Rapid #: -22327871

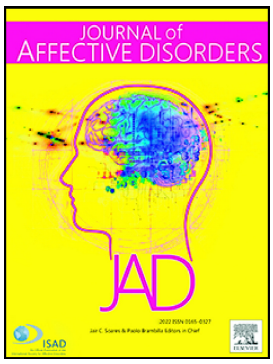
CROSS REF ID: **733845**

LENDER: **GZN (University of Wisconsin - Milwaukee) :: Ejournals**

BORROWER: **MYG (MIT (Massachusetts Inst. of Technology)) :: Main Library**

TYPE:	Article CC:CCG
JOURNAL TITLE:	Journal of affective disorders
USER JOURNAL TITLE:	Journal of Affective Disorders
ARTICLE TITLE:	Predicting response to a smartphone-based cognitive-behavioral therapy for body dysmorphic disorder
ARTICLE AUTHOR:	Greenberg
VOLUME:	
ISSUE:	
MONTH:	
YEAR:	2024
PAGES:	-
ISSN:	0165-0327
OCLC #:	38911953

Processed by RapidX: 3/29/2024 9:24:39 AM



Predicting response to a smartphone-based cognitive-behavioral therapy for body dysmorphic disorder

Jennifer L. Greenberg, Hilary Weingarden, Susanne S. Hoeppner, Rebecca M. Berger-Gutierrez, Dalton Klare, Ivar Snorrason, Omar Costilla-Reyes, Morgan Talbot, Katharine E. Daniel, Rachel C. Vanderkruik, Armando Solar-Lezama, Oliver Harrison, Sabine Wilhelm

PII: S0165-0327(24)00460-9

DOI: https://doi.org/10.1016/j.jad.2024.03.044

Reference: JAD 17303

To appear in:

Received date: 13 October 2023

Revised date: 26 February 2024

Accepted date: 9 March 2024

Please cite this article as: J.L. Greenberg, H. Weingarden, S.S. Hoeppner, et al., Predicting response to a smartphone-based cognitive-behavioral therapy for body dysmorphic disorder, (2024), https://doi.org/10.1016/j.jad.2024.03.044

RUNNING HEAD: Predictors of app-based CBT for BDD

Tables: 3

Figures: 1

Supplemental materials: 2

Predicting response to a smartphone-based cognitive-behavioral therapy for body dysmorphic disorder

Jennifer L. Greenberg[1], Hilary Weingarden[1], Susanne S. Hoeppner[1], Rebecca M. Berger-Gutierrez[1], Dalton Klare[1], Ivar Snorrason[1], Omar Costilla-Reyes[2], Morgan Talbot[2,3], Katharine E. Daniel[1], Rachel C. Vanderkruik[1]
Armando Solar-Lezama[2], Oliver Harrison[4], Sabine Wilhelm[1]
[1]Massachusetts General Hospital and Harvard Medical School, 185 Cambridge Street, Boston, MA, USA
[2]Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA
[3]Boston Children's Hospital and Harvard Medical School, 300 Longwood Avenue, Boston, MA, USA
[4]Koa Health, 55 Baker Street, London, UK

Corresponding Author:
Jennifer L. Greenberg
Department of Psychiatry
Massachusetts General Hospital
185 Cambridge Street, Suite 2000
Boston, MA, 02114, USA
Phone: 617-726-5374
Fax: 617-643-3080
Email: jlgreenberg@mgh.harvard.edu

Highlights

- Machine learning and regressions were used to examine predictors of app-CBT for BDD
- Immediate CBT, credibility, and sexual minority status predicted better outcomes
- App-CBT may be particularly helpful for historically marginalized populations
- Making treatment more readily available and credible could improve outcomes

Abstract

Background: Body dysmorphic disorder (BDD) is a severe, chronic disorder if untreated. Smartphone cognitive behavioral therapy (CBT) for BDD is efficacious and can reduce key treatment barriers (e.g., lack of clinicians, cost, stigma). While promising, little is known about who is more or less likely to benefit from this approach.

Methods: This is a secondary data analysis of a randomized, waitlist-controlled trial of smartphone CBT for BDD. Participants (*N*=80) were recruited nationally and randomized to receive a 12-week, coach-guided CBT for BDD app, either immediately or after a 12-week waitlist. The main outcome for this analysis was BDD severity (BDD-YBOCS) over time (baseline, week 6, week 12) during the active app use phase in each randomized group (n=74). Secondary outcomes included treatment response (≥30% reduction in BDD-YBOCS) and remission (total BDD-YBOCS ≤16) at end-of-treatment.

Results: Immediate (vs. delayed) CBT predicted better outcomes (symptom improvement), as did gender identity (symptom improvement), higher baseline treatment credibility and expectancy (response, remission), lower baseline BDD severity (remission), and sexual minority status (vs. heterosexual; response, remission).

Limitations: Limitations include the relatively small sample, drop-out rate of 22%, and limited gender and racial-ethnic diversity. **Conclusions:** These results highlight a potential advantage of smartphone CBT in historically marginalized populations, and the importance of efforts to hasten treatment access, bolster confidence in the treatment at treatment onset, and develop stratified care models to optimize treatment allocation and efficacy.

Keywords: body dysmorphic disorder; cognitive behavioral therapy; predictors; app; digital mental health

Predicting response to a smartphone-based cognitive-behavioral therapy for body dysmorphic disorder

Introduction

Body dysmorphic disorder (BDD) is a severe mental disorder characterized by excessive preoccupation with one or more perceived appearance flaws and time-consuming appearance-related rituals intended to hide, fix, or check perceived flaws (APA, 2013). BDD is common, affecting 1.7-2.9% of the general population, chronic, and associated with significant impairment in psychosocial functioning, poor quality of life, psychiatric comorbidity, and high rates of suicide (Angelakis et al., 2016; Buhlmann et al., 2010; Koran et al., 2008; Phillips, 2000; Phillips et al., 2005; Schieber et al., 2015).

Cognitive behavioral therapy (CBT) is the first-line psychological treatment for BDD. Randomized controlled trials (RCTs) have demonstrated the efficacy of CBT for BDD when delivered in face-to-face, internet, and smartphone-based app formats (Enander et al., 2016; Harrison et al., 2016; Wilhelm et al., 2019; Wilhelm et al. 2022). Yet, only 17.4% of individuals with BDD receive CBT (Marques et al., 2011). Digital interventions for BDD, including internet-based and coach-guided smartphone-based CBT for BDD, are scalable and cost-effective, thereby addressing many existing barriers to treatment such as shortages of available clinicians, high cost, long wait times, and stigma (Enander et al., 2016; Flygare et al., 2023; Wilhelm et al., 2022). The relatively anonymous and flexible delivery format of app-based interventions for BDD may also facilitate greater treatment readiness in those who would otherwise be hesitant to establish care with a face-to-face provider and allows individuals to use skills

when and where they need it most (Weingarden et al., 2020; Wilhelm et al., 2022). Still, not all individuals will benefit equally from a smartphone-based intervention. Understanding who is more or less likely to benefit from a light touch, coach-guided smartphone-based CBT for BDD would facilitate more personalized and efficient allocation of scarce clinical resources. Those who are more likely to benefit from this light touch program could be diverted away from resource-heavy alternatives, whereas those who are less likely to benefit from this format could be prioritized for face-to-face treatment.

Existing research on predictors of CBT for BDD outcomes in face-to-face and internet-based studies is limited and has yielded mixed results. Greater baseline treatment credibility (Flygare et al., 2020; Phillips et al., 2021), treatment expectancy and readiness to change (Greenberg et al., 2019), high working alliance (Flygare et al., 2020), and obsessive compulsive personality disorder (Phillips et al., 2021) have predicted better CBT outcomes in one to two studies each. In a secondary analysis of data drawn from three studies of CBT for BDD (*N*=90), those with a significant early reduction in BDD severity had a 92% chance of responding at end-of treatment; however, importantly, minimal early symptom change was not indicative of eventual non-response (Greenberg et al. 2022). Other studies have shown that worse CBT outcomes were predicted by greater BDD severity (Flygare et al., 2020), depressive symptom severity (Flygare et al., 2020), longer duration of BDD (Flygare et al., 2020), poorer BDD-related insight (Greenberg et al., 2019; Neziroglu,et al., 2001) and use of serotonin reuptake inhibitors (Phillips et al., 2021); however, an RCT of CBT versus anxiety management for BDD found no evidence that any baseline predictors examined

(duration of BDD, BDD-related insight, depression) significantly predicted outcomes (Veale et al., 2014). Notably, a 2016 meta-analysis of seven RCTs of CBT for BDD and a 2023 systematic review of predictors and moderators of treatment response in CBT for BDD found no consistent predictors of treatment outcome across studies (Harrison et al., 2016; Hogg et al., 2023). These mixed findings may be due to relatively small sample sizes and methodological differences across studies, thereby underscoring the need for continued and innovative investigation into who is most likely benefit from BDD treatment and under what conditions.

Machine learning (ML) models are an underutilized but powerful approach to increase this understanding. Despite ML's ability to identify complex and non-linear predictive patterns among variables that might be difficult to discover through human inspection or traditional statistical analyses alone, all but one previous study (Flygare et al., 2020) has relied on traditional regression models to examine predictors of CBT for BDD outcomes. Non-linear ML models such as decision trees can also generate decision rules (e.g., "rules-of-thumb") that could be used to guide clinical decision making. Thus, in addition to probing potential predictors of treatment outcomes using traditional statistical models (e.g., linear regression), it is also important to begin leveraging ML models in BDD research.

The current study examined predictors of smartphone app-based CBT for BDD using data from a 12-week randomized, waitlist-controlled trial of Perspectives (*N*=74) (Wilhelm et al., 2022) and is the first study of predictors of treatment outcomes in app-delivered CBT for BDD. We used traditional (i.e., statistical) regression models as well as machine learning approaches to examine a range of demographic, clinical, and

treatment-related factors as potential predictors of symptom improvement during treatment as well as for treatment response and remission at post-treatment. Given the largely equivocal findings on predictors of CBT for BDD treatment outcomes evidenced across delivery methods, our aims were exploratory, and we made no *a priori* hypotheses. However, we did expect that higher levels of treatment credibility/expectancy would be associated with larger improvements in BDD symptoms because treatment credibility has emerged as a leading predictor of post-treatment outcomes across psychotherapies (Constantino et al., 2018).

Methods

Overview and Study Design

This was a secondary analysis of data from a randomized, waitlist-controlled trial of coach-guided, smartphone-delivered CBT for BDD (Perspectives; Clinicaltrials.gov ID NCT03673046). Participants (n=80) were randomly assigned to receive the 12-week CBT for BDD app either immediately or after a 12-week waitlist. The present study uses data from the active treatment phase of both groups (n=74). Detailed information about study procedures, study aims, eligibility, and the treatment rationale and procedures are described in the primary manuscript (Wilhelm et al., 2022).

Participants

Participants in this secondary data analysis included all participants randomized to receive the immediate app-based CBT treatment (*n*=40), as well as those participants who started app-based CBT after the 12-week waitlist and who had not achieved remission status prior to starting treatment (*n*=34). Participants were adults with a primary DSM-5 diagnosis of BDD, living in the United States, and recruited nationally

from July 2019 through March 2021. Participants taking psychotropic medications were required to be on a stable dose for at least two months before initiating the study. Exclusion criteria included current severe substance use, severe depression, acute suicidal ideation, lifetime bipolar or psychosis, concurrent therapy, or ≥ 4 previous sessions of CBT for BDD.

Procedures

Study procedures were approved by the Massachusetts General Hospital institutional review board and participants gave informed consent prior to participation. Participants were randomly assigned at a 1:1 ratio to receive the Perspectives app immediately or after a 12-week wait and randomization was stratified by medication status. Clinician-administered measures were conducted via secure video call by blinded, doctoral-level independent evaluators and self-report measures were completed via survey links to Research Electronic Data Capture (REDCap; Harris et al., 2009). For those in the waitlist condition, we used baseline demographics and clinical history data from the RCT-phase baseline assessment, but clinical symptom predictor data from the end-of-waitlist (prior to beginning app-based CBT).

Measures

The Yale-Brown Obsessive-Compulsive Scale Modified for BDD (BDD-YBOCS; Phillips et al. 1997; Phillips et al., 2014), a 12-item semi-structured clinician-administered measure of past week BDD symptom severity, was the primary outcome measure. BDD-YBOCS total scores range from 0-48, with higher scores indicating more severe BDD. Internal consistency in the current sample was α=.76 at baseline. Measures of baseline characteristics included demographics, the number of current

psychiatric comorbidities (assessed with the Mini International Neuropsychiatric Interview (MINI 7.02; Sheehan et al. 1998), BDD-related insight (measured by the Brown Assessment of Beliefs Scale [BABS]; Eisen et al., 1998; Phillips et al., 2013; current sample α=.78), depressive symptoms (Quick Inventory of Depressive Symptomatology [QIDS-SR]; Rush et al., 2003; current sample α=.69), readiness to change (University of Rhode Island Change Assessment Questionnaire [URICA]; McConnaughy et al., 1983; current sample α=.76 (pre-contemplation) to α=.89 (maintenance)), and treatment credibility and expectancy (Credibility and Expectancy Questionnaire [CEQ]; Devilly & Borkovec, 2000; current sample α=.73 (credibility), α=.86 (expectancy)).

Treatment

All participants received Perspectives, a 12-week, coach-guided smartphone app-based CBT for BDD (Wilhelm et al., 2020; Wilhelm et al., 2022) either immediately or following a 12-week wait. Perspectives includes brief psychoeducation and interactive, skills-based exercises covering the core components of CBT for BDD: psychoeducation, cognitive restructuring, exposure and ritual prevention, mindfulness and perceptual (mirror) retraining, improving self-esteem through core belief exercises and enhancing valued activities, and relapse prevention. Treatment was facilitated by bachelors-level coaches who were available to participants through asynchronous in-app secure messaging. Coaches also conducted two, brief phone calls with participants: one at baseline to orient participants to treatment and another at mid-treatment to promote engagement and motivation and to answer questions. Participants were instructed to use the app daily.

Data Analyses

Data preparation and treatment group comparison. Demographic variables and some clinical variables were recoded to binary variables for analyses as follows: For gender identity, female vs. male and non-binary; for race and ethnicity, non-Hispanic White vs. all other; for sexual orientation, heterosexual vs. all other; for educational attainment we coded two binary variables, high school degree or less vs. all other and post-college education vs. all others; for QIDS-SR depression scores we also coded two binary variables, no or mild depression (QIDS-SR scores ≤ 10) vs. all others and severe and very severe depression (QIDS-SR scores ≥ 16) vs. all others; for psychotropic medication use, baseline use of serotonin reuptake inhibitor (SRI) medication vs. not; two participants started SRIs while on the waitlist and at least one month prior to delayed treatment and were both coded as using SRIs at baseline. For comorbid psychiatric disorders, we coded a binary variable of 1 or more comorbid diagnoses vs. none. Continuous variables were z-transformed prior to analysis to make the model estimates more interpretable. To determine group differences between participants in the immediate versus delayed app-CBT groups, we used independent samples t-tests for continuous variables and either chi-square tests of independence (if cell sizes ≥10) or Fisher's exact tests (if cell sizes <10) for categorical variables.

Definitions of outcome variables. The main outcome was BDD symptom severity over time (treatment baseline, week 6, week 12), as measured by BDD-YBOCS total scores. Secondary outcomes included binary treatment response and remission status at end-of-treatment (week 12). Treatment response was defined as a 30% or greater reduction in BDD-YBOCS scores from baseline to end-of-treatment (Phillips et

al., 1997), and remission status as a total score of ≤16 on the BDD-YBOCS at end-of-treatment. The remission definition combines partial and full remission (Fernandez de la Cruz et al., 2019).

Missing data and multiple imputation strategy. Baseline data was missing for two participants for BDD duration (2.7%), CEQ-credibility (2.7%), CEQ-expectancy (2.7%), and for one participant for the URICA readiness to change (RTC) score (1.4%). Seventeen participants had missing data for the BDD-YBOCS at one or both assessments after the baseline assessment (11% (*n*=8) at week 6, 23% (*n*=17) at week 12). To account for missingness, we used multiple imputation with predictive mean matching to produce 100 imputed datasets. The imputation model included the following variables: all hypothesized predictor variables (see Table 1), treatment group, as well as secondary outcome variables for depression, quality of life, and functional impairment used in the main outcome paper (Wilhelm et al., 2022). Missing data models for longitudinal outcomes used all prior and concurrent available data at any given timepoint.

Statistical analyses. Using the imputed datasets, we performed screenings of eighteen variables (including group assignment), one at a time, to identify variables that moderated BDD-YBOCS symptom change (i.e., [moderator]*time interaction terms; main outcome) or predicted treatment response or remission (secondary outcomes). Variables that passed the screening step at $p<.25$ (Hosmer et al., 2013) were evaluated for multicollinearity, and the final set of moderators or predictors (all with $r<.6$) was then used in the final generalized linear mixed model (GLMM; main outcome) or multiple logistic regression models (secondary outcomes). Each of the multiple moderator

models was also adjusted for baseline values of each moderator ~~predictor~~ included. Model estimates for each moderator or prediction model were estimated for each imputed dataset and then combined using Rubin's method (Rubin, 1976, 2004). Significance in the final model was evaluated at $p<.05$. Effect sizes (ES) of the moderators were estimated as $ES = \beta_M * 12/SD_{Bsl}$, in which β_M was the estimated effect of the moderator (yes/no for categorical moderator; 1 SD increase for continuous moderator) on the weekly slope and SD_{Bsl} was the standard deviation of BDD-YBOCS scores at baseline in the combined sample; these effect sizes can be interpreted in the same way as Cohen's *d*. Baseline differences in BDD-YBOCS scores in different moderator groups are presented as model-estimated marginal means (EMM) with 95% confidence intervals. All analyses were conducted in SAS 9.4 for Windows.

Machine Learning Analyses. Using the imputed data sets, we separately predicted treatment response and BDD remission status using two linear ML models (logistic regression, support vector machines) and three non-linear ML models (k-nearest-neighbors, decision trees, and random forests). We evaluate the algorithms based on how accurately they predict performance (e.g., whether a participant remitted during treatment) and interpretability (i.e., how readily a human can understand the decisions made within the algorithm to arrive at a given prediction). We measured performance using Area Under the ROC Curve (AUC), which is bounded between 0 and 1; an AUC of .5 indicates prediction performance at around chance levels, while higher scores indicate better performance.

Seventeen variables were available as features that ML models could use for prediction; the two depression binary variables were combined into one ordinal variable

(low, moderate, and severe depression). All non-binary features were z-transformed. We used a forward selection procedure to find the final set of features included in each model, where new features were added one by one in a stepwise fashion until the AUC did not increase by more than .05 with any one additional feature. We used this threshold because each added feature increases the risk of overfitting, and we reasoned that AUC increases smaller than 0.05 are unlikely to be clinically meaningful. The order in which the feature set is selected can be interpreted as a rank order of feature importance.

Each model was validated using a 5-fold cross-validation procedure (James et al., 2013), which was repeated for each of the 100 imputed datasets, resulting in 500 AUC values. The mean of these values was the final AUC estimate for a given model configuration. All models were implemented using the Scikit-Learn version 1.2.1 Python library (Pedregosa et al., 2011).

Results

Participant Characteristics. Participants were predominantly female (85%, *n*=63), White (73%, *n*=54), non-Hispanic (88%, *n*=65), and were on average 27.1(10.0) (M(SD)) years old. Only one participant identified with a non-binary gender identity. Forty-one percent (*n*=30) of the sample identified as a sexual minority, including participants endorsing bisexual (24%), lesbian or gay (3%), or other (14%, including unknown and choosing not to disclose) sexual orientations. Table 1 shows a comparison of baseline characteristics between participants who were randomized to receive app-based CBT immediately compared to those who received it after the 12-week waitlist. More detailed sample characteristics are available in the main outcome

paper (Wilhelm et al., 2022). Participants in the delayed CBT group were more likely to be of a non-White race or of Hispanic ethnicity (p=.026) and provided slightly lower credibility ratings for the app CBT treatment (p=.016) compared to those who were randomized to receive the app CBT immediately (Table 1). Two participants changed SRI medication during the treatment period: one stopped and one increased the dosage of an SRI; both participants noted that these changes did not impact their BDD symptoms.

Statistical Models. Of the eighteen moderators of symptom change screened, eight moderated symptom change in univariate predictor models at a significance level of p<.25 and thus were included in subsequent analyses. Gender identity other than female (ES: -0.83; p=.175), sexual minority status (ES: -0.75; p=.092), racial or ethnic minority status (ES: 0.67; p=.154), no or mild depression (ES: 0.63; p=.147), severe depression (ES: -0.78; p=.175), greater treatment credibility (ES: -0.46; p=.048), greater treatment outcome expectancy (ES: -0.40; p=.072) and being in the immediate app-based CBT group (ES: -1.08; p=.014) were all univariately associated with greater BDD-YBOCS improvements during treatment (Supplemental Table S1). Due to the moderate to strong correlation between credibility and outcome expectancy ratings (r=.65; Supplemental Table S2), we only selected one of these two variables for the multiple moderator model and chose treatment credibility due to its slightly stronger association with the outcome (Supplemental Table S1). In the multivariable moderator model that combined the remaining seven moderators, both being in the immediate treatment group and non-female gender identity remained significant moderators of BDD symptom change over time. Being in the delayed treatment group (after a 12-week waitlist period)

was associated with significantly lower BDD symptom severity at treatment baseline (EMM [95%CI]: -3.3 [-5.4, -1.2], ES=0.75, *p*=.003) and smaller BDD symptom improvements (less negative symptom slopes; ES=0.91, *p*=.036; Table 2) than being in the immediate app-based CBT group. Participants with a non-female gender identity experienced a greater improvement in BDD-YBOCS symptoms (ES=-1.20, *p*=.034; Table 2); we were unable to detect a baseline difference between people with female or non-female gender identities (EMM [95%CI]: 0.2 [-2.6, 3.0], *p*=.866, ES=0.066).

In the logistic regression models predicting treatment response, seven of the eighteen examined predictors were univariately associated with treatment response at a significance level of *p*<.25 and thus were included in subsequent analyses. Gender identity other than female (OR [95% CI]: 6.15 [0.69, 55.16], *p*=.104), sexual minority status (OR [95% CI]: 3.08 [0.96, 9.88], *p*=.058), higher ratings for treatment credibility (OR [95% CI]: 1.75 [0.94, 3.26], *p*=.081), and higher ratings for treatment outcome expectancy (OR [95% CI]: 1.66 [0.92, 2.96], *p*=.090) were univariately associated with greater odds of treatment response by end of treatment (Supplemental Table S1). Age (OR [95% CI]: 0.74 [0.45, 1.22], *p*=.240), BDD duration (OR [95% CI]: 0.73 [0.44, 1.23], *p*=.237), and being in the delayed treatment group (OR [95% CI]: 0.52 [0.18, 1.55], *p*=.242) were associated with lower odds of treatment response (Supplemental Table S1). Due to the high correlation between credibility and expectancy noted above, we selected credibility as the only of these two variables for the multivariable predictor model. Similarly, age and BDD duration were highly correlated (r=0.91), and we chose BDD duration for the multivariable predictor model. In the multivariable predictors model of treatment response, only sexual minority emerged as a significant predictor, such that

participants with sexual minority status were more likely to experience a treatment response by end of treatment than other participants (OR [95% CI]: 3.98 [1.00, 15.78], p=.049; Table 2).

In the logistic regression models predicting remission from BDD symptoms, three of the eighteen examined predictors were univariately associated with symptom remission at a significance level of p<.25 and thus were included in subsequent analyses. Higher ratings for treatment credibility (OR [95% CI]: 2.22 [1.02, 4.86]; p=.046) and higher ratings for treatment outcome expectancy (OR [95% CI]: 1.61 [0.92, 2.82]; p=.094) were univariately associated with greater odds of BDD symptom remission by end of treatment (Supplemental Table S1). Higher baseline BDD symptom severity was univariately associated with lower odds of symptom remission (OR [95% CI]: 0.54 [0.30, 0.98]; p=.044). As before, we selected credibility over expectancy for the multivariable predictor model. In the multivariable predictors model of BDD symptom remission, both treatment credibility and baseline BDD symptom severity predicted remission, such that higher treatment credibility increased the odds of remission, while higher baseline BDD symptom severity decreased the odds of remission (Table 2).

Machine Learning Models. All ML models demonstrated AUC estimates well above the 0.5 chance level for predicting both response and remission. The different models evidenced relatively similar performance (AUCs ranging .61 to .73), with logistic regression, support vector machines, and random forests tending to perform the best overall (Table 3).

Treatment credibility was always the first chosen feature out of the forward selection procedure used across all models and for both outcomes (Table 3). This

suggests that treatment credibility is the most important, informative single variable for guiding predictions about treatment response and remission in our data, regardless of ML model choice. There was variation among the models regarding the features chosen after credibility. For predicting response, gender identity was selected as the second or third feature by all five model types and added more than 0.05 to the AUC in all except the logistic regression model; sexual minority status was only selected by two of these models but added 0.08 and 0.09 to the AUC when selected. For predicting remission, treatment credibility was the sole feature in three of the five models. However, baseline BDD-YBOCS and post-college education evidenced some predictive value depending on the model type.

Finally, although decision trees were not the highest-performing models in terms of overall AUC, decision trees trained to predict treatment response offered a novel, clinically interpretable insight. In a simplified decision tree model of treatment response with only credibility as a predictor (which decreased AUC by .053 compared to the model presented in Table 3), this algorithm consistently chose two decision thresholds on the credibility scale: scores above 22 strongly predicted treatment response (OR [95% CI]: 16.44 [.94, 286.39]) and remission (OR [95% CI]: 10.03 [1.276, 78.90]), while scores of 16 or lower predicted non-response (OR [95% CI]: 0.30 [0.091, 1.01]), and non-remission (OR [95% CI]: 0.31 [0.083, 1.14]). Intermediate scores (17-22) did not yield accurate predictions using credibility scores alone. See Figure 1 for a detailed visualization.

Discussion

In this exploratory analysis of predictors of treatment outcomes following 12-weeks of app-based CBT for BDD, we identified receiving treatment immediately and non-female gender identity as moderators of BDD symptom improvement, sexual minority status as a predictor of treatment response and trend-level moderator of BDD symptom improvement, and treatment credibility and baseline BDD symptom severity as predictors of remission in statistical models. In machine learning models, baseline treatment credibility was the most informative feature in predicting treatment outcomes; clinically meaningfully high and low credibility scores were also uncovered. No other variables examined moderated or predicted symptom improvement across analyses.

Predictors of Better App-based CBT for BDD Outcomes

While not significant in all statistical models, treatment credibility and expectancy were consistent univariate outcome moderators or predictors at a trend-level for all three outcomes (i.e., BDD symptom improvement, treatment response, and remission). Our finding that greater treatment credibility and expectancy predicted better outcomes is consistent with prior studies examining predictors of BDD outcomes in trials of face-to-face CBT versus waitlist (Greenberg et al., 2019), face-to-face CBT versus supportive psychotherapy for BDD (Phillips et al., 2021), and internet-based CBT for BDD versus internet-based supportive psychotherapy for BDD (Flygare et al., 2020). Although credibility was only significantly associated with remission in our multivariable statistical models, ML results found that credibility was the most important feature for correctly determining which participants were most likely to respond favorably and remit. This result is promising, because treatment credibility has been shown to be independently modifiable (Arch et al., 2015). The apparent importance of treatment

credibility is also encouraging from a treatment access perspective; early research suggests that digital CBT for BDD is regarded as similarly credible to in-person CBT for BDD (Bernstein et al., 2023). Notably, results from our decision tree ML model also offered clinically actionable guidelines with respect to pre-treatment credibility ratings: credibility scores at or below 16 on the CEQ were associated with an increased likelihood of non-response, while scores above 22 were associated with an increased likelihood of treatment response. Thus, for people with low credibility scores at treatment onset, additional efforts to increase their belief in the treatment's efficacy by building in onboarding materials that emphasize the treatment's efficacy through data and user testimonials may be helpful. Overall, our findings underscore the importance of pre-treatment credibility and, although our design prevents causal claims, suggest that enhancing treatment credibility prior to treatment onset may improve post-treatment outcomes.

Receiving app-based CBT for BDD immediately, as compared to delayed treatment after a 12-week waiting list, was associated with greater reductions in BDD symptom severity, even when included among other significant moderators. Of note, the delayed treatment group tended to have slightly lower BDD-YBOCS scores at baseline (p=.068) and a slightly larger proportion of participants of racial-ethnic minorities (p<.026). Thus, it is possible this effect may be explained by a delayed group sample that was less acutely severe and therefore had less room for improvement. Structural inequities in health care leading to understandably lower trust in mental health care interventions among People of Color (McGuire & Miranda, 2014) may also be relevant to this finding—especially given that the delayed group also reported significantly lower

treatment credibility at baseline (p<.016). However, it is also possible that motivation, insight, or other practical considerations (e.g., availability, commitment) may wane over time in treatment-seeking individuals having to wait long periods before receiving treatment. Indeed, delayed treatment emerged as a significant predictor even after including treatment credibility as a simultaneous predictor. This finding is important, given that treatment-seeking individuals often wait long periods between the time of seeking and receiving treatment. Thus, hastening access may be critical to maximizing treatment outcomes.

We also found that non-female gender identity was associated with greater BDD symptom improvement and self-reported sexual minority status (e.g., gay, lesbian, bisexual, other; 41% of the sample) was associated with greater BDD symptom improvement (single-moderator model only) and treatment response (multivariable logistic regression model). Digital health interventions, including smartphone-based apps, have been highlighted as a key pathway to improving healthcare disparities among historically non-help-seeking (e.g., non-female gender) and marginalized populations, such as sexual minority individuals. People who identify as a sexual minority tend to already be adept, pervasive users of digital technologies to explore their identities, seek out like-minded peers, and access resources (Gilbey et al., 2020; Lucassen et al., 2014). In addition to being cost-effective, self-guided smartphone-based apps afford greater privacy and minimize stigma when accessing care, are available whenever and wherever users need them, and empower users by providing greater control over their own care (Aguilera, 2015; Bowen et al. 2016; Gilbey et al., 2020; Yousaf et al., 2015). This is especially important for people who identify as sexual

minority identities, and other historically marginalized populations, who are less able to access safe, affirming, and adequate mental health services. More data is needed to see if this level of treatment uptake and enhanced positive treatment outcomes can be replicated in larger samples or with other app-based psychological interventions.

In univariate predictor and moderator models, racial-ethnic minority status and mild and moderate depression were significantly associated with greater BDD symptom improvement, and age and BDD duration were associated with lower odds of remission; however, none of these remained significant in multivariable models. Depression has emerged as a significant predictor of outcome in some, but not all studies, with greater baseline depression associated with poorer outcome (Hogg et al., 2023). Prior studies have not found age or racial-ethnic minority status to predict BDD outcome (Hogg et al., 2023). These discrepant findings may be due to relatively small sample sizes in other studies or methodological differences; the current study used a digitally-based treatment with national recruitment, which may have helped to enhance age and racial-ethnic minority status diversity.

Potential Predictors Yielding Null Effects

Despite some previous findings to the contrary (e.g., Greenberg et al., 2019; Neziroglu et al., 2001; Phillips et al., 2021), BDD-related insight, readiness to change, or SRI use did not moderate symptom improvement or predict treatment response or remission. Education or baseline comorbidity also did not moderate symptom improvement or predict treatment response or remission. It may be that these relationships exist in our data but are too weak to be detected in our somewhat small sample. However, we were primarily interested in identifying the strongest predictors of

CBT for BDD outcomes, because those are likely to best inform clinical decision making.

Promisingly, results from the current study also suggest that individuals regardless of their baseline severity can achieve treatment response with app-based CBT for BDD; however, those with greater BDD severity may require more intensive or longer interventions to achieve remission status. Earlier studies of face-to-face CBT for BDD have shown continued, linear progress and more robust response and remission rates with longer vs. shorter treatment duration (Greenberg et al., 2022; Weingarden et al., 2021; Wilhelm et al., 2019). It is also possible that those with greater BDD severity would benefit more quickly with adjunctive support from a clinician either instead of or alongside an app-based CBT. Additional research is needed to facilitate a stratified care model for BDD as has been tested in depression and anxiety (Wolitzky-Taylor et al., 2023).

Limitations

Limitations of the current study included a relatively small sample (especially for our ML models) and a drop-out rate of 22%, which limited the power available to detect moderators and predictors. In addition, we ran 18 univariate tests for each of our three treatment outcomes with an alpha of .25, making it highly plausible that we would observe a false positive. However, all the moderator and predictor variables chosen were based on prior research or of clinical relevance. Once identified in univariate models, we kept all potential moderators and predictors in their respective multivariable models even if no longer significant to avoid over-identifying spurious predictors. Based on this inclusivity-vs-specificity trade-off, we consider our analyses to be more

exploratory in nature than confirmatory, and any moderator or predictor identified in these analyses warrants further investigation. Our sample was also limited in diversity with respect to non-female gender identities and minority racial-ethnic backgrounds. Due to low prevalence, we grouped one person identifying with a non-binary gender identity with male gender identities, all non-heterosexual sexual orientations together, and all non-White and Hispanic ethnic/racial backgrounds together, respectively, in analyses. While this decision allowed us to explore whether belonging to any minoritized group along a given identity dimension was associated with different treatment responses, more research is needed to examine potential differences in treatment outcomes between patients holding different racial, ethnic, or sexual orientation identities.

Conclusion

App-based CBT for BDD is a promising solution for closing the access to care gap. Coach-guided smartphone-based CBT for BDD is efficacious, even for very ill patients, and apps can be accessed widely and at little cost, preserving more scarce and costly clinician time for individuals who may not respond to app-based treatment. App-based CBT for BDD may be particularly helpful when accessed immediately and by historically non-help-seeking (e.g., non-female gender) and marginalized communities, including individuals who identify as sexual orientation minorities. Treatment developers and those providing human support alongside the app should make efforts to enhance treatment credibility, as greater credibility robustly predicted better outcomes. Additional studies are needed to further understand predictors and moderators of treatment outcomes, including understanding the varying levels of support (self-guided, coach-

guided app, clinician) needed for different patients to maximize personalized, effective treatment of BDD.

References

Aguilera, A. (2015). Digital technology and mental health interventions: opportunities and challenges. *Arbor, 191*, a210. doi: 10.3989/arbor.2015.771n1012.

Angelakis, I., Gooding, P. A., & Panagioti, M. (2016). Suicidality in body dysmorphic disorder (BDD): A systematic review with meta-analysis. *Clinical Psychology Review, 49*, 55-66. doi:10.1016/j.cpr.2016.08.002

Arch, J. J., Twohig, M. P., Deacon, B. J., Landy, L. N., & Bluett, E. J. (2015). The credibility of exposure therapy: Does the theoretical rationale matter? *Behaviour Research and Therapy,* 72, 81–92. https://doi.org/10.1016/j.brat.2015.05.008

Bernstein, E. E., Weingarden, H., Greenberg J. L., Williams, J., Hoeppner, S. S., Snorrason, I., Phillips, K. A., Harrison, O., & Wilhelm S. (2023). Credibility and expectancy of smartphone-based cognitive behavioral therapy among adults with body dysmorphic disorder. *Journal of Obsessive-Compulsive and Related Disorders, 36,* https://doi.org/10.1016/j.jocrd.2023.100781

Bowen, D., Jabson, J., & Kamen, C. (2016). mHealth: an avenue for promoting health among sexual and gender minority populations?. *mHealth*, *2*, 36. https://doi.org/10.21037/mhealth.2016.09.01

Buhlmann, U., Glaesmer, H., Mewes, R., Fama, J. M., Wilhelm, S., Brähler, E., & Rief, W. (2010). Updates on the prevalence of body dysmorphic disorder: A population-based survey. *Psychiatry Research, 178,* 171–175. https://doi.org/10.1016/j.psychres.2009.05.002

Constantino, M. J., Coyne, A. E., Boswell, J. F., Iles, B. R., & Vîslă, A. (2018). A meta-analysis of the association between patients' early perception of treatment

credibility and their posttreatment outcomes. Psychotherapy (Chicago, Ill.), 55(4), 486–495. https://doi.org/10.1037/pst0000168

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 31*(2), 73-86. doi:10.1016/s0005-7916(00)00012-4

Eisen, J., Phillips, K. A., Baer, L., Beer, D., Atala, K., & Rasmussen, S. (1998). Brown Assessment of Beliefs Scale: Reliability and Validity. *American Journal of Psychiatry, 155*, 102-108. doi:10.1176/ajp.155.1.102

Enander, J., Andersson, E., Mataix-Cols, D., Lichtenstein, L., Alström, K., Andersson, G., … Rück, C. (2016). Therapist guided internet based cognitive behavioural therapy for body dysmorphic disorder: Single blind randomised controlled trial. *BMJ (Online)*, *4*, e005923. https://doi.org/10.1136/bmj.i241

Fernández de la Cruz, L., Enander, J., Rück, C., Wilhelm, S., Phillips, K. A., Steketee, G., Sarvode Mothi, S., Krebs, G., Bowyer, L., Monzani, B., Veale, D., & Mataix-Cols, D. (2021). Empirically defining treatment response and remission in body dysmorphic disorder. *Psychological Medicine*, *51*, 83–89. https://doi.org/10.1017/S0033291719003003

Flygare, O., Enander, J., Andersson, E., Ljótsson, B., Ivanov, V. Z., Mataix-Cols, D., & Rück, C. (2020). Predictors of remission from body dysmorphic disorder after internet-delivered cognitive behavior therapy: a machine learning approach. *BMC Psychiatry, 20*, 247. doi:10.1186/s12888-020-02655-4

Gilbey, D., Morgan, H., Lin, A., & Perry, Y. (2020). Effectiveness, Acceptability, and

Feasibility of Digital Health Interventions for LGBTIQ+ Young People: Systematic Review. *Journal of Medical Internet Research*, *22*, e20158. https://doi.org/10.2196/20158

Greenberg, J. L., Jacobson, N. C., Hoeppner, S. S., Bernstein, E. E., Snorrason, I., Schwartzberg, A., Steketee, G., Phillips, K. A., & Wilhelm, S. (2022). Early response to cognitive behavioral therapy for body dysmorphic disorder as a predictor of outcomes. *Journal of psychiatric research*, *152*, 7–13. https://doi.org/10.1016/j.jpsychires.2022.06.001 PMCID: PMC9447469

Greenberg, J. L., Phillips, K. A., Steketee, G., Hoeppner, S. S., & Wilhelm, S. (2019). Predictors of response to cognitive-behavioral therapy for body dysmorphic disorder. *Behavior Therapy, 50*, 839-849. doi:10.1016/j.beth.2018.12.008. PMCID: PMC6582981

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*, Article 2. https://doi.org/10.1016/j.jbi.2008.08.010

Harrison, A., Fernández de la Cruz, L., Enander, J., Radua, J., & Mataix-Cols, D. (2016). Cognitive-behavioral therapy for body dysmorphic disorder: A systematic review and meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *48*, 43–51. https://doi.org/10.1016/j.cpr.2016.05.007

Hogg, E; Adamopoulos, P; Krebs, G; (2023) Predictors and moderators of treatment

response in cognitive behavioural therapy for body dysmorphic disorder: A systematic review. *Journal of Obsessive-Compulsive and Related Disorders*, *38*. 10.1016/j.jocrd.2023.100822.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*, vol. 112. Springer.

Koran, L. M., Abujaoude, E., Large, M. D., & Serpe, R. T. (2008). The prevalence of body dysmorphic disorder in the United States adult population. *CNS Spectrums*, *13*, 316–322. https://doi.org/10.1017/s1092852900016436

Lucassen, M. F. G., Hatcher, S., Stasiak, K., Fleming, T., Shepherd, M., & Merry, S. N. (2013). The views of lesbian, gay and bisexual youth regarding computerised self-help for depression: An exploratory study. *Advances in Mental Health, 12*, 22–33. doi: 10.5172/jamh.2013.12.1.22.

Marques, L., Weingarden, H. M., Leblanc, N. J., & Wilhelm, S. (2011). Treatment utilization and barriers to treatment engagement among people with body dysmorphic symptoms. *Journal of Psychosomatic Research*, *70*, 286–293. https://doi.org/10.1016/j.jpsychores.2010.10.002

McConnaughy, E. A., Prochaska, J. O., & Velicer, W. F. (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research, and Practice, 20*, 368-375

McGuire, T. G., & Miranda, J. (2008). New evidence regarding racial and ethnic

disparities in mental health: policy implications. *Health Affairs (Project Hope)*, *27*, 393–403. https://doi.org/10.1377/hlthaff.27.2.393

Neziroglu, F., Stevens, K. P., McKay, D., & Yaryura-Tobias, J. A. (2001). Predictive validity of the overvalued ideas scale: outcome in obsessive-compulsive and body dysmorphic disorders. *Behaviour Research and Therapy, 39*, 745-756. doi:10.1016/s0005-7967(00)00053-x

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825-2830.

Phillips, K. A. (2000). Quality of life for patients with body dysmorphic disorder. *Journal of Nervous and Mental Disease*, *188*, 170–175. https://doi.org/10.1097/00005053-200003000-00007

Phillips, K. A., Greenberg, J. L., Hoeppner, S. S., Weingarden, H., O'Keefe, S., Keshaviah, A., ... & Wilhelm, S. (2021). Predictors and moderators of symptom change during cognitive-behavioral therapy or supportive psychotherapy for body dysmorphic disorder. *Journal of Affective Disorders*, *287*, 34-40. PMCID: PMC8276884 https://doi.org/10.1016/j.jad.2021.03.011

Phillips, K. A., Hart, A., & Menard, W. (2014). Psychometric evaluation of the Yale–Brown Obsessive-Compulsive Scale Modified for Body Dysmorphic Disorder (BDD-YBOCS). *Journal of Obsessive-Compulsive and Related Disorders, 3*, 205-208. doi:10.1016/j.jocrd.2014.04.004

Phillips, K. A., Hart, A. S., Menard, W., & Eisen, J. L. (2013). Psychometric evaluation of the Brown Assessment of Beliefs Scale in body dysmorphic disorder. *The Journal of Nervous and Mental Disease, 201*, 640-643. doi:10.1097/NMD.0b013e3182983041

Phillips, K. A., Hollander, E., Rasmussen, S. A., Aronowitz, B. R., DeCaria, C., & Goodman, W. K. (1997). A severity rating scale for body dysmorphic disorder: Development, reliability, and validity of a modified version of the Yale-Brown Obsessive Compulsive Scale. *Psychopharmacology Bulletin*, *33*, 17-22.

Phillips, K. A., Menard, W., Fay, C., & Weisberg, R. (2005). Demographic characteristics, phenomenology, comorbidity, and family history in 200 individuals with body dysmorphic disorder. *Psychosomatics*, *46*, 317–325. https://doi.org/10.1176/appi.psy.46.4.317

Rubin, D. B. (1976). Inference and missing data. Biometrika, 63, 581–592. https://doi.org/10.1093/biomet/63.3.581

Rubin, D. B. (2004). Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons, Inc.

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., … Keller, M. B. (2003). The 16-Item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, *54*, 573–583. https://doi.org/10.1016/S0006-3223(02)01866-8

Schieber, K., Kollei, I., de Zwaan, M., & Martin, A. (2015). Classification of body

dysmorphic disorder - what is the advantage of the new DSM-5 criteria? *Journal of Psychosomatic Research*, *78*, 223–227. https://doi.org/10.1016/j.jpsychores.2015.01.002

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). 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. *The Journal of Clinical Psychiatry*, *59 Suppl 20*, 22-33;quiz 34-57.

Veale, D., Anson, M., Miles, S., Pieta, M., Costa, A., & Ellison, N. (2014). Efficacy of cognitive behaviour therapy versus anxiety management for body dysmorphic disorder: a randomised controlled trial. *Psychotherapy and Psychosomatics, 83*(6), 341-353. doi:10.1159/000360740

Weingarden, H., Hoeppner, S. S., Snorrason, I., Greenberg, J. L., Phillips, K. A., & Wilhelm, S. (2021). Rates of remission, sustained remission, and recurrence in a randomized controlled trial of cognitive behavioral therapy versus supportive psychotherapy for body dysmorphic disorder. *Depression and Anxiety*, 10.1002/da.23148. Advance online publication. https://doi.org/10.1002/da.23148 PMCID: PMC8443701

Weingarden, H., Matic, A., Calleja, R. G., Greenberg, J. L., Harrison, O., & Wilhelm, S. (2020). Optimizing Smartphone-Delivered Cognitive Behavioral Therapy for Body Dysmorphic Disorder Using Passive Smartphone Data: Initial Insights From an Open Pilot Trial. *JMIR MHealth and UHealth*, *8*, Article 6. https://doi.org/10.2196/16350

Wilhelm, S., Phillips, K. A., Greenberg, J. L., O'Keefe, S. M., Hoeppner, S. S., Keshaviah, A., … Schoenfeld, D. A. (2019). Efficacy and Posttreatment Effects of Therapist-Delivered Cognitive Behavioral Therapy vs Supportive Psychotherapy for Adults with Body Dysmorphic Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*, *76*, 363–373. https://doi.org/10.1001/jamapsychiatry.2018.4156 PMCID: PMC6450292

Wilhelm, S., Weingarden, H., Greenberg, J. L., Hoeppner, S. S., Snorrason, I., Bernstein, E. E., McCoy, T. H., & Harrison, O. T. (2022). Efficacy of App-Based Cognitive Behavioral Therapy for Body Dysmorphic Disorder with Coach Support: Initial Randomized Controlled Clinical Trial. *Psychotherapy and Psychosomatics*, *91*, 277–285. https://doi.org/10.1159/000524628

Wilhelm, S., Weingarden, H., Greenberg, J. L., McCoy, T. H., Ladis, I., Summers, B. J., Matic, A., & Harrison, O. (2020). Development and Pilot Testing of a Cognitive-Behavioral Therapy Digital Service for Body Dysmorphic Disorder. *Behavior Therapy, 51,* 15–26. https://doi.org/10.1016/j.beth.2019.03.007

Wolitzky-Taylor, K., LeBeau, R., Arnaudova, I., Barnes-Horowitz, N., Gong-Guy, E., Fears, S., Congdon, E., Freimer, N., & Craske, M. (2023). A Novel and Integrated Digitally Supported System of Care for Depression and Anxiety: Findings From an Open Trial. *JMIR Mental Health*, *10*, e46200. https://doi.org/10.2196/46200

Yousaf, O., Grunfeld, E. A., & Hunter, M. S. (2015). A systematic review of the factors associated with delays in medical and psychological help-seeking among men. *Health Psychology Review*, *9*, 264–276.

Author Statement

Contributors

JLG: Conceptualization, design, data acquisition, interpretation, original draft preparation
HW: Conceptualization, design, data acquisition, interpretation, review & editing
SSH: Data acquisition, analysis, interpretation, original draft preparation
RBG: Data acquisition, interpretation, original draft preparation
DK: Analysis, interpretation, original draft preparation
IS: Data acquisition, interpretation, review & editing
OCR: Analysis, interpretation, original draft preparation
MT: Analysis, interpretation, original draft preparation
KED: Interpretation, review & editing
RCV: Interpretation, review & editing
AS: Interpretation, review & editing
OH: Resources, review & editing
SW: Conceptualization, funding acquisition, design, data acquisition, interpretation, review & editing

Acknowledgements

None

Role of the Funding Source

Data collection was supported by Koa Health. Investigators from MGH (SW, HW, and JG) developed the Perspectives app in collaboration with technologists and designers from Koa Health via a user-centered, iterative design approach. The MGH investigators were responsible for the study design and for the execution of the study; Koa Health supported the technical activities related to the development and deployment of the app and provided funding for the project. Koa Health had no role in the recruitment of participants and had no access to data during the trial. Dr. Weingarden was supported in part by the National Institute of Mental Health of the National Institutes of Health under Award Number K23MH119372 (Weingarden). Mr. Talbot was supported by in part by the National Institute of General Medical Sciences Award Number T32GM144273. The content is solely the responsibility of the authors

and does not necessarily represent the official views of the National Institutes of Health. Drs. Costilla-Reyes and Dr. Solar-Lezama were supported in part by the National Science Foundation under Award Number 1918839.

Declaration of Interest

Dr. Greenberg has received research support from Koa Health and is a presenter for the Massachusetts General Hospital Psychiatry Academy in educational programs supported through independent medical education grants from pharmaceutical companies. She has received speaking honoraria from L'Oreal (for a presentation at a SkinCeuticals cosmetic surgery and dermatology conference) and RBC Consultants for the CeraVe Psychodermatology Advisory Board. Dr. Weingarden receives research support from Koa Health and is a presenter for the Massachusetts General Hospital Psychiatry Academy in educational programs supported through independent medical education grants from pharmaceutical companies. Additionally, Dr. Weingarden has a consulting agreement with Hello Therapeutics, Inc. Dr. Wilhelm is a presenter for the Massachusetts General Hospital Psychiatry Academy in educational programs supported through independent medical education grants from pharmaceutical companies; she has received royalties from Elsevier Publications, Guilford Publications, New Harbinger Publications, Springer, and Oxford University Press. Dr. Wilhelm has also received speaking honoraria from various academic institutions and foundations, including the International Obsessive Compulsive Disorder Foundation, the Tourette Association of America, and the Centers for Disease Control and Prevention. In addition, she received payment from the Association for Behavioral and Cognitive

Therapies for her role as Associate Editor of the Behavior Therapy journal and John Wiley & Sons, Inc. for her role as Associate Editor of the journal Depression & Anxiety. Dr. Wilhelm has also received honoraria from One-Mind for her role on PsyberGuide Scientific Advisory Board. Dr. Wilhelm is also on the Scientific Advisory Board for Koa Health, Inc, Jimini Health, and Noom, Inc. Dr. Wilhelm has received research support from Koa Health, Inc. Drs. Hoeppner, Snorrasson, and Vanderkruik, Ms. Berger-Gutierrez, and Mr. Klare received research support from Koa Health. Dr. Solar-Lezama receives salary support from IBM, Intel and Siemens has a consulting agreement with Google. Dr. Harrison is Founder/CEO of Koa Health, a digital mental health company that collaborated with Dr Wilhelm and her team at MGH to build Perspectives. Dr Harrison also serves on the WHO Roster of Experts for Digital Health, sits on the Board of EMPOWER (a non-profit organization promoting the training of community health workers to provide mental healthcare), and is a member of the Expert Panel for implementing the Wellcome Trust's mental health strategy. Dr. Costilla-Reyes, Mr. Talbot, and Ms. Daniel have no conflicts of interest to report.

Table 1. Baseline characteristics of participants with body dysmorphic disorder enrolled in an app-based CBT treatment study.

Variable	Immediate app-CBT n=40	Waitlisted app-CBT n=34	*p*
Demographics			
Age, y, mean (SD)	27.8 (9.9)	26.2 (10.2)	0.505
Male and other [a], % (n)	10.0 (4)	20.6 (7)	0.326
Non-White race or Hispanic ethnicity, % (n)	25.0 (10)	50.0 (17)	0.026
Sexual minority, % (n)	40.0 (16)	41.2 (14)	0.918
Education, % (n)			
<= High school graduate	20.0 (8)	17.7 (6)	1.000
Graduate or professional school	25.0 (10)	26.5 (9)	1.000
Clinical Characteristics			
BDD-YBOCS total score, mean (SD)	29.9 (4.0)	28.0 (4.7)	0.068
BABS total score, mean (SD)	15.1 (3.2)	13.5 (4.6)	0.095
QIDS-SR, % (n)			
No/Mild depression	47.5 (19)	50.0 (17)	0.830
Severe/Very severe depression	17.5 (7)	11.8 (4)	0.533
BDD duration, y, mean (SD)	14.0 (9.9)	12.8 (12.3)	0.639
Any comorbidity, % (n)	67.5 (27)	67.7 (23)	0.989
SRI use, % (n)	22.5 (9)	29.4 (10)	0.596
Treatment Expectations and Experience			
CEQ expectancy total score, mean (SD)	14.8 (4.6)	13.7 (4.6)	0.320
CEQ credibility total score, mean (SD)	19.4 (3.5)	17.1 (4.6)	0.016
URICA RTC, mean (SD)	10.2 (1.3)	9.6 (1.9)	0.135
COVID-19 Impact			
BDD symptom worsening, mean (SD)	-0.03 (1.1)	0.3 (1.0)	0.199

Notes: [a] includes 12 males and 1 genderqueer or nonbinary person; CBT = cognitive behavioral therapy; BDD-YBOCS = Body Dysmorphic Disorder modification of the Yale-Brown Obsessive Compulsive Scale; BABS = Brown Assessment of Beliefs Scale; QIDS-SR = Quick Inventory of Depressive Symptomatology, self-report; BDD = body dysmorphic disorder; SRI = Serotonin reuptake inhibitor medication; CEQ = Credibility and Expectancy Questionnaire; URICA RTC = University of Rhode Island Change Assessment Scale readiness to change index; missing data: BDD age of onset (n=2), treatment credibility & expectancy (n=2), URICA (n=1).

Table 2. Summary of final models examining moderators of body dysmorphic disorder symptom change, treatment response, or remission with multiple independent variables.

Outcome	Moderator/Predictor	Estimate[a]	[95% CI]	ES	p
BDD-YBOCS change over time (baseline to week 12)					
	Group (delayed vs. immediate app-CBT)	0.33	[0.02, 0.65]	0.91	0.036
	Gender identity (male & other vs. female)	-0.44	[-0.84, -0.03]	-1.20	0.034
	Sexual minority (yes vs. no)	-0.30	[-0.60, 0.01]	-0.81	0.057
	Racial or ethnic minority (yes vs. no)	0.12	[-0.19, 0.44]	0.33	0.449
	No or mild depression (yes vs. no)	0.16	[-0.16, 0.47]	0.43	0.331
	Severe depression (yes vs. no)	-0.18	[-0.60, 0.23]	-0.50	0.388
	Treatment credibility (1 SD increase)	-0.12	[-0.29, 0.05]	-0.33	0.161
Treatment response, week 12 (end of treatment)					
	Group (delayed vs. immediate app-CBT)	0.52	[0.14, 1.87]		0.316
	Gender identity (male & other vs. female)	9.60	[0.85, 107.92]		0.067
	Sexual minority (yes vs. no)	3.98	[1.00, 15.78]		0.049
	BDD duration (1 SD increase)	0.85	[0.45, 1.61]		0.619
	Treatment credibility (1 SD increase)	1.78	[0.80, 3.95]		0.157
Remission, week 12 (end of treatment)					
	Treatment credibility (1 SD increase)	2.36	[1.07, 5.21]		0.034
	BDD-YBOCS total score (1 SD increase)	0.49	[0.25, 0.94]		0.033

Notes: [a] Estimates are regression coefficients for moderator effects in the model with BDD-YBOCS scores over time as the outcome and odds ratios for the outcomes treatment response and remission; CI = Confidence Interval; ES = effect size, calculated for moderator effects only; BDD-YBOCS = Body Dysmorphic Disorder modification of the Yale-Brown Obsessive Compulsive Scale; CBT = cognitive behavioral therapy. Continuous variables were z-transformed prior to analysis, so that estimates reflect the effect of a one standard deviation change in the moderator or predictor.

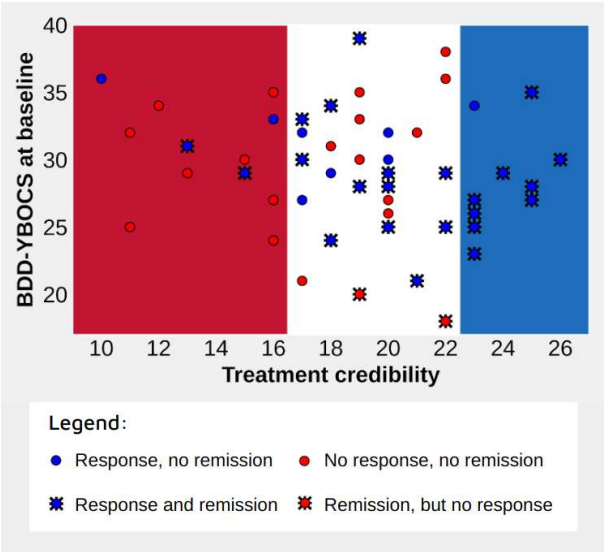
Table 3. Machine learning models predicting treatment response and remission

Response			
Model type	**AUC**	**Max tree depth**	**Features in order of forward selection**
Logistic Regression	0.726	N/A	credibility, sexual minority status
SVM	**0.733**	N/A	credibility, sexual minority status, gender identity, BABS
Decision Tree	0.696	5	credibility, gender identity
Random Forest	0.723	5	credibility, gender identity
KNN	0.682	N/A	credibility, gender identity

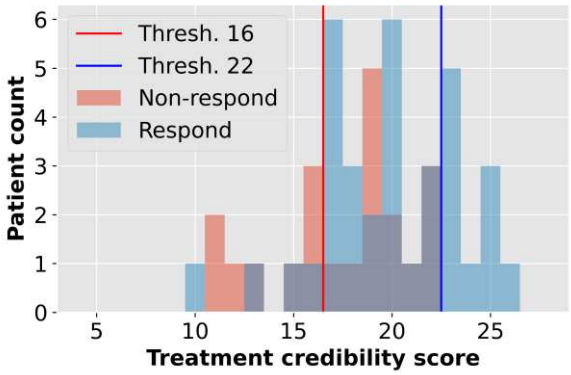
Remission			
Model type	**AUC**	**Max tree depth**	**Features in order of forward selection**
Logistic Regression	0.697	N/A	credibility
SVM	**0.718**	N/A	credibility, BDD-YBOCS
Decision Tree	0.616	2	credibility
Random Forest	0.709	4	credibility
KNN	0.688	N/A	credibility, post-graduate education status

Notes: SVM = Support Vector Machines; KNN = K-Nearest-Neighbors; AUC = Area under the ROC Curve (average); BABS = Brown Assessment of Beliefs Scale; BDD_-YBOCS = Body Dysmorphic Disorder modification of the Yale-Brown Obsessive Compulsive Scale. The highest AUC value is in bold for each outcome. The right-most column shows the features chosen by forward selection in the order in which they were selected (can be thought of as a rank order list in terms of feature importance). All models were implemented using Scikit-learn. Max tree depth is a tuned decision tree/random forest hyperparameter that limits the maximum number of sequential binary decisions allowed in the learned tree(s). At each forward selection step, max tree depths of 1-6 were tried for decisions trees (and 2-7 for random forests), with the best model used for decisions about variable choices and stopping the forward selection. Logistic regression used the “liblinear” solver in Scikit-learn. All other model hyperparameters were set to the default values for model classes in Scikit-learn version 1.2.1 (documentation accessible at https://scikit-learn.org/1.2/).

Figure 1. Analysis of a decision tree classification strategy using only a single threshold on treatment credibility

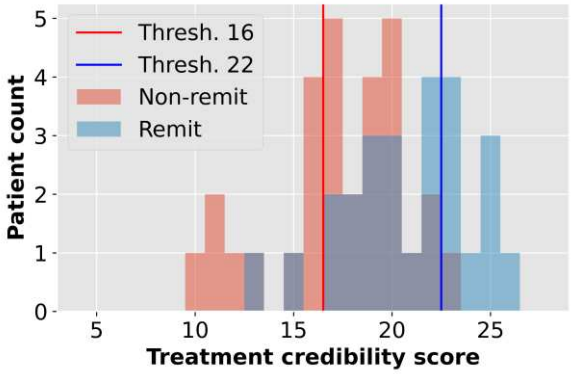


(a) Visualization of decision tree boundaries reveals striking effects at extreme values of credibility.



(b) Histogram of credibility scores, separated by treatment response

Notes: BDD-YBOCS = Body Dysmorphic Disorder modification of the Yale-Brown Obsessive Compulsive Scale. In panel (a), each patient is one point plotted by their credibility score and baseline BDD-YBOCS score. The red region contains all patients with credibility scores of 16 or lower, and the blue region contains those with scores above 22. Panels (b) and (c) show histograms of credibility scores, with histogram colors separated by response and remission.



(c) Histogram of credibility scores, separated by BDD remission

Supplemental Table 1. Univariate tests of baseline moderators of body dysmorphic disorder (BDD) symptom change over time or predictors of treatment response or BDD symptom remission.

	Moderator: BDD-YBOCS change				Predictor: Treatment response at week 12			Predictor: BDD remission at week 12		
Predictor/moderator	**Estimate**	**[95% CI]**	**ES**	**Pr > |t|**	**OR**	**[95% CI]**	**Pr > |t|**	**OR**	**[95% CI]**	**Pr > |t|**
Age	-0.005	[-0.159, 0.150]	-0.01	0.954	**0.74**	**[0.45, 1.22]**	**0.240**	0.78	[0.46, 1.33]	0.355
Male or other gender identity[a]	**-0.304**	**[-0.744, 0.136]**	**-0.83**	**0.175**	**6.15**	**[0.69, 55.16]**	**0.104**	2.24	[0.49, 10.27]	0.299
Racial or ethnic minority	**0.245**	**[-0.092, 0.582]**	**0.67**	**0.154**	0.85	[0.28, 2.60]	0.775	0.54	[0.17, 1.67]	0.282
Sexual minority	**-0.276**	**[-0.597, 0.045]**	**-0.75**	**0.092**	**3.08**	**[0.96, 9.88]**	**0.058**	1.64	[0.55, 4.86]	0.371
Highschool education or less	0.134	[-0.272, 0.539]	0.36	0.518	0.93	[0.25, 3.43]	0.907	0.63	[0.16, 2.50]	0.508
College education or more	0.024	[-0.349, 0.397]	0.07	0.900	1.12	[0.33, 3.86]	0.857	1.09	[0.33, 3.62]	0.883
BDD duration	-0.004	[-0.161, 0.153]	-0.01	0.957	**0.73**	**[0.44, 1.23]**	**0.237**	0.79	[0.46, 1.36]	0.391
Mild depression	**0.231**	**[-0.081, 0.542]**	**0.63**	**0.147**	0.76	[0.26, 2.18]	0.605	1.74	[0.62, 4.90]	0.295
Severe depression	**-0.288**	**[-0.704, 0.128]**	**-0.78**	**0.175**	1.53	[0.36, 6.54]	0.563	0.75	[0.19, 2.95]	0.680
Treatment credibility	**-0.168**	**[-0.335, -0.002]**	**-0.46**	**0.048**	**1.75**	**[0.94, 3.26]**	**0.080**	**2.22**	**[1.02, 4.86]**	**0.046**
Treatment expectancy	**-0.145**	**[-0.304, 0.013]**	**-0.40**	**0.072**	**1.66**	**[0.92, 2.96]**	**0.090**	**1.61**	**[0.92, 2.82]**	**0.094**
URICA readiness to change	-0.073	[-0.245, 0.100]	-0.20	0.409	1.15	[0.64, 2.06]	0.639	1.31	[0.69, 2.51]	0.405
Worse BDD due to COVID-19	0.022	[-0.143, 0.188]	0.06	0.792	0.83	[0.48, 1.42]	0.491	0.84	[0.49, 1.46]	0.542
SRI medication use	-0.104	[-0.471, 0.263]	-0.28	0.579	1.47	[0.42, 5.17]	0.550	0.76	[0.23, 2.49]	0.649
Any current comorbid diag. (vs. 0)	-0.156	[-0.496, 0.183]	-0.43	0.367	1.38	[0.46, 4.10]	0.565	1.77	[0.58, 5.41]	0.318
BDD-YBOCS total score	n/a	n/a	n/a	n/a	1.19	[0.70, 2.03]	0.527	**0.54**	**[0.30, 0.98]**	**0.044**
BABS total score	-0.083	[-0.245, 0.078]	-0.23	0.311	1.34	[0.78, 2.33]	0.291	0.84	[0.49, 1.44]	0.525
Delayed treatment group	**0.396**	**[0.082, 0.711]**	**1.08**	**0.014**	**0.52**	**[0.18, 1.55]**	**0.242**	0.60	[0.21, 1.74]	0.349

Note: BDD-YBOCS = Body Dysmorphic Disorder modification of the Yale-Brown Obsessive Compulsive Scale; BABS = Brown Assessment of Beliefs Scale; ES = effect size, calculated for moderator effects only; diag. = diagnosis. Moderators of symptom change (BDD-YBOCS change models) or predictors of outcomes (treatment response or remission at week 12 models) are bolded if they were significant at p<.1. Continuous variables were z-transformed prior to analysis, so that estimates reflect the effect of a one standard deviation change in the predictor.

Supplemental Table 2. Correlations between baseline variables examined as moderators of body dysmorphic disorder symptom severity change or predictors of treatment response or remission.

Baseline variable	1	2	3	4	5	6	7	8	9
1 Age									
2 Male or other gender identity	-0.10								
3 Racial or ethnic minority	**-0.36** **	0.00							
4 Sexual minority	**-0.27** *	-0.07	-0.09						
5 Highschool education or less	**-0.36** **	0.35	0.36	**0.41** *					
6 Graduate or professional school	0.18	-0.16	-0.34	0.03	-0.98 n/a				
7 BDD duration	**0.91** ***	-0.15	**-0.40** ***	-0.14	**-0.27** *	0.15			
8 Mild depression	-0.07	0.39	0.08	-0.23	0.15	0.08	-0.05		
9 Severe depression	0.20	-0.18	-0.33	-0.23	-0.97 n/a	-0.38	0.13	-0.99 n/a	
10 Treatment credibility	-0.09	0.02	0.05	-0.09	-0.04	-0.08	-0.12	-0.12	0.08
11 Treatment expectancy	-0.03	0.16	0.08	-0.05	0.20	-0.02	-0.11	-0.24 *	-0.08
12 URICA readiness to change	0.15	0.00	-0.03	-0.02	-0.02	0.03	0.15	-0.20	0.05
13 Worse BDD due to COVID-19	-0.14	-0.07	0.25	**0.36** **	**0.45** **	-0.02	-0.15	0.02	-0.02
14 SRI medication use	-0.05	0.03	0.22	0.24	-0.09	0.14	-0.09	-0.34	0.20
15 Any current comorbid diagnoses	-0.10	0.09	0.08	0.17	-0.06	-0.10	-0.12	-0.22	**0.48** *
16 BDD-YBOCS total score	**0.27** *	-0.16	-0.01	0.00	-0.13	-0.14	**0.24** *	**-0.37** **	**0.41** ***
17 BABS total score	0.08	-0.16	-0.12	-0.06	0.00	0.09	0.10	-0.12	**0.29** *
18 Delayed treatment group	-0.08	0.28	**0.40** *	0.02	-0.05	0.03	-0.06	0.04	-0.16

Notes: * = p<.05, ** = p<.01, *** = p<.001; BDD = body dysmorphic disorder; URICA = University of Rhode Island Change Assessment Scale; SRI = Serotonin reuptake inhibitor medication; BDD-YBOCS = Body Dysmorphic Disorder modification of the Yale-Brown Obsessive Compulsive Scale; BABS = Brown Assessment of Beliefs Scale; correlation coefficients were based on Pearson correlation for correlations between two continuous measures or continuous measures and ordinal measures (i.e., worse BDD due to COVID-19), point biserial correlations for correlations between continuous variables with binary variables, and tetrachoric correlations for correlations between binary variables.

Supplemental Table 2 (Continued). Correlations between baseline variables examined as moderators of body dysmorphic disorder symptom severity change or predictors of treatment response or remission.

Baseline variable	10	11	12	13	14	15	16	17
1 Age								
2 Male or other gender identity								
3 Racial or ethnic minority								
4 Sexual minority								
5 Highschool education or less								
6 Graduate or professional school								
7 BDD duration								
8 Mild depression								
9 Severe depression								
10 Treatment credibility								
11 Treatment expectancy	**0.65** ***							
12 URICA readiness to change	**0.45** ***	**0.39** ***						
13 Worse BDD due to COVID-19	-0.01	0.03	-0.03					
14 SRI medication use	0.11	0.17	0.03	0.24				
15 Any current comorbid diagnoses	0.16	0.09	0.14	0.10	**0.52** **			
16 BDD-YBOCS total score	0.03	-0.07	-0.07	0.16	0.04	0.05		
17 BABS total score	-0.08	-0.18	-0.22	-0.02	-0.05	-0.07	**0.47** ***	
18 Delayed treatment group	**-0.28** *	-0.12	-0.18	0.20	0.13	0.00	-0.21	-0.20

Notes: * = p<.05, ** = p<.01, *** = p<.001; BDD = body dysmorphic disorder; URICA = University of Rhode Island Change Assessment Scale; SRI = Serotonin reuptake inhibitor medication; BDD-YBOCS = Body Dysmorphic Disorder modification of the Yale-Brown Obsessive Compulsive Scale; BABS = Brown Assessment of Beliefs Scale; correlation coefficients were based on Pearson correlation for correlations between two continuous measures or continuous measures and ordinal measures (i.e., worse BDD due to COVID-19), point biserial correlations for correlations between continuous variables with binary variables, and tetrachoric correlations for correlations between binary variables.