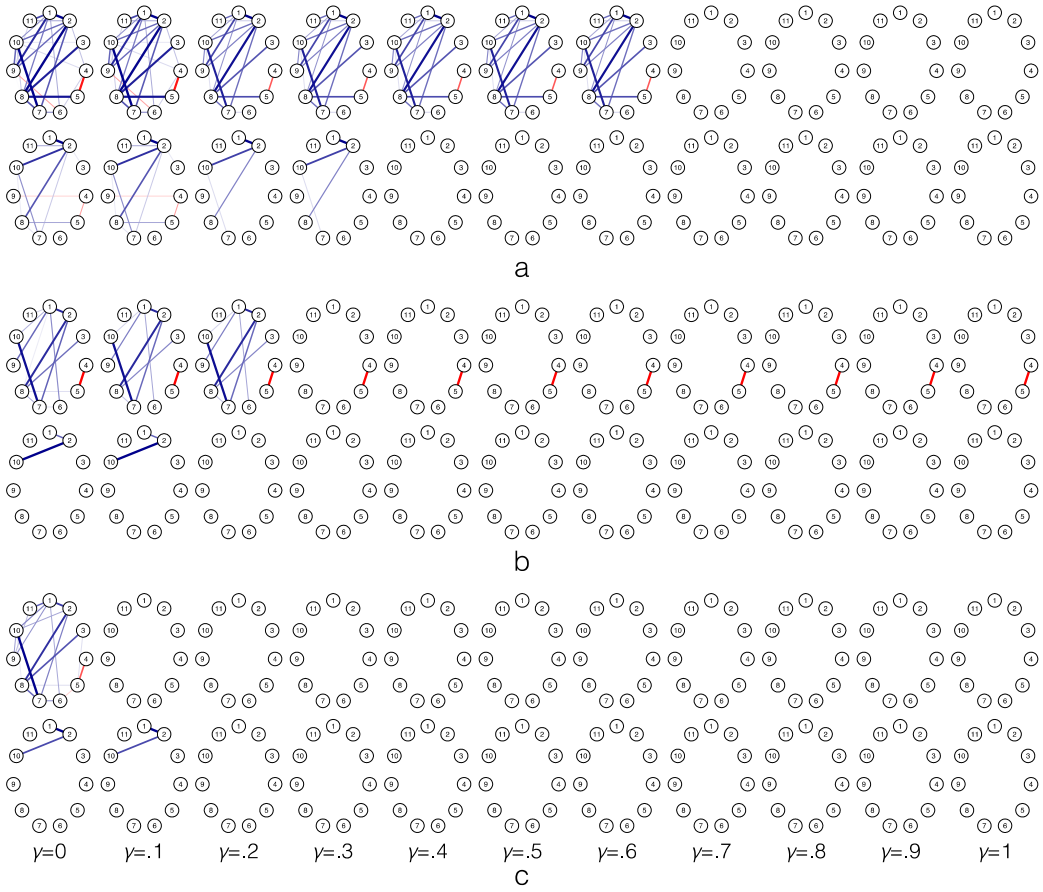
Third, groups were matched by controlling for (regressing or partialling out) general level of functioning as an indicator of severity with WHODAS. By regressing each depression variable separately on the WHODAS sum score, the variance of the depression variable that is not explained by severity is contained in the residuals. The residuals are used to determine the networks. Note that all patients (253 persisters and 262 remitters) can be retained with this strategy. It is perhaps useful to note that it is not possible to partial out the IDS sum score itself (instead of the WHODAS disability score); the IDS sum score is a deterministic function of the variables in the network and conditioning on this sum score leads to strong artificial negative correlations between variables (see eFigure 3).

eMethods 2. Quantifying Importance of Symptoms

Analysis of symptoms that distinguish most between persisters and remitters, were conducted by calculating the difference in centrality of symptoms between persisters and remitters. Symptoms with the highest effect size for increase in centrality in persisters, compared to remitters across all four measures, were considered most distinctive in persisters.

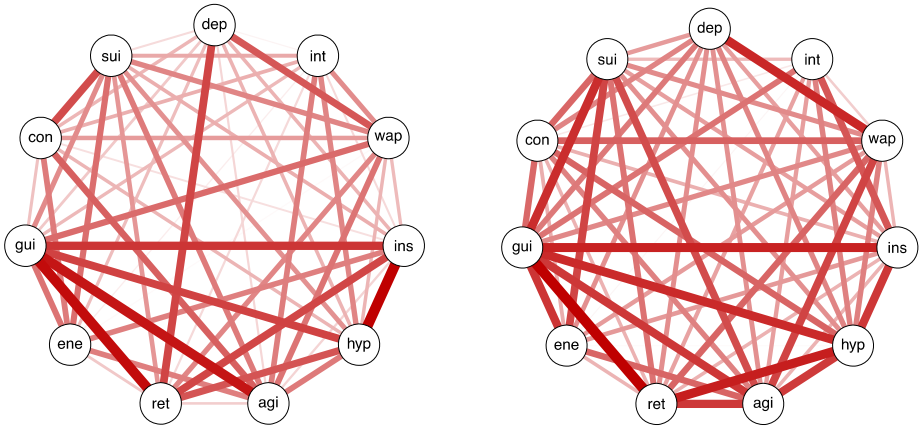
Effect sizes for difference in centrality were obtained with the bootstrap method[e2]. Data of persisters and remitters were resampled 1000 times, resulting in distributions of all symptom centrality measures in both groups. Effect size Cohen's d were calculated for difference in mean centrality measure of persisters and remitters[e3]. Since we considered four centrality measures, effect sizes were averaged (see eTable 4 for the resulting average effect sizes).

eFigure 1. Networks Across the Entire Range of *γ*



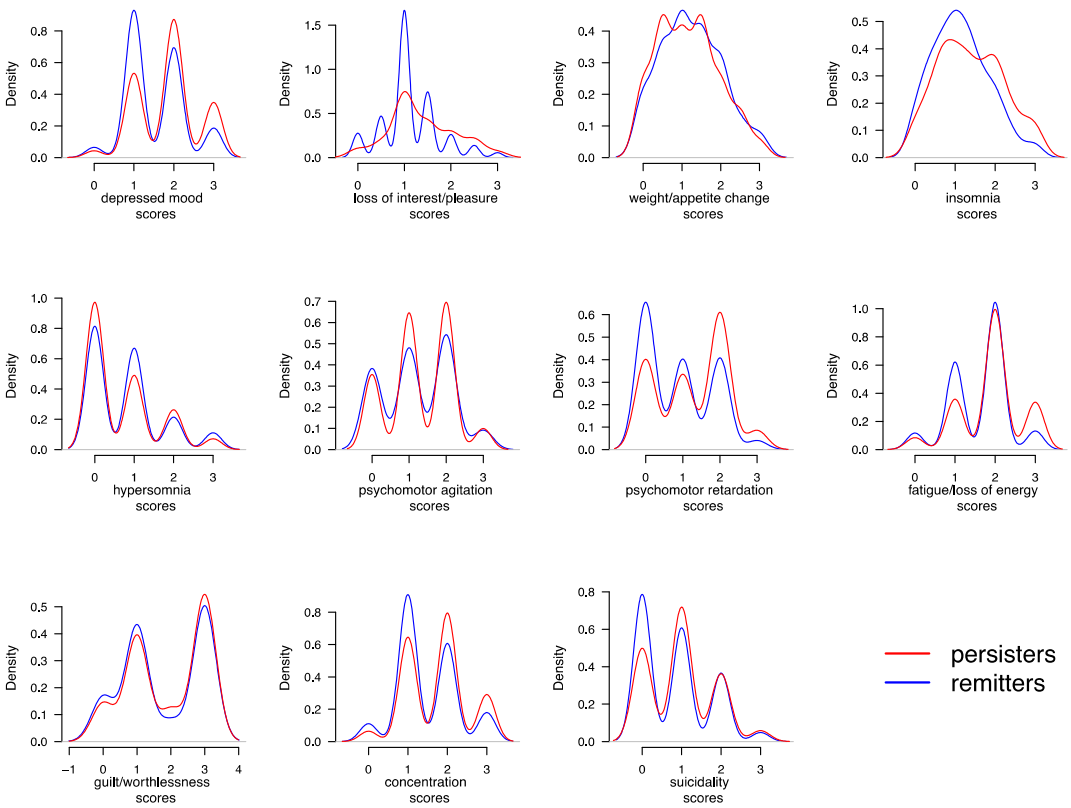
Network structures across the entire range of gamma (0, .1, .2, ..., 1) with and without controlling for severity. Networks of persisters (upper panel) and remitters (lower panel) with original data (a), with data after matching on IDS sum score (b), and data after partialling out general level of functioning (c). Symptom numbers: 1-depressed mood, 2-loss of interest/pleasure, 3-weight/appetite change, 4-insomnia, 5-hypersomnia, 6-psychomotor retardation, 7-psychomotor agitation, 8-fatigue/loss of energy, 9-guilty/worthlessness, 10-concentration, 11-suicidality.

