Borderline personality disorder classification based on brain network measures during emotion regulation

Cremers, H.; van Zutphen, L.; Duken, S.; Domes, G.; Sprenger, A.; Waldorp, L.; Arntz, A.

ORIGINAL PAPER

Borderline personality disorder classification based on brain network measures during emotion regulation

Henk Cremers[1] · Linda van Zutphen[2] · Sascha Duken[1] · Gregor Domes[3] · Andreas Sprenger[4] · Lourens Waldorp[5] · Arnoud Arntz[1,2]

Abstract

Borderline Personality Disorder (BPD) is characterized by an increased emotional sensitivity and dysfunctional capacity to regulate emotions. While amygdala and prefrontal cortex interactions are regarded as the critical neural mechanisms underlying these problems, the empirical evidence hereof is inconsistent. In the current study, we aimed to systematically test different properties of brain connectivity and evaluate the predictive power to detect borderline personality disorder. Patients with borderline personality disorder ($n=51$), cluster C personality disorder ($n=26$) and non-patient controls ($n=44$), performed an fMRI emotion regulation task. Brain network analyses focused on two properties of task-related connectivity: phasic refers to task-event dependent changes in connectivity, while tonic was defined as task-stable background connectivity. Three different network measures were estimated (strength, local efficiency, and participation coefficient) and entered as separate models in a nested cross-validated linear support vector machine classification analysis. Borderline personality disorder vs. non-patient controls classification showed a balanced accuracy of 55%, which was not significant under a permutation null-model, $p=0.23$. Exploratory analyses did indicate that the tonic strength model was the highest performing model (balanced accuracy 62%), and the amygdala was one of the most important features. Despite being one of the largest data-sets in the field of BPD fMRI research, the sample size may have been limited for this type of classification analysis. The results and analytic procedures do provide starting points for future research, focusing on network measures of tonic connectivity, and potentially focusing on subgroups of BPD.

✉ Henk Cremers
    h.r.cremers@uva.nl

1   Department of Clinical Psychology, University of Amsterdam, Nieuwe Achtergracht 129B, 1001 NK Amsterdam, The Netherlands

2   Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands

3   Department of Biological and Clinical Psychology, University of Trier, Trier, Germany

4   Department of Neurology, University of Lübeck, Lübeck, Germany

5   Department of Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands

Introduction

Borderline Personality Disorder (BPD) is characterized by a pervasive pattern of instability of interpersonal relationships, self-image, affect, and impulse control [1]. Particularly, an increased sensitivity to emotions and a dysfunctional capacity to regulate emotions is considered to be one of the hallmark features of BPD [2]. Research generally focuses on increased amygdala and reduced prefrontal activity as the neural mechanism underlying these processes [3]. Furthermore, previous research suggests a disruption of functional connectivity between these brain areas (e.g., [4, 5]). While some research findings indeed fit this pattern, the inconsistency of results on this topic is perhaps even more notable [3]. For example, there is a large discrepancy between studies on the involvement of lateral and medial prefrontal regions in BPD during emotion regulation [3].

There are several potential reasons for this, such as small samples (low statistical power) and flexibility in analyses, which we aim to address in the current study.

Graph theory is a powerful framework to evaluate the highly interconnected nature of the brain [6], particularly by addressing the balance between the integration of a network on the one hand, and the segregation into modules on the other [6]. Network measures capture to what extend brain regions form important hubs that connect different submodules of a network [7] and alterations in these hub regions characterize several psychiatric disorders [8]. While network analyses are often based on resting-state functional Magnetic Resonance Imaging (fMRI) data, the principles can be extended to capture task-related network reconfiguration of a network [9]. Here we aim to investigate the predictive power of these different network properties to classify BPD patients.

Predictive modeling through machine learning algorithms [10] provides a framework to quantify how well the different network measures can classify BPD. It can be considered a two-step approach, where the first question is if there is any pattern to detect and classify a target variable. Subsequently, one can aim to assess the contributions (feature weights) of the input data, at least in case of linear classification models, which are most commonly used in fMRI [11]. To date, one study has applied a brain network approach to classifying BPD [12] and displayed a promising classification accuracy of 80%. Increased efficiency of the amygdala, entorhinal cortex, and temporal pole in BPD were some of the main discriminative features, which may reflect clinically well-observed borderline characteristics of emotion processing dysfunction. Yet, this study contained a small sample (20 patients, 10 controls), which necessitates further investigation, and whether or not such models can distinguish BPD from other personality disorders also remains to be investigated. Additionally, this work utilized resting-state connectivity, which does show overlap with task-related activation patterns [13], but does not directly tap into emotion regulation processes.

Here, we investigated whole-brain connectivity during an emotion regulation task, to assess task-state related shifts in network configuration [9, 14]. Specifically, we focused on delineating two characteristics: phasic connectivity: the task-event related changes in brain connectivity (see [9, 15, 16 for comparable approaches] and tonic connectivity: background connectivity of brain regions, unrelated to task-events. We build on our previous work [17, 18] and compare BPD to non-patient controls (NPC) and cluster-C personality disorder patients as a relevant clinical control group. Due to the importance of emotion regulation in BPD, we expected that the phasic connectivity shifts would show the highest classification accuracy. We further examined which brain regions were most important in the classification, particularly the

critical functional subnetwork modules related to psychiatric disorders: the emotion (including the amygdala), motivation, cognitive control, and default mode modules of the brain [19]. We hypothesized that BPD would show decreased cognitive control-related phasic connectivity during cognitive emotion regulation compared to controls.

Materials and methods

A detailed outline of the participant inclusion, the task, as well as fMRI data acquisition and preprocessing is described elsewhere [17, 18]. Below are summaries for each section.

Participants

Patients with BPD or Cluster-C personality disorders were recruited from mental health clinics at two sites in the Netherlands (Maastricht, Heerlen) and three sites in Germany (Freiburg, Lübeck, Hamburg), see "Brain parcellation and subject inclusion" of the results for the subject inclusion. Non-patient controls (NPC) were recruited from the general population at each site. Participants had to be hetero- or bisexual women aged 18–65 years, and were assessed by trained interviewers according to the DSM-IV criteria using the Structural Clinical Interview (SCID) II and I [20]. All participants provided written informed consent. The study was approved by the local ethical committees.

Task

Participants performed an adapted version of an emotion regulation paradigm, which involved the presentation of pictures that were preceded by a safe (emotion regulation) or look instruction. During the safe trials, participants were asked to imagine themselves as being in a safe situation. This regulation strategy was based on a central element of schema therapy, one of the main therapies for BPD [17]. There were 4 categories of picture stimuli (negative, neutral, positive, and erotic), and the task consisted of 96 trials divided into 4 runs of 24 trials each (see Fig. 1a for an outline of a trial). As part of the scanning session, participants also underwent resting-state scans before and after the task.

fMRI data

Acquisition and preprocessing

Functional and structural MRI data were acquired with 3 T scanners. The Functional images were acquired with a T2*-weighted echo planar imaging (EPI) sequence, with the following parameters: TR = 2000 ms, TE = 27 ms, flip angle = 90°, FoV = 192 × 192 mm, voxel
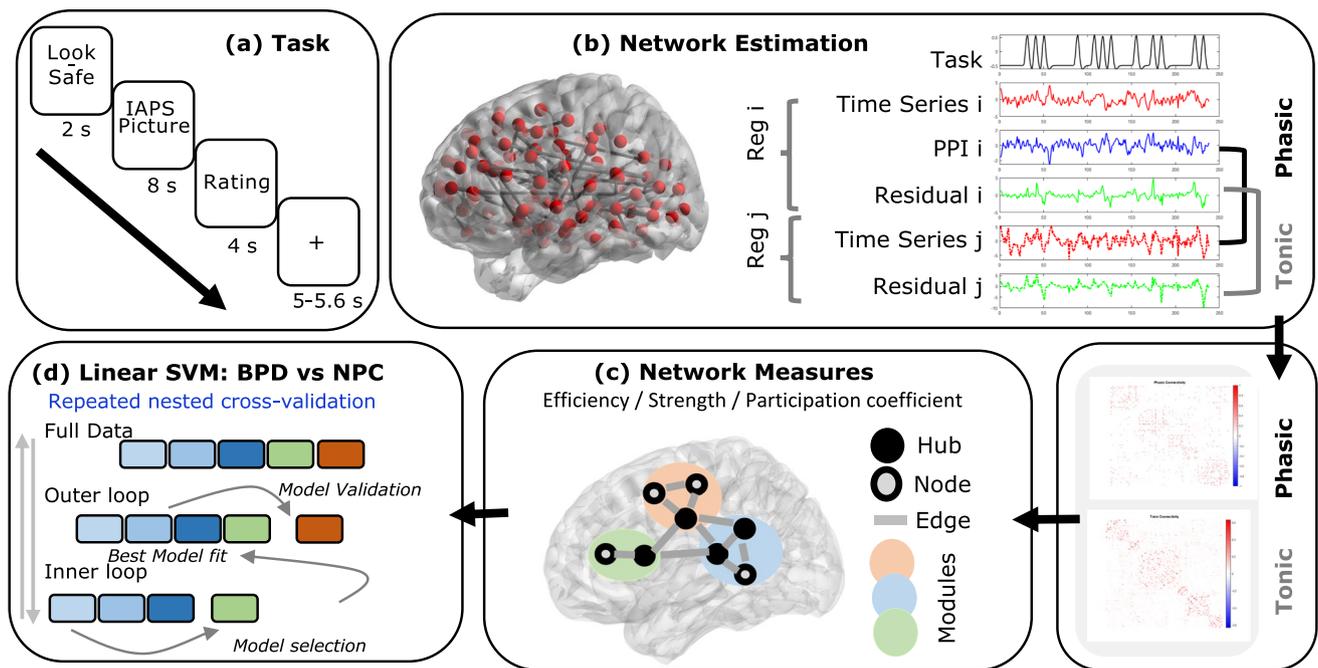
Fig. 1 Emotion regulation task and analysis pipeline. **a** Outline of the task: each trial consisted of an instruction (look or safe), IAPS picture presentation, rating period, and fixation-cross screen. **b** Brain regions (Reg) included in the analyses and the statistical models testing the connectivity between two regions (i and j). For each region, the time-series was extracted and multiplied with the task regressors (only one shown), which resulted in the psychophysiological interaction (PPI) terms. The phasic connectivity analyses referred to a regression of the PPI term on the time-series of another region while controlling for the task regressors, time-series, and confounds (not shown). The tonic connectivity analyses referred to the graphical lasso estimation of the residual times series of all regions, after correcting for the task regressors and confounding effects. Each connectivity analysis resulted in a connectivity matrix, which was averaged across runs. **c** Overview of the basic properties of a network consisting of nodes connected by edges and forming subnetworks (modules). Three network measures are used to assess hub functioning: strength, local efficiency, and the participation coefficient. **d** Linear support vector machine classification in a one vs. one fashion (e.g., BPD vs. NPC) using the different network measures from the phasic and tonic connectivity as input features. The different models (e.g., phasic strength, phasic local efficiency, tonic participation coefficient) were assed in a repeated (200 times) nested cross-validation procedure. In the inner loop, the best performing model was selected, re-fitted to the test data of the outer loop, and evaluated on the balanced accuracy of the validation data

size $= 3 \times 3 \times 3$ mm, and matrix $= 64 \times 64$. In Maastricht 240 images and in Freiburg and Lübeck 252 images were collected. The number of interleaved axial slices in one volume was 32 in Maastricht and 34 in Freiburg and Lübeck. In Maastricht and Freiburg, the T2*-weighted slices were adjusted with a negative tilt of 30°, with the goal of minimizing susceptibility and distortion artifacts within the amygdala. The anatomical images were acquired with a T1-weighted sequence, with the following parameters: TR $= 2250$ ms, TE $= 2.6$ ms, flip angle $= 9°$, Field of View (FoV) $= 256 \times 256$ mm, voxel size $1 \times 1 \times 1$ mm. In total, 192 images were obtained in Maastricht, 160 in Freiburg, and 170 in Lübeck. The preprocessed images from a previous study were used (see [17] for details of the preprocessing pipeline) and

transformed from BrainVoyager format into nifti format, using Neuroelf (www.neuroelf.net).

Analyses

Network estimation Phasic brain connectivity networks were estimated by performing a whole-brain psychophysiological interaction (PPI) analysis. PPI analyses aim to identify brain regions, where the time-series (physiological signal) connectivity is moderated by a task condition ("the psychological variable") [21]. Traditionally, this approach has consisted of predefining a source (or seed) region and then estimating its connectivity with other (target) regions. We extended this rationale to a whole-brain network method (see [15, 16] for similar approaches). Here, each brain node is once considered to be the source, while the other nodes

are the target. The PPI terms were estimated separately for each task regressor, but since PPI analyses are in general less powerful compared to estimating activation (i.e., the main effect of a task regressor) [22], we opted for a summary model that collapsed the different emotional valences (negative, neutral, positive), and included the general look and safe condition. Each model contained the time-series of the source region, the task regressors (safe and look), the interaction terms (PPIs) and the confound regressors (motion parameters, cue presentation, and ratings) (Fig. 1b). This procedure resulted in two different brain connectivity matrices for each run, consisting of contrast estimates for the safe and look condition. The connectivity matrices were then made symmetric by averaging corresponding (a-b, and b-a) parameter estimates, and were subsequently averaged across the four different runs. Subsequently the absolute values of the connectivity matrices were taken, and the connectivity matrices were additionally thresholded at 5%, (i.e., retaining only the strongest 5% of the connection).

To estimate the tonic brain connectivity an approach was followed as described by [15, 23]. For each time-series the task-related variance was "removed" by regressing a model containing task regressors and confounds (motion parameter, cue presentation, and ratings) on the time-series, and subsequently using the residual time-series for further analyses (Fig. 1b). The connectivity matrix was estimated by applying a graphical lasso [24] to the residuals of all nodes with the function graphicallasso.m (https://statweb.stanford.edu/~tibs/glasso) for a range of regularization parameters, lambda. The optimal regularization parameter lambda was estimated for each network (each run and each subject) by minimizing the Bayesian Information Criterion [25, 26]. These connectivity matrices were then averaged across the four different runs. The absolute value of the connectivity matrix was taken and entered in a graph theoretical analysis.

As a control to the connectivity analyses, the task-related activity per region was also estimated by a basic regression analyses of the task condition (safe and look, and safe > look) on the time-series of each region.

