Supplementary material

1.0 Supplementary methods

1.1. Network stability

The details of network stability checks were described earlier (Epskamp et al., 2016). In short, although several studies applied network models to psychopathology, fewer studies described possible statistical methods to test the stability of generated networks. 'Stability' refers to the degree of certainty at which the structure of a network and its interpretation would be comparable when estimating the same network in another (comparable) sample. Epskamp et al. (2016) described statistical methods to further explore this issue in networks. Applying such stability checks is needed to improve reproducibility of networks. With respect to the current study, given the wide variety of the symptoms, we wanted to test the robustness of our network. Therefore, we estimated the accuracy of edge-weights, by drawing bootstrapped confidence intervals (CIs) and performed the 'bootstrapped difference test' for edge-weights and centrality measure 'node strength'.

1.2. Bootstrapped confidence intervals

For both stability checks the R package 'bootnet' was used (Epskamp et al., 2016; R Core Team, 2016). In a generated network an edge between two symptoms has a given weight. However, we want to know whether an edge-weight has the same value when constructing the same network in another sample. We therefore utilized bootstrapping methods to construct a 95% CI around the regularized edge weight (termed as bootstrapped CI; Epskamp et al., 2016). A wide interval indicates that the stability is low and a narrow interval indicates that the stability is high.

When 'bootstrapping' (Efron, 1979), a new dataset is constructed by multiplying (here 1000 iterations) the sample to population size and sample from this bootstrapped population. Then this population is sampled again; called resampling. These new samples will show a certain variation with a normal distribution. The 95% - bootstrapped CI is the variation that covers 95% of the cases. Of note, the results of the bootstrapped CIs should not be interpreted as a significance tests, but only as a method to estimate the accuracy of the network, generated from this specific sample.

1.3. Bootstrapped difference test

We also applied bootstrapping to investigate whether edge-weights or centrality measures differ significantly from each other. A difference score between the bootstrap values of a certain edge-weight and bootstrap values of another edge-weight is calculated. When constructing a bootstrapped CI around this difference score, a null-hypothesis test can be performed. Moreover, when the constructed CI includes 'zero', the edge-weights do not differ significantly from each other. In the same way, we applied the bootstrapped difference test to the centrality measure 'node strength' (i.e., the sum of the weighted number and strength of all connections of a specific node to all other nodes).

Of note, when performing multiple significant tests, the problem of 'multiple testing' exists (i.e., due to performing many significance tests, significant results could be find purely by chance). To correct for these Type I Errors, the 'Bonferroni corrections' (Bland and Altman, 1995) are frequently applied. However, applying Bonferroni corrections to the network approach will be too conservative and result in very low significance levels, which appears unfavorable (Epskamp et al., 2016). However, there are currently no other methods of correcting for the 'multiple testing problem' or other stability checks. Therefore, the results of the bootstrapped difference test should be interpreted with caution.

2.0. Supplementary Figures

2.1. Figure S1: Stability of the network: bootstrapped confidence intervals

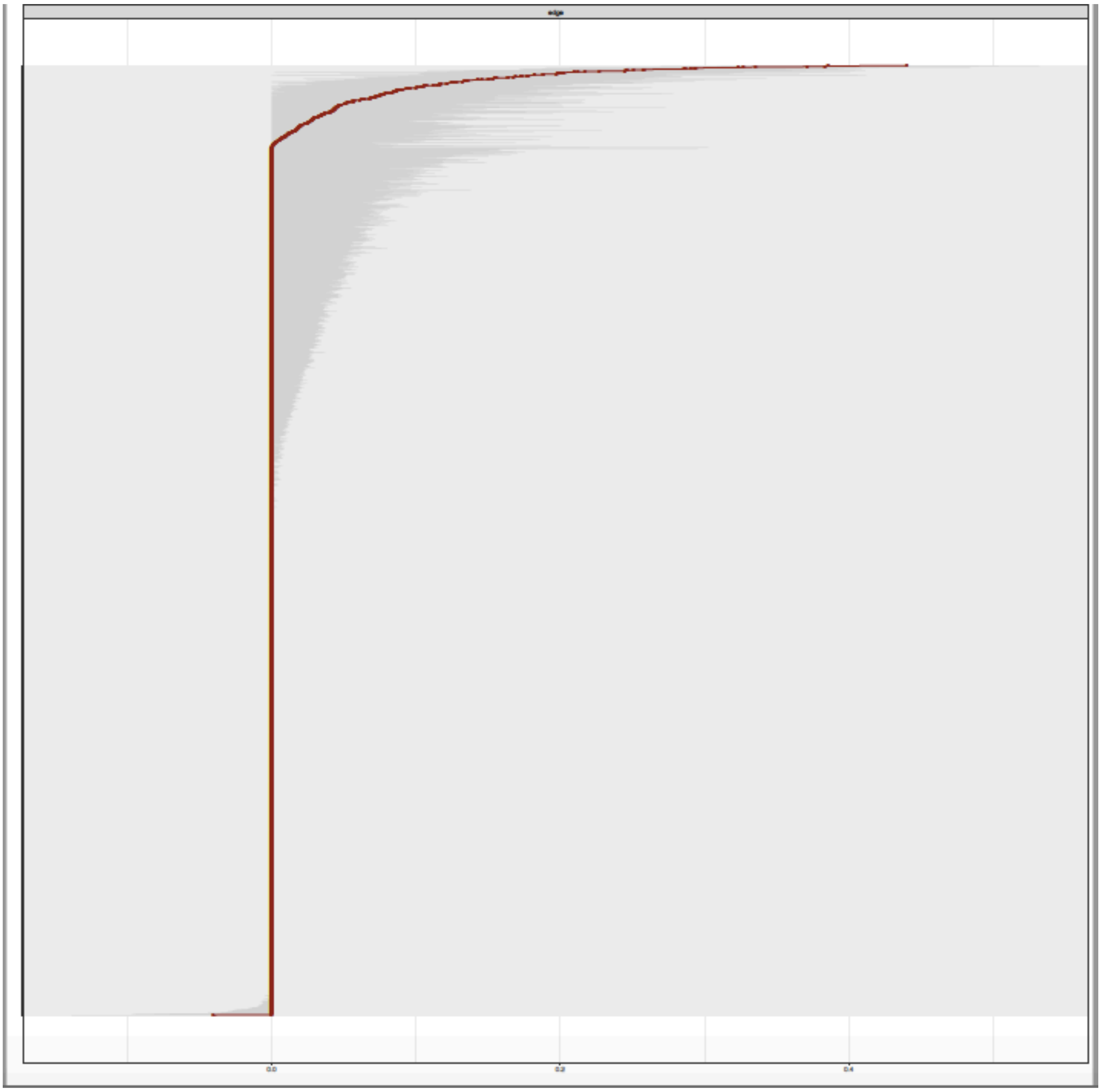

Figure showing the bootstrapped confidence intervals (CIs) of estimated edge-weights of all 79 symptoms (as assessed with the CASH), in grey. On the x-axis we show the distribution of the bootstrapped estimations of the CIs; on the y-axis all edges are shown, however, the labels of the edges were deleted to prevent cluttering. The edges are arranged such that the one with the highest edge-weight is at the top and the lowest edge-weight at the bottom. The red line shows the calculated edge-weight in our network, while the grey area – surrounding the red line – indicates the width of the bootstrapped CIs. Of note, overlapping CIs means that edge-weights likely do not significantly differ from one-another, in that case the order of the edges (i.e., the applied top-down ordering) should be interpreted with caution.
When interpreting this figure please note: i) many edges are consistently estimated as zero, ii) some edges are larger then zero, but the bootstrapped CIs contain zero and iii) some edges are larger than 0 with CIs not including zero. The fact that a large number of edges with a weight

larger than zero have large CIs with zero in them means that we should interpret the network with caution (i.e., the figure suggest that bootstrapped edge-weights might differ from each other, but this might not be significant).