eFigure 2. Networks When IDS Sum Score Is Partialed Out



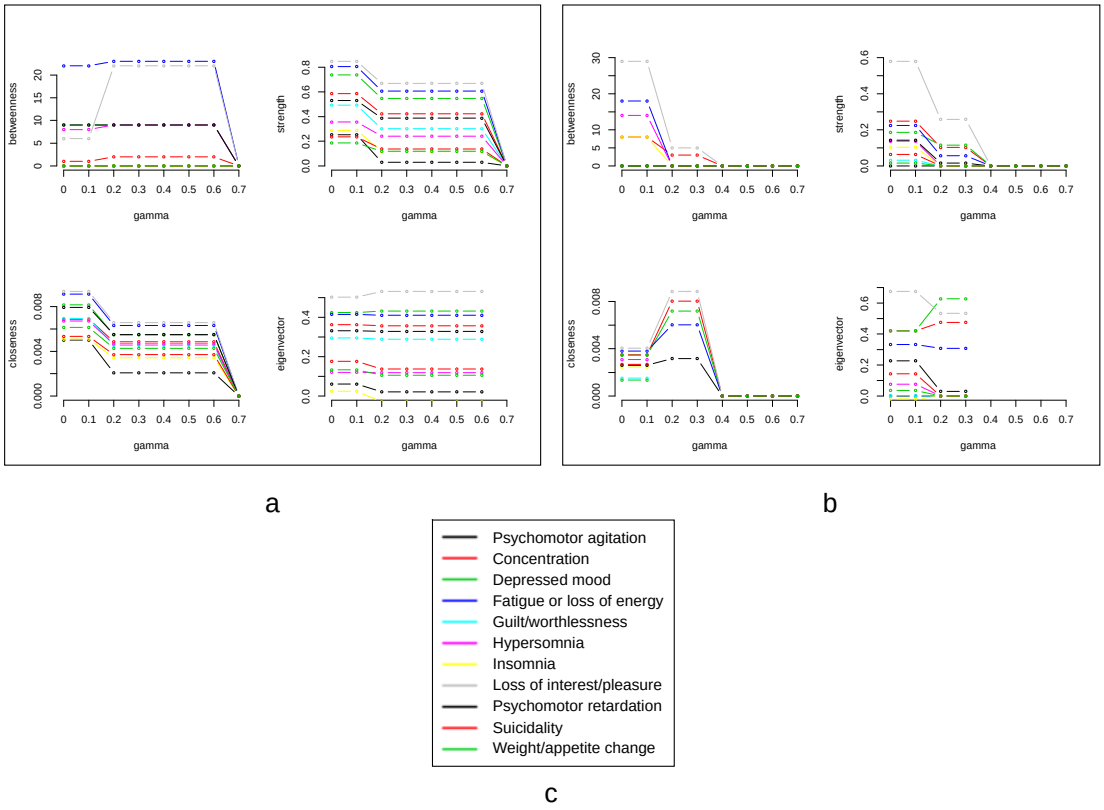
Persisters (left, n=253) and remitters (right, n=262). See eTable 1 for definitions of abbreviated terms.

eFigure 3. Symptom Score Distributions



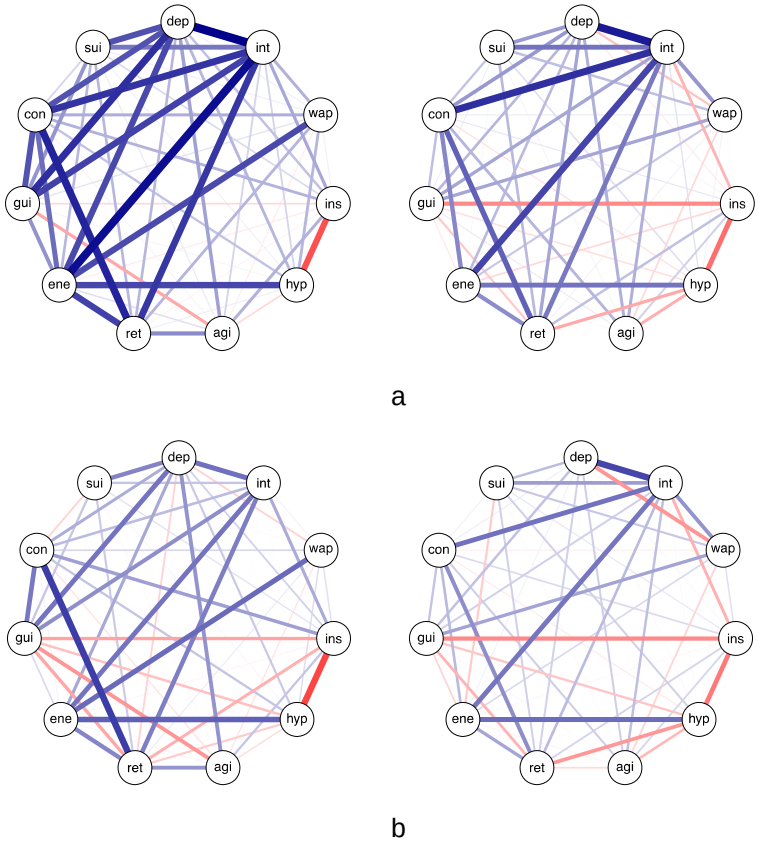
Density plots of symptom scores of persisters (red) and remitters (blue).

eFigure 4. Stability Analysis of Centrality Measures



Area Under Curves (AUC) for centrality measures betweenness (upper left panels), strength (upper right panels), closeness (lower left panels), and eigenvector centrality (lower right panels) for persisters (a) and remitters (b). For every value of gamma, the value of the centrality measure is displayed; each color represents a symptom in the network as indicated in the legend (c).

eFigure 5. Network Structures Based on Ordinary Analyses



Pearson's correlation (a) and unregularized partial correlation networks (b) of persisters (left) and remitters (right). Blue lines represent positive (partial) correlations, whereas red lines represent negative (partial) correlations. See eTable 1 for definitions of abbreviated terms.

eTable 1. Analysis of Item Scores After Matching on IDS Sum Scores

Symptom (abbreviation)	Mean (*sd*)		Statistic[a]	*P* Value
	Persisters (n=172)	Remitters (n=171)		
Depressed mood (dep)	1.67 (0.71)	1.63 (0.76)	14236	0.51
Loss of interest or pleasure (int)	1.21 (0.67)	1.21 (0.63)	14957	0.85
Weight/appetite change (wap)	1.06 (0.76)	1.28 (0.79)	17227	0.01
Insomnia (ins)	1.28 (0.79)	1.18 (0.71)	13790	0.27
Hypersomnia (hyp)	0.58 (0.78)	0.81 (0.93)	16714	0.02
Psychomotor agitation (agi)	1.25 (0.83)	1.28 (0.89)	15094	0.73
Psychomotor retardation (ret)	1.13 (0.95)	1.01 (0.90)	13773	0.24
Fatigue or loss of energy (ene)	1.74 (0.77)	1.67 (0.72)	14076	0.39
Feeling guilty (gui)	1.78 (1.16)	1.88 (1.16)	15488	0.42
Concentration/decision making (con)	1.60 (0.76)	1.53 (0.81)	14009	0.36
Suicidality (sui)	0.87 (0.79)	0.83 (0.84)	14304	0.57

[a] The test statistic from Wilcoxon's rank sum test.

eTable 2. Results of NCT With Networks Across Range of *γ*

	P Value	Statistic[a]
Original data	0.026	11.98
Matched on IDS sum score	0.375	1.69
WHODAS partialled out	0.055	13.24

[a] The test statistic from NCT (difference in network connectivity)

eTable 3. Additional Indicators for Weighted Network Density

Network metric	Persisters	Remitters
Average shortest path length	1.58	3.93
Transitivity	0.49	0.41
Diameter	0.30	0.41

The average shortest path length (or characteristic path length) is the average number of steps in the shortest path between all possible pairs of symptoms (the lower, the more densely connected[e4]). Transitivity is defined as the number of *triangles* (i.e., when symptoms A, B, and C are all connected, they form a triangle) proportional to the possible number of triangles (the higher, the more densely connected[e5]). Diameter is defined as the largest number of connections (the longest path length) between any two symptoms (the lower, the more densely connected[e6]).

eTable 4. Average effect sizes for difference between mean centrality in persisters and remitters

Symptoms	Average effect size	Average difference
Depressed mood	0.73	1.88
Loss of interest or pleasure	-0.33	4.44
Weight/appetite change	0.54	0.24
Insomnia	0.52	0.66
Hypersomnia	0.65	1.33
Psychomotor agitation	0.50	0.19
Psychomotor retardation	0.90	1.13
Fatigue or loss of energy	1.13	3.02
Feeling guilty	1.18	0.71
Concentration/decision making	0.25	1.21
Suicidality	0.39	0.21

eReferences

1. Foygel R, Drton M. Extended bayesian information criteria for gaussian graphical models. In *Adv Neural Inf Process Syst*. 2010;604-612.
2. Efron H, Tibshirani RJ. *An introduction to the bootstrap*. New York, Chapman & Hall; 1993.
3. Cohen J. *Statistical power analysis for the behavioral sciences (2nd ed)*. Hillsdale, NJ: Erlbaum; 1988.
4. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. *Nature*. 1998;393:440–442.
5. Wasserman S, Faust K. *Social network analysis: Methods and applications.* Cambridge university press; 1994.
6. Weisstein W. "Graph Diameter." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/GraphDiameter.html