Supplementary Online Content

Wilhelm S, Phillips KA, Greenberg JL, et al. 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.* Published online February 20, 2019. doi:10.1001/jamapsychiatry.2018.4156

eMethods. Recruitment, Demographics, Description of the Treatments, Training and Fidelity of Therapists, Training and Fidelity of Independent Evaluators, Randomization Procedures, Blinding Procedures, Adverse Events Monitoring, Survival Analysis of Drop-out, Assessment of Missingness Methods

eResults. Treatment Adherence and Competence, Blinding Effectiveness, Primary and Secondary Outcomes During Follow-up: Table of Model Estimates, Survival Analysis of Dropout, Assessment of Missingness Results

eTable 1. Tests of Intercept Differences, Overall Slope Effects, and Slope Differences in Random Intercept Random Coefficient Models of Primary and Secondary Treatment Outcomes During Follow-up (Weeks 24-50)

eTable 2. Baseline Characteristics Evaluated as Risk Factors for Early Drop-out from Treatment

eFigure. Survival Curves of Time to Drop-out from Therapy During the Treatment Period

This supplementary material has been provided by the authors to give readers additional information about their work.

1a. eMethods

Recruitment

Participants were recruited through advertisement, clinician referral, and self-referral to the BDD programs at MGH and RIH. Both sites mailed study letters and brochures to local mental health and medical providers (e.g., dermatologists, plastic surgeons). Study flyers were posted in the community (e.g., hospitals, gyms, salons). Ads were run on radio stations, on television, in local and college newspapers, on social media, and on public transportation. Information about the study was posted on mental health websites (e.g., National Alliance on Mental Illness).

Demographics

All demographic information was collected via participant self-report. Demographic data collected included gender, sexual orientation, race, ethnicity, religious affiliation, marital status, living situation (e.g., supervised living vs. living with partner vs. alone), highest education level completed, why participants did not complete their education degrees (only if no degree completion was indicated), occupational status, number of weeks of unemployment if unemployed, and economic support (e.g., employment, welfare, disability). The options for race (American Indian/Alaska Native/First Nations, Asian, Black/African American, Native Hawaiian/Other Pacific Islander, White, more than one race, other) and ethnicity (Hispanic or Latino vs. Not Hispanic or Latino) were investigator-defined and based on the National Institutes of Health (NIH) policy on reporting race and ethnicity data. Demographic data about sