BMC Medicine

The importance of transdiagnostic symptom level assessment to understanding prognosis for depressed adults: analysis of data from six randomized control trials.

--Manuscript Draft--

Manuscript Number:	BMED-D-20-03223R2
Full Title:	The importance of transdiagnostic symptom level assessment to understanding prognosis for depressed adults: analysis of data from six randomized control trials.
Article Type:	Research article
Section/Category:	Psychiatry
Funding Information:	

Wellcome Trust (201292/Z/16/Z)	Dr Joshua E J Buckman
MQ: Transforming Mental Health (MQDS16/72)	Dr Zachary D Cohen
UCLH Biomedical Research Centre	Dr Rob Saunders Prof Stephen Pilling
National Institute for Health Research	Prof David Kessler Prof Nicola Wiles
University College London	Dr Ciaran O'Driscoll Dr Gareth Ambler Prof Glyn Lewis
Vanderbilt University	Prof Steven D Hollon
University of Southampton	Prof Tony Kendrick
University of Exeter	Prof Edward Watkins
University of York	Prof Simon Gilbody
University of Pennsylvania (US)	Prof Robert J DeRubeis

Abstract:

Background: Depression is commonly perceived as a single underlying disease with a number of potential treatment options. However, patients with major depression differ dramatically in their symptom presentation and comorbidities, e.g. with anxiety disorders. There are also large variations in treatment outcomes and associations of some anxiety comorbidities with poorer prognoses, but limited understanding as to why, and little information to inform the clinical management of depression. There is a need to improve our understanding of depression, incorporating anxiety comorbidity, and consider the association of a wide range of symptoms with treatment outcomes.
Method: Individual patient data from six RCTs of depressed patients (total n=2858) were used to estimate the differential impact symptoms have on outcomes at three post intervention timepoints using individual items and sum scores. Symptom networks (Graphical Gaussian Model) were estimated to explore the functional relations among symptoms of depression and anxiety and compare networks for treatment remitters and those with persistent symptoms to identify potential prognostic indicators.
Results: Item-level prediction performed similarly to sum scores when predicting outcomes at 3 to 4 months and 6 to 8 months, but outperformed sum scores for 9 to 12 months. Pessimism emerged as the most important predictive symptom (relative to all other symptoms), across these time points. In the network structure at study entry, symptoms clustered into physical symptoms, cognitive symptoms, and anxiety symptoms. Sadness, pessimism, and indecision acted as bridges between communities, with sadness and failure/worthlessness being the most central (i.e. interconnected) symptoms. Connectivity of networks at study entry did not differ for future remitters vs. those with persistent symptoms.
Conclusion: The relative importance of specific symptoms in association with outcomes and the interactions within the network highlight the value of transdiagnostic assessment and formulation of symptoms to both treatment and prognosis. We discuss the potential for complementary statistical approaches to improve our understanding of psychopathology.

Corresponding Author:	Joshua Buckman University College London UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University College London
Corresponding Author's Secondary Institution:	
First Author:	Ciaran O'Driscoll, D.Clin.Psy
First Author Secondary Information:	
Order of Authors:	Ciaran O'Driscoll, D.Clin.Psy
	Joshua E J Buckman, D.Clin.Psy; PhD
	Eiko I Fried, PhD
	Rob Saunders, PhD
	Zachary D Cohen, PhD
	Gareth Ambler, PhD
	Robert J DeRubeis, PhD
	Simon Gilbody, DPhil
	Steven D Hollon, PhD
	Tony Kendrick, MD
	David Kessler, MD
	Glyn Lewis, PhD
	Edward Watkins, PhD
	Nicola Wiles, PhD
	Stephen Pilling, PhD
Order of Authors Secondary Information:	
Response to Reviewers:	Thank you for the comments. We have addressed the revisions and hope the manuscript is now suitable for acceptance. Reviewer reports: Reviewer #1: The authors have adequately addressed my comments. Reviewer #4: The authors have sufficiently addressed my concerns. The following are optional suggestions. I found this phrasing much clearer than the phrasing used in the manuscript; using this direct phrasing in the manuscript could clear up a lot of potential confusion about the sample. "There is an ongoing debate in the field whether the most central items derived from network models offer predictive utility" I think I made my point poorly in my previous review. I fully agree with the authors that the usefulness of network centrality is under contention. My concern was just a minor phrasing issue - all relevant items offer some predictive utility, hence central items (and noncentral items) offer predictive utility in the abstract. The debate is over whether central items are different than noncentral items in any meaningful way. I thought the authors point would be clearer phrased in a different way, e.g.:

	"There is an ongoing debate in the field whether central items derived from network models have any special predictive utility" [any predictive utility beyond other items / etc] We appreciate the clarification and have altered the sentence. Line 389: “There is an ongoing debate in the field whether central items derived from network models offer predictive utility beyond other items (71–73).” "The colours correspond to symptoms. We have tried a side by side bar graph but with so many variables we found that to be more difficult to read." I couldn't find Figure 1 in the revised version, but if the colours correspond to symptoms, a legend should be added. While we appreciate the suggestion, the symptoms are clearly labelled in each plot in Figure 1. Adding a legend would appear to be superfluous, crowd the figure, and conversely reduce the readability. Editorial requests: - Figure 1 is missing - unless 'Figure 2' is a combination of both figures - needs clarification Apologies for the omission, it appeared on the pdf proof we had from the last submission, immediately before Figure 2, but we have uploaded afresh to ensure it is in the system. - The following authors have not yet confirmed authorship: Robert J DeRubeis - derubeis@psych.upenn.edu Stephen Pilling - s.pilling@ucl.ac.uk Can you please ask them to respond to our automated email, or to contact me directly, to confirm whether they agree to be listed as co-authors? (alessandro.recchioni@biomedcentral.com) The authors have notified us that they have confirmed their authorship by contacting you directly. - Please make the following change to your section headings: change ‘Introduction’ to ‘Background’. We have made this change.

The importance of transdiagnostic symptom level assessment to understanding prognosis for depressed adults: analysis of data from six randomized control trials.

O'Driscoll, C.,[1] Buckman, J. E. J.,[1,2] Fried, E. I.,[3] Saunders, R.,[1] Cohen, Z. D.,[4] Ambler, G.,[5] DeRubeis, R. J.,[6] Gilbody, S.,[7] Hollon, S. D.,[8] Kendrick, T.,[9] Kessler. D.,[10] Lewis, G.,[11] Watkins, E.,[12] Wiles. N.,[13] & Pilling, S.[1,14]

1 – Centre for Outcomes Research and Effectiveness (CORE), Research Department of Clinical, Educational & Health Psychology, University College London, 1-19 Torrington Place, London WC1E 7HB, UK

2- iCope – Camden & Islington Psychological Therapies Services, Camden & Islington NHS Foundation Trust, St Pancras Hospital, London NW1 0PE, UK.

3 - Department of Clinical Psychology, Leiden University, Leiden, The Netherlands.

4 – Department of Psychiatry, University of California, Los Angeles, Los Angeles, CA, USA

5 – Statistical Science, University College London, 1-19 Torrington Place, London WC1E 7HB

6 – School of Arts and Sciences, Department of Psychology, 425 S. University Avenue, Philadelphia PA, 19104-60185, USA

7 - Department of Health Sciences, University of York, Seebohm Rowntree Building, Heslington, York YO10 5DD, UK

8 –Department of Psychology, Vanderbilt University, Nashville, TN, USA

9 – Primary Care, Population Sciences and Medical Education, Faculty of Medicine, University of Southampton, Aldermoor Health Centre, Southampton SO16 5ST.

10 – Centre for Academic Primary Care, Population Health Sciences, Bristol Medical School, University of Bristol, Canynge Hall, Bristol, UK

11 – Division of Psychiatry, University College London, Maple House, London W1T 7NF, UK

12- Department of Psychology, University of Exeter, Sir Henry Wellcome Building for Mood Disorders Research, Perry Road, Exeter EX4 4QG

13 - Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield House, Bristol, UK

14 - Camden & Islington NHS Foundation Trust, St Pancras Hospital, 4 St Pancras Way, London, NW1 0PE, UK

Corresponding Authors: Ciarán O'Driscoll & Joshua Buckman

Abstract

Background: Depression is commonly perceived as a single underlying disease with a number of potential treatment options. However, patients with major depression differ dramatically in their symptom presentation and comorbidities, e.g. with anxiety disorders. There are also large variations in treatment outcomes and associations of some anxiety comorbidities with poorer prognoses, but limited understanding as to why, and little information to inform the clinical management of depression. There is a need to improve our understanding of depression, incorporating anxiety comorbidity, and consider the association of a wide range of symptoms with treatment outcomes.

Method: Individual patient data from six RCTs of depressed patients (total n=2858) were used to estimate the differential impact symptoms have on outcomes at three post intervention timepoints using individual items and sum scores. Symptom networks (Graphical Gaussian Model) were estimated to explore the functional relations among symptoms of depression and anxiety and compare networks for treatment remitters and those with persistent symptoms to identify potential prognostic indicators.

Results: Item-level prediction performed similarly to sum scores when predicting outcomes at 3 to 4 months and 6 to 8 months, but outperformed sum scores for 9 to 12 months. Pessimism emerged as the most important predictive symptom (relative to all other symptoms), across these time points. In the network structure at study entry, symptoms clustered into physical symptoms, cognitive symptoms, and anxiety symptoms. Sadness, pessimism, and indecision acted as bridges between communities, with sadness and failure/worthlessness being the most central (i.e. interconnected) symptoms. Connectivity of networks at study entry did not differ for future remitters vs. those with persistent symptoms.

Conclusion: The relative importance of specific symptoms in association with outcomes and the interactions within the network highlight the value of transdiagnostic assessment and formulation of symptoms to both treatment and prognosis. We discuss the potential for complementary statistical approaches to improve our understanding of psychopathology.

Keywords: item level analysis, network modelling, comorbidity, depression, anxiety

Background

Psychological therapies and medication are effective treatments for depression (e.g., 1,2). However, effect sizes have been modest and gains in treatment outcomes have plateaued (3). Interventions for depression target a broad range of symptoms, and knowledge of 'what' is being intervened upon is not necessary to the delivery of most treatments, and poses problems for causal inference (4). To improve interventions, we may need to improve our knowledge of the structure of depression (5).

Depression is heterogeneous in terms of aetiology and symptom profile (6–8). Mood disorders are highly comorbid with anxiety disorders, and may share psychological and biological vulnerabilities (9,10). The risk of one disorder can increase the risk of another (11) and the same end state may be achieved via many different paths (equifinality) (12,13).These disorders are not discrete entities and as such, neglecting the symptomatic heterogeneity discards potential insights (14).

There is strong evidence that different symptoms are not equivalent or interchangeable (15) and studies of individual symptoms in the last decade have brought important understanding. For example, individual symptoms may differ in response to treatment (16,17), and have been shown to have a differential impact on functioning (18–20). Depression is a recurrent disorder with the probability of relapse strongly associated with the presence of residual depressive symptoms at the end of treatment (21,22). Comorbid anxiety disorders are related both to worse treatment outcomes (23) and to an increased risk of relapse (21). An assumed unidimensional view of depression, characterized by sum score (sum of symptom severity scores) measurement and prediction models conceals the variability within depression (24) Understanding the relative importance of comorbid symptoms may offer more information than severity of disorder alone and provide additional treatment and prognostic information (25). Large-scale, multisite clinical trial data, coupled with innovative statistical methods can provide categorisation and treatment optimisation to provide immediate benefits by informing clinical decisions (26–28).

There is also value in studying the relations among these symptoms. Network theory posits that the relationships between common affective, cognitive, and somatic symptoms of these disorders, may reflect potential causal pathways and elucidate maintenance mechanisms (29). Depression and anxiety have been modelled as symptom networks using cross-sectional and longitudinal data, demonstrating the interrelation between the symptoms of each disorder, where comorbidity results from mutually reinforcing interrelation between symptoms of each disorder (30,31). Anhedonia, anxiety, worry, fatigue and sadness are predominantly influential symptoms in these networks (5,32,33). The relationship between symptoms / mechanisms can help to predict outcome and potentially inform treatment targets and the development of treatments targeting specific mechanisms (34).

There are inconsistencies in the network literature exploring depression and anxiety, due to design, sampling, and variability arising from differing measurement (15,35). When attempting to discriminate between groups for the purposes of identifying whom may benefit from treatment (prognosis at group level), there are varying results from network comparison studies, where it has been suggested that densely connected networks may be less likely to recover (36). However, these differences are not always observed (37) and require large sample sizes to detect any effect. It is also unclear how these networks generalize to idiographic networks at the present stage Past research has been conducted on small samples with low quality assessment of patients (or non-clinical samples) and lack of adequate consideration of comorbidity (missing out on the wider spectrum of anxiety disorders).

In this study we aim to:

1) Identify important symptoms for outcome by examining the (differential) impact of individual symptoms on prognosis for adults with depression that took part in randomized controlled trials after seeking treatment in primary care; and assess whether individual symptoms offer predictive value above sum scores.
2) Discern the functional relations among symptoms and clarify the interplay between highly comorbid symptoms of depression and anxiety disorders.
3) Consider whether there are differences in the baseline symptom networks of patients that remitted vs those whose depression persisted, after treatment.

Method

Datasets

Data were drawn from a subset of the Dep-GP individual patient data (IPD) database (36). The formation of the Dep-GP IPD dataset has been described elsewhere (36). Bibliographic databases were searched up to 29th April 2020 for RCTs of unipolar depressed adults seeking treatment for depression or with depressive symptoms significant enough for them to seek treatment, recruited from primary care; had at least one active treatment arm; and used the CIS-R at baseline.

Studies were excluded if they were studies of: patients with depression secondary to a diagnosis of personality disorder, psychotic conditions, or neurological conditions; bi-polar or psychotic depressions; children or adolescents; feasibility studies; or were studies of adults with either depression or an anxiety disorder, rather than a primary depression with or without comorbid anxiety. Additional inclusion criteria for the present study were the use of the Beck Depression Inventory (2nd Edition) (BDI-II) (37) at study entry. The inclusion criteria ensured uniformity in the measurement of depressive and anxiety symptoms, chronicity of problems and determination of diagnoses including anxiety comorbidities.

Data on all individual patients from all six eligible RCTs were included in the current study, these were: COBALT (38), GENPOD (39), PANDA (40), TREAD (41), MIR (42) and IPCRESS (43).

Measures

Individual items from the BDI-II (37); and individual symptom subscales of the CIS-R (44), including duration of depression and anxiety which have been shown to be independently associated with prognosis for depressed adults (45).

Outcomes

The primary outcome was endpoint depressive symptoms at three to four months post-study entry. Five of the studies used the BDI-II at three to four months, and one used the PHQ-9. A continuous ‘depression severity’ score was developed by converting the responses on each measure to a latent trait depressive symptom severity score (PROMIS T-Score) (46), using the expected a posteriori parameter from a multidimensional item-response theory based score conversion tool (47). Depressive symptoms (PROMIS T-Score) at six to eight months post-study entry, and nine to twelve months were secondary outcomes.

As a sensitivity analysis, the BDI-II scores were used as outcomes for the three timepoints; (five studies at three to four months; four studies at six to eight months, and three studies at nine to twelve months).

Data analysis

All analyses were performed in R 3.6 (48) and Stata 16.0 (49). Analysis code is available from https://osf.io/wck6b/. The data that support the findings of this study are available from the lead author of the Dep-GP (JB) subject to agreement from the chief investigators or data controllers of the individual RCTs. Restrictions apply to the availability of these data, which were used under license for this study.

Pre-processing

Datasets were combined and pre-processed together. There was no missing data at study entry. All items were investigated to ensure they met assumptions for inclusion in the network models, including assessing for: near zero variance; roughly equal variance of items; asymmetrical distributions; and topological overlap (50). Items were removed if they violated assumptions across all studies. We aimed to address topological overlap using the 'goldbricker' function in R (51) with a threshold of 25% (correlations between items should have significantly different correlations with 25% of the other items), accepting minimal correlation of 0.5.The respective pair of items were combined into a single variable using Principal Component Analysis (PCA) if reasonable to combine from a clinical perspective. Items were afterwards rescaled to their original Likert scale values to make variances comparable across items (52).

Association with outcomes

We aimed to examine the differential impact of individual symptoms on outcomes; and assess whether individual symptoms offer predictive value above sum scores.Sum score totals were entered into a linear regression model, while the item severity scores were entered into an elastic net generalized linear model (ENR) (53). ENR, a statistical method combining lasso and ridge regression approaches, minimizes overfitting and the use of ten separate, ten-fold repeated cross validation aids in assessing the effectiveness of the model. The item-level and sum-score models were compared using root mean squared error, mean absolute error and R^2.

As the item-level predictors were assumed to be correlated and that we wished to assess the explanatory power of individual predictors, we estimated the contribution of each item to the outcome prediction using Shapley Additive exPlanations (54), following ENR model estimation. Five hundred Monte Carlo repetitions were used to estimate each Shapley value. This metric is more accurate than other variable importance metrics when predictors are dependent (55). Items with large Shapley values are 'important', indicating the relative contribution of an item to the model while accounting for correlated features in the data.

Network modelling

A Graphical Gaussian Model (GGM) aims to capture the direct effects (edges) between items while controlling for all other items in the network. A network was estimated by combining data from the six RCTs. The sample was then split into two networks (those with persistent symptoms vs. remitters: BDI-II score <10 at 3-4 months), the networks were re-estimated and compared using the network comparison test with 1000 iterations (56).

We performed a number of analyses to test the robustness of the networks we estimated. While lasso (57), regularized GGMs (58) are most frequently reported in the network literature, lasso specificity has recently been shown to be lower than expected in dense networks with many small edges, leading to an increase in false positives (59). We also estimated an unregularized GGM using an iterative modelling procedure: the Extended Bayesian

Information Criterion (EBIC). Selecting unregularized GGMs according to EBIC has been shown to converge to the true model (60). The algorithm runs 100 glasso models, re-fits all models without regularization, and subsequently adds and removes edges until EBIC can no longer be improved. The best performing model (EBIC parameter) was selected to provide a conservative GGM estimation (high specificity).

Chronicity of disorders has been shown to interact with symptom severity (45,61). We corrected for the potential confounding effects of duration of depression and anxiety within the network models.

Combining data obtained from different studies holds the potential for between-study differences to influence estimation. A network estimation procedure (fused graphical lasso: FGL) (62) has been designed to manage this issue, however, this involves estimating networks individually and penalizing between study differences. Where study size affects the estimation of edges, this can lead to penalization based on sample size rather than on true differences between the network structures (63). As such, it was decided to estimate based on the combined sample and to compare this to the FGL network (joint estimation using a fused penalty, and 10-fold cross validation), to assess the potential influence of group level differences.

Finally, the network model was tested for the stability of expected influence centrality and the accuracy of interrelations using a nonparametric bootstrapping procedure (1000 iterations) (64). For details of these see the Supplementary material.

We obtained two types of information from the resulting network structures. First, symptoms can form clusters or communities with other symptoms to which they are connected reflecting commonalities between them. We estimated the community structure by using a bootstrapped walktrap algorithm (65), investigated for item stability before selecting communities. Second, the overall connectivity of a symptom, i.e. its connection to other symptoms, can be quantified in a number of ways and is referred to as centrality. Some scholars have argued that activation of a central symptom has the potential to activate associated symptoms in the network (66), where symptom centrality is then interpreted as symptom importance, given that identifying such symptoms may have the potential to elucidate the processes underlying comorbidity and implications for treatment. Within the context of communities specifically, symptoms which connected to more than one community are referred to as bridge symptoms. Within cross-sectional networks (as explored here), we refer to centrality as a statistical parameter, i.e. the strength of predictive associations between symptoms. Centrality does not automatically translate into clinical relevance (67) and cautious interpretation is warranted (63). It requires consideration of: how the symptoms activate within the network (flow or transfer); the conceptual similarity between symptoms; and whether there is missing information on the shared variance (68). Symptom centrality was calculated using: Expected Influence (EI: strength of the relationships a given node has with other node); and the geometric mean of the Participation Ratio (PR) and Participation Coefficient (PC); and normalized bridge expected influence centrality (69). The PR quantifies the number and strength of edges, while the PC takes the community structure into account (70).

Results

*Table 1: Descriptive table of studies included in the dataset. Summary of included variables provided in supplementary materials. * International Baccalaureate equivalent ** High school diploma equivalent.*

Demographic details for the studies are presented in Table 1. Overall samples were comparable. The severity of depressive symptoms captured by BDI-II scores at baseline in the PANDA sample was lower than the other trials. Descriptive results are reported in the supplementary materials.

Association with outcomes

In order to assess the utility of item level models, we compared them to sum score models. For all item level models (Table 2), the optimal shrinkage parameters for the elastic net regression model were selected via minimum cross-validated error criterion ($\alpha = 0.1$ and $\lambda = 0.05$). While models performed similarly at three to four months and six to eight months, the item level elastic net regression model outperformed linear regression with BDI-II and CIS-R (sum of anxiety items) totals at the nine to twelve month time point. The sensitivity analysis performed similarly. Due to the absence of two studies (IPCRESS and PANDA) at the nine to twelve month endpoint, we reran the analyses for the earlier time points without these studies. This sensitivity analysis did not reveal any difference in the pattern of model performance.

Pessimism (Figure 1) was consistently the most important item; health anxiety was in the upper quartile at each time point; and concentration, failure/worthlessness, also in the upper quartile at three to four months; guilt and sleep at six to eight months; and somatic symptoms at nine to twelve months.

Table 2: Performance of the regression models. Sum scores: BDI-II and CIS-R; RMSE root mean squared error; MAE mean absolute error; R^2 proportion of the variance explained.

Figure 1: Shapley values for variable importance are plotted: (showing the difference contribution of items to predictions).

Network Modelling

For the individual items in the network model, near zero variance (e.g. due to floor and ceiling effects) was not observed. However, we saw asymmetric distributions (skew) on a number of items. As such, a Spearman covariance matrix was estimated and used to estimate the network model. Multi-collinearity was identified for two pairs of items (loss of pleasure with loss of interest, failure with worthlessness). New items were constructed using PCA for each

pair. The optimal model for the network analysis was an unregularized Graphical Gaussian Model using the EBIC.

A walktrap algorithm identified three, stable, symptom communities (median = 3, SD = 0.15, 95% CI [2.71,3.29]). The three communities split into anxiety items, depressive cognitions and depressive physical symptoms. Bridging EI elucidated three bridging symptoms between the communities: sadness and indecisiveness (from the physical symptoms community); and pessimism (cognitive symptoms community);

Figure 2: Network plot (top) with communities. Bridge symptoms are categorized separately, however sadness and indecisiveness fall into community 1, and pessimism into community 3. The thickness of the edges indicates to what degree items are related, and the colour of the edges indicates the relationship sign (i.e. positive = blue, negative = red). Centrality estimates: PC/PR and EI (bottom).

Centrality estimates (i.e. measures of the strength of connection to other items) are reported in figure 2. The EI correlation stability coefficient was high (0.75), suggesting that the ordering of items based on centrality remained the same after re-estimating the network with fewer cases (the probability the correlation between original centrality indices and centrality of networks based on subsets was 0.7 or higher) and can be reliably interpreted.

The estimates from the different metrics (EI and PC/PR) were correlated (r = 0.58). The most central symptoms were Sadness (PC/PR) and Failure/Worthlessness (EI). Failure/Worthlessness had a significantly higher EI centrality than twenty-one other symptoms (see supplementary material). The next most central nodes (EI) were sadness, self-criticism, and loss of energy (all z-score > 1), followed by concentration, loss of pleasure/interest, and fatigue (z-score > 0.96). While the next most central nodes when using PC/PR were pessimism, failure/worthlessness, and punishment (all z-score > 1), then guilt, indecisiveness, and suicidal thoughts (all z-score > 0.80). Notably, while suicidal thoughts were highly central according the PC/PR metric (z-score = 0.80) it was much less central using EI (z-score = -0.67). Loss of energy displayed the opposite relationship, more central for EI (z-score = 1.01) than PC/PR (z-score = -2.03). Loss of energy and obsessions were jointly the least central nodes using PC/PR, and obsessions was also the least central when using EI.

Robustness checks suggest the resulting Graphical Gaussian Model was stable and accurate. Stability and accuracy plots, individual networks (with the fused penalty) and the fused network model are supplied in the supplementary materials. Mean severity was not significantly correlated ($p < 0.05$) with EI ($r = 0.21$) or PC/PR ($r = -0.05$), while the standard deviation was significantly correlated for both EI ($r = -0.56$) and PC/PR ($r = -0.41$). Symptom severity was not associated with nodes being interconnected. Lower variability was associated with variability, which is the reverse of a more typical concern: differential variability driving the centrality of nodes (52).

The interrelation of the network and the FGL network were compared (r = 0.72), suggesting that between study differences had a small effect on network estimation. The network was corrected for the influence of duration of depression and anxiety, however the overall influence on edge estimation was negligible (interrelation between the corrected network and a network estimated without duration variables: r = 0.997). Overall, the resulting network model can be considered robust.

Network Comparison Test

Networks (unregularized) were compared (1000 iterations) for those who were classified as in remission (n= 956) and those who were not (n = 1466). Mean severity differences at baseline were significant for all items ($p<0.001$). The correlation between networks was high ($r=0.67$). While there were significant difference between edges, the overall networks (see supplementary material) did not differ in connectivity (global strength invariance: $p < 0.08$) and post hoc tests were not warranted. There was only evidence of one difference in centrality between the networks: somatic symptoms were more connected in the remitter network than the persister network ($p<0.001$).

Discussion

Individuals with depression also present with comorbidity and this could present an issue for depression treatment. Understanding how symptoms influence one another across traditional diagnostic boundaries, and how they influence important outcomes, may provide insights relevant to the assessment and treatment of mood disorders. In this study we initially examined the differential impact of individual symptoms on prognosis and assessed whether individual symptoms offer predictive value above sum scores. The item level models of outcomes post-treatment and the sum score models were similarly associated with outcomes at three to four and six to eight months but explained considerably more variance at nine to twelve months. Pessimism was consistently the most important predictor of future outcomes (independent of its mean), indicating that experiencing pessimism rather than severity of the symptom is responsible for this association. Secondly, we explored the functional relations among comorbid symptoms of depression and anxiety disorders using network analysis. The symptom network comprised of three communities clearly clustering into: anxiety items; depressive cognitions; and depressive physical symptoms. The primary bridge symptoms between communities were sadness; pessimism; and indecision. The most central symptoms across both centrality metrics were sadness and failure/worthless. Finally, we analysed differences in the symptom networks at study entry for patients that remitted vs. those whose depression persisted, after treatment. Network comparison revealed no overall differences in connectivity. Together, the present findings suggest the utility of item-level analysis in informing the content of assessments and consideration of individual items over and above scale scores when predicting prognosis.

Findings in context

Exploring the associations with treatment outcomes revealed that item-level prediction methods performed similarly to sum scores, and outperformed sum score models at the nine to twelve month endpoint. It's not clear why there is a difference at this timepoint, while it was not due to attrition between endpoints, it could be due to random variation. It may also reflect the course of depression following intervention, or the cyclical nature of depression such that individual items are better at predicting the relapse or maintenance of symptoms after benefits of treatment have faded, or where an amelioration of symptoms occurred due to further treatment post randomisation. There is an ongoing debate in the field whether central items derived from network models offer predictive utility beyond other items (71–73). Pessimism was not only the best predictor across outcomes, it was a central item (ranked 2nd on PC/PR and 6th on EI centrality) that acted as a bridge between communities and showed strong associations with sadness and failure/worthlessness. Sadness, comparatively, did not predict well across time points. It is worth noting, that sadness falls within the physical symptom community and pessimism within the cognitive community. The amenability to act on an emotion (sadness), is understandably less than that of a cognition (pessimism), a target of

cognitive therapy. While pessimism in association with a sense of failure / worthlessness may negatively impact treatment engagement (i.e. the motivation to sustain goal pursuit in the face of obstacles) (74). Given the central role and prognostic value of pessimism we might speculate that it is associated with treatment factors, where pessimism hinders some people making progress and may not be directly addressed by some psychological treatments.

