Differentiating weight-restored anorexia nervosa and body dysmorphic disorder using neuroimaging and psychometric markers

Don A. Vaughn[1☯], **Wesley T. Kerr**[2,3,4☯], **Teena D. Moody**[5], **Gigi K. Cheng**[1],
Francesca Morfini[5], **Aifeng Zhang**[6], **Alex D. Leow**[6], **Michael A. Strober**[5], **Mark S. Cohen**[1,5,7],
Jamie D. Feusner[5]*

1 Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, California, United States of America, **2** David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, United States of America, **3** Department of Biomathematics, University of California Los Angeles, Los Angeles, California, United States of America, **4** Department of Internal Medicine, Eisenhower Medical Center, Rancho Mirage, California, United States of America, **5** Department of Psychology, Harvard University, Cambridge, Massachusetts, United States of America, **6** Department of Psychiatry, University of Illinois College of Medicine, Chicago, Illinois, United States of America, **7** Departments of Neurology, Radiology, Biomedical Physics, Psychology, Bioengineering and California Nanosystems Institute, University of California Los Angeles, Los Angeles, California, United States of America

☯ These authors contributed equally to this work.
* JFeusner@mednet.ucla.edu

Abstract

Anorexia nervosa (AN) and body dysmorphic disorder (BDD) are potentially life-threatening conditions whose partially overlapping phenomenology—distorted perception of appearance, obsessions/compulsions, and limited insight—can make diagnostic distinction difficult in some cases. Accurate diagnosis is crucial, as the effective treatments for AN and BDD differ. To improve diagnostic accuracy and clarify the contributions of each of the multiple underlying factors, we developed a two-stage machine learning model that uses multimodal, neurobiology-based, and symptom-based quantitative data as features: task-based functional magnetic resonance imaging data using body visual stimuli, graph theory metrics of white matter connectivity from diffusor tensor imaging, and anxiety, depression, and insight psychometric scores. In a sample of unmedicated adults with BDD (n = 29), unmedicated adults with weight-restored AN (n = 24), and healthy controls (n = 31), the resulting model labeled individuals with an accuracy of 76%, significantly better than the chance accuracy of 35% ($\hat{p} < 10^{-4}$). In the multivariate model, reduced white matter global efficiency and better insight were associated more with AN than with BDD. These results improve our understanding of the relative contributions of the neurobiological characteristics and symptoms of these disorders. Moreover, this approach has the potential to aid clinicians in diagnosis, thereby leading to more tailored therapy.

Introduction

Anorexia nervosa (AN) and body dysmorphic disorder (BDD) are psychiatric disorders that share several phenomenological characteristics such as distorted perception of appearance [1], obsessive and compulsive symptoms [2], and poor insight [3,4]. BDD is characterized principally by an individual's preoccupation with a perceived defect in their appearance [1]. Prevalence in the general population is estimated between 0.7% [5,6] and 2.9% [7,8]. AN is characterized by restriction of energy intake, fear of gaining weight or becoming fat, disturbance in the individual's perception of body weight, and significantly low BMI [1]. The prevalence is estimated to be 0.9% and 0.3% in females and males, respectively [9].

Both disorders are very serious conditions that confer an elevated risk of mortality. Of all psychiatric disorders, AN has the highest risk of mortality: 0.6% per year [10,11]. Medical morbidity and mortality related to the starvation state are major risks in affected persons [12]. During the lifetime of an individual with a BDD diagnosis, 24–28% will attempt suicide [13]; the suicide-related mortality rate in individuals with BDD is roughly 0.3% per year [14]. The alarmingly frequent mortalities and severe risks associated with AN and BDD underscore that it is imperative to recognize these disorders and to intervene with treatment to minimize poor outcomes.

The selection of an effective treatment for AN or BDD is aided by proper diagnosis. Cognitive behavioral therapy (CBT) can be effective in BDD [15], but its usefulness in AN is minimal [16]. BDD can respond to treatment with selective serotonin reuptake inhibitors (SSRIs), but evidence for this in AN is lacking [17,18]. Family-based psychotherapy in adolescents with AN has been shown to be effective [19].

Clinically, it can be difficult in some cases to distinguish between AN and BDD. Similarities in symptom characteristics are such that some researchers have raised the question of whether one disorder should be considered a subtype of the other, or at least have noted that they share pathological features [20,21]. For example, 30% of individuals with BDD have significant weight-related appearance concerns (e.g., concern about their cheeks being too fat), even if they do not have a comorbid eating disorder [22]. Consequently, individuals with BDD may restrict calorie intake and engage in appearance-related compulsive behaviors. In this example, Diagnostic Statistical Manual (DSM-5 or DSM-IV) criteria would categorize the person as having AN or AN plus BDD if the body weight was less than 85% of that expected, and BDD alone if 85% or above. However, the behavioral manifestations and subjective experience for the patient may be otherwise identical.

In fact, individuals frequently meet criteria for both AN and BDD [23,24]. Of those with AN, 25–39% are diagnosed with lifetime BDD; 32% of those with BDD will have a lifetime eating disorder [23–25]. In studies that directly compared individuals with AN to those with BDD, the two groups shared lower body satisfaction, lower appearance evaluation, lower attitudes toward one's self, more body areas of dissatisfaction, and greater preoccupation with being overweight, compared with healthy controls [26–28]. In terms of clinically distinguishing features, aside from the low body weight criteria in AN, the sex distribution is notably more skewed toward females in AN (90% female), whereas BDD has a more equal prevalence across sexes [7,29]. While insight tends to be poor in both conditions, it is worse on average in BDD [4]. With these data in mind, we believe that it would be advantageous to have additional quantitative means of distinguishing between these conditions.

Neuroimaging has provided several features that may help distinguish between individuals with AN, those with BDD, and healthy controls. Recent research suggests that distorted perceptions of appearance may be linked to abnormalities in the visual system [30–32]. A combined analysis of data from task-based functional magnetic resonance imaging (fMRI) and

EEG showed that, relative to healthy controls, individuals with AN and those with BDD exhibited hypoactivity in both early secondary visual regions and the dorsal stream when viewing low-spatial frequency face and house stimuli [33]. An event-related potential (ERP) analysis of the same EEG data likewise revealed that the amplitude of P100 responses was smaller in individuals with AN than in those with BDD and in healthy controls, whereas N170 latencies were longer and N170 amplitudes were smaller in AN and BDD compared with healthy controls [34].

Diffusion tensor imaging (DTI) additionally has demonstrated that brain white matter connectivity patterns differ in individuals with BDD and healthy controls [35], and between individuals with AN and healthy controls [36]. Connectivity pattern abnormalities in white matter networks have been found within the visual system of individuals with BDD [3,37]. Among the few studies that have compared AN and BDD directly, correlations were present between psychometric insight and ERP amplitude in BDD but not in AN [34]. Insight was correlated with the network measure path length in white matter in AN but not in BDD [3]. Despite these known neurological differences, a differential diagnosis currently is made using criteria based principally on weight, observed behaviors, and reported symptoms.

It is important also to distinguish psychiatrically ill individuals from healthy controls. This is an issue particularly in adolescents, the age group most likely to develop AN and BDD. The role of clinicians is to identify patients who would benefit from treatment while avoiding a pathologic diagnosis for normal adolescents. Mislabeling developmentally-normal, self-conscious behavior (e.g., overvaluing appearance, mirror-checking) as a disorder can have detrimental effects on an individual's perceived identity (Rosenhan, 1973). On the other hand, mislabeling a disorder as normal development could normalize the pathologic behaviors and experiences of AN and BDD (e.g., purging, compulsive behaviors, significant interference in functioning, emotional distress). BDD often goes undetected in adolescents [38], yet the suicide attempt rate is possibly higher than in adults (21–44%) [39].

With this background in mind, our current study had two goals: 1) to understand mechanistically the multivariate differences in neurobiology and psychometric scales that map to the diagnoses, and 2) to aid clinicians by developing an enhanced and more objective means of distinguishing between AN, BDD, and healthy controls. Both of these goals were achieved by combining multimodal neuroimaging and psychometric data into a multivariate logistic regression model. The feature weights of the multivariate model address the first goal by providing an evidence-based magnitude, direction, and significance as to how each modality relates to a particular diagnosis. The model's probabilistic outputs address the second goal by providing diagnostic predictions. Although many studies of psychiatric populations collect more than one modality of neurobiological data and multiple psychometric scales, few have combined these in fully integrated models to enhance the accuracy of predictive classification (or for other purposes).

Materials and methods

Participants

Our study included data from 84 participants (age 22 +/- 5 years, 88% female): 24 diagnosed with weight-restored AN, 29 diagnosed with BDD, and 31 healthy controls. All participants were unmedicated. The University of California, Los Angeles (UCLA) Institutional Review Board approved the protocol, and all participants provided written informed consent. Capacity to provide informed consent was determined by the study physician (Dr. Feusner). For minors who participated we obtained signed parental permission and youth assent.

For semantic ease, we defined AN/BDD to be all of the participants with either AN or BDD; AN/BDD does not indicate that the participants were comorbid with both disorders, as

Table 1. Demographics.

	AN	BDD	CTL	Test statistic	p
Age (yrs)	21 ± 5	23 ± 5	21 ± 5	$F = 2.2$, df = 2	0.12
Female	23/24 (96%)	26/29 (90%)	25/31 (81%)	$x^2 = 3.1$, df = 2	0.21
BMI	20 ± 2	22 ± 3	22 ± 3	$F = 4.1$, df = 2	0.02
Illness duration (months)	72 ± 63	118 ± 70	N/A	$t = 2.6$, df = 51	0.03
Lowest lifetime BMI	16 ± 2	N/A	N/A	N/A	N/A
Number of participants	24	29	31	N/A	N/A

Healthy controls (denoted as CTL in this table), participants with AN, and participants with BDD were matched across age, gender, and body mass index (BMI). Participant age and the proportion of females were not significantly different across groups; BMI and illness duration differed significantly between groups. Errors are standard deviation.

participants with both AN and BDD were excluded. Participants were of equivalent age and gender (Table 1).

We excluded participants who met any of the following criteria: (1) medical disorders affecting cerebral metabolism (e.g., diabetes mellitus or hyperthyroidism); (2) pregnancy; (3) neurological disorders; (4) concurrent Axis I disorders other than dysthymia, major depressive disorder, or generalized anxiety disorder; (5) score >44 (very severe) on the Montgomery-Åsberg Depression Rating Scale (MADRS) [40] and/or major depressive disorder with psychotic features; (6) an assessment of imminent suicidality (by Dr. Feusner); (7) current or past diagnosis of AN (in the BDD group) or a current or past diagnosis of BDD (in the AN group); and (8) current treatment with CBT. Participants were required to be free from psychoactive medications for at least 8 weeks prior to entering the study. To avoid the confounds of acute starvation on brain structure and function, all participants with AN had to be weight-restored during data collection (as defined by BMI ≥ 18.5) despite having previously met full DSM-IV criteria for AN. BDD participants had to score ≥ 20 on the BDD version of the Yale-Brown Obsessive Compulsive Scale (BDD-YBOCS) [41].

Psychometrics

To establish diagnoses, all participants were administered the Mini International Neuropsychiatric Interview (MINI) [42] and the BDD Diagnostic Module [43], modeled after the DSM-IV. Symptom severity was quantified in AN with modified versions of the Yale-Brown-Cornell Eating Disorder Scale (YBC-EDS) [44] and the Eating Disorder Evaluation (EDE) Edition 16.0D [45], and in BDD with the BDD-YBOCS [41]. Participants with AN and those with BDD, were administered the Brown Assessment of Beliefs Scale (BABS) [46], the Hamilton Anxiety Rating Scale (HAM-A) [47], and the Montgomery-Asberg Depression Scale (MADRS) [40].

Overview and rationale of feature selection

The principal goal of the study was to create a classification model to help distinguish AN/BDD from healthy controls, and AN from BDD. To make use of a range of neurobiological and psychometric data, we chose structural (DTI) and functional (fMRI) data (acquired in all groups), psychometric ratings measuring anxiety and depression (acquired in all groups), and insight (acquired in AN and BDD). From the DTI data, we chose to use the graph theory metric of *normalized path length* (NPL) because 1) this global network connectivity measure provides a single summarized metric of white matter network connectivity for each participant, and 2) our previous univariate analysis suggested differing patterns in AN and BDD [3]. We chose to include fMRI task data from brain responses to both low and high spatial frequency bodies and faces. We

hypothesized that bodies and faces (rather than the third stimuli class, houses) would best distinguish both controls from AN/BDD and distinguish AN from BDD, because bodies were likely to be more salient for people with AN and faces for those with BDD [48].

fMRI data

We collected fMRI data as participants engaged in simple tasks of matching photographs of others' bodies and of others' faces [49]. Each run contained blocks of images with low-only, normal, or high-only spatial-frequency content, and a control task of matching shapes. Previous analyses suggested that the stimuli types that showed the least overlap between participants with AN/BDD and controls were low and high spatial-frequency images; thus, we used responses to high and low spatial-frequency but not normal spatial-frequency images from the bodies and faces data as features in our model. In the interest of deriving a signal metric per network, we executed a network coherence analysis of the functional data. Based on previous results, we hypothesized that the most informative coherence values would come from three networks of interest: primary visual, higher order visual, and salience networks [50]. For further detail, see *supplementary material*.

DTI data

The neuroimaging data included DTI in the form of 64 gradient direction diffusion-weighted images with b = 1000 s/mm^2 and one minimally diffusion-weighted scan (the B_0 image). Graph theory metric data were obtained from a processing pipeline previously described [3]. Briefly, we derived 87 cortical and subcortical structures using Freesurfer (Martinos Center for Biomedical Imaging, USA) from each participant's T1-weighted scan data, and we performed continuous fiber tracking from the diffusion-weighted data using DTIStudio software (http://www.mristudio.org). Combining the parcellation and fiber tractography results yielded an 87 x 87 connectivity matrix of tractography relationships between regions. For each participant's data, we calculated the shortest path length between each pair of nodes and then averaged this path length over all nodes to determine the characteristic path length (CPL). We calculated the normalized path length (NPL) as the ratio of observed CPL to the CPL of an identically sized but randomly connected network [51].

Handling the missing data

As is a challenge with many multimodal datasets, we had missing data for some participants across psychometric and neuroimaging modalities. We addressed this problem using multiple imputation [52]. Due to a modified initial study design, the first eight participants with AN did not complete the BABS. None of the control participants completed the BABS because they, by definition, had no disorder for which they could provide insight. We did not use functional data from 27 of the participants because of their excessive head movement. Of our 84 participants, we had data for 100% of HAM-A/MADRS, 85% (45/53) of the BABS, 100% of DTI, and 68% (57/84) of fMRI.

Additionally, we also performed a complete case analysis using the 57 participants for whom we had no missing data (17 AN, 19 BDD, 21 control). The complete case results—in terms of model performance, feature significance, and feature magnitude—were not dissimilar from the imputation model and thus are not discussed further.

Diagnostic statistical modeling

We predicted the diagnoses using a two-step multiple-imputation logistic regression model; we first predicted whether a participant was a healthy control or had AN/BDD, and for the

latter group, we predicted the disorder (AN or BDD). As features, we used coherence values from fMRI data in primary visual, higher order visual, and salience networks using body and face high- and low spatial-frequency visual stimuli; white matter NPL from DTI data; and anxiety, depression, and insight psychometric scores. We trained these models using cyclical leave-one-out cross-validation and assessed their statistical significance via permutation testing, which yielded an empirical, two-tailed p-value indicated by $\hat{p}$. All confidence intervals (CI) are 95% standard error of the mean (SEM), unless otherwise stated.

Results

Dataset information

There were significant univariate differences between participants with AN, participants with BDD, and healthy controls. Participants with BDD had significantly higher BABS scores (4.6 ± 1.5) than participants with AN ($t = 3.1$, df = 43, $p = 0.004$) (Fig 1A). Participants with AN, those with BDD, and healthy controls had unequal mean NPL ($F = 4.8$, df = 2, $p = 0.01$, Fig 1B). This difference in central tendency was driven by the AN group (μ_{AN}-μ_{BDD} = 0.11 ± 0.03, $t = 3.3$, df = 51, $p = 0.002$), specifically a bimodal distribution of AN participants with NPLs that were either similar or much longer than the BDD and control groups. Note that this bimodality was not explained by large fluctuations in the denominator (CPL of an equivalent random network) (Fig 1C). Healthy controls scored significantly lower than participants with AN/BDD on both the HAM-A ($\mu_{A/B}$-μ_{C} = 7.8 ± 1.2, $t = 6.3$, df = 82, $p < 10^{-4}$, Fig 1D) and MADRS ($\mu_{A/B}$-μ_{C} = 13.5 ± 1.6, $t = 8.2$, df = 82, $p < 10^{-4}$, Fig 1E). We observed a high Pearson correlation between the HAM-A and MADRS scores of all three participant groups (r = 0.85 ± 0.05). Thus, we collapsed both scales into a single combined metric, MH, by extracting the first principal component, which explained 93% of the variance in the data (Fig 1F).

Because of excessive head movement, we had to exclude fMRI data from 29% (7/24) of participants diagnosed with AN, 34% (10/29) of participants diagnosed with BDD, and 32% (10/31) of controls. There was no significant difference in the proportion with missing data across the three groups (Chi-squared, 0.17, df = 3, p = 0.92).

Given that only 68% of coherence values were available, we conducted a preliminary analysis to evaluate whether imputed coherence explained a significant amount of the residual variance in the models. We used a regularized logistic regression model to evaluate the marginal effect of including a combination of high and low spatial frequency stimuli of faces and bodies for each of the three networks. The regularized model did not significantly explain more variation in the data, when considering the number of variables added (deviance difference 12.8, df = 12, p = 0.3). Additionally, the overall accuracy of this overall model decreased with this additional information, likely due to overfitting.

This suggests that the variation accounted for by coherence was explained by other variables in the limited model, and thus, the inclusion of coherence did not improve our prediction of the outcome. Given this result, all subsequent analyses excluded coherence as a feature.

Distinguishing healthy controls from AN/BDD

When distinguishing healthy controls from participants with AN or BDD, the first step predictive model had a leave-one-out accuracy of 89% (75/84, permutation mean 41%, $\hat{p} < 0.01$), a positive predictive value (PPV) of 92% (48/52, permutation mean 53%, $\hat{p} < 0.01$), a negative predictive value (NPV) of 84% (27/32, permutation mean 33%, $\hat{p} < 0.01$), and an area under the receiver operator characteristic (AUC-ROC) of 93% (Fig 2A). The overall deviance difference was 66.2 (df = 2, $p < 10^{-4}$), and this was driven by a significant association of lower MH
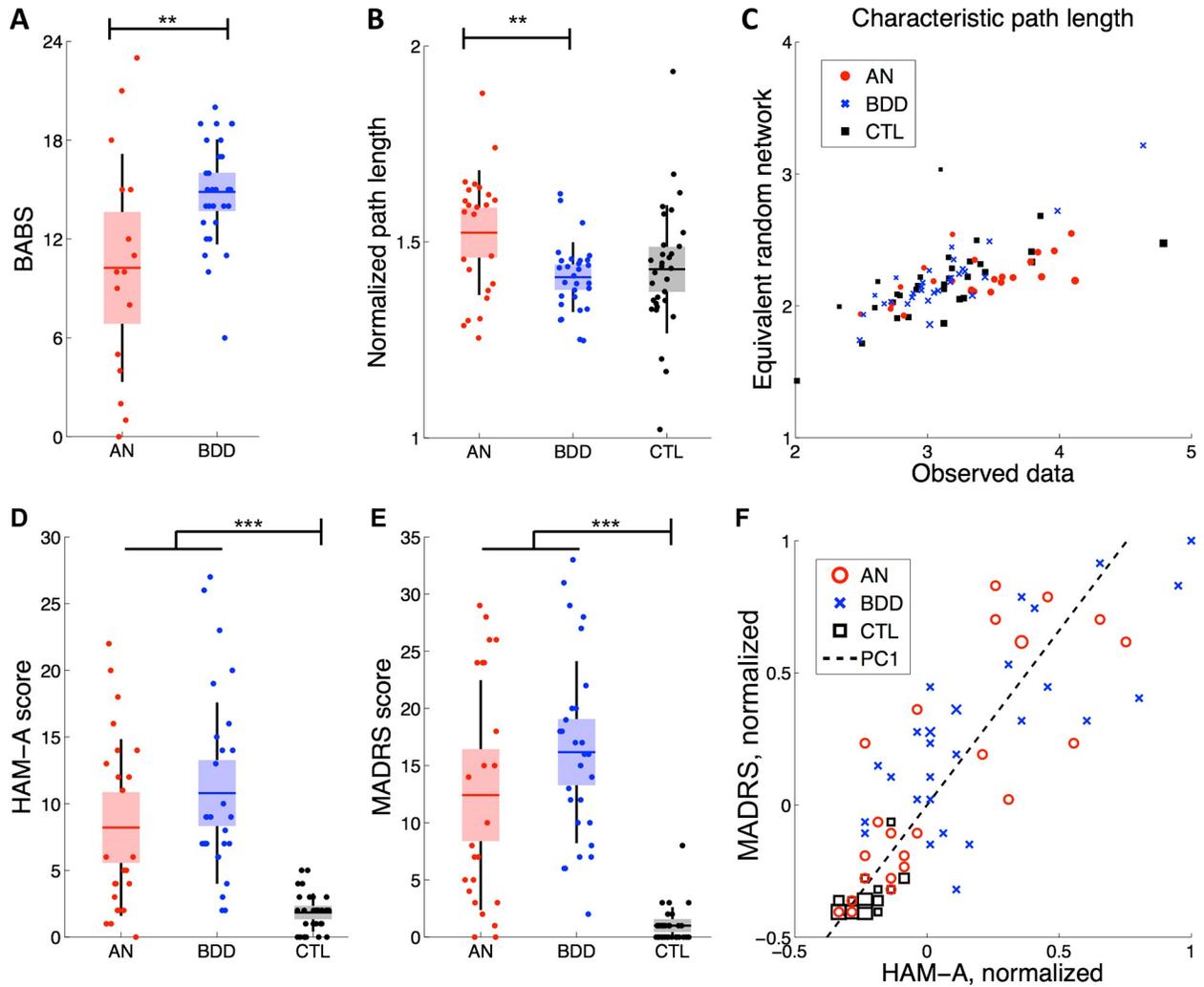
Fig 1. Dataset information. (A) Participants with BDD had significantly different BABS scores than those with AN (p = 0.004). In this and all remaining panels, the horizontal line is the mean, box is the SEM 95% CI, black lines demarcate one standard deviation, and dots reflect individual participants. **(B)** NPL differed significantly between AN and BDD groups (p = 0.002). **(C)** The NPL comprised the observed CPL and the CPL of an equivalently-sized random network. This plot demonstrates that the denominator of NPL—the CPL of an equivalent random network—does not show large fluctuations that would explain the bimodal distribution in AN as presented in panel B. (D) Participants with AN/BDD scored significantly higher on the HAM-A scale than healthy controls ($p<$ 0.001). (E) Participants with AN/BDD scored significantly higher on the MADRS scale than healthy controls ($p<$ 0.001). (F) The first single principal component (PC1) of HAM-A and MADRS explained 93% of the variance among all participants with HAM-A and MADRS scores (n = 84). Larger markers represent the presence of more than one participant. BABS = Brown Assessment of Beliefs Scale; HAM-A = Hamilton Anxiety Rating Scale; MADRS = Montgomery-Asberg Depression Scale.

https://doi.org/10.1371/journal.pone.0213974.g001

with healthy controls (estimate 11.82 per PCA-unit, SE 3.00, p<10⁻⁴, Fig 2B). NPL was not associated significantly with either group (estimate -2.5 per unit, SE 2.5, $\hat{p}$ = 0.31).

AN vs BDD prediction

The second step predictive model—designed to distinguish participants with AN from those with BDD—had a leave-one-out accuracy of 74% (39/53, permutation mean 40%, $\hat{p} < 0.001$), an AN predictive value of 78% (14/18, permutation mean 40%, $\hat{p} < 0.01$), a BDD predictive value of 71% (25/35, permutation mean 41%, $\hat{p} < 0.001$), and an AUC-ROC of 67% (Fig 2C). The overall difference in deviance was 10.0 (df 3, $\hat{p}$ = 0.006) with a positive association of
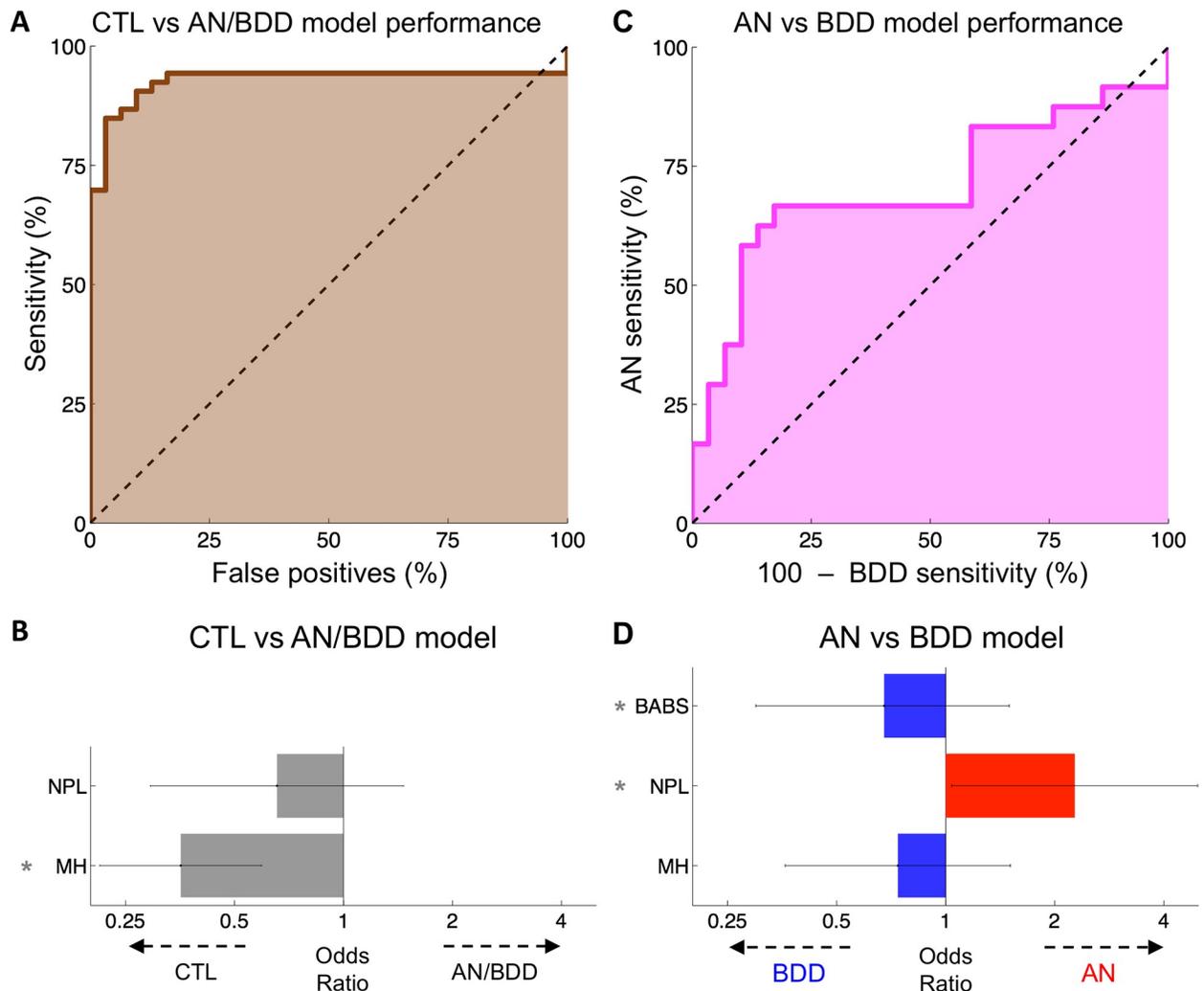
Fig 2. Classifier performance. (A) The receiver operator characteristic of the healthy controls (denoted as CTL) vs AN/BDD model. Our model distinguished healthy controls from participants with AN or BDD with an AUC-ROC of 93%. **(B)** Bars reflect the weight of each factor used in the CTL vs AN/BDD classification, expressed as the odds ratio per standard deviation. Error bars are SEM and asterisks demarcate statistically significant features ($\hat{p} < 0.05$). **(C)** Our AN vs BDD model distinguished participants with AN from those with BDD with an AUC-ROC of 67%. **(D)** The weight of each factor used in the AN vs BDD classification, expressed as an odds ratio per standard deviation. Error bars are (Gaussian) SEM and thus for reference only; significance values were calculated from the (non-Gaussian) permutations.

https://doi.org/10.1371/journal.pone.0213974.g002

higher NPL (lower globally efficient white matter connectivity) with AN (log odds 6.0 per unit, SE 2.9, $\hat{p} = 0.02$), a positive association of higher BABS scores (lower insight) with BDD (log odds -0.07 per unit, SE 0.07, $\hat{p} = 0.011$ with a non-Gaussian distribution), and no significant association of MH with either group (-0.63 per unit, SE 0.75, $\hat{p} = 0.40$) (Fig 2D).

Multistep prediction

The two-step combination of the above models had a three-group leave-one-out accuracy of 76% (64/84, 95% CI: 67–85%, permutation mean 35%, $\hat{p} < 10^{-4}$) with AN predictive value of 63% (12/19, permutation mean 40%, $\hat{p} < 0.001$), BDD predictive value of 76% (25/33, permutation mean 41%, $\hat{p} < 0.001$), and healthy control predictive value of 84% (27/32, permutation mean 33%, $\hat{p} < 0.001$).

Discussion

By combining multimodal neuroimaging and psychometric data, we were able to differentiate accurately between healthy controls, participants with weight-restored AN, and participants with BDD. This represents the first study to apply this approach to AN and BDD for classification purposes and to understand relative contributions of factors for distinguishing these DSM-defined syndromes. These results shed light on a more dimensional understanding of the neurobiological and psychometric variables that contribute to these disorders.

Classification performance

While anxiety and depression scores were sufficient to distinguish healthy controls from those with AN/BDD, a combination of psychometric and neuroimaging data best distinguished individuals with weight-restored AN from those BDD. Yet, even with more predictive features, our model distinguishing weight-restored AN from BDD was less accurate than our model distinguishing AN/BDD from healthy controls. As observed clinically, AN and BDD show phenomenological overlap; thus, it is not surprising that differentiating healthy controls from participants with either disorder was much more accurate than distinguishing participants with weight-restored AN from those with BDD. Even though we excluded participants with a DSM diagnosis of both conditions, our lower classification accuracy between weight-restored AN and BDD, as compared to AN/BDD and healthy controls, supports previous findings that some individuals with AN may share symptomatic and behavioral elements of BDD, and vice versa [26–28]. The predictive power our model achieved in distinguishing two phenomenologically overlapping disorders highlights the utility of integrating functional and structural neurobiological data with clinical symptoms and behavior-based data.

Our two-way classification accuracies—89% in distinguishing controls from individuals with AN/BDD, and 74% in differentiating weight-restored AN from BDD—are similar to previous works in other disorders. Using neuroimaging data, other groups have distinguished healthy controls from individuals with social anxiety disorder [53], ADHD [54], major depressive disorder [55], autism [56], and Alzheimer's disease [57] with an accuracy of 80–90%. Fewer groups, however, have attempted to separate individuals with related but distinct diagnoses. The closest comparable studies classified disorder subtypes or predicted the progression of a disorder. For example, functional connectivity distinguished subtypes of ADHD with 65% accuracy [54], regional white matter volume discriminated between treatment-resistant and treatment-sensitive depression with an accuracy of 66% [55], and fMRI and PET data from individuals with mild cognitive impairment predicted who would develop Alzheimer's disease within 18 months and who would not with 68% accuracy [57].

Combining both classification models yields a two-step clinical diagnostic flowchart for a potential patient: (1) Does this individual potentially have an appearance-related psychiatric disorder? (2) If so, how strongly does the evidence suggest that the individual has symptoms of AN or BDD? By combining our participants' results for both models, the relationship among these conditions is more clear (Fig 3).

There was fairly strong separability between groups. Although the classifier confused some healthy controls and participants with weight-restored AN, none of the participants with BDD were misclassified as healthy controls. This suggests that if the two-stage classifier were to classify an individual as likely having AN/BDD, and then were to assign a final prediction of BDD, there would be a high likelihood that the individual has an actual clinical diagnosis of BDD.
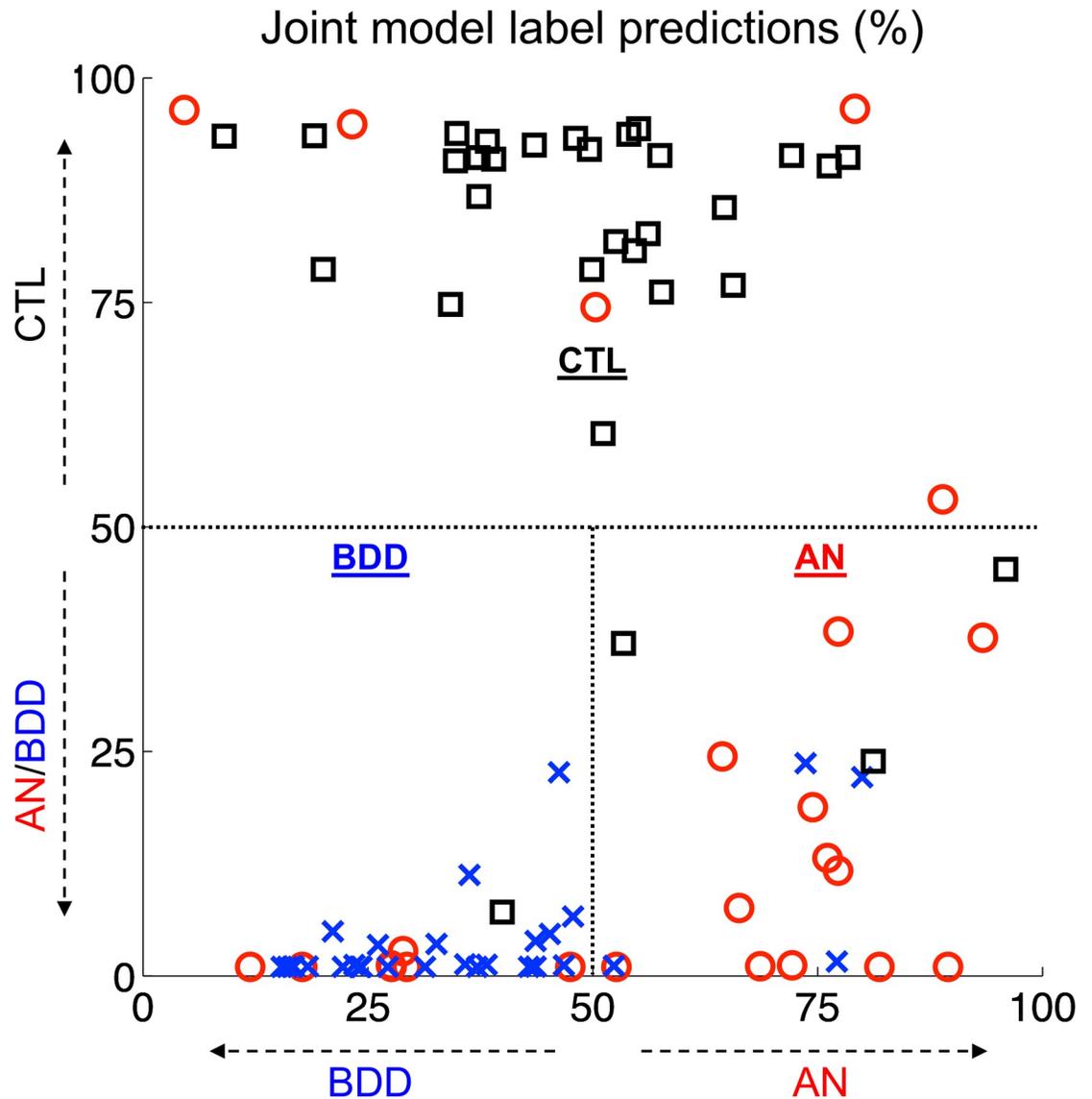
Fig 3. Joint predictions of both models. The synthesis of the control (CTL) vs AN/BDD model and the AN vs BDD model reveals a relatively clear delineation between healthy control participants (black), participants with AN (red), and participants with BDD (blue).

https://doi.org/10.1371/journal.pone.0213974.g003

Neurobiological interpretations

In our model, lower insight contributed significantly to the distinction between BDD and AN. This accords with a previous study that found that, on average, individuals with BDD have lower insight than individuals with AN [4]. Clinical experience also suggests that in AN, more frequently than in BDD, a greater proportion of individuals possess the insight that their body weight is low, yet they nevertheless desire to remain in an underweight state or achieve an even lower weight, accounting for many recalcitrant cases [58,59]. Individuals with AN have compulsive regimens of exercise, food restriction, or selectivity that remain ego-syntonic.

The result that the white matter network connectivity measure (NPL) contributed significantly to the distinction between AN and BDD is in line with our previous finding of higher

NPL in AN than in BDD [3]. Our related previous research, however, was univariate, single-modal, and contained no predictive modeling. In other words, whereas our previous univariate analyses could distinguish significant differences among AN, BDD, and healthy controls for each of the neuroimaging and psychometric data separately, the current results provide information about how each measure contributes to differentiating the disorders when accounting for the effects of the others. In the multimodal approach we developed here, insight and NPL both contributed significantly to distinguishing AN from BDD in a multivariate context, prompting the question of how these features might interact with each other to contribute to brain-behavior relationships.

Indeed, the phenomenon of poor insight may itself be linked to NPL in AN. In prior work, we found that poor insight in AN was associated with longer NPL in the anterior and posterior cingulate cortices [3]. Because this finding was specific to these cingulate subregions, it seems to support a relationship between insight and information transfer between these regions (involved in error-detection and conflict monitoring [60]) and the rest of the brain, rather than global network integration. Future studies are warranted to uncover potential relationships between insight and regional or global brain network topology. This will help clarify underlying connectivity patterns associated with the phenomenon of insight, which has important clinical implications for if, and how, patients engage in treatment.

Interestingly, NPL in the AN group demonstrated properties of a bimodal distribution; one subgroup had a clustering of NPL values that resembled those of the BDD and healthy control cohorts, while another subgroup showed much longer NPL values. Thus, longer average NPL, which suggests a less globally-integrated white matter connectivity network across the brain, may be a useful distinguishing feature that separates AN from BDD. The differing network topology of a subgroup of AN raises the question of whether classification methods alternative to DSM might better capture natural "cut-points" of the neuropathology of AN. This was outside of the focus and scope of the present work, and future studies are required to investigate this further.

Our results show that healthy controls could be differentiated easily from individuals with appearance-related psychiatric disorders using a single psychometric metric of anxiety/depression. While this is informative, it illustrates that differentiating healthy controls from weight-restored AN or BDD can be ascertained readily by a clinical assessment; neurobiological data from neuroimaging is unnecessary. The more difficult distinction, in some clinical cases, is between AN and BDD; our model also performed well with this classification.

Limitations and future directions

In this study, we based classification accuracy results on DSM criteria, though we acknowledge the ongoing debate about the utility of DSM or International Classification of Diseases (ICD) versus dimensional approaches, as suggested by the NIH Research Domain Criteria (RDoC), in the classification of psychiatric disorders. The benefit of developing more advanced algorithms for classification is predicated on the existence of at least some applications for diagnostic classification. Despite limitations of DSM- or ICD-defined diagnostic categories, there is practical value in distinguishing the clinical syndromes of AN from BDD, namely, for the sake of clinical decisions about specific pharmacotherapy and psychotherapy approaches. As outlined above, these are notably different for the diagnostic categories of AN and BDD, and—amongst other important implications of these distinctions—early interventions to prevent the starvation state are imperative for the prevention of medical morbidity and mortality.

On an individual level, our statistical classification approach yielded different results from the gold-standard (clinical diagnoses) in several cases. While this may reflect a limitation of

the present method, it could also indicate different interpretations of clinical criteria when applied to specific patients. With a modest sample of the present size, it would not likely be informative to evaluate the classification differences on individuals, but future work may want to consider whether such differences arise from factors that may ultimately prove useful in interventions. There also may be yet-unidentified methodological and clinical variables that moderate our classifier's performance.

Our classification model was developed with information from participants from the local community. Variations in methods of test administration, neuroimaging, clinical diagnostics, and underlying heterogeneity of these disorders may differ slightly from location to location, so these results need to be further validated on larger, independent datasets from different regions that capture the full range of the disease before direct application to clinical medicine.

Prior to any considerations of applying this method in clinical diagnostics, several additional follow up studies are needed to 1) replicate the findings in independent and larger samples, and 2) test and validate the approach in populations for whom this tool would be most useful clinically. These populations include the following: 1) those with early stages of illness, namely, adolescents who have not yet experienced significant weight loss and stand to benefit from targeted interventions that could prevent the most impactful manifestations of the illness; 2) those at risk of developing either disorder, perhaps because of a strong family history and/or anxious temperament in childhood; and 3) those who fall on the border between AN and BDD, as exemplified in the Introduction.

Additional testing and validation in acute, rather than weight-restored, AN populations is necessary to understand its generalizability. A strong caveat of the current dataset, therefore, is that the informative neurobiological feature of NPL, and conceivably even insight, may have been influenced by previous starvation states in AN participants. It is necessary to determine the utility of this classification procedure using the current features through future validation, which can be accomplished through the longitudinal study of individuals at risk or in very early pre-starvation stages of symptom manifestation. Similarly, testing in normative adolescent populations is also necessary given the more nuanced differences in appearance concerns and behaviors between healthy adolescents and individuals who meet specific clinical criteria. Additionally, given the increased likelihood of acquisition heterogeneity in multimodal data, these results must be tested for generalizability across worldwide research centers.

If validated successfully, our multimodal, two-stage classification model may aid in clinical disambiguation between healthy individuals and individuals with AN or BDD. Critically, if the methodology proves accurate in populations that are at risk or very early in the course of illness, then it may contribute to the prevention of the starvation state. This step not only would help reduce the risk of medical morbidity and mortality, but also could improve the effectiveness of other treatments that have difficulty taking hold in the context of extreme malnutrition [61,62].

Supporting information

S1 File. Detailed methods. Additional information regarding fMRI methods, imputations, and statistical modeling.
(PDF)

Author Contributions

Conceptualization: Don A. Vaughn, Michael A. Strober, Jamie D. Feusner.

Data curation: Jamie D. Feusner.

Formal analysis: Don A. Vaughn, Wesley T. Kerr, Teena D. Moody, Gigi K. Cheng, Francesca Morfini, Aifeng Zhang, Alex D. Leow.

Funding acquisition: Jamie D. Feusner.

Investigation: Jamie D. Feusner.

Methodology: Don A. Vaughn, Wesley T. Kerr, Mark S. Cohen, Jamie D. Feusner.

Project administration: Jamie D. Feusner.

Resources: Mark S. Cohen, Jamie D. Feusner.

Supervision: Jamie D. Feusner.

Writing – original draft: Don A. Vaughn, Wesley T. Kerr, Jamie D. Feusner.

Writing – review & editing: Don A. Vaughn, Wesley T. Kerr, Michael A. Strober, Mark S. Cohen, Jamie D. Feusner.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5). American Psychiatric Pub; 2013.

2. Woodside BD, Staab R. Management of psychiatric comorbidity in anorexia nervosa and bulimia nervosa. CNS Drugs. 2006; 20: 655–663. https://doi.org/10.2165/00023210-200620080-00004 PMID: 16863270

3. Zhang A, Leow A, Zhan L, GadElkarim J, Moody T, Khalsa S, et al. Brain connectome modularity in weight-restored anorexia nervosa and body dysmorphic disorder. Psychol Med. 2016; 46: 2785–2797. https://doi.org/10.1017/S0033291716001458 PMID: 27429183

4. Hartmann AS, Thomas JJ, Wilson AC, Wilhelm S. Insight impairment in body image disorders: delusionality and overvalued ideas in anorexia nervosa versus body dysmorphic disorder. Psychiatry Res. 2013; 210: 1129–1135. https://doi.org/10.1016/j.psychres.2013.08.010 PMID: 23992792

5. Faravelli C, Salvatori S, Galassi F, Aiazzi L, Drei C, Cabras P. Epidemiology of somatoform disorders: a community survey in Florence. Soc Psychiatry Psychiatr Epidemiol. 1997; 32: 24–29. PMID: 9029984

6. Otto MW, Wilhelm S, Cohen LS, Harlow BL. Prevalence of Body Dysmorphic Disorder in a Community Sample of Women. Am J Psychiatry. 2001; 158: 2061–2063. https://doi.org/10.1176/appi.ajp.158.12.2061 PMID: 11729026

