Journal of Affective Disorders 189 (2016) 314–320

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research report

What are 'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis

Eiko I. Fried [a,*], Sacha Epskamp [b], Randolph M. Nesse [c], Francis Tuerlinckx [a], Denny Borsboom [b]

[a] Faculty of Psychology and Educational Sciences, University of Leuven, Leuven, Belgium
[b] Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
[c] School of Life Sciences, Arizona State University, Tempe, USA

ARTICLE INFO

Article history:
Received 3 July 2015
Received in revised form
20 August 2015
Accepted 5 September 2015
Available online 1 October 2015

Keywords:
Centrality
Depression symptoms
Major depression
Network analysis

ABSTRACT

Background: The symptoms for Major Depression (MD) defined in the DSM-5 differ markedly from symptoms assessed in common rating scales, and the empirical question about core depression symptoms is unresolved. Here we conceptualize depression as a complex dynamic system of interacting symptoms to examine what symptoms are most central to driving depressive processes.
Methods: We constructed a network of 28 depression symptoms assessed via the Inventory of Depressive Symptomatology (IDS-30) in 3,463 depressed outpatients from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. We estimated the centrality of all IDS-30 symptoms, and compared the centrality of DSM and non-DSM symptoms; centrality reflects the connectedness of each symptom with all other symptoms.
Results: A network with 28 intertwined symptoms emerged, and symptoms differed substantially in their centrality values. Both DSM symptoms (e.g., sad mood) and non-DSM symptoms (e.g., anxiety) were among the most central symptoms, and DSM criteria were not more central than non-DSM symptoms.
Limitations: Many subjects enrolled in STAR*D reported comorbid medical and psychiatric conditions which may have affected symptom presentation.
Conclusion: The network perspective neither supports the standard psychometric notion that depression symptoms are equivalent indicators of MD, nor the common assumption that DSM symptoms of depression are of higher clinical relevance than non-DSM depression symptoms. The findings suggest the value of research focusing on especially central symptoms to increase the accuracy of predicting outcomes such as the course of illness, probability of relapse, and treatment response.

1. Introduction

Reliable diagnosis is an essential prerequisite for the study of mental disorders. The question how to reliably measure Major Depression (MD) is unresolved: depression biomarkers have very limited explanatory power (Cai et al., 2015; Schmaal et al., 2015), and MD was among the least reliable diagnoses in the DSM-5 field trials (Regier et al., 2013).

When assessing depression, specific symptoms are used as indicators for a presumed underlying disorder. While the DSM-5 (APA, 2013) relies on nine criterion symptoms for MD, common rating scales comprise multiple items not part of the DSM criteria. For instance, the Beck Depression Inventory (BDI) (Beck et al., 1996) includes irritability, pessimism, and feelings of being punished, the Hamilton Rating Scale for Depression (HRSD) (Hamilton, 1960) covers anxiety, genital symptoms, hypochondriasis, and insights into the depressive illness, and the Center for Epidemiological Studies Depression Scale (CESD) (Radloff, 1977) includes frequent crying, talking less, and perceiving others as unfriendly. This inconsistency implies a lack of consensus regarding the construct and measurement of depression.

In this article, we attempt to provide a new theoretical and empirical perspective on the question of 'good' depression symptoms. Symptoms are commonly understood as passive indicators *of* some condition or disease, implying that depression symptoms cluster because they stem from a common cause (Fried, 2015; Schmittmann et al., 2013). This view has recently been challenged by the network framework that conceptualizes depression and other mental disorders as webs of causally connected symptoms: insomnia can cause fatigue which in turn triggers concentration and psychomotor problems (Borsboom and Cramer, 2013; Van de

* Correspondence to: University of Leuven, Faculty of Psychology and Educational Sciences, Research Group of Quantitative Psychology and Individual Differences, Tiensestraat 102, 3000 Leuven, Belgium.
E-mail address: Eiko.Fried@gmail.com (E.I. Fried).

http://dx.doi.org/10.1016/j.jad.2015.09.005
0165-0327/

Leemput et al., 2014). Such problems can be organized in feedback loops and create so-called attractor states—highly stable networks—that are hard to escape.

Instead of asking whether a symptom indicates the underlying disorder well, we aim to understand how closely interconnected a symptom is with all other symptoms in the psychopathological network. This metric, known as *centrality* (Opsahl et al., 2010), indicates the overall connectivity of a symptom, and has gained substantial attention in the clinical literature (Bringmann et al., 2015; Robinaugh et al., 2014; Wigman et al., 2015). Centrality is easy to understand from the perspective of social networks: if a celebrity or major newspaper shares news on Twitter, the information will likely spread quickly and widely through the social network; a peripheral person with very few connections is much less likely to impact on the network. For depression, the activation of a highly central symptom means that impulses will spread through the network and activate a large number of other symptoms, whereas a peripheral symptom is less relevant from a dynamic systems perspective because it has few means to influence the network.

The main goals of this article are (A) to explore the centrality of a large number of depressive symptoms, and (B) to compare the centrality of the DSM criteria with the centrality of non-DSM symptoms such as anxiety and irritability that are highly prevalent in depressed samples and associated with worse clinical trajectories (Fava et al., 2008; Judd et al., 2013).

2. Method

2.1. STAR*D protocol

We reanalyzed the version 3.0 dataset from the NIH-supported Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study (Fava et al., 2003; Rush et al., 2004). STAR*D was a multisite randomized clinical trial conducted in the USA. In the first treatment stage, 4,041 patients were enrolled, and all participants received the selective serotonin reuptake inhibitor citalopram. Data were collected via telephone interviews; interviewers had received sufficient training and were masked to treatment. STAR*D was monitored and approved by the institutional review boards of all participating institutions, and after complete description of the study to the subjects, written informed consent was obtained after the study had been fully explained.

2.2. Participants

STAR*D participants had to be between 18 and 75 years, fulfill DSM-IV criteria for single or recurrent nonpsychotic MD, and exhibit a score of at least 14 points on the HRSD. Exclusion criteria were a history of bipolar disorder, schizophrenia, schizoaffective disorder, or psychosis, or current anorexia, bulimia, or primary obsessive compulsive disorder. Further exclusion criteria and details about the study design are described elsewhere (Fava et al., 2003; Rush et al., 2004).

From the 4,041 participants originally enrolled into STAR*D, 3,867 (95.69%) patients provided data during the first measurement point of the first treatment stage. Of these, 10.45% had to be removed due to missing values on the IDS-C or demographic variables, leaving 3,463 depressed patients in the final sample.

2.3. Outcome measures

We analyzed the clinician-rated version of the IDS-C (Rush et al., 1996) assessed at the first measurement point in the first treatment stage of STAR*D. The IDS-C encompasses 30 depression symptoms, both DSM and non-DSM symptoms; it also covers most DSM-5 criterion symptoms in disaggregated form. Disaggregated information was not available for the two symptom domains weight problems (increase vs. decrease) and appetite problems (increase vs. decrease). In line with the manual of the scale, we constructed the aggregated domains 'weight problems' and 'appetite problems'. This led to a total of 28 individual symptoms (Table 1): 15 symptoms that are part of the DSM criteria for MD, and 13 non-DSM symptoms.

Table 1
IDS-C depression symptoms.

#	IDS-C symptoms	Short-code	DSM-5 symptoms	Disaggregated symptoms	Mean	SD
1	Early insomnia	In1	x	x	1.62	1.26
2	Mid insomnia	In2	x	x	2.00	1.17
3	Late insomnia	In3	x	x	1.17	1.23
4	Hypersomnia	Hyp	x	x	0.44	0.87
5	Sadness	Sad	x		2.01	0.74
6	Irritability	Irr			1.33	0.86
7	Anxious/tense	Anx			1.41	0.89
8	Mood reactivity	Rea			1.37	1.05
9	Diurnal variation	Var			0.92	1.17
10	Mood quality	Qua			1.63	1.21
11	Appetite change	App	x	x	1.28	1.11
12	Weight change	Wei	x	x	1.14	1.21
13	Concentration/ decisions	Con	x		1.96	0.94
14	Self-blame/ worthless	Bla	x		1.88	1.17
15	Pessimism	Pes			1.35	0.96
16	Suicidal ideation	Sui	x		0.64	0.74
17	Interest loss	Int	x	x	1.74	1.01
18	Energy loss	Ene	x		1.79	0.94
19	Pleasure loss	Ple	x	x	1.43	1.15
20	Loss of sexual interest	Sex			1.57	1.30
21	Psychomotor retardation	Ret	x	x	0.70	0.59
22	Psychomotor agitation	Agi	x	x	0.87	0.79
23	Somatic complaints	Som			1.34	1.01
24	Sympathetic arousal	Sym			0.94	0.81
25	Panic / phobia	Pan			0.65	0.95
26	Gastrointestinal problems	Gas			0.65	0.88
27	Interpersonal sensitivity	Inp			1.29	1.22
28	Paralysis	Par			0.84	1.09

2.4. Statistical analysis

Overall, we performed three groups of analyses. In a first step, we used the *R*-package *qgraph* (Epskamp et al., 2012) to estimate the network structure of the 15 DSM symptoms, and the network structure of all 28 IDS-C symptoms (both networks are undirected due to the cross-sectional nature of the data). Such networks contain nodes (symptoms) and edges (associations among symptoms). We employed the *glasso* (or graphical lasso) procedure that estimates a network in which the edges are partial correlation coefficients. This means each edge represents the relationship between two variables, controlling for all other relationships in the network. We control for false positive edges using the least absolute shrinkage and selection operator (lasso) (Tibshirani, 1996). As a result, very small edges (likely due to noise) are set exactly to zero. The shrinkage parameter is chosen to minimize the extended Bayesian Information Criterion (Chen and Chen, 2008), and can accurately recover underlying network structures (Van Borkulo

et al., 2014). The graphical representation of networks is based on the Fruchterman–Reingold algorithm that places nodes with stronger and/or more connections closer together. Since IDS-C symptoms are ordered-categorical, analyses were based on polychoric correlations. We tested the robustness of the 28-symptom network and of the graph theoretical measures derived from the network using a bootstrap sampling procedure that is described in detail in the supplementary materials.

Second, we estimated the centrality of all symptoms, which represents the connectedness of a given symptom with all other symptoms in the network. Our main focus in this report lies on *node strength centrality*, a common and stable centrality metric defined as the sum of all associations a given symptom exhibits with all other nodes (Opsahl et al., 2010). We estimated confidence intervals (CI) of the node strength for each symptom by drawing 2,000 bootstrap samples of the data and recalculating the node strength for each resampling of the participants; to do so, we used the *R*-package *bootnet* (Epskamp, 2015) developed for this report. Apart from node strength, other centrality metrics such as betweenness centrality (based on the concept of shortest path length connecting any two symptoms; a symptom with a high betweenness lies along the shortest path connecting many other symptoms) and closeness centrality (a measure of how close a symptom is to all other symptoms) are available (Opsahl et al., 2010). Since closeness and betweenness centrality were substantially correlated with node strength centrality in the networks presented here, we focus on node strength in the main report and present betweenness and closeness results in the supplementary materials.

Third, we performed a number of tests to compare the centrality values or edges across different symptom groups (e.g., DSM vs. non-DSM symptoms). Since network metrics are related in complex ways and do not satisfy the assumptions of *t*-tests, we employed permutation tests that compare the observed variable of interest—e.g., centrality differences across two groups—to a distribution of possible differences between groups. We created the distribution by assigning symptoms randomly to the two groups 100,000 times, and estimated the difference between groups each time. If the observed difference between two groups was within the 2.5% on either side of the distribution, we considered the test significant at the 5% level.

3. Results

3.1. Demographic characteristics

The 3,463 participants included in the final sample were on average 41 years old (*SD*=13), and about 63% of the sample was female. The mean IDS-C score was 36 (*SD*=12; range 1–74), indicating moderately severe depression. Table 1 provides an overview of all symptoms along with their descriptive statistics.

3.2. Network analysis of 15 DSM symptoms

In a first step, we constructed a psychopathological network consisting of the 15 IDS-C symptoms featured in the DSM diagnostic criteria for MD (Fig. 1, left); 71 of all possible 105 edges (68%) were estimated to be above zero. The network revealed strong associations among the sleep symptoms (*hyp, in1*, *in2*, *in3*), a strong connection between weight (*wei*) and appetite problems (*app*), and a close bond between loss of interest (*int*) and loss of pleasure (*ple*) which represent the two disaggregated items of the DSM core criterion symptom 'diminished interest or pleasure'. Interestingly, psychomotor agitation (*agi*) and retardation (*ret*) were weakly positively connected. Overall, symptoms seemed organized in roughly three clusters (sleep, *wei*/*app*, rest), with *agi* being largely isolated.

When inspecting the node strength of the DSM symptoms (Fig. 1, right), we found a smooth decline and no abrupt changes in symptom importance. Since psychomotor agitation (*agi*) is nearly unconnected in the network, it is not surprising to find it has the lowest node strength. Loss of energy (*ene*), on the other hand, is situated in the center of the network, and thus also exhibits the largest symptom importance in the network. A visualization of the betweenness and closeness centralities for the 15 DSM symptoms can be found in supplementary Fig. S1.

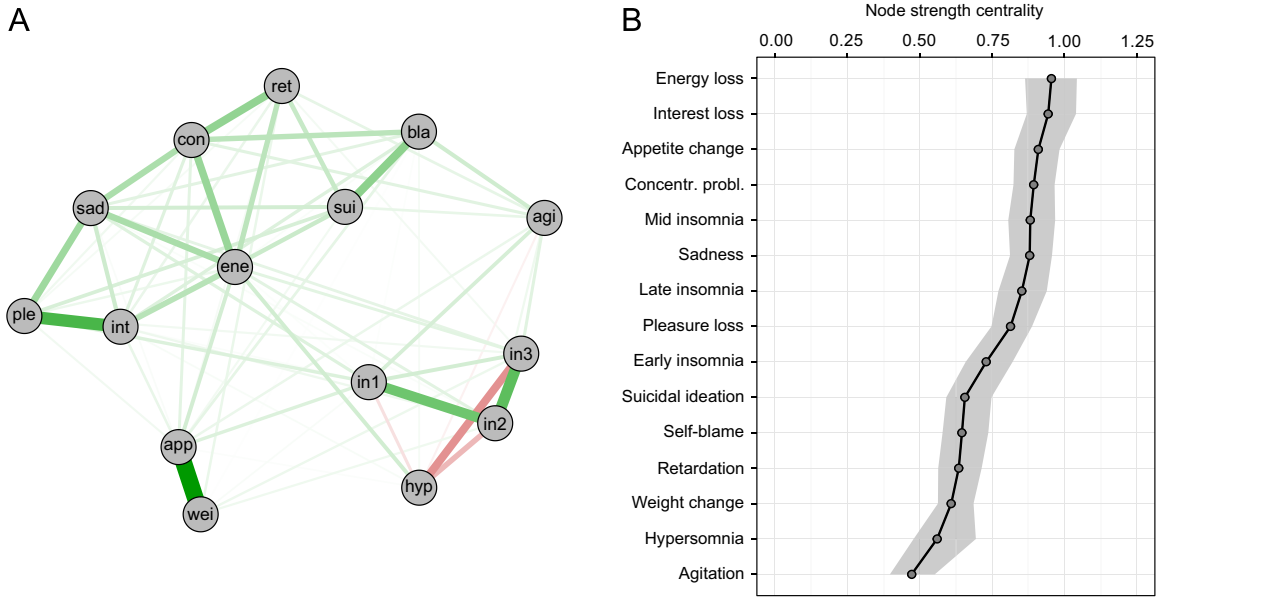

Fig. 1. A: Network containing 15 IDS-C DSM criterion symptoms of Major Depression. Green lines (solid lines in print version) represent positive associations, red lines (dotted lines in print version) negative ones, and the thickness and brightness of an edge indicate the association strength. The layout is based on the Fruchterman–Reingold algorithm that places the nodes with stronger and/or more connections closer together and the most central nodes into the center. See Table 1 for symptom short-codes. B: node strength centrality estimates of the 15 IDS-C DSM criterion symptoms of Major Depression, including 95% confidence intervals.