2.2 Figure S2: Stability of the network: bootstrapped stability tests for node strength

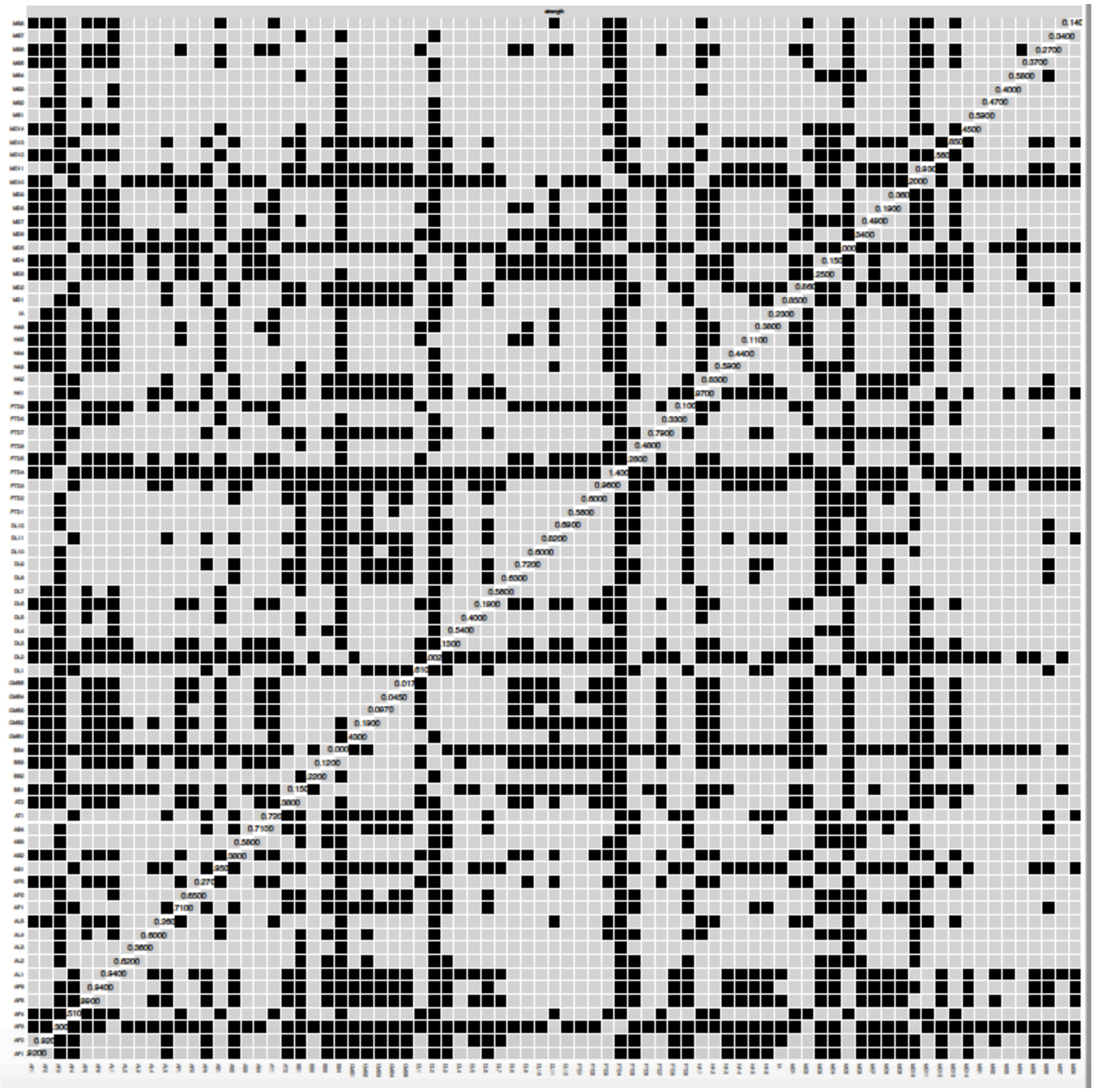

Figure showing bootstrapped stability difference tests (α = 0.05) of the centrality measure 'node strength' (i.e., the sum of the weighted number and strength of all connections of a specific node to all other nodes) in order to show whether the centrality measure node strength of the symptoms differs significantly between nodes. Both axes list all 79 nodes. The gray boxes represent node strengths that do not differ significantly from each other, while black boxes indicate node strengths that do differ significantly. The numbers in the white boxes (i.e., the diagonal line) represent the value of the centrality measure 'node strength' of the specific node. Here, a high number of black boxes can be observed, suggesting significant differences between the centrality measures of ‘node strength’ within our network.