7. Koran LM, Abujaoude E, Large MD, Serpe RT. The prevalence of body dysmorphic disorder in the United States adult population. CNS Spectr. 2008; 13: 316–322. PMID: 18408651

8. Schieber K, Kollei I, de Zwaan M, Martin A. Classification of body dysmorphic disorder—what is the advantage of the new DSM-5 criteria? J Psychosom Res. 2015; 78: 223–227. https://doi.org/10.1016/j.jpsychores.2015.01.002 PMID: 25595027

9. Hudson JI, Hiripi E, Pope HG Jr, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. Biol Psychiatry. 2007; 61: 348–358. https://doi.org/10.1016/j.biopsych.2006.03.040 PMID: 16815322

10. Sullivan PF. Mortality in anorexia nervosa. Am J Psychiatry. search.proquest.com; 1995; 152: 1073–1074. https://doi.org/10.1176/ajp.152.7.1073 PMID: 7793446

11. Hoek HW. Incidence, prevalence and mortality of anorexia nervosa and other eating disorders. Curr Opin Psychiatry. 2006; 19: 389–394. https://doi.org/10.1097/01.yco.0000228759.95237.78 PMID: 16721169

12. Keel PK, Dorer DJ, Eddy KT, Franko D, Charatan DL, Herzog DB. Predictors of mortality in eating disorders. Arch Gen Psychiatry. 2003; 60: 179–183. PMID: 12578435

13. Phillips KA. Suicidality in Body Dysmorphic Disorder. Prim psychiatry. 2007; 14: 58–66. PMID: 18449358

14. Phillips KA, Menard W. Suicidality in body dysmorphic disorder: a prospective study. Am J Psychiatry. 2006; 163: 1280–1282. https://doi.org/10.1176/appi.ajp.163.7.1280 PMID: 16816236

15. Harrison A, Fernández de la Cruz L, Enander J, Radua J, Mataix-Cols D. Cognitive-behavioral therapy for body dysmorphic disorder: A systematic review and meta-analysis of randomized controlled trials. Clin Psychol Rev. 2016; 48: 43–51. https://doi.org/10.1016/j.cpr.2016.05.007 PMID: 27393916

16. Hartmann A, Weber S, Herpertz S, Zeeck A, German Treatment Guideline Group for Anorexia Nervosa. Psychological treatment for anorexia nervosa: a meta-analysis of standardized mean change. Psychother Psychosom. 2011; 80: 216–226. https://doi.org/10.1159/000322360 PMID: 21494063

17. Holtkamp K, Konrad K, Kaiser N, Ploenes Y, Heussen N, Grzella I, et al. A retrospective study of SSRI treatment in adolescent anorexia nervosa: insufficient evidence for efficacy. J Psychiatr Res. 2005; 39: 303–310. https://doi.org/10.1016/j.jpsychires.2004.08.001 PMID: 15725429

18. Williams J, Hadjistavropoulos T, Sharpe D. A meta-analysis of psychological and pharmacological treatments for Body Dysmorphic Disorder. Behav Res Ther. 2006; 44: 99–111. https://doi.org/10.1016/j.brat.2004.12.006 PMID: 16301017

19. Lock J, Stewart Agras W, Bryson S, Kraemer HC. A Comparison of Short- and Long-Term Family Therapy for Adolescent Anorexia Nervosa. J Am Acad Child Adolesc Psychiatry. 2005; 44: 632–639. https://doi.org/10.1097/01.chi.0000161647.82775.0a PMID: 15968231

20. Grant JE, Phillips KA. Is Anorexia Nervosa a Subtype of Body Dysmorphic Disorder? Probably Not, but Read On. . .. Harv Rev Psychiatry. Taylor & Francis; 2004; 12: 123–126. https://doi.org/10.1080/10673220490447236 PMID: 15204807

21. Cororve MB, Gleaves DH. Body dysmorphic disorder: a review of conceptualizations, assessment, and treatment strategies. Clin Psychol Rev. 2001; 21: 949–970. PMID: 11497214

22. Kittler JE, Menard W, Phillips KA. Weight concerns in individuals with body dysmorphic disorder. Eat Behav. Elsevier; 2007; 8: 115–120. https://doi.org/10.1016/j.eatbeh.2006.02.006 PMID: 17174859

23. Grant JE, Kim SW, Eckert ED. Body dysmorphic disorder in patients with anorexia nervosa: prevalence, clinical features, and delusionality of body image. Int J Eat Disord. 2002; 32: 291–300. https://doi.org/10.1002/eat.10091 PMID: 12210643

24. Ruffolo JS, Phillips KA, Menard W, Fay C, Weisberg RB. Comorbidity of body dysmorphic disorder and eating disorders: severity of psychopathology and body image disturbance. Int J Eat Disord. 2006; 39: 11–19. https://doi.org/10.1002/eat.20219 PMID: 16254870

25. Rabe-Jablonska J, Sobow Tomasz M. The links between body dysmorphic disorder and eating disorders. Eur Psychiatry. 2000; 15: 302–305. PMID: 10954874

26. Hrabosky JI, Cash TF, Veale D, Neziroglu F, Soll EA, Garner DM, et al. Multidimensional body image comparisons among patients with eating disorders, body dysmorphic disorder, and clinical controls: a multisite study. Body Image. 2009; 6: 155–163. https://doi.org/10.1016/j.bodyim.2009.03.001 PMID: 19410528

27. Hartmann AS, Thomas JJ, Greenberg JL, Elliott CM, Matheny NL, Wilhelm S. Anorexia nervosa and body dysmorphic disorder: A comparison of body image concerns and explicit and implicit attractiveness beliefs. Body Image. 2015; 14: 77–84. https://doi.org/10.1016/j.bodyim.2015.03.013 PMID: 25920044

28. Kollei I, Brunhoeber S, Rauh E, de Zwaan M, Martin A. Body image, emotions and thought control strategies in body dysmorphic disorder compared to eating disorders and healthy controls. J Psychosom Res. 2012; 72: 321–327. https://doi.org/10.1016/j.jpsychores.2011.12.002 PMID: 22405229

29. Buhlmann U, Glaesmer H, Mewes R, Fama JM, Wilhelm S, Brähler E, et al. Updates on the prevalence of body dysmorphic disorder: a population-based survey. Psychiatry Res. 2010; 178: 171–175. https://doi.org/10.1016/j.psychres.2009.05.002 PMID: 20452057

30. Madsen SK, Bohon C, Feusner JD. Visual processing in anorexia nervosa and body dysmorphic disorder: Similarities, differences, and future research directions. J Psychiatr Res. 2013; 47: 1483–1491. https://doi.org/10.1016/j.jpsychires.2013.06.003 PMID: 23810196

31. Pietrini F, Castellini G, Ricca V, Polito C, Pupi A, Faravelli C. Functional neuroimaging in anorexia nervosa: A clinical approach. Eur Psychiatry. 2011; 26: 176–182. https://doi.org/10.1016/j.eurpsy.2010.07.011 PMID: 20934859

32. Beilharz F, Castle DJ, Grace S, Rossell SL. A systematic review of visual processing and associated treatments in body dysmorphic disorder. Acta Psychiatr Scand. 2017; 136: 16–36. https://doi.org/10.1111/acps.12705 PMID: 28190269

33. Li W, Lai TM, Bohon C, Loo SK, McCurdy D, Strober M, et al. Anorexia nervosa and body dysmorphic disorder are associated with abnormalities in processing visual information. Psychol Med. Cambridge Univ Press; 2015; 45: 2111–2122. https://doi.org/10.1017/S0033291715000045 PMID: 25652023

34. Li W, Lai TM, Loo SK, Strober M, Mohammad-Rezazadeh I, Khalsa S, et al. Aberrant early visual neural activity and brain-behavior relationships in anorexia nervosa and body dysmorphic disorder. Front Hum Neurosci. ncbi.nlm.nih.gov; 2015; 9: 301. https://doi.org/10.3389/fnhum.2015.00301 PMID: 26082703

35. Grace SA, Labuschagne I, Kaplan RA, Rossell SL. The neurobiology of body dysmorphic disorder: A systematic review and theoretical model. Neurosci Biobehav Rev. 2017; 83: 83–96. https://doi.org/10.1016/j.neubiorev.2017.10.003 PMID: 29017915

36. King JA, Frank GKW, Thompson PM, Ehrlich S. Structural Neuroimaging of Anorexia Nervosa: Future Directions in the Quest for Mechanisms Underlying Dynamic Alterations. Biol Psychiatry. 2017; https://doi.org/10.1016/j.biopsych.2017.08.011 PMID: 28967386

37. Arienzo D, Leow A, Brown JA, Zhan L, Gadelkarim J, Hovav S, et al. Abnormal brain network organization in body dysmorphic disorder. Neuropsychopharmacology. 2013; 38: 1130–1139. https://doi.org/10.1038/npp.2013.18 PMID: 23322186

38. Mataix-Cols D, Fernández de la Cruz L, Isomura K, Anson M, Turner C, Monzani B, et al. A Pilot Randomized Controlled Trial of Cognitive-Behavioral Therapy for Adolescents With Body Dysmorphic Disorder. J Am Acad Child Adolesc Psychiatry. 2015; 54: 895–904. https://doi.org/10.1016/j.jaac.2015.08.011 PMID: 26506580

39. Phillips KA, Didie ER, Menard W, Pagano ME, Fay C, Weisberg RB. Clinical features of body dysmorphic disorder in adolescents and adults. Psychiatry Res. Elsevier; 2006; 141: 305–314. https://doi.org/10.1016/j.psychres.2005.09.014 PMID: 16499973

40. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry. 1979; 134: 382–389. PMID: 444788

41. Phillips KA, Hollander E, Rasmussen SA, Aronowitz BR, DeCaria C, Goodman WK. A severity rating scale for body dysmorphic disorder: development, reliability, and validity of a modified version of the Yale-Brown Obsessive Compulsive Scale. Psychopharmacol Bull. 1997; 33: 17–22. PMID: 9133747

42. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998; 59 Suppl 20: 22–33;quiz 34–57.

43. Phillips KA, Atala KD, Pope HG. Diagnostic instruments for body dysmorphic disorder. American Psychiatric Association 148th Annual Meeting. 1995.

44. Mazure CM, Halmi KA, Sunday SR, Romano SJ, Einhorn AM. The Yale-Brown-Cornell Eating Disorder Scale: development, use, reliability and validity. J Psychiatr Res. 1994; 28: 425–445. PMID: 7897615

45. Fairburn, Cooper, Connor. Cognitive Behavior Therapy and Eating Disorders (16.0D ed.). New York, NY: Guilford Press; 2008.

46. Eisen JL, Phillips KA, Baer L, Beer DA, Atala KD, Rasmussen SA. The Brown Assessment of Beliefs Scale: reliability and validity. Am J Psychiatry. 1998; 155: 102–108. https://doi.org/10.1176/ajp.155.1.102 PMID: 9433346

47. Hamilton M. Diagnosis and rating of anxiety. Br J Psychiatry. 1969; 3: 76–79.

48. Phillips KA. The Broken Mirror: Understanding and Treating Body Dysmorphic Disorder. Oxford University Press; 2005.

49. Feusner JD, Townsend J, Bystritsky A, Bookheimer S. Visual information processing of faces in body dysmorphic disorder. Arch Gen Psychiatry. jamanetwork.com; 2007; 64: 1417–1425. https://doi.org/10.1001/archpsyc.64.12.1417 PMID: 18056550

50. Moody TD, Sasaki MA, Bohon C, Strober MA, Bookheimer SY, Sheen CL, et al. Functional connectivity for face processing in individuals with body dysmorphic disorder and anorexia nervosa. Psychol Med. Cambridge Univ Press; 2015; 45: 3491–3503. https://doi.org/10.1017/S0033291715001397 PMID: 26219399

51. Sporns O, Chialvo D, Kaiser M, Hilgetag C. Organization, development and function of complex brain networks. Trends Cogn Sci. 2004; 8: 418–425. https://doi.org/10.1016/j.tics.2004.07.008 PMID: 15350243

52. Rubin DB. Multiple Imputation after 18+ Years. J Am Stat Assoc. 1996; 91: 473–489.

53. Liu F, Guo W, Fouche J-P, Wang Y, Wang W, Ding J, et al. Multivariate classification of social anxiety disorder using whole brain functional connectivity. Brain Struct Funct. 2015; 220: 101–115. https://doi.org/10.1007/s00429-013-0641-4 PMID: 24072164

54. Fair DA, Nigg JT, Iyer S, Bathula D, Mills KL, Dosenbach NUF, et al. Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. Front Syst Neurosci. 2012; 6: 80. https://doi.org/10.3389/fnsys.2012.00080 PMID: 23382713

55. Liu F, Guo W, Yu D, Gao Q, Gao K, Xue Z, et al. Classification of different therapeutic responses of major depressive disorder with multivariate pattern analysis method based on structural MR scans. PLoS One. 2012; 7: e40968. https://doi.org/10.1371/journal.pone.0040968 PMID: 22815880

56. Chen H, Duan X, Liu F, Lu F, Ma X, Zhang Y, et al. Multivariate classification of autism spectrum disorder using frequency-specific resting-state functional connectivity—A multi-center study. Prog Neuropsychopharmacol Biol Psychiatry. 2016; 64: 1–9. https://doi.org/10.1016/j.pnpbp.2015.06.014 PMID: 26148789

57. Liu F, Wee C-Y, Chen H, Shen D. Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification. Neuroimage. 2014; 84: 466–475. https://doi.org/10.1016/j.neuroimage.2013.09.015 PMID: 24045077

58. Guarda AS. Treatment of anorexia nervosa: Insights and obstacles. Physiol Behav. 2008; 94: 113–120. https://doi.org/10.1016/j.physbeh.2007.11.020 PMID: 18155737

59. Nordbø RHS, Espeset EMS, Gulliksen KS, Skårderud F, Holte A. The meaning of self-starvation: qualitative study of patients' perception of anorexia nervosa. Int J Eat Disord. 2006; 39: 556–564. https://doi.org/10.1002/eat.20276 PMID: 16845678

60. Agam Y, Hämäläinen MS, Lee AKC, Dyckman KA, Friedman JS, Isom M, et al. Multimodal neuroimaging dissociates hemodynamic and electrophysiological correlates of error processing. Proc Natl Acad Sci U S A. 2011; 108: 17556–17561. https://doi.org/10.1073/pnas.1103475108 PMID: 21969565

61. Marzola E, Nasser JA, Hashim SA, Shih P-AB, Kaye WH. Nutritional rehabilitation in anorexia nervosa: review of the literature and implications for treatment. BMC Psychiatry. 2013; 13: 290. https://doi.org/10.1186/1471-244X-13-290 PMID: 24200367

62. Tchanturia K, Davies H, Roberts M, Harrison A, Nakazato M, Schmidt U, et al. Poor cognitive flexibility in eating disorders: examining the evidence using the Wisconsin Card Sorting Task. PLoS One. 2012; 7: e28331. https://doi.org/10.1371/journal.pone.0028331 PMID: 22253689