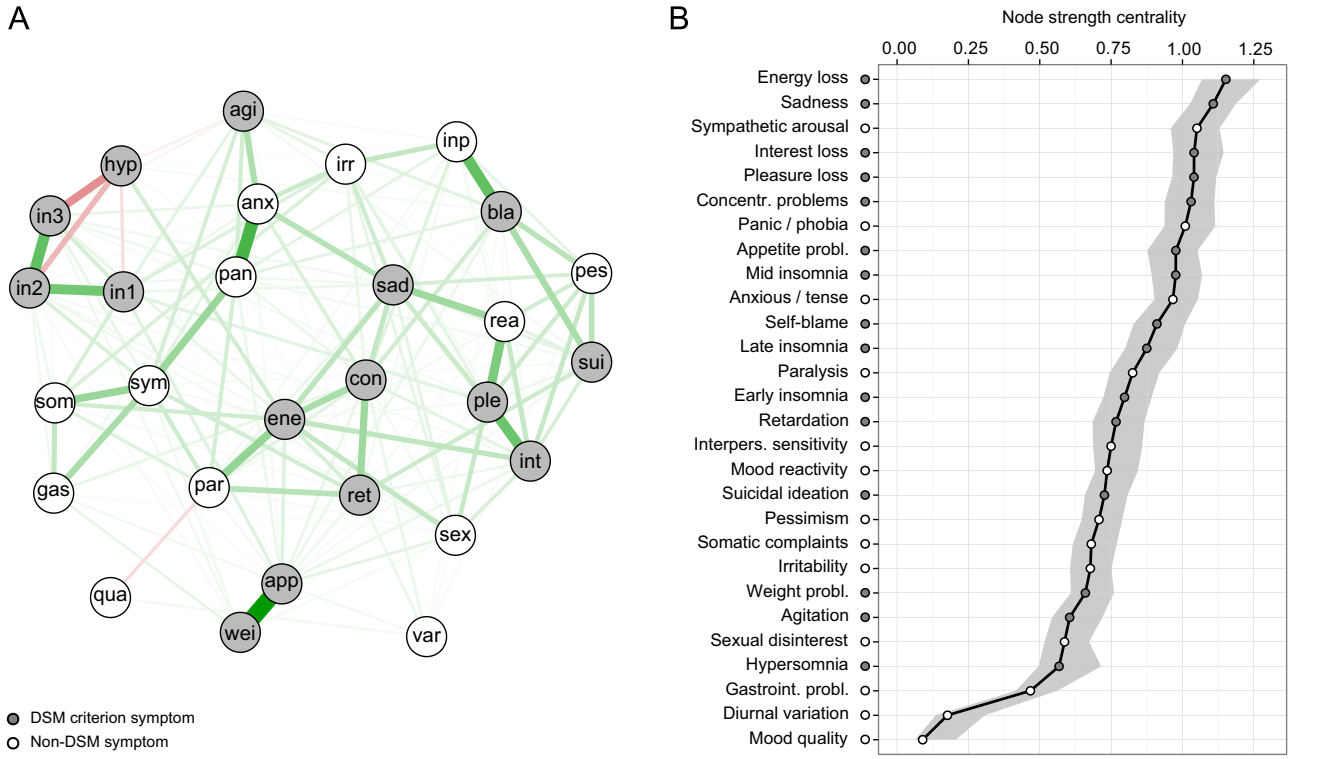

Fig. 2. A: Network containing 28 IDS-C depression symptoms. Green lines (solid lines in print version) represent positive associations, red lines (dotted lines in print version) negative ones, and the thickness and brightness of an edge indicate the association strength. The layout is based on the Fruchterman–Reingold algorithm that places the nodes with stronger and/or more connections closer together and the most central nodes into the center. See Table 1 for symptom short-codes. B: node strength centrality estimates of the 28 IDS-C depression symptoms, including 95% confidence intervals.

3.3. Global network analysis of 28 symptoms

In a second step, we included all 28 IDS-C symptoms in the network (Fig. 2, left). The resulting network featured no unconnected nodes, and 185 of all possible 378 edges (49%) were estimated to be above zero. The four sleep symptoms (*hyp*, *in1*, *in2*, *in3*) were closely connected, and only exhibited weak associations with other symptoms. Similar to the DSM-network in Fig. 1, appetite problems (*app*) and weight problems (*wei*) were closely associated, the disaggregated items of the DSM core symptom 'diminished interest or pleasure' (*int*, *ple*) formed a strong bond, and psychomotor agitation (*agi*) and retardation (*ret*) exhibited no strong negative relationship. In addition, panic/phobia (*pan*) and anxious/tense (*anx*) were highly related, as were interpersonal sensitivity (*inp*) and self-blame (*bla*). The two items diurnal variation (*var*; no relationship between mood and time of day indicating no depression, a specific relationship indicating depression) and mood quality (*qua*; identical to grief indicating no depression, distinct from grief indicating depression) showed few connections.

Inspecting the node strength (Fig. 2, right) revealed that both DSM and non-DSM symptoms were among the ten most central nodes. The DSM core symptoms diminished interest / pleasure (*int*, *ple*) as well as sad mood (*sad*) were highly central, and the two anxiety symptoms (*anx*, *pan*) were also of considerable importance in the network. Diurnal variation (*var*) and mood quality (*qua*) can be considered centrality outliers. A visualization of the betweenness and closeness centralities for the 28 symptoms can be found in supplementary Fig. S2.

Overall, it is obvious from Fig. 2 that there are no fundamental differences between DSM and non-DSM symptoms from a network perspective. The picture emerging from both figures is that the two sets of symptoms are strongly intertwined. The bootstrap sampling procedure revealed that node strength estimates were very robust, and neither the particular symptoms included in the IDS-C, nor the specific number of nodes in the network, biased the node strength estimates (see supplementary materials).

3.4. Comparison of DSM and non-DSM symptoms

In a last analytic step, we used a permutation test to statistically compare the centrality estimates of DSM and non-DSM symptoms. Groups did not differ regarding betweenness centrality ($p=0.12$) and closeness centrality ($p=0.64$). For node strength, DSM criteria were significantly more central than the non-DSM symptoms ($p=0.03$), although the evidence was not very strong and would not survive controlling for multiple testing via Bonferroni correction ($p=0.08$).

We evaluated the robustness of these findings in several ways. First, DSM and non-DSM symptoms did not differ regarding their means ($W=121$, $p=0.30$) or standard deviations ($W=89$, $p=0.72$) using Mann–Whitney *U* tests. This implies that symptoms within the two groups were not differentially severe or variable, ruling out concerns that centrality results were biased by, for instance, DSM symptoms being more severe, or non-DSM having a higher variability. Second, a permutation test revealed that the 10 disaggregated symptoms were not more central than the other 18 symptoms (node strength: $p=0.86$; betweenness and closeness: $p=1$); this is relevant because all disaggregated symptoms were exclusively DSM symptoms, which could have potentially confounded the analysis. Finally, both symptoms identified as centrality outliers, mood variability (*var*) and mood quality (*qua*), are non-DSM symptoms, potentially biasing the comparison of symptom groups. When we repeated the comparison excluding *var* and *qua*, the previous suggestive evidence of a centrality difference between DSM and non-DSM symptoms disappeared for node strength centrality (permutation test; $p=0.13$) and remained non-significant for betweenness ($p=0.28$) and closeness ($p=1$).

4. Discussion

To our knowledge, we provide the first network analysis of 15 disaggregated DSM criterion symptoms of depression, along with the first analysis of a large number of both DSM and non-DSM depressive symptoms. Symptoms differed markedly in their centrality estimates, and DSM criteria were not more central than non-DSM symptoms. This implies that the symptoms featured in the DSM-5 are no more appropriate as indicators of depression than non-DSM symptoms, and that *particular* symptoms (both DSM and non-DSM symptoms) may hold special clinical significance.

4.1. Detailed discussion of the results

The variability of node strength estimates in the 28-symptom network analysis was considerable: while some DSM symptoms such as hypersomnia and psychomotor agitation were among the least central symptoms (ranked 23 and 25), the three IDS-C items representing the DSM core criteria for MD were among the top 5 central symptoms in the network (ranked 1, 4, and 5 out of 28). The DSM-5 assigns a diagnosis of MD if a patient exhibits five or more symptoms, at least one of which has to be one of the two core symptoms depressed mood and diminished interest or pleasure (APA, 2013). The results underline the potential clinical importance of these core criteria, and support a prior study documenting that DSM core symptoms were among the depression symptoms with the highest impact on impairment of psychosocial functioning (ranked 1 and 4 out of 14 symptoms) (Fried and Nesse, 2014). Sad mood and anhedonia have also been shown to outperform other depression symptoms, and in some cases even the sum of all depression symptoms, in predicting depression diagnosis (Rosenström et al., 2015).

The most central non-DSM symptom in our report was sympathetic arousal (palpitations, tremors, blurred vision, sweating), featuring strong connections with somatic complaints (limb heaviness, pain, headaches), gastrointestinal problems, and panic/phobia. While it is well-established that somatic symptoms are prevalent in depressed individuals (Fried and Nesse, 2015a; Zimmerman et al., 2006), we are unaware of research specifically examining the role of somatic depression symptoms in psychopathological networks.

Apart from sympathetic arousal, two anxiety symptoms, panic/phobia and anxious/tense, exhibited high node strength values (ranks 7 and 10). A host of studies have documented the important role of anxiety in depressed patients, which predicts reduced treatment efficacy (Fava et al., 2008; Gollan et al., 2012) as well as chronicity of MD, hospitalization, and disability (Van Loo et al., 2014). High comorbidity rates between mood and anxiety disorders (Kessler et al., 2005) are well established and traditionally understood as a patients having two distinct diseases. An alternative hypothesis supported by network studies is that depression and anxiety symptoms do not form distinct symptom clusters—they substantially overlap and are organized within a larger psychopathological network (Cramer et al., 2010; Fried, 2015; Goekoop and Goekoop, 2014). This means that once a few specific symptoms are activated, this activation can spread from anxiety to depression symptoms (and vice versa) via highly central symptoms.

Diurnal variation and mood quality were largely isolated, meaning that they are unlikely to worsen other MD symptoms once activated. In contrast to most other MD symptoms, they do not range from absent to present: the item diurnal variation lies between “no regular relationship between mood and time of day” and “mood clearly better / worse at a fixed time”, while quality of mood is assessed on a scale ranging from “mood undisturbed or identical to bereavement” to “mood qualitatively distinct from grief”. It may thus not be surprising that these symptoms are only very weakly associated with other symptoms.

4.2. Conclusions and implications

The network perspective does not support the integrity of the DSM criteria, and we see three implications. First, it is of note that the reasons why particular symptoms are featured in the DSM seem to be based more on history than evidence. In 1957, Cassidy et al. (Cassidy et al., 1957) put together a list of symptoms for manic-depressive disorders which was based on the cardinal symptoms proposed by Kraepelin. In Cassidy's report, a diagnosis required the presence of low mood along with six out of ten secondary symptoms (thinking slowly, agitation, insomnia, fatigue, poor appetite, weight loss, constipation, problems concentrating, suicidal thoughts, decreased libido). This list was adapted in 1972 by Feighner et al. (Feighner et al., 1972): constipation was removed, and hypersomnia, guilt, worthlessness, anhedonia, and indecisiveness were added. The criteria have remained largely unchanged in the last 4 decades, and predominant depression scales are just about as old. In a recently published list of the hundred most-cited papers in science (Van Noorden et al., 2014), ranks 51, 53, and 54 were rating scales for depression—the HRSD (1960), the BDI (1961), and the CES-D (1977). What we know about depression today is, to a large degree, based on studies using these instruments, with many results likely idiosyncratic to the particular scales used in particular studies (Santor et al., 2009; Snaith, 1993). The assessment of a large number of symptoms in future studies—across different diagnoses to provide insights on the mechanisms underlying comorbidities—could generate data that may move the field forward substantially (Fried and Nesse, 2015b).

Second, our results are consistent with a previous study examining the impact of 14 partially disaggregated DSM depression symptoms on impairment of psychosocial functioning (Fried and Nesse, 2014), in which the two DSM core symptoms, along with energy loss and concentration problems, were the four most impairing symptoms. These symptoms were also the four most central DSM symptoms in the global network analysis presented here, and among the most central symptoms in the DSM network. This finding supports the notion of centrality as measure of clinical significance: highly central symptoms are likely to activate other symptoms in a network, which may lead to increased levels of overall impairment caused by these specific problems. Future research on centrality may allow prevention and intervention strategies to target specific symptoms before these impact on the rest of the network and lead to a full-fledged depression. It is of note that the DSM core criteria of MD do not receive any special attention in common depression rating scales (such as the HRDS, BDI, or CES-D). This also holds for standard psychometric models in which symptoms are used as equivalent indicators of MD (Schmittmann et al., 2013), and we are not aware of any model that allows for a differentiation between two hierarchical levels of symptoms. While the distinction between core and secondary symptoms may be somewhat arbitrary—concentration problems, for instance, were highly central and also among the most impairing depression symptoms (ranked 2 out of 14) (Fried and Nesse, 2014)—we believe that a focus on severe and central symptoms, especially in the context of dynamic network models, may reveal important insights in future studies.

The third implication is for the common notion of *symptom equivalence* implicit in research studies and statistical models of depression—the idea that symptoms are interchangeable indicators of the same underlying disorder (Fried, 2015; Schmittmann et al., 2013). Our findings of differential symptom centrality support the growing chorus of voices suggesting that depression

symptoms differ in important aspects such as biomarkers and risk factors, and that we should pay special attention to particular symptoms (for a review, see Fried and Nesse, 2015b). This also implies that commonly used sum-scores obfuscate reciprocal interactions among symptoms (Faravelli, 2004; Fried, 2015), and we believe that important insights can be gained from analyzing individual symptoms. A focus on disaggregated symptoms is of particular importance: psychomotor retardation and agitation, for instance, connect to very different symptoms in the network analysis presented above, and may play substantially different roles in depression.

Finally, it is important to point out that the present work should not be misunderstood as critique of the DSM that already has taken a heavy beating in the last years (e.g., Insel, 2013). Our goal is to encourage researchers and clinicians to start thinking about the importance of individual symptoms and their associations, and move beyond the specific symptoms listed in the DSM (Fried, 2015). At this point, it is too early to suggest potential revisions for future iterations of diagnostic systems, and the question how network approaches can inform nosology is beyond the scope of this paper. Nonetheless, if the basic principle of dynamic systems theory applies to mental disorders, if symptoms trigger subsequent symptoms in causal processes of mutual influences, and if these influences happen not only within, but also across diagnoses, the current routine approach to add a small number of symptoms—many of which are not specific to MD—to a sum-score to reflect depression severity may require fundamental revisions.

4.3. Limitations

This study has to be interpreted in the light of a number of limitations. First, STAR*D is a highly representative sample of individuals diagnosed with MD because it allows for certain comorbidities . The sample thus reflects clinical reality and increases the generalizability of results, seeing that more than half of all depressed patients suffer from at least one comorbid diagnosis (Kessler et al., 2005). At the same time, our findings may not generalize to other depression trials because most exclude participants with comorbid conditions.

Second, some IDS symptoms, for instance the three insomnia items, are substantially inter-correlated, which may have increased their centrality estimates. The question arises whether such items should be combined into one node instead of keeping them separately. In other fields of network science such as gene co-expression, this problem has been addressed from the perspective of *topological overlap* (Oldham et al., 2008; Zhang and Horvath, 2005). If nodes such as insomnia items are distinct phenomena that are correlated, they likely exhibit differential patterns of relations to other nodes and should be retained in the network. If, on the other hand, such nodes are just differently worded items that measure the same construct, they will show similar relations to other items, and should be combined because they complicate the network with redundant information. In the absence of definitive work on topological overlap for psychological variables (cf. Costantini, 2014), we decided to retain all nodes instead of arbitrarily combining some, but not others (appetite and weight problems were also substantially related, as were panic/phobia and anxious/tense). Instead, we performed a number of robustness analyses to ensure the stability of the results.

Third, due to the cross-sectional nature of the data, the estimated networks are undirected, and centrality estimates do not provide information whether a symptom mostly actively triggers other symptoms (*out*degree centrality), or whether a symptom mostly is triggered by other nodes (*in*degree centrality). Longitudinal studies allow for differentiating between these two types of centrality (e.g., Bringmann et al., 2014), and future work focusing on outdegree centrality specifically as an indicator for clinical relevance promises important insights.

Finally, participants enrolled into the STAR*D study had to fulfill DSM-IV criteria for single or recurrent nonpsychotic MD and exhibit a score of at least 14 points on the HRSD. This means that patients were selected, among other criteria, based on the presence of DSM symptoms, which may have led to an increased severity and variability of these symptoms compared to non-DSM symptoms. This, in turn, may have biased centrality estimates due to restriction of range: symptoms with a smaller mean and decreased variability are unlikely to exhibit pronounced associations with other symptoms. However, the two symptom groups did not differ significantly in their means or standard deviations, making such a bias very unlikely.

4.4. Conclusion

Measures of symptoms centrality derived from network analysis provide new insights regarding the clinical significance of specific depression symptoms. These insights have major clinical implications and suggest new approaches that may better predict outcomes such as the course of illness, probability of relapse, and treatment response.

Declaration of interests

The authors declare that they have no conflicts of interest with respect to their authorship or the publication of this article.

Role of funding source

The research leading to the results reported in this paper was sponsored by the Research Foundation Flanders (G.0806.13), the Belgian Federal Science Policy within the framework of the Interuniversity Attraction Poles program (IAP/P7/06), and the grant GOA/15/003 from University of Leuven. The STAR*D study was supported by NIMH Contract #N01MH90003 to the University of Texas Southwestern Medical Center (http://www.nimh.nih.gov). The ClinicalTrials.gov identifier is NCT00021528. This manuscript reflects the views of the authors and may not reflect the opinions or views of the STARnD study investigators or the NIMH.

Conflict of interest

None.

Acknowledgment

We are grateful to all participants of the STAR*D study.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jad.2015.09.005.

References

APA, 2013. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition American Psychiatric Association, Washington, DC.

Beck, A.T., Steer, R., Brown, G., 1996. Manual for the Beck Depression Inventory-II. Psychological Corporation, San Antonio, TX.

Borsboom, D., Cramer, A.O.J., 2013. Network analysis: an integrative approach to the structure of psychopathology. Annu. Rev. Clin. Psychol. 9, 91–121. http://dx.doi.org/10.1146/annurev-clinpsy-050212-185608.

Bringmann, L.F., Lemmens, L.H.J.M., Huibers, M.J.H., Borsboom, D., Tuerlinckx, F., 2015. Revealing the dynamic network structure of the Beck Depression Inventory-II. Psychol. Med. 45, 747–757. http://dx.doi.org/10.1017/S0033291714001809.