2.3 Figure S3: Stability of the network: bootstrapped stability tests for edge-weights

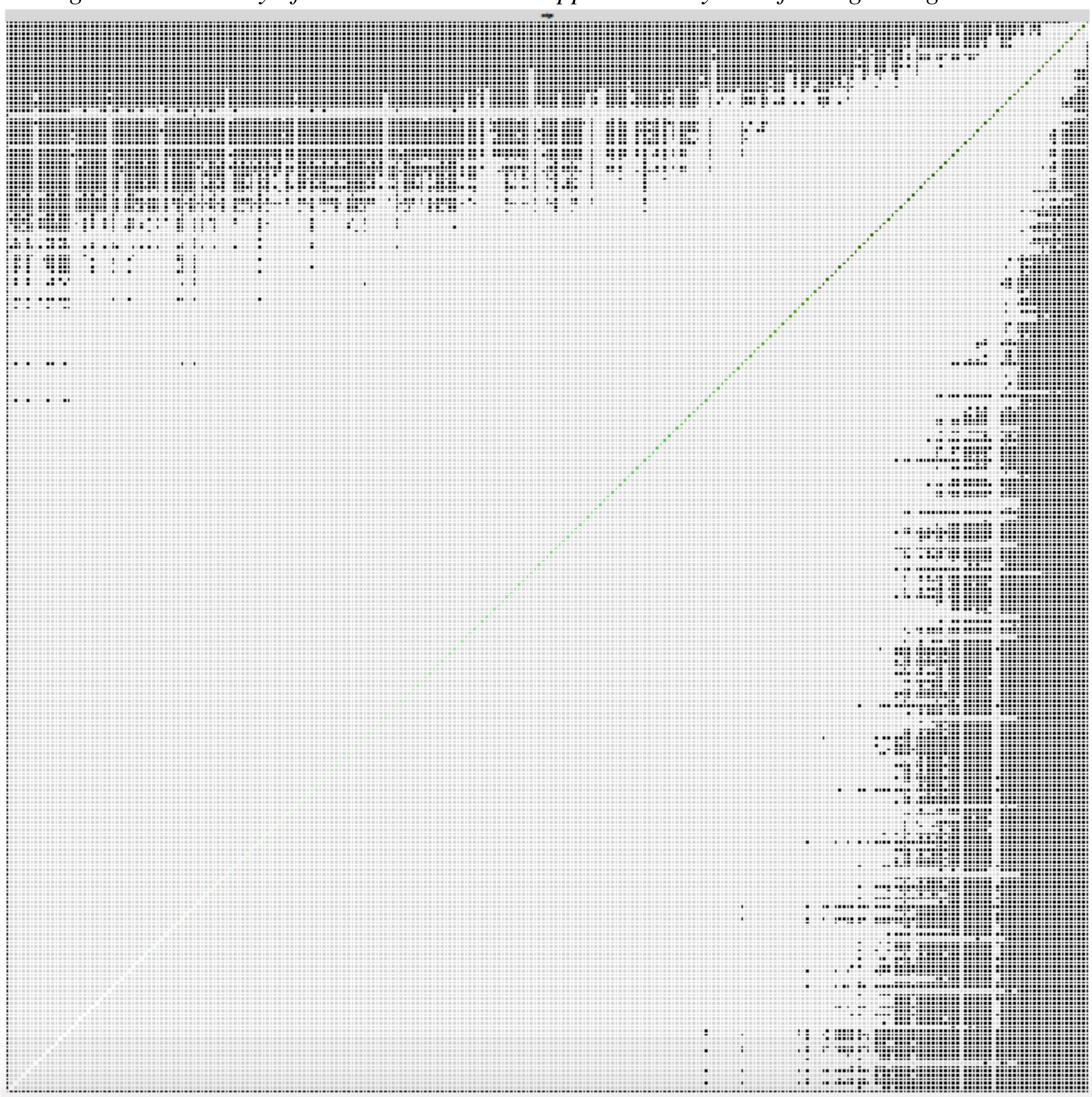

Figure showing bootstrapped stability difference tests ($\alpha = 0.05$) for edge-weights to investigate whether edge-weights differ significantly from each other. At both axes are all edges (i.e., n*(n-1)/2= 3081) between all 79 symptoms, listed. The gray boxes represent edge weights that do not differ significantly from each other, while black boxes indicate edge-weights that do differ significantly. The diagonal represents the strength of the edge weights, changing from white (i.e., weaker edges) to dark green (i.e., stronger weight).