Symptoms of anxiety and depression clustered into separate communities with certain symptoms acting as bridges between communities. The bridge symptoms are statistically relevant and theoretically linked: indecision is a symptom in the classifications of both depression and generalized anxiety disorder; pessimism overlaps with worry (75); and the strong cross-community edge of sadness to worry, was similar to findings in other studies (32,76). The results therefore provide evidence that these bridging symptoms may be important in the emergence of comorbidity between anxiety disorders and depression.

Planned comparisons of networks at study entry for those whose depression would persist versus those who would be in remission, revealed no overall difference in connectivity, in contrast to Van Borkulo et al. (77), but similar to Schweren et al. (78).

Overall, we found no correlation between centrality metrics and Shapley values. This extends prior work on the association between centrality and the prognostic utility of items (71). Failure / worthlessness was predictively important at three to four months, displayed high centrality and is suggested to be a key symptom in depression and anxiety (30). The predictive utility of health anxiety and somatic concerns may be considered alongside the observation from the network comparison where there was a difference in centrality with somatic concerns more connected in the remitter network. Health anxiety was in the upper quintile of variable importance across timepoints, but relatively unimportant in terms of centrality. Not surprisingly, given the conceptual overlap, with health anxiety, the strongest edge was with somatic concerns. As such, the degree of concern for one's health, or attention to somatic symptoms, whilst not playing a significant role within the maintenance of depression, may act as a motivational spur to engage with treatment (in this way enabling rather than disabling the individual). The absence of this anxiety may reflect an apathy about one's health which is not captured by the motivational item in the BDI. While the predictive modelling did consider the influence of each item independent of the other items, modelling the predictive value of individual items may be improved by examining the association between the changes at symptom level and the overall network (79,80).

The network derived in this study provides empirical phenomena that can be explained by principles of network theory. This requires interpreting the network as a causal system, even though we cannot infer temporal relationship between symptoms and there is an absence of causal mechanisms within the external field (e.g. environmental factors) (29). These limitations apply to most of the findings in the network literature, although overinterpretation is common (81). Holding this in mind, we can consider possible pathways and mediating role of symptoms through the network. For example, taking suicidal ideation as a clinically severe symptom, we can identify the shortest path from worry (82) passing through sadness (bridge), and from loss of pleasure/interest (83) to suicidal thoughts, passed through pessimism (bridge). It is possible that any causal effect between these connections may be part of a longer pathway within the network highlighting a need for greater attention to be given to symptom interactions.

The statistical model investigates a symptom level, transdiagnostic conceptualization of the symptom interactions for individuals diagnosed with depression participating in RCTs. These interventions are based on biological or psychological theories, most notably Beck's cognitive of theory of depression (84). Clinically, pragmatism trumps theoretical completeness; simple

interventions which achieve rapid change do not require a detailed appreciation of the potential underlying mechanisms. However, oversimplified theories may restrict the ability to identify causal patterns; and gaps emerge in practice where the model is suggested to not fit the patient (85). More process-driven interventions targeting shared features of disorders have been developed (86,87), yet there is no unifying theory. The findings presented may help bridge the gap between disorder-specific theories and more transdiagnostic theories. Considering how symptoms may interact can help clinicians and researchers to understand underlying processes, and in turn to conceptualise their patients' difficulties in a way that supplements existing knowledge. A functional analysis which integrates the association between sadness and worry does not need to conceptualise the individual as having two disorders, but can consider how, for the individual, this interaction is being fueled and may be contributing to their distress.

The journey to develop models that provide both explanatory and predictive utility, will lead to greater understanding of psychopathology (88). While the analysis presented is primarily exploratory, it sets up clear testable hypotheses. These can be derived by examining the central structures within the network, formulating hypotheses and testing on an independent sample (89). For instance, whether the bridge edges belonging to pessimism, sadness and indecisiveness re-emerge in an independent sample, or whether a discrete intervention targeting pessimism would alter the network structure and lead to improved outcome. These statistical methods may help inform how identifying pathways and targets may lead to improved treatments all dependent on better assessment of symptoms.

Strengths and Limitations

This study has clear strengths, making use of a large sample of individuals participating in RCTs for depression in primary care. The use of same assessment measures at study entry removed the need to harmonise data across different measures for the network. While this is less true of outcomes where issues of measurement errors arise from the use of PROMIS T-Score, the sensitivity analyses provided confidence in the results.

The demographic balance across samples may affect generalisability however five of the six trials were pragmatic trials more closely representative of patient populations. Most cases of depression are treated in primary care, and the studies being set in primary care, improve the potential generalisability to patients seen in this setting (90).

This study was limited to the use of aggregate/group level findings to inform within person processes. However, the presence of an RCT outcome variable affords us the ability to detect changes from one state (e.g. depressed) to another (e.g. remitted), which is typically not the case with idiographic research studies that collect cross-sectional data. Exploring the prognostics value of networks on deterioration of symptoms would extend the utility of network analysis. This would however require generating idiographic networks, where reliable estimation necessitates many timepoints (low sensitivity at 100 timepoints; (91).

The accuracy of the network is limited by the items included and those omitted. The network does not cover the breadth of comorbidity of symptoms across psychopathology and is missing other environmental variables. Social adversity is associated with worse treatment outcomes for some patients with depression, it can be important to assess for and address these issues in clinic, where possible, to mitigate the risks of poor prognoses (92). There is also the possibility that the centrality of sadness particularly, represents a strong association with a latent variable rather than a specific role within the network (93).

The network models adjusted for duration of depression and anxiety, and a sensitivity analysis assessed for the influence of between study variability, adding robustness to the findings. While RCTs are used in the analysis, treatment arms were not factored in, and treated as equivalent when estimating outcome. This may make the findings generalizable where

findings are applicable regardless of treatment offered especially as the treatments included reflect those commonly available in primary care. Controlling for treatment group within the outcome modelling and controlling for relevant covariates (e.g. age, gender and social economic status) would also have improved the robustness of the findings. Such adjustments would have been fitting where the emphasis was on developing the best predictive model, instead of comparing the predictive ability of symptoms vs. total scores. More comprehensive prediction modelling using the Dep-GP dataset has been conducted (94). Additionally, our modelling did not include train/test split, as the whole sample was used in estimation of the network models. While a true out-of-sample 'holdout' dataset would have provided an unbiased evaluation of model fit, and is the preferred method for evaluating such models (95), the internal cross-fold validation employed in the symptom level model offers a layer of robustness supporting the final model estimates (where overfitting presents an issue). This study focussed on item-level analysis in comparison to sum-scores, future comparisons with models which may measure latent constructs in other ways, could be informative.

Single item symptom measurement will have unknown reliability and construct validity. Equally, the restricted range (e.g. a four-point scale) may not adequately capture the range of symptom variance occurring in the sample. Symptom measurement on a broader scale may improve the prediction of changes over time.

Conclusions

Our study used samples from high-quality randomised controlled trials, and the findings can be generalised to adults with depression being treated in primary care. This study has reiterated the importance of assessing for both depressive and anxious symptoms among adults seeking treatment for depression, and that valuable information about prognosis can be gained by understanding the interrelations between individual symptoms; information which is not available when considering sum scores or baseline symptom severity alone. This may be particularly important to longer term outcomes from treatment. Treatment selection and application is often hampered by comorbid symptoms and considered to introduce 'complexity' (96). Considering the bidirectional relationship between symptoms, and associations which may be mediated by another symptom (e.g. a bridge symptom) may help to consider comorbidity as normative.

While specific symptoms and associations have been highlighted, the aim is not to offer simple heuristics to inform clinical judgement and decision making. The relative importance of the highlighted associations should not be overweighed. The aim is not to identify individual items, but to consider the network of interactions. The critical role of individual symptoms and their interactions give rise to the activation of the network through pathways and anxiety and depressive cognitive and physical symptoms may activate one another via these pathways. This network highlights how symptoms of depression and anxiety disorders influence one another. Clinically, there is a need for treatments to adequately assess and address comorbidity.

List of abbreviations

BDI-II – Beck depression invention (2nd edition)

CIS-R - revised clinical interview schedule

EBIC – extended bayesian information criterion

EI – expected influence

ENR – Elastic net regression

FGL – fused graphical lasso

GGM – gaussian graphical model

IPD – individual patient data

MAE – mean absolute error

PC - participation coefficient

PCA – principal component analysis

PHQ-9 – patient health questionnaire 9

PR - participation ratio

RMSE – root mean squared error

Declarations

Ethics approval and consent to participate.

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The data that support the findings of this study are available from the authors of the individual trial studies. Restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

Competing interests

Not applicable

Funding

This work was supported by the Wellcome Trust through a Clinical Research Fellowship to JEJB (201292/Z/16/Z), MQ Foundation (for ZDC: MQDS16/72), the Higher Education Funding Council for England, the National Institute of Health Research (NIHR), NIHR University College London Hospitals Biomedical Research Centre (RS and SP), NIHR Biomedical Research Centre at the University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol (NW and DK), University College London (GA, GL), University of Pennsylvania (RDR), Vanderbilt University (SDH), University of Southampton (TK), University of Exeter (EW), and University of York (SG). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

The included studies that make up the Dep-GP IPD database were funded by:

1. COBALT: The National Institute for Health Research Health Technology Assessment (NIHR HTA) programme (project number 06/404/02).

2. GENPOD: Medical Research Council and supported by the Mental Health Research Network.

3. IPCRESS: BUPA Foundation

4. MIR: National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme (project 11/129/76) and supported by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol

5. PANDA: NIHR Programme Grant for Applied Research (RP-PG-0610-10048).

6. TREAD: National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme.

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors were fully independent of their respective funders and had responsibility for the decision to submit for this manuscript for publication. The guarantor accepts full responsibility for the work and the conduct of the study, had access to the data, and controlled the decision to publish.

Authors' contributions

Conceptualization (CO'D, JB, EF, RS), data curation (JB), statistical analysis (CO'D, JB), draft preparation (CO'D), review and editing (JB, EF, RS, ZC, GA, RD, SG, SH, TK, DK, GL, EW, NW, SP). All authors contributed to and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Arroll B, Chin WY, Martis W, Goodyear-Smith F, Mount V, Kingsford D, et al. Antidepressants for treatment of depression in primary care: A systematic review and meta-analysis. J Prim Health Care. 2016;8(4):325–34.

2. Santoft F, Axelsson E, Öst LG, Hedman-Lagerlöf M, Fust J, Hedman-Lagerlöf E. Cognitive behaviour therapy for depression in primary care: Systematic review and meta-analysis. Psychol Med. 2019;49(8):1266–74.

3. Cuijpers P, van Straten a, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med. 2010 Feb;40(2):211–23.

4. Eronen MI. Causal discovery and the problem of psychological interventions. New Ideas Psychol [Internet]. 2020;59(March):100785. Available from: https://doi.org/10.1016/j.newideapsych.2020.100785

5. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D. What are "good" depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. J Affect Disord. 2016 Jan;189:314–20.

6. Cuijpers P, Reynolds CF, Donker T, Li J, Andersson G, Beekman A. Personalized treatment of adult depression: Medication, psychotherapy, or both? A systematic review. Depress Anxiety. 2012;29(10):855–64.

7. Hardeveld F, Spijker J, De Graaf R, Nolen WA, Beekman ATF. Recurrence of major depressive disorder and its predictors in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Psychol Med. 2013 Jan 31;43(1):39–48.

8. Simon GE, Perlis RH. Personalized Medicine for Depression: Can We Match Patients With Treatments? Am J Psychiatry [Internet]. 2010 Dec;167(12):1445–55. Available from: http://psychiatryonline.org/doi/abs/10.1176/appi.ajp.2010.09111680

9. Cummings CM, Caporino NE, Kendall PC. Comorbidity of anxiety and depression in children and adolescents: 20 years after. Psychol Bull [Internet]. 2014 May;140(3):816–45. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0034733

10. Brown TA, Barlow DH. A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: Implications for assessment and treatment. Psychol Assess. 2009;21(3):256–71.

11. Kessler RC, Ormel J, Petukhova M, McLaughlin KA, Green JG, Russo LJ, et al. Development of Lifetime Comorbidity in the World Health Organization World Mental Health Surveys. Arch Gen Psychiatry. 2011 Jan 3;68(1):90.

12. Kendler KS. What psychiatric genetics has taught us about the nature of psychiatric illness and what is left to learn. Mol Psychiatry. 2013 Oct 30;18(10):1058–66.

13. Wichers M. The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. Psychol Med. 2014;44(7):1349–60.

14. Fried EI, Nesse RM. Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. J Affect Disord. 2015 Feb;172:96–102.

15. Fried EI. The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. J Affect Disord [Internet]. 2017

Jan;208(September 2016):191–7. Available from: http://dx.doi.org/10.1016/j.jad.2016.10.019

16. Bekhuis E, Schoevers R, De Boer M, Peen J, Dekker J, Ven H, et al. Symptom-Specific Effects of Psychotherapy versus Combined Therapy in the Treatment of Mild to Moderate Depression: A Network Approach. Psychother Psychosom. 2018;87(2):121–3.

17. Boschloo L, Cuijpers P, Karyotaki E, Berger T, Moritz S, Meyer B, et al. Symptom-specific effectiveness of an internet-based intervention in the treatment of mild to moderate depressive symptomatology: The potential of network estimation techniques. Behav Res Ther [Internet]. 2019 Nov;122:103440. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0005796719301263

18. Fried EI, Nesse RM. The Impact of Individual Depressive Symptoms on Impairment of Psychosocial Functioning. Gong Q, editor. PLoS One. 2014 Feb 28;9(2):e90311.

19. Fried E, von Stockert S, Haslbeck JMB, Lamers F, Schoevers RA, Penninx BWJH. Using network analysis to examine links between individual depressive symptoms, inflammatory markers, and covariates. Psychol Med [Internet]. 2019 Oct 28;1–9. Available from: https://www.cambridge.org/core/product/identifier/S0033291719002770/type/journal_article

20. Gollan JK, Fava M, Kurian B, Wisniewski SR, Rush AJ, Daly E, et al. What are the clinical implications of new onset or worsening anxiety during the first two weeks of SSRI treatment for depression? Depress Anxiety. 2012 Feb;29(2):94–101.

21. Buckman JEJ, Underwood A, Clarke K, Saunders R, Hollon SD, Fearon P, et al. Risk factors for relapse and recurrence of depression in adults and how they operate: A four-phase systematic review and meta-synthesis. Clin Psychol Rev. 2018 Aug;64(7):13–38.

22. Judd LL. The Role of Residual Subthreshold Depressive Symptoms in Early Episode Relapse in Unipolar Major Depressive Disorder. Arch Gen Psychiatry. 1999 Aug 1;56(8):764–5.

23. Webb CA, Cohen ZD, Beard C, Forgeard M, Peckham AD, Björgvinsson T. Personalized prognostic prediction of treatment outcome for depressed patients in a naturalistic psychiatric hospital setting: A comparison of machine learning approaches. J Consult Clin Psychol [Internet]. 2020 Jan;88(1):25–38. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/ccp0000451

24. Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. BMC Med. 2015 Dec 6;13(1):72.

25. Fried E. Moving forward: how depression heterogeneity hinders progress in treatment and research. Expert Rev Neurother [Internet]. 2017 May 4;17(5):423–5. Available from: http://dx.doi.org/10.1080/14737175.2017.1307737

26. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annu Rev Clin Psychol. 2018 May 7;14(1):91–118.

27. DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. The personalized advantage index: Translating research on prediction into individualized treatment recommendations. A demonstration. PLoS One. 2014;9(1):1–8.

28. Fernandez KC, Fisher AJ, Chi C. Development and initial implementation of the

dynamic assessment treatment algorithm (DATA). PLoS One. 2017;12(6):1–16.

29. Borsboom D. A network theory of mental disorders. World Psychiatry. 2017 Feb;16(1):5–13.

30. Heeren A, Bernstein EE, McNally RJ. Deconstructing trait anxiety: a network perspective. Anxiety, Stress Coping. 2018;31(3):262–76.

31. Cramer AOJ, Waldorp LJ, van der Maas HLJ, Borsboom D. Comorbidity: A network perspective. Behav Brain Sci. 2010 Jun 29;33(2–3):137–50.

32. Beard C, Millner AJ, Forgeard MJC, Fried EI, Hsu KJ, Treadway MT, et al. Network analysis of depression and anxiety symptom relationships in a psychiatric sample. Psychol Med. 2016 Dec 14;46(16):3359–69.

33. Boschloo L, van Borkulo CD, Borsboom D, Schoevers RA. A Prospective Study on How Symptoms in a Network Predict the Onset of Depression. Psychother Psychosom. 2016;85(3):183–4.

34. Kendler KS, Campbell J. Interventionist causal models in psychiatry: repositioning the mind–body problem. Psychol Med. 2009 Jun 10;39(6):881–7.

35. Newson JJ, Hunter D, Thiagarajan TC. The Heterogeneity of Mental Health Assessment. Front Psychiatry. 2020;11(February):1–24.

36. Buckman JEJ, Saunders R, Cohen ZD, Clarke K, Ambler G, DeRubeis RJ, et al. What factors indicate prognosis for adults with depression in primary care? A protocol for meta-analyses of individual patient data using the Dep-GP database. Wellcome Open Res [Internet]. 2020 Apr 1;4:69. Available from: https://wellcomeopenresearch.org/articles/4-69/v3

37. Beck AT, Steer RA, Brown GK. Beck depression inventory-II. San Antonio. 1996;78(2):490–8.

38. Thomas LJ, Abel A, Ridgway N, Peters T, Kessler D, Hollinghurst S, et al. Cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment resistant depression in primary care: The CoBalT randomised controlled trial protocol. Contemp Clin Trials [Internet]. 2012;33(2):312–9. Available from: http://dx.doi.org/10.1016/j.cct.2011.10.016

39. Lewis G, Mulligan J, Wiles N, Cowen P, Craddock N, Ikeda M, et al. Polymorphism of the 5-HT transporter and response to antidepressants: Randomised controlled trial. Br J Psychiatry. 2011;198(6):464–71.

40. Lewis G, Duffy L, Ades A, Amos R, Araya R, Brabyn S, et al. The clinical effectiveness of sertraline in primary care and the role of depression severity and duration (PANDA): a pragmatic, double-blind, placebo-controlled randomised trial. The Lancet Psychiatry. 2019 Nov;6(11):903–14.

41. Chalder M, Wiles NJ, Campbell J, Hollinghurst SP, Searle A, Haase AM, et al. A pragmatic randomised controlled trial to evaluate the cost-effectiveness of a physical activity intervention as a treatment for depression: The treating depression with physical activity (TREAD) trial. Health Technol Assess (Rockv). 2012;16(10).

42. Kessler, David S., MacNeill SJ, Tallon D, Lewis G, Peters TJ, Hollingworth W, et al. Mirtazapine added to SSRIs or SNRIs for treatment resistant depression in primary care: phase III randomised placebo controlled trial (MIR). BMJ [Internet]. 2018 Oct 31;k4218. Available from: http://www.bmj.com/lookup/doi/10.1136/bmj.k4218

43. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, et al. Therapist-delivered

internet psychotherapy for depression in primary care: a randomised controlled trial. Lancet. 2009 Aug;374(9690):628–34.

44. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. Psychol Med [Internet]. 1992 May 9;22(02):465. Available from: http://www.journals.cambridge.org/abstract_S0033291700030415

45. Buckman JEJ, Saunders R, Cohen ZD, Barnett P, Clarke K, Ambler G, et al. Indicators of Prognosis Independent of Treatment for Adults with Depression in Primary Care, Going Beyond Baseline Symptom-Severity: A Systematic Review and Individual Patient Data Meta-Analysis. SSRN [Internet]. 2020; Available from: https://www.ssrn.com/abstract=3520082

46. Choi SW, Schalet B, Cook KF, Cella D. Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS Depression. Psychol Assess [Internet]. 2014;26(2):513–27. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0035768

47. Fischer HF, Rose M. Www.common-metrics.org: A web application to estimate scores from different patient-reported outcome measures on a common scale. BMC Med Res Methodol [Internet]. 2016;16(1):1–5. Available from: http://dx.doi.org/10.1186/s12874-016-0241-0

48. R Core Team. R: A language and environment for statistical computing. [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2017. Available from: https://www.r-project.org

49. StataCorp. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.; 2019.

50. Fried EI, Cramer AOJ. Moving Forward: Challenges and Directions for Psychopathological Network Theory and Methodology. Perspect Psychol Sci [Internet]. 2017 Nov 5;12(6):999–1020. Available from: http://journals.sagepub.com/doi/10.1177/1745691617705892

51. Jones P. Tools for Identifying Important Nodes in Networks. CRAN; 2018.

52. Terluin B, De Boer MR, De Vet HCW. Differences in connection strength between mental symptoms might be explained by differences in variance: Reanalysis of network data did not confirm staging. PLoS One. 2016;11(11):1–12.

53. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw. 2010;33(1):1–22.

54. Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: Advances in Neural Information Processing Systems 30. .; 2017. p. 4765–74.

55. Molnar C. Interpretable machine learning. Leanpub; 2019.

56. van Borkulo C, Boschloo L, Borsboom D, Penninx BWJH, Waldorp LJ, Schoevers RA. Package 'NetworkComparisonTest.' JAMA Psychiatry. 2015;

57. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B. 2016;58:267–88.

58. Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph : Network Visualizations of Relationships in Psychometric Data. J Stat Softw [Internet]. 2012;48(4). Available from: http://www.jstatsoft.org/v48/i04/

59. Williams DR, Rast P. Back to the basics: Rethinking partial correlation network

methodology. Br J Math Stat Psychol. 2020;73(2):187–212.

60. Foygel R, Drton M. Extended Bayesian information criteria for Gaussian graphical models. Adv Neural Inf Process Syst. 2010;604–12.

61. Lorenzo-Luaces L, Rodriguez-Quintana N, Bailey AJ. Double trouble: Do depression severity and duration interact to predicting treatment outcomes in adolescent depression? Behav Res Ther. 2020;103637.

62. Costantini G, Richetin J, Preti E, Casini E, Epskamp S, Perugini M. Stability and variability of personality networks . A tutorial on recent developments in network psychometrics. Pers Individ Dif [Internet]. 2019;136:68–78. Available from: https://doi.org/10.1016/j.paid.2017.06.011

63. Fried EI, Eidhof MB, Palic S, Costantini G, Huisman-van Dijk HM, Bockting CLH, et al. Replicability and Generalizability of Posttraumatic Stress Disorder (PTSD) Networks: A Cross-Cultural Multisite Study of PTSD Symptoms in Four Trauma Patient Samples. Clin Psychol Sci [Internet]. 2018 May 5;6(3):335–51. Available from: http://journals.sagepub.com/doi/10.1177/2167702617745092

64. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. Behav Res Methods [Internet]. 2018 Feb 24;50(1):195–212. Available from: http://link.springer.com/10.3758/s13428-017-0862-1

65. Golino H, Shi D, Christensen AP, Garrido LE, Nieto MD, Sadana R, et al. Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. Psychol Methods [Internet]. 2020 Mar 19; Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/met0000255

66. Borsboom D, Cramer AOJ. Network Analysis: An Integrative Approach to the Structure of Psychopathology. Annu Rev Clin Psychol [Internet]. 2013 Mar 28;9(1):91–121. Available from: http://www.annualreviews.org/doi/10.1146/annurev-clinpsy-050212-185608

67. Bringmann LF, Elmer T, Epskamp S, Krause RW, Schoch D, Wichers M, et al. What do centrality measures measure in psychological networks? J Abnorm Psychol [Internet]. 2019 Nov;128(8):892–903. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/abn0000446

68. Robinaugh DJ, LeBlanc NJ, Vuletich HA, McNally RJ. Network analysis of persistent complex bereavement disorder in conjugally bereaved adults. J Abnorm Psychol [Internet]. 2014;123(3):510–22. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/abn0000002

69. Jones P, Ma R, McNally RJ. Bridge Centrality: A Network Approach to Understanding Comorbidity. Multivariate Behav Res [Internet]. 2019;0(0):1–15. Available from: https://doi.org/10.1080/00273171.2019.1614898

70. Letina S, Blanken TF, Deserno MK, Borsboom D. Expanding Network Analysis Tools in Psychological Networks: Minimal Spanning Trees, Participation Coefficients, and Motif Analysis Applied to a Network of 26 Psychological Attributes. Complexity. 2019;2019.

71. Elliott H, Jones PJ, Schmidt U. Central Symptoms Predict Posttreatment Outcomes and Clinical Impairment in Anorexia Nervosa: A Network Analysis. Clin Psychol Sci. 2020 Jan 11;8(1):139–54.

72. Rodebaugh TL, Tonge NA, Piccirillo ML, Fried E, Horenstein A, Morrison AS, et al.

Does centrality in a cross-sectional network suggest intervention targets for social anxiety disorder? J Consult Clin Psychol [Internet]. 2018 Oct;86(10):831–44. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/ccp0000336

73. Spiller TR, Levi O, Neria Y, Suarez-Jimenez B, Bar-Haim Y, Lazarov A. On the validity of the centrality hypothesis in cross-sectional between-subject networks of psychopathology. BMC Med. 2020;18(1):1–14.

74. Dickson JM, Moberly NJ, O'Dea C, Field M. Goal fluency, pessimism and disengagement in depression. PLoS One. 2016;11(11):1–11.

75. MacLeod AK, Williams JM, Bekerian DA. Worry is reasonable: The role of explanations in pessimism about future personal events. J Abnorm Psychol [Internet]. 1991;100(4):478–86. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/0021-843X.100.4.478

76. Price M, Legrand AC, Brier ZMF, Hébert-Dufresne L. The symptoms at the center: Examining the comorbidity of posttraumatic stress disorder, generalized anxiety disorder, and depression with network analysis. J Psychiatr Res [Internet]. 2019;109(June 2018):52–8. Available from: https://doi.org/10.1016/j.jpsychires.2018.11.016

77. Van Borkulo C, Boschloo L, Borsboom D, Penninx BWJH, Lourens JW, Schoevers RA. Association of symptom network structure with the course of longitudinal depression. JAMA Psychiatry. 2015;72(12):1219–26.

78. Schweren L, Van Borkulo CD, Fried E, Goodyer IM. Assessment of symptom network density as a prognostic marker of treatment response in adolescent depression. JAMA Psychiatry. 2018;75(1):98–100.

79. Robinaugh DJ, Millner AJ, McNally RJ. Identifying highly influential nodes in the complicated grief network. J Abnorm Psychol [Internet]. 2016;125(6):747–57. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/abn0000181

80. Papini S, Rubin M, Telch MJ, Smits JAJ, Hien DA. Pretreatment Posttraumatic Stress Disorder Symptom Network Metrics Predict the Strength of the Association Between Node Change and Network Change During Treatment. J Trauma Stress. 2020;33(1):64–71.

81. Fried EI. Lack of theory building and testing impedes progress in the factor and network literature. PsyArXiv. 2020;

82. Law KC, Tucker RP. Repetitive negative thinking and suicide: a burgeoning literature with need for further exploration. Curr Opin Psychol [Internet]. 2018;22:68–72. Available from: http://dx.doi.org/10.1016/j.copsyc.2017.08.027

83. Ducasse D, Loas G, Dassa D, Gramaglia C, Zeppegno P, Guillaume S, et al. Anhedonia is associated with suicidal ideation independently of depression: A meta-analysis. Depress Anxiety. 2018;35(5):382–92.

84. Beck AT. Depression: Clinical, experimental, and theoretical aspects. Harper & Row; 1967.

85. Roth A, Fonagy P. What works for whom?: a critical review of psychotherapy research. Guilford Press; 2006.

86. Barlow DH, Allen LB, Choate ML. Toward a Unified Treatment for Emotional Disorders – Republished Article. Behav Ther [Internet]. 2016;47(6):838–53. Available from: http://dx.doi.org/10.1016/j.beth.2016.11.005

87. Hvenegaard M, Moeller SB, Poulsen S, Gondan M, Grafton B, Austin SF, et al. Group rumination-focused cognitive-behavioural therapy (CBT) v. group CBT for depression: phase II trial. Psychol Med [Internet]. 2020 Jan 11;50(1):11–9. Available from: https://www.taylorfrancis.com/books/9781000102147/chapters/10.4324/9781003076391-177

88. Yarkoni T, Westfall J. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. Perspect Psychol Sci. 2017;12(6):1100–22.

89. Williams DR, Rast P, Pericchi LR, Mulder J. Comparing Gaussian graphical models with the posterior predictive distribution and Bayesian model selection. Psychol Methods [Internet]. 2020 Oct;25(5):653–72. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/met0000254

90. McManus S, Bebbington P, Jenkins R, Brugha T. Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014. Leeds; 2016.

91. Mansueto, A. C., Wiers, R., van Weert, J., Schouten, B. C., Epskamp S. Investigating the Feasibility of Idiographic Network Models. 2020;

92. Buckman JEJ, Saunders R, O'Driscoll C, Cohen ZD, Stott J, Ambler G, et al. Is social support pre-treatment associated with prognosis for adults with depression in primary care? Acta Psychiatr Scand [Internet]. 2021 Feb 16;acps.13285. Available from: https://onlinelibrary.wiley.com/doi/10.1111/acps.13285

93. Hallquist MN, Wright AGC, Molenaar PCM. Problems with Centrality Measures in Psychopathology Symptom Networks: Why Network Psychometrics Cannot Escape Psychometric Theory. Multivariate Behav Res [Internet]. 2019;0(0):1–25. Available from: https://doi.org/10.1080/00273171.2019.1640103

94. Buckman JEJ, Cohen ZD, O'Driscoll C, Fried EI, Saunders R, Ambler G. . ., et al. Predicting prognosis for adults with depression using individual symptom data: a comparison of modelling approaches. 2020; Available from: https://doi.org/10.31219/osf.io/xkwdc

95. Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal–external, and external validation. J Clin Epidemiol [Internet]. 2016 Jan;69(1):245–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0895435615001754

96. Kuyken W, Tsivrikos D. Therapist competence, comorbidity and cognitive-behavioral therapy for depression. Psychother Psychosom. 2008;78(1):42–8.

Table 1 Click here to access/download;Table;Table1 .xlsx

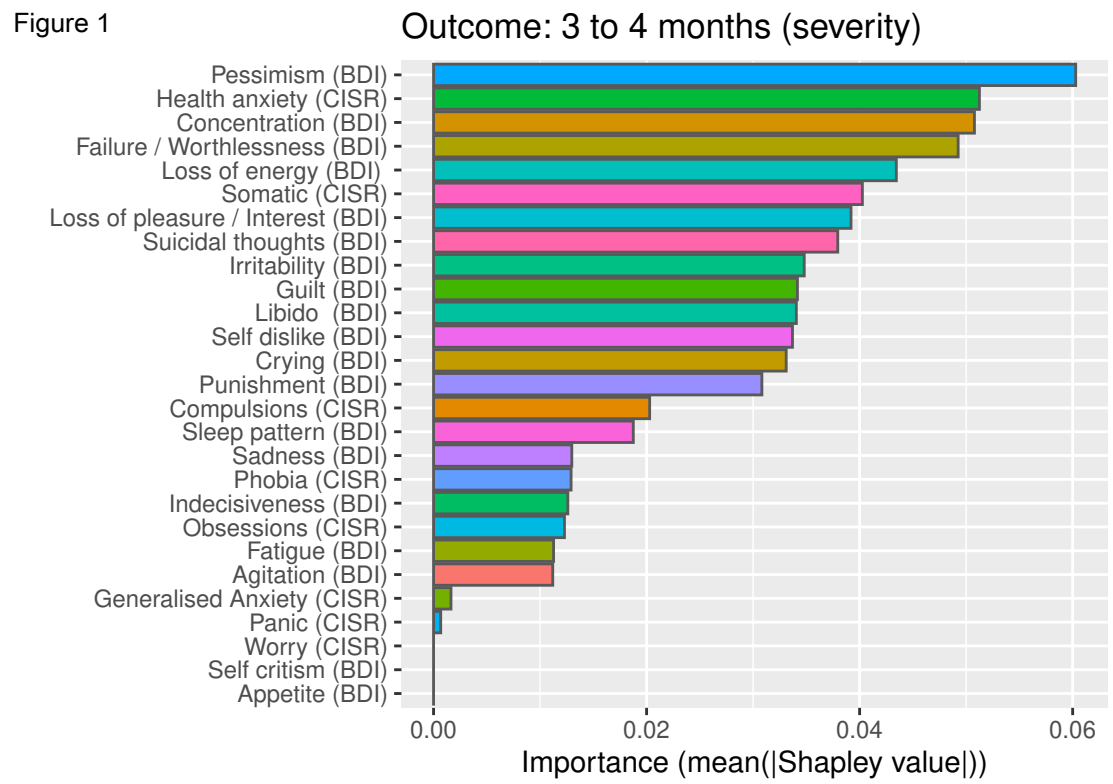
	COBALT (N=469)	GENPOD (N=601)	IPCRESS (N=295)
Baseline BDI-II total			
Mean (SD)	31.8 (10.7)	33.7 (9.67)	33.2 (8.80)
Median [Min, Max]	30.0 [14.0, 60.0]	33.0 [15.0, 60.0]	33.0 [15.0, 58.0]
Gender			
Female	339 (72.3%)	408 (67.9%)	200 (67.8%)
Male	130 (27.7%)	193 (32.1%)	95 (32.2%)
Age			
Mean (SD)	49.6 (11.7)	38.8 (12.4)	34.9 (11.6)
Median [Min, Max]	50.0 [18.0, 74.0]	38.0 [18.0, 74.0]	34.0 [18.8, 74.6]
Employment Status			
Employed	206 (43.9%)	357 (59.4%)	178 (60.3%)
Seeking employment	151 (32.2%)	123 (20.5%)	35 (11.9%)
Not seeking employment	112 (23.9%)	121 (20.1%)	82 (27.8%)
Education			
Degree or higher	95 (20.3%)	0 (0%)	102 (34.6%)
A-level or Diplomas*	123 (26.2%)	0 (0%)	88 (29.8%)
GCSE**	131 (27.9%)	0 (0%)	62 (21.0%)
None or Other	120 (25.6%)	0 (0%)	43 (14.6%)
Missing	0 (0%)	601 (100%)	0 (0%)
Ethnicity			
White	459 (97.9%)	575 (95.7%)	281 (95.3%)
Non-White	10 (2.1%)	26 (4.3%)	14 (4.7%)
Diagnoses			
Number of Comorbid Diagnoses	2.40 (1.09)	2.39 (0.92)	2.32 (0.99)
Generalized Anxiety Disorder	312 (66.52%)	410 (68.22%)	186 (63.05%)
OCD	79 (16.84%)	114 (18.97%)	62 (21.02%)
Panic Disorder	67 (14.29%)	51 (8.49%)	16 (5.42%)
Agoraphobia	61 (13.01%)	75 (12.48%)	28 (9.49%)
Social Phobia	64 (13.65%)	64 (10.65%)	44 (14.92%)
Specific Phobias	91 (19.40%)	127 (21.13%)	46 (15.59%)
Chronic Fatigue Syndrome	343 (73.13%)	476 (79.20%)	220 (74.58%)

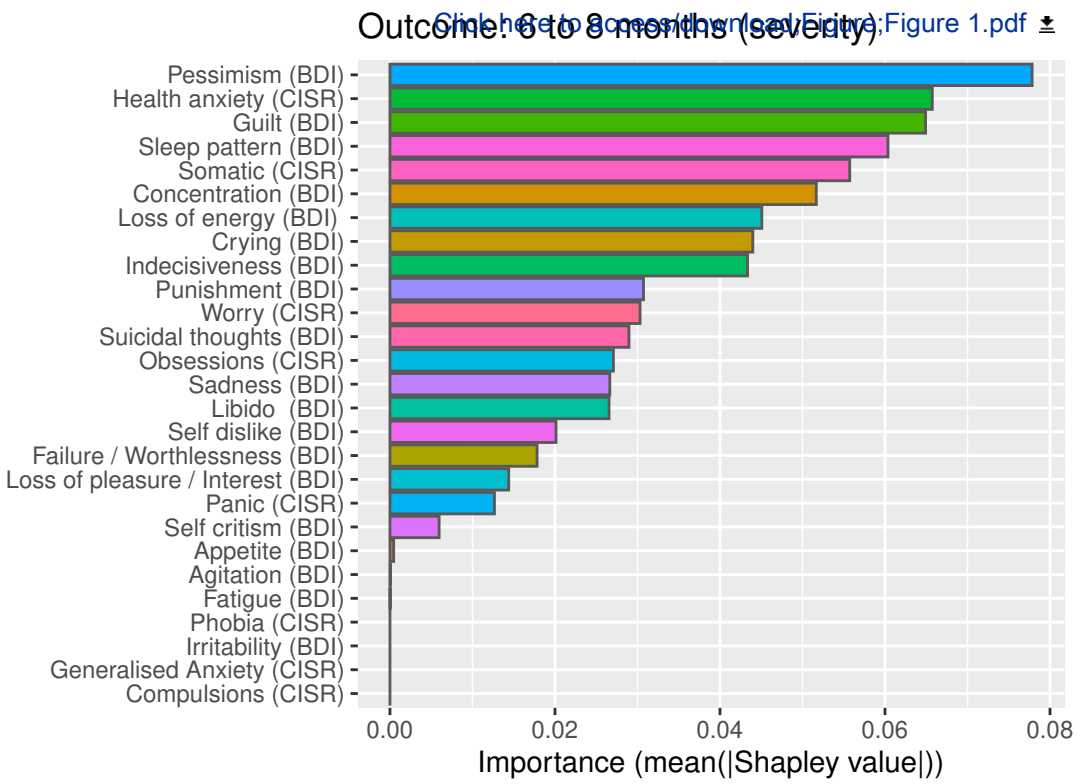
MIR (N=480)	PANDA (N=652)	TREAD (N=361)	Overall (N=2858)
31.1 (9.91)	23.9 (10.3)	32.1 (9.24)	30.4 (10.5)
30.0 [14.0, 58.0]	23.0 [2.00, 54.0]	31.0 [14.0, 57.0]	30.0 [2.00, 60.0]
332 (69.2%)	384 (58.9%)	239 (66.2%)	1902 (66.6%)
148 (30.8%)	268 (41.1%)	122 (33.8%)	956 (33.4%)
50.7 (13.2)	39.7 (15.0)	39.8 (12.6)	42.5 (14.1)
51.0 [19.0, 84.0]	38.5 [18.0, 73.0]	39.0 [18.0, 69.0]	42.0 [18.0, 84.0]
237 (49.4%)	433 (66.4%)	230 (63.7%)	1641 (57.4%)
102 (21.2%)	73 (11.2%)	48 (13.3%)	532 (18.6%)
141 (29.4%)	146 (22.4%)	83 (23.0%)	685 (24.0%)
95 (19.8%)	230 (35.3%)	87 (24.1%)	609 (21.3%)
135 (28.1%)	220 (33.7%)	104 (28.8%)	670 (23.4%)
150 (31.2%)	145 (22.2%)	102 (28.3%)	590 (20.6%)
100 (20.8%)	57 (8.7%)	68 (18.8%)	388 (13.6%)
0 (0%)	0 (0%)	0 (0%)	601 (21.0%)
469 (97.7%)	579 (88.8%)	336 (93.1%)	2699 (94.4%)
11 (2.3%)	73 (11.2%)	25 (6.9%)	159 (5.6%)
2.10 (0.97)	1.43 (1.18)	2.20 (1.17)	2.09 (1.12)
219 (45.63%)	299 (45.86%)	238 (65.93%)	1664 (58.2%)
62 (12.92%)	52 (7.98%)	50 (13.85%)	419 (14.7%)
45 (9.38%)	42 (6.44%)	14 (3.88%)	235 (8.2%)
81 (16.88%)	42 (6.44%)	35 (9.70%)	322 (11.3%)
58 (12.08%)	68 (10.43%)	52 (14.40%)	350 (12.2%)
62 (12.92%)	98 (15.03%)	61 (16.90%)	485 (17%)
311 (64.79%)	288 (44.17%)	257 (71.19%)	1895 (66.3%)

Table 2

		PROMIS T-Score		
		RMSE	R^2	MAE
3 to 4 months N=2646	Items	0.925	0.146	0.73
	Sum scores	0.926	0.143	0.73
6 to 8 months N=1297	Items	0.926	0.147	0.734
	Sum scores	0.924	0.146	0.735
9 to 12 months N=1110	Items	0.919	0.161	0.744
	Sum scores	0.935	0.126	0.753

Figure 1


Outcome: 3 to 4 months (severity)
Pessimism (BDI)
Health anxiety (CISR)
Concentration (BDI)
Failure / Worthlessness (BDI)
Loss of energy (BDI)
Somatic (CISR)
Loss of pleasure / Interest (BDI)
Suicidal thoughts (BDI)
Irritability (BDI)
Guilt (BDI)
Libido (BDI)
Self dislike (BDI)
Crying (BDI)
Punishment (BDI)
Compulsions (CISR)
Sleep pattern (BDI)
Sadness (BDI)
Phobia (CISR)
Indecisiveness (BDI)
Obsessions (CISR)
Fatigue (BDI)
Agitation (BDI)
Generalised Anxiety (CISR)
Panic (CISR)
Worry (CISR)
Self critism (BDI)
Appetite (BDI)
0.00
0.02
0.04
0.06
Importance (mean(|Shapley value|))


Outcome: 6 to 8 months (severity)
Pessimism (BDI)
Health anxiety (CISR)
Guilt (BDI)
Sleep pattern (BDI)
Somatic (CISR)
Concentration (BDI)
Loss of energy (BDI)
Crying (BDI)
Indecisiveness (BDI)
Punishment (BDI)
Worry (CISR)
Suicidal thoughts (BDI)
Obsessions (CISR)
Sadness (BDI)
Libido (BDI)
Self dislike (BDI)
Failure / Worthlessness (BDI)
Loss of pleasure / Interest (BDI)
Panic (CISR)
Self critism (BDI)
Appetite (BDI)
Agitation (BDI)
Fatigue (BDI)
Phobia (CISR)
Irritability (BDI)
Generalised Anxiety (CISR)
Compulsions (CISR)
0.00
0.02
0.04
0.06
0.08
Importance (mean(|Shapley value|))

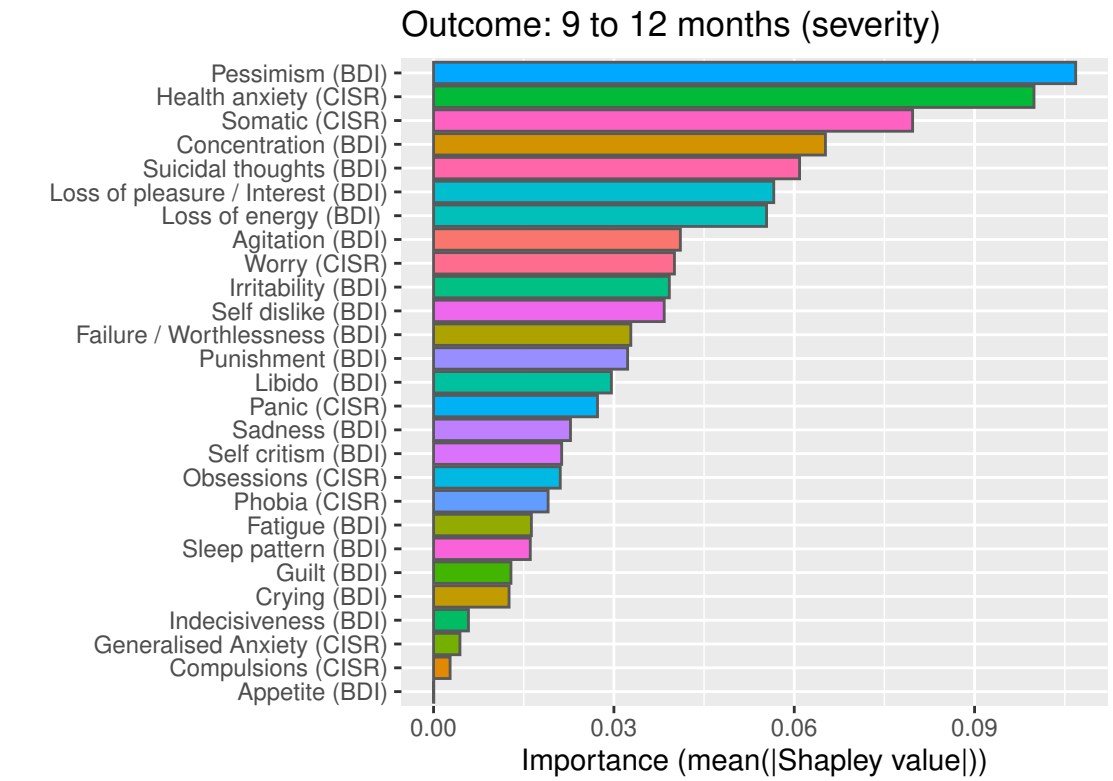

Outcome: 9 to 12 months (severity)
Pessimism (BDI)
Health anxiety (CISR)
Somatic (CISR)
Concentration (BDI)
Suicidal thoughts (BDI)
Loss of pleasure / Interest (BDI)
Loss of energy (BDI)
Agitation (BDI)
Worry (CISR)
Irritability (BDI)
Self dislike (BDI)
Failure / Worthlessness (BDI)
Punishment (BDI)
Libido (BDI)
Panic (CISR)
Sadness (BDI)
Self critism (BDI)
Obsessions (CISR)
Phobia (CISR)
Fatigue (BDI)
Sleep pattern (BDI)
Guilt (BDI)
Crying (BDI)
Indecisiveness (BDI)
Generalised Anxiety (CISR)
Compulsions (CISR)
Appetite (BDI)
0.00
0.03
0.06
0.09
Importance (mean(|Shapley value|))

Figure 2

Click here to access/download;Figure;Figure 2.png