Cai, N., Bigdeli, T.B., Kretzschmar, W., Li, Y., Liang, J., Song, L., Hu, J., Li, Q., Jin, W., Hu, Z., Wang, G., Wang, L., Qian, P., Liu, Y., Jiang, T., Lu, Y., Zhang, X., Yin, Y., Li, Y., Xu, X., Gao, J., Reimers, M., Webb, T., Riley, B., Bacanu, S., Peterson, R.E., Chen, Y.,

Zhong, I., Liu, Z., Wang, G., Sun, J., Sang, H., Jiang, G., Zhou, X., Li, Y., Li, Y., Zhang, W., Wang, X., Fang, X., Pan, R., Miao, G., Zhang, Q., Hu, J., Yu, F., Du, B., Sang, W., Li, K., Chen, G., Cai, M., Yang, L., Yang, D., Ha, B., Hong, X., Deng, H., Li, G., Li, K., Song, Y., Gao, S., Zhang, J., Gan, Z., Meng, H., Pan, J., Gao, C., Zhang, K., Sun, N., Li, Y., Niu, Q., Zhang, Y., Liu, T., Hu, C., Zhang, Z., Lv, L., Dong, J., Wang, X., Tao, M., Wang, X., Xia, J., Rong, H., He, Q., Liu, T., Huang, G., Mei, Q., Shen, Z., Liu, Y., Shen, J., Tian, T., Liu, X., Wu, W., Gu, D., Fu, G., Shi, J., Chen, Y., Gan, X., Liu, L., Wang, L., Yang, F., Cong, E., Marchini, J., Yang, H., Wang, J., Shi, S., Mott, R., Xu, Q., Wang, J., Kendler, K.S., Flint, J., 2015. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 523, 588–591. http://dx.doi.org/10.1038/nature14659.

Cassidy, W.L., Planagan, N.B., Spellman, M., Cohen, M.E., 1957. Clinical observations in manic-depressive disease; a quantitative study of one hundred manic-depressive patients and fifty medically sick controls. J. Am. Med. Assoc. 164, 1535–1546.

Chen, J., Chen, Z., 2008. Extended Bayesian information criteria for model selection with large model spaces. Biometrika 95, 759–771. http://dx.doi.org/10.1093/biomet/asn034.

Costantini G., 2014, Network Analysis : A New Perspective on Personality Psychology.

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. http://dx.doi.org/10.1017/S0140525X09991567, discussion 150–93.

Epskamp S., 2015. bootnet: Bootstrap methods for various network estimation routines.

Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D., Borsboom, D., 2012. qgraph: Network Visualizations of Relationships in Psychometric Data. J. Stat. Softw. 48, 1–18.

Faravelli, C., 2004. Assessment of psychopathology. Psychother. Psychosom. 73, 139–141. http://dx.doi.org/10.1159/000076449.

Fava, M., Rush, A.J., Alpert, J.E., Balasubramani, G.K., Wisniewski, S.R., Carmin, C.N., Biggs, M.M., Zisook, S., Leuchter, A., Howland, R., Warden, D., Trivedi, M.H., 2008. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. Am. J. Psychiatry 165, 342–351. http://dx.doi.org/10.1176/appi.ajp.2007.06111868.

Fava, M., Rush, A.J., Trivedi, M.H., Nierenberg, A., Thase, M.E., Sackeim, H.A., Quitkin, F.M., Wisniewski, S., Lavori, P.W., Rosenbaum, J.F., Kupfer, D.J., 2003. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. Psychiatr. Clin. North Am. 26, 457–494.

Feighner, J.P., Robins, E., Guze, S.B., Woodruff, R.A., Winokur, G., Munoz, R., 1972. Diagnostic criteria for use in psychiatric research. Arch. Gen. Psychiatry 26, 57–63.

Fried, E.I., 2015. Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. Front. Psychol. 6, 1–11. http://dx.doi.org/10.3389/fpsyg.2015.00309.

Fried, E.I., Nesse, R.M., 2014. The Impact of Individual Depressive Symptoms on Impairment of Psychosocial Functioning. PLoS One 9, e90311. http://dx.doi.org/10.1371/journal.pone.0090311.

Fried, E.I., Nesse, R.M., 2015a. Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. J. Affect. Disord. 172, 96–102. http://dx.doi.org/10.1016/j.jad.2014.10.010.

Fried, E.I., Nesse, R.M., 2015b. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. BMC Med. 13, 1–11. http://dx.doi.org/10.1186/s12916-015-0325-4.

Goekoop, R., Goekoop, J.G., 2014. A Network View on Psychiatric Disorders: Network Clusters of Symptoms as Elementary Syndromes of Psychopathology. PLoS One 9, e112734. http://dx.doi.org/10.1371/journal.pone.0112734.

Gollan, J.K., Fava, M., Kurian, B., Wisniewski, S.R., Rush, A.J., Daly, E., Miyahara, S., Trivedi, M.H., 2012. What are the clinical implications of new onset or worsening anxiety during the first two weeks of SSRI treatment for depression? Depress. Anxiety 29, 94–101. http://dx.doi.org/10.1002/da.20917.

Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 23, 56–62.

T.R. Insel, Transforming Diagnosis [WWW Document], Natl. Inst. Ment. Heal 2013, URL ⟨http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml⟩.

Judd, L.L., Schettler, P.J., Coryell, W., Akiskal, H.S., Fiedorowicz, J.G., 2013. Overt Irritability/Anger in Unipolar Major Depressive Episodes: Past and Current Characteristics and Implications for Long-term Course. JAMA Psychiatry 70, 1171–1180. http://dx.doi.org/10.1001/jamapsychiatry.2013.1957.

Kessler, R.C., Chiu, W.T., Demler, O., Merikangas, K.R., Walters, E.E., 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch. Gen. Psychiatry 62, 617–627. http://dx.doi.org/10.1001/archpsyc.62.6.617.

Oldham, M.C., Konopka, G., Iwamoto, K., Langfelder, P., Horvath, S., Geschwind, D.H., 2008. Functional organization of the transcriptome in human brain. Nat. Neurosci. 11, 1271–1282. http://dx.doi.org/10.1038/nn.2207l.

Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks: Generalizing degree and shortest paths. Soc. Netw. 32, 245–251. http://dx.doi.org/10.1016/j.socnet.2010.03.006.

Radloff, L.S., 1977. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Appl. Psychol. Meas. 1, 385–401. http://dx.doi.org/10.1177/014662167700100306.

Regier, D.A., Narrow, W.E., Clarke, D.E., Kraemer, H.C., Kuramoto, S.J., Kuhl, E.A., Kupfer, D.J., 2013. DSM-5 Field Trials in the United States and Canada, Part II: Test-Retest Reliability of Selected Categorical Diagnoses. Am. J. Psychiatry 170 (1), 59–70. http://dx.doi.org/10.1176/appi.ajp.2012.12070999.

Robinaugh, D.J., Leblanc, N.J., Vuletich, H.A., McNally, R.J., 2014. Network Analysis of Persistent Complex Bereavement Disorder in Conjugally Bereaved Adults. J. Abnorm. Psychol. 123, 510–522. http://dx.doi.org/10.1037/abn0000002.

Rosenström, T., Elovainio, M., Jokela, M., Pirkola, S., Seppo, K., Lindfors, O., Keltikangas-Järvinen, L., 2015. Concordance between Composite International Diagnostic Interview and self-reports of depressive symptoms: a re-analysis TOM. Int. J. Methods Psychiatr. Res. 20, 1–5. http://dx.doi.org/10.1002/mpr.1478.

Rush, A.J., Fava, M., Wisniewski, S.R., Lavori, P.W., Trivedi, M.H., Sackeim, H.A., Thase, M.E., Nierenberg, A., Quitkin, F.M., Kashner, T.M., Kupfer, D.J., Rosenbaum, J.F., Alpert, J.E., Stewart, J.W., McGrath, P.J., Biggs, M.M., Shores-Wilson, K., Lebowitz, B.D., Ritz, L., Niederehe, G., 2004. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. Control. Clin. Trials 25, 119–142. http://dx.doi.org/10.1016/S0197-2456(03)00112-0.

Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., Trivedi, M.H., 1996. The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychol. Med. 26, 477–486.

Santor, D.A., Gregus, M., Welch, A., 2009. Eight Decades of Measurement in Depression. Measurement 4, 135–155. http://dx.doi.org/10.1207/s15366359mea0403.

Schmaal, L., Veltman, D.J., van Erp, T.G.M., Sämann, P.G., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W.J., Vernooij, M.W., Ikram, M.A., Wittfeld, K., Grabe, H.J., Block, A., Hegenscheid, K., Völzke, H., Hoehn, D., Czisch, M., Lagopoulos, J., Hatton, S.N., Hickie, I.B., Goya-Maldonado, R., Krämer, B., Gruber, O., Couvy-Duchesne, B., Rentería, M.E., Strike, L.T., Mills, N.T., de Zubicaray, G.I., McMahon, K.L., Medland, S.E., Martin, N.G., Gillespie, N.A., Wright, M. J., Hall, G.B., MacQueen, G.M., Frey, E.M., Carballedo, A., van Velzen, L.S., van Tol, M.J., van der Wee, N.J., Veer, I.M., Walter, H., Schnell, K., Schramm, E., Normann, C., Schoepf, D., Konrad, C., Zurowski, B., Nickson, T., McIntosh, A.M., Papmeyer, M., Whalley, H.C., Sussmann, J.E., Godlewska, B.R., Cowen, P.J., Fischer, F.H., Rose, M., Penninx, B.W.J.H., Thompson, P.M., Hibar, D.P., 2015. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol. Psychiatry, 1–7. http://dx.doi.org/10.1038/mp.2015.69.

Schmittmann, V.D., Cramer, A.O.J., 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. http://dx.doi.org/10.1016/j.newideapsych.2011.02.007.

Snaith, P., 1993. What do depression rating scales measure? Br. J. Psychiatry 163, 293–298. http://dx.doi.org/10.1192/bjp.163.3.293.

Tibshirani, R., 1996. Regression Shrinkage and Selection via the Lasso. J. R. Stat. Soc. Ser. B 58, 267–288.

Van Borkulo, C.D., Borsboom, D., Epskamp, S., Blanken, T.F., Boschloo, L., Schoevers, R.A., Waldorp, L.J., 2014. A new method for constructing networks from binary data. Sci. Rep. 4, 1–10. http://dx.doi.org/10.1038/srep05918.

Van de Leemput, I.A., Wichers, M.C., Cramer, A.O.J., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E.H., Viechtbauer, W., Giltay, E.J., Aggen, S.H., Derom, C., Jacobs, N., Kendler, K.S., van der Maas, H.L.J., Neale, M.C., Peeters, F., Thiery, E., Zachar, P., Scheffer, M., 2014. Critical slowing down as early warning for the onset and termination of depression. Proc. Natl. Acad. Sci. U. S. A. 111, 87–92. http://dx.doi.org/10.1073/pnas.1312114110.

Van Loo, H.M., Cai, T., Gruber, M.J., Li, J., de Jonge, P., Petukhova, M., Rose, S., Sampson, N. a, Schoevers, R. a, Wardenaar, K.J., Wilcox, M. a, Al-Hamzawi, A.O., Andrade, L.H., Bromet, E.J., Bunting, B., Fayyad, J., Florescu, S.E., Gureje, O., Hu, C., Huang, Y., Levinson, D., Medina-Mora, M.E., Nakane, Y., Posada-Villa, J., Scott, K.M., Xavier, M., Zarkov, Z., Kessler, R.C., 2014. Major Depressive Disorder Subtypes To Predict Long-Term Course. Depress. Anxiety 13, 1–13. http://dx.doi.org/10.1002/da.22233.

Van Noorden, R., Maher, B., Nuzzo, R., 2014. The top 100 papers. Nature 514, 550–553. http://dx.doi.org/10.1038/514550a.

Wigman, J.T.W., van Os, J., Borsboom, D., Wardenaar, K.J., Epskamp, S., Klippel, A., MERGE, Viechtbauer, W., Myin-Germeys, I., Wichers, M., 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., 1–13. http://dx.doi.org/10.1017/S0033291715000331.

Zhang, B., Horvath, S., 2005. A general framework for weighted gene co-expression network analysis. Stat. Appl. Genet. Mol. Biol. 4 (Article17). http://dx.doi.org/10.2202/1544-6115.1128.

Zimmerman, M., McGlinchey, J.B., Young, D., Chelminski, I., 2006. Diagnosing major depressive disorder I: A psychometric evaluation of the DSM-IV symptom criteria. J. Nerv. Ment. Dis. 194, 158–163. http://dx.doi.org/10.1097/01.nmd.0000202239.20315.16.