Functional module assignment Each of the resulting brain nodes was assigned to a higher level functional module as proposed in [19]. The functional module assignment was performed as follows: the term related to each functional module was entered in Neurosynth [27], an automated meta-analysis tool; "Emotion", "Motivation", "Cognitive Control", "Default Mode", resulting in four ("association test") images. The cortical region related to the term "emotion" that showed overlap with the cognitive control regions was assigned to the latter. Finally, the correspondence of the brain regions of the above-described parcellation and the four functional modules was estimated by testing the over-

lap of each region with any voxel of the resulting Neurosynth maps, see Fig. 4.

Network measures The resulting phasic and tonic networks (absolute weighted graphs), where used to estimative several network measures [6], using the Brain Connectivity Toolbox (https://sites.google.com/site/bctnet/). For each node, the strength (the sum of connectivity values), local efficiency (average inverse shortest path length of a node), and the participation coefficient (diversity of intermodular connections of individual nodes) [28] were estimated (Fig. 1c). These measures were selected to capture some of the essential properties of network nodes (i.e., integration and segregation) which are suggested to be related to personality [29].

Classification: linear support vector machine To estimate the predictive accuracy in classifying borderline personality disorder, a linear support vector machine (SVM) function implemented in Matlab (The MathWorks, Inc) was applied to the brain network measures and the task activity contrast estimates as a set of features. The advantage of using the network measures as features is that it yields a reduced set (here 121 features) of variables for classification, compared to using the element-wise connectivity (here 7260), and can thus also be regarded as a dimensionality reduction method while allowing inference on the brain node level.

The different network measures (strength, participation coefficient, local efficiency) for the phasic (safe > look, and the safe + look contrast) and tonic connectivity, as well as the main effect of the task (also safe > look, and the safe + look contrast) where entered as feature models (resulting in 11 different models). Before the classification procedure, the input feature data was "corrected" for the different sites by performing a regression analysis with dummy coded site regressors and then using the resulting residuals.

Each group was compared to each other group (one vs. one classification) in a nested-cross-validation procedure, where the balanced accuracy (the average of the sensitivity and specificity) served as the key outcome measure (effect size) of the classification performance. The data was divided in tenfold validation and test/train data. In the inner loop of the nested cross-validation, the test/train data was further divided in a fivefold cross-validation train and test data to estimate the balanced accuracy per model. The feature model with the highest balanced accuracy was then fitted to all test/training data of the inner loop, and the model was evaluated on the performance of the validation data of the outer loop [see Fig. 1 for an illustration of the procedure]. This process was repeated 200 times to obtain a stable estimate of the mean balanced accuracy and a 95% confidence interval.

The SVM classification was then repeated 1000 times with randomly permuted labels [25], to obtain a permutation null distribution. The p-value was defined as the number of times the null distribution showed a balanced accuracy higher than the average balanced accuracy of the validation data divided by the total number of permutations. Follow-up analyses were subsequently performed to aid the functional interpretation of the best performing model of the inner loop: (1) the SVM model weights were averaged across folds and repetitions resulting in an averaged model of feature weights, indicating the relative (by ranking [30]) contribution of each brain region in the classification (2) groups were tested on the difference of the network measures per functional module (emotion, motivation, cognitive control, default mode). The Matlab code for the network and SVM prediction analyses is available at https://github.com/henkc remers/NetworkAnalysis.

Results

Brain parcellation and subject inclusion

51 borderline patients, 26 cluster-C patients, and 44 non-patients were included in the current analyses, see Supplementary material 1 for details and Fig. S1.2 for the included brain nodes.

Classification accuracy

The classification results for each group compared to each other group showed a limited cross-validated balanced accuracy, 55% CI95 [45 63] $p = 0.23$, for the BPD vs. NPC, 50% CI95 [39 59] $p = 0.46$ for the NPC vs. CLC, and 48%, CI95 [38 59], $p > 0.5$ for BPD vs. CLC, see Fig. 2 for an overview of these and other classification metrics. Thus, only the BPD vs. NPC classification showed a very modest, yet non-significant, effect.

Exploratory analysis

We did opt to explore the results further to guide potential future research. Figure 3 shows the results of the balanced accuracy of the inner loop, which indicates that the *tonic strength* model was most often selected (40%), and showed a classification accuracy of 62%. Figure 4 then illustrates the rank-ordered features of the tonic strength BPD vs. NPC classification, the functional module assignment, and the averaged strength data per group. It is of note that the left amygdala was one of the 5 highest-ranking features and showed an increased strength (stronger connections) in the BPD group than the NPC group. The strength centrality was further tested on group differences between BPD and NPC for the four functional defined modules; emotion, motivation, cognitive control, and default mode. The results did not indicate any substantial difference (all Cohen's $d < 0.12$, and all $p > 0.5$) between BPD and NPC (see Fig. 5).
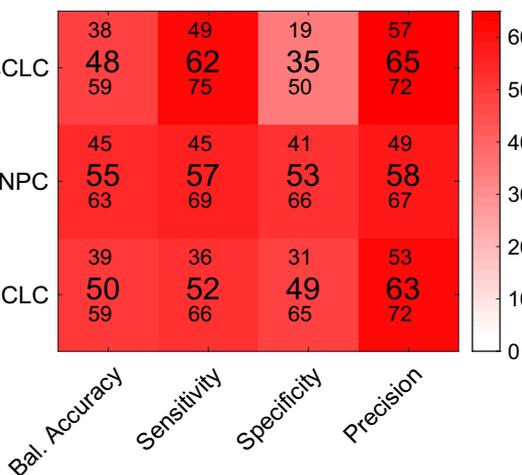

Fig. 2 Performance of the SVM classification analysis. Each row represents the group comparison, and the columns indicate the outcome metrics. Values in the middle of each cell are the mean for a metric, the value above and below indicate the lower and upper bound of the 95% confidence interval. *Bal. Accuracy* balanced accuracy
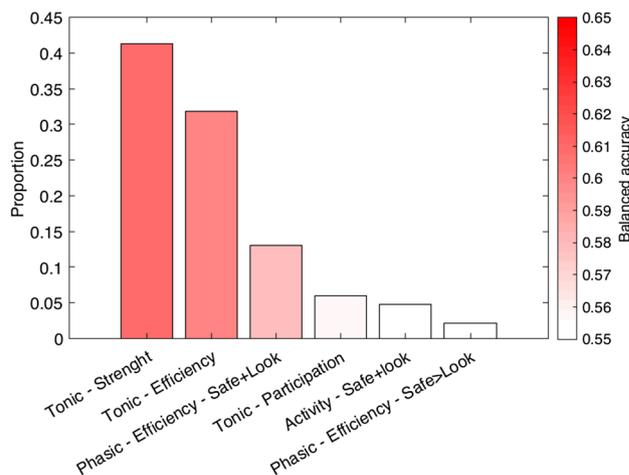

Fig. 3 Results of the inner loop of the BPD vs. NPC model. The y-axis depicts the proportion of times the model was the highest performing model in the inner loop. The color-coding indicates the balanced accuracy of the inner loop of each model

◄**Fig. 4** BPD vs. NPC Tonic-strength model. Upper panel the SVM weights plotted using BrainNet Viewer [31]. Red means a node has higher centrality in BPD, and blue represents higher centrality in NPC. The size of the nodes is proportional to the SVM weights. Middle panel from outside to inside: (A) rank-ordered averaged weights of the linear support vector machine, Color bar (bottom of figure), and node size indicates the relative contribution. (B) Functional module assignment, see bottom of figure for legend. (C) Node strength averaged per group, see lower panel for the legend. See supplementary Table S1 for a complete list of the nodes. Different nodes can contain the same label, see supplementary material

Discussion

The current study used network measures in a machine learning approach to classify BPD patients when engaged in an emotion regulation task. There was a small, yet not statistically significant, classification accuracy for BPD vs. NPC of 55%. Within the nested classification procedure, the tonic strength model showed the highest balanced classification accuracy of 62%. Contrary to our expectations, the network measure models based on phasic connectivity performed worse in classifying BPD vs. NPC. While statistical significance testing is arbitrary and heavily debated [32, 33], the effects found in our study (classification accuracy) are undoubtedly small. First, we will consider possible explanations for this and recommendations for future research; then we will evaluate the potential of the analytic procedure.

The current data set is the largest BPD study [18] in the published literature [3, 34]. At the same time, the sample size is still modest considering statistical power in general [35], and arguably even more so for machine learning analyses [36]. On the one hand, machine learning analyses focuses on the classification performance on an entire model and hence consist of a single (or a few) outcome measures which improve power over mass-univariate testing [35]. However, there is a large error surrounding the

cross-validation accuracy in relatively small samples [36]. Indeed, the estimated confidence intervals here (± 8–10%) indicate a large variability of the balanced accuracy, and the permutation distributions showed a large standard deviation. Moreover, in the nested cross-validation, only a subset of the data is used for model testing, and we did observe a substantial drop in balanced accuracy from the inner to the outer loop (see [11] for a discussion). The subsampling analyses did display an increase in balanced accuracy as a function of sample size for the BPD vs. NPC classification (see supplementary material 3), showing that larger samples could lead to increased classification accuracy. In sum, these limitations underscore the necessity to further study this topic, preferably in a larger (for example, aggregated) sample.

Previous machine learning fMRI research on BPD [12, 37, 38] did show a substantially higher classification accuracy, yet there are some methodological factors to consider. While machine learning approaches, through cross-validation, provide an estimate of the generalization of a model [35], they are nonetheless also susceptible to methodological variability and researcher degrees of freedom [39]. Furthermore, the cross-validation procedures used in other studies (leave-one-out cross-validation) tend to overestimate accuracy [11]. Lastly, previous BPD classification studies contained much smaller sample sizes, which, in combination with a certain degree of flexibility in analytic strategies can lead to possibly inflated effect sizes [35] and the discussed large confidence intervals surrounding the estimates. In general, an overview of fMRI based classification of psychiatric disorders showed that when tested on validation/external data, classification accuracy drops substantially [40].

Aside from the above-mentioned methodological issues, there are critical clinical and biological aspects to deliberate when evaluating the balanced classification accuracy. BPD is, like most psychiatric disorders, a heterogeneous construct: patients with different symptom profiles can be

Fig. 5 Manhattan style bar plot of strength centrality per module and group. The two adjacent black symbols per bar indicate the mean value of the subject per group (across regions; left of each pair) and mean of the different regions (across subjects; right of each pair)
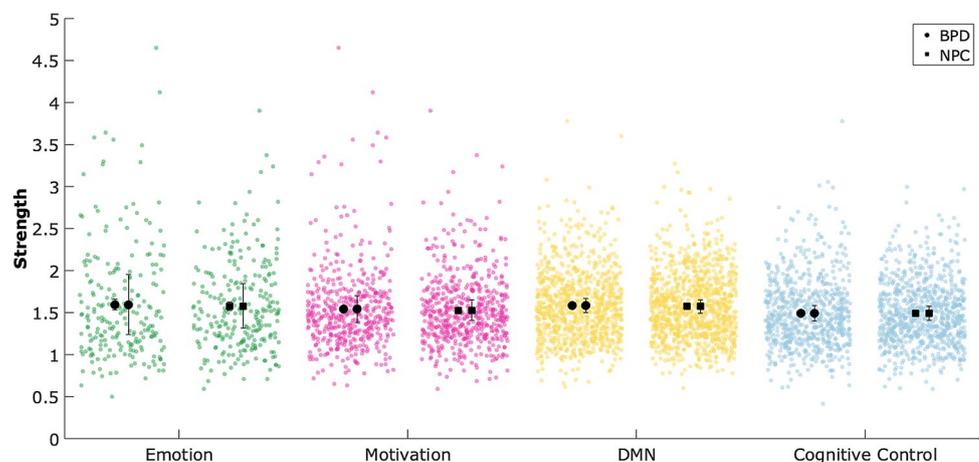
diagnosed with BPD. In addition, the BPD diagnostic criteria do not incorporate any brain markers—for good reasons, since no reliable biomarker has been identified yet [41]. Clinical and biological heterogeneity of BPD imposes a low ceiling on the maximum classification accuracy one can realistically expect. That is, while in principle, a heterogeneous clinical construct could have a homogenous biological basis (i.e., multifinality [42]), we regard this to be unlikely in BPD. A potential future research avenue is to test if clinically identified subtypes of BPD [43] are also potentially more biologically homogenous in terms of brain network organization. In this case, classifying these subtypes with network centrality might be possible with much higher accuracy.

While the overall classification accuracy thus was small, we did further explore the highest performing model (tonic strength) to indicate the possibilities of network measures as features in classification analyses. As noted in the method section, a benefit of network measures is the dimensionality reduction it involves and making inferences possible at the node level. For example, the amygdala was one of the five highest-ranking positive features in the classification of BPD vs. NPC. The amygdala is considered to be a critical region in almost any form of psychopathology, including BPD [3]. The BPD patients displayed higher amygdala connectivity strength than the NPC. This indicates that the amygdala in BPD is more interconnected with the rest of the brain, which could signal an increase in arousal [44] to other brain regions. At the same time, the follow-up analyses of the strength centrality of regions in the functional modules (emotion, motivation, cognitive control, and default mode network), did not show any substantial difference between the BPD and NPC. This could imply that there is no particular functional network that differentiates BPD from NPC, and that distributed and subtle differences across brain regions represent the neural mechanism underlying BPD. However, it is also possible that there is considerable variation among individual subjects concerning which cognitive control regions are most involved in this emotion regulation task, which would also create this observed scattered pattern.

Contrary to our expectations, as mentioned, we found that the phasic connectivity (connectivity that depends on specific task-events) models were not able to detect BPD vs. NPC. Psychophysiological interaction analysis is a well-established approach to assess phasic task-related connectivity. However, a principal shortcoming of this analysis (and more generally, the testing of interaction effects [45]) is that by controlling for the "main" effects of task and time-series, the psychophysiological interaction term needs to demonstrate a large effect above and beyond these main effects [22]. It is unknown if this could explain the limited success in classifying BPD vs. NPC based on the phasic connectivity, but it is a principle difficulty in assessing phasic

connectivity changes. Also, here, more research is needed to evaluate the predictive power of phasic connectivity in classifying BPD patients and other psychiatric disorders. In that sense, this discussion can be regarded as part of a broader discussion on the added benefit of investigating dynamic vs. stationary connectivity [46, 47].

A fundamental limitation is the difference between the participating sites in the angle (tilt) of the field of view in fMRI data acquisition, and consequently, the loss of data. We have addressed this problem by performing a trade-off analysis of minimal node intensity as an index of data quality and subject inclusion yet aiming to incorporate important subcortical and cortical nodes. Nonetheless, seven subjects were excluded from the current analyses, and a part of the brain was not included in the analyses. This might have limited the classification accuracy, yet to the best of our knowledge, not biased the results.

Conclusion

The current study found a small and non-significant effect of BPD classification based on the brain network measures. Larger samples are needed to more accurately assess how well BPD patients can be classified. Future research should, furthermore, focus on clinically or biologically defined potential subtypes of BPD. Finally, a better understanding of the neural network organization and connectivity of borderline personality disorder could eventually have the potential to optimize individually tailored treatments.

Compliance with ethical standards

Conflict of interest The authors report no biomedical financial interest or potential conflicts of interest.

Code availability Code: https://github.com/henkcremers/NetworkAnalysis.

References

1. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Association, Arlington, VA

2. Crowell SE, Beauchaine TP, Linehan MM (2009) A biosocial developmental model of borderline personality: Elaborating and extending linehan's theory. Psychol Bull 135(3):495–510. https://doi.org/10.1037/a0015616

3. van Zutphen L, Siep N, Jacob GA et al (2015) Emotional sensitivity, emotion regulation and impulsivity in borderline personality disorder: a critical review of fMRI studies. Neurosci Biobehav Rev 51:64–76. https://doi.org/10.1016/j.neubiorev.2015.01.001

4. Krause-Utz A, Veer IM, Rombouts SARB et al (2014) Amygdala and anterior cingulate resting-state functional connectivity in borderline personality disorder patients with a history of interpersonal trauma. Psychol Med 44:2889–2901. https://doi.org/10.1017/S0033291714000324

5. Krause-Utz A, Winter D, Niedtfeld I, Schmahl C (2014) The latest neuroimaging findings in borderline personality disorder. Curr Psychiatry Rep 16:438–513. https://doi.org/10.1007/s11920-014-0438-z

6. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10:186–198. https://doi.org/10.1038/nrn2575

7. Power JD, Schlaggar BL, Lessov-Schlaggar CN, Petersen SE (2013) Evidence for hubs in human functional brain networks. Neuron 79:798–813. https://doi.org/10.1016/j.neuron.2013.07.035

8. Fornito A, Bullmore ET, Zalesky A (2017) opportunities and challenges for psychiatry in the connectomic era. Biol Psychiatry Cogn Neurosci Neuroimaging 2:9–19. https://doi.org/10.1016/j.bpsc.2016.08.003

9. Cole MW, Reynolds JR, Power JD et al (2013) Multi-task connectivity reveals flexible hubs for adaptive task control. Nat Publ Group 16:1348–1355. https://doi.org/10.1038/nn.3470

10. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference, and prediction Springer Series in Statistics. Springer New York, New York

11. Varoquaux G, Raamana PR, Engemann DA et al (2017) Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. NeuroImage 145:166–179. https://doi.org/10.1016/j.neuroimage.2016.10.038

12. Xu T, Cullen KR, Mueller B et al (2016) Network analysis of functional brain connectivity in borderline personality disorder using resting-state fMRI. Neuroimage (Amst) 11:302–315. https://doi.org/10.1016/j.nicl.2016.02.006

13. Smith SM, Fox PT, Miller KL et al (2009) Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci 106:13040–13045. https://doi.org/10.1073/pnas.0905267106

14. Cole MW, Bassett DS, Power JD et al (2014) Intrinsic and task-evoked network architectures of the human brain. Neuron 83:238–251. https://doi.org/10.1016/j.neuron.2014.05.014

15. Fornito A, Harrison BJ, Zalesky A, Simons JS (2012) Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. Proc Natl Acad Sci 109:12788–12793. https://doi.org/10.1073/pnas.1204185109

16. Gerchen MF, Bernal-Casas D, Kirsch P (2014) Analyzing task-dependent brain network changes by whole-brain psychophysiological interactions: a comparison to conventional analysis. Hum Brain Mapp 35:5071–5082. https://doi.org/10.1002/hbm.22532

17. van Zutphen L, Siep N, Jacob GA et al (2017) Always on guard: emotion regulation in women with borderline personality disorder compared to nonpatient controls and patients with cluster-C personality disorder. J Psychiatry Neurosci 43:170008–170047. https://doi.org/10.1503/jpn.170008

18. Baczkowski BM, van Zutphen L, Siep N et al (2016) Deficient amygdala–prefrontal intrinsic connectivity after effortful emotion regulation in borderline personality disorder. Eur Arch Psychiatry Clin Neurosci 267:1–15. https://doi.org/10.1007/s00406-016-0760-z

19. Cremers HR, Roelofs K (2016) Social anxiety disorder: a critical overview of neurocognitive research. WIREs Cogni Sci. https://doi.org/10.1002/wcs.1390

20. First MB, Spitzer RL, Gibbon M, Williams J (1997) Structured Clinical Interview for DSM-IV® Axis I Disorders (SCID-I), Clinician Version, Administration Booklet. Biometrics Research, New York State Psychiatric Institute, New York

21. Friston KJ, Buechel C, Fink GR et al (1997) Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6:218–229. https://doi.org/10.1006/nimg.1997.0291

22. O'Reilly JX, Woolrich MW, Behrens TEJ et al (2012) Tools of the trade: psychophysiological interactions and functional connectivity. Soc Cogn Affect Neurosci. https://doi.org/10.1093/scan/nss055

23. Fair DA, Schlaggar BL, Cohen AL et al (2007) A method for using blocked and event-related fMRI data to study "resting state" functional connectivity. NeuroImage 35:396–405. https://doi.org/10.1016/j.neuroimage.2006.11.051

24. Friedman J, Hastie T, Tibshirani R (2008) Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9:432–441. https://doi.org/10.1093/biostatistics/kxm045

25. Rosa MJ, Portugal L, Hahn T et al (2015) Sparse network-based models for patient classification using fMRI. NeuroImage 105:493–506. https://doi.org/10.1016/j.neuroimage.2014.11.021

26. Cribben I, Wager TD, Lindquist MA (2013) Detecting functional connectivity change points for single-subject fMRI data. Front Comput Neurosci 7:143. https://doi.org/10.3389/fncom.2013.00143

27. Yarkoni T, Poldrack RA, Nichols TE et al (2011) Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670. https://doi.org/10.1038/nmeth.1635

28. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. NeuroImage 52:1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

29. Tompson SH, Falk EB, Vettel JM, Bassett DS (2018) Network approaches to understand individual differences in brain connectivity: opportunities for personality neuroscience. Personal Neurosci 1:1125. https://doi.org/10.1017/pen.2018.4

30. Chang YW, Lin C-J (2008) Feature ranking using linear SVM. In: Proceedings of the workshop on the causation and prediction challenge at WCCI 2008. http://proceedings.mlr.press/v3/chang08a.html

31. Xia M, Wang J, He Y (2013) BrainNet Viewer: a network visualization tool for human brain connectomics. PLoS ONE 8:e68910. https://doi.org/10.1371/journal.pone.0068910

32. Amrhein V, Greenland S, McShane B (2019) Scientists rise up against statistical significance. Nature 567:305–307. https://doi.org/10.1038/d41586-019-00857-9

33. Gelman A, Stern H (2012) The difference between "significant" and 'not significant' is not itself statistically significant. Am Stat 60:328–331. https://doi.org/10.1198/000313006X152649

34. Marceau EM, Meuldijk D, Townsend ML et al (2018) Biomarker correlates of psychotherapy outcomes in borderline personality disorder: a systematic review. Neurosci Biobehav Rev 94:166–178. https://doi.org/10.1016/j.neubiorev.2018.09.001

35. Cremers HR, Wager TD, Yarkoni T (2017) The relation between statistical power and inference in fMRI. PLoS ONE 12:e0184923. https://doi.org/10.1371/journal.pone.0184923

36. Varoquaux G (2018) Cross-validation failure: small sample sizes lead to large error bars. NeuroImage 180:68–77. https://doi.org/10.1016/j.neuroimage.2017.06.061

37. Tingting Xu, Cullen KR, Houri A, et al Classification of borderline personality disorder based on spectral power of resting-state fMRI. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE pp 5036–5039

38. Sato JR, de Araujo Filho GM, de Araujo TB et al (2012) Can neuroimaging be used as a support to diagnosis of borderline personality disorder? an approach based on computational neuroanatomy and machine learning. J Psychiatr Res 46:1126–1132. https://doi.org/10.1016/j.jpsychires.2012.05.008

39. Yarkoni T, Westfall J (2017) Choosing prediction over explanation in psychology: lessons from machine learning. Perspect Psychol Sci 11:1745691617693393–1122. https://doi.org/10.1177/1745691617693393

40. Woo C-W, Chang LJ, Lindquist MA, Wager TD (2017) Building better biomarkers: brain models in translational neuroimaging. Nat Publ Group 20:365–377. https://doi.org/10.1038/nn.4478

41. Kapur S, Phillips AG, Insel TR (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 17:1174–1179. https://doi.org/10.1038/mp.2012.105

42. Cicchetti D, Rogosch FA (1996) Equifinality and multifinality in developmental psychopathology. Dev Psychopathol 8:597–600. https://doi.org/10.1017/S0954579400007318

43. Smits ML, Feenstra DJ, Bales DL et al (2017) Subtypes of borderline personality disorder patients: a cluster-analytic approach. Borderline Personal Disord Emot Dysregul 4:1. https://doi.org/10.1186/s40479-017-0066-4

44. Davis M, Whalen PJ (2001) The amygdala: vigilance and emotion. Mol Psychiatry 6:13–34

45. McClelland GH, Judd CM (1993) Statistical difficulties of detecting interactions and moderator effects. Psychol Bull 114(2):376–390. https://doi.org/10.1037/0033-2909.114.2.376

46. Hutchison RM, Womelsdorf T, Allen EA et al (2013) Dynamic functional connectivity: promise, issues, and interpretations. NeuroImage 80:360–378. https://doi.org/10.1016/j.neuroimage.2013.05.079

47. Laumann TO, Snyder AZ, Mitra A et al (2016) On the stability of BOLD fMRI correlations. Cereb Cortex 103:213. https://doi.org/10.1093/cercor/bhw265