2.4 Figure S4: Centrality measures of all symptoms

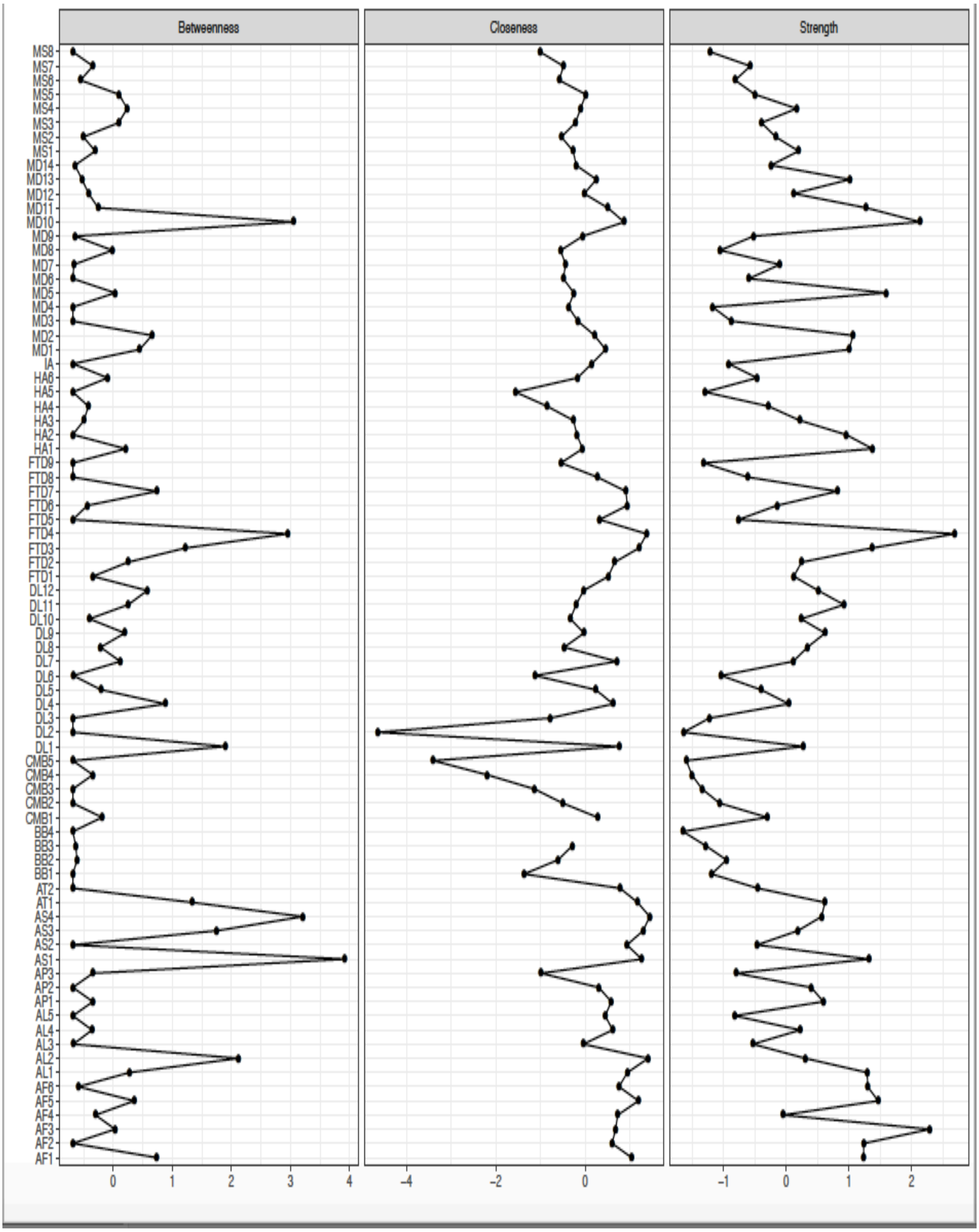

Figure showing three centrality measures: betweenness, node strength and closeness of all symptoms as assessed with the CASH (Barrat et al., 2004; Boccaletti et al., 2006; Opsahl et al., 2010). 'Betweenness' is measured by calculating how often a particular symptom lies on the shortest path between any combination of two nodes. 'Node strength' is calculated as the sum of the weighted number and strength of all connections of a specific node in relation to all other nodes. ‘Closeness’ is the average distance from the node of interest to all other nodes. Centrality indices are shown as standardized z-scores. For abbreviations of symptoms see Table 1.

Reference list

Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. Proc. Natl. Acad. Sci. U. S. A. 101, 3747–52. doi:10.1073/pnas.0400087101

Bland, J. M. and Altman, D. G. (1995). Multiple significance tests: the bonferroni method. Bmj, 310(6973):170.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U., 2006. Complex networks: Structure and dynamics. Phys. Rep. 424, 175–308. doi:10.1016/j.physrep.2005.10.009

Efron, 1979 Efron, B. (1979). Bootstrap methods: another look at the jackknife. The Annals of Statistics, 7(1):1–26

Epskamp, S., Borsboom, D., and Fried, E.I., 2016. Estimating Psychological Networks and their Accuracy: A Tutorial Paper. Retrieved from https://arxiv.org/pdf/1604.08462v3.

Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks Generalizing degree and shortest paths. Soc. Netw. 32, 245–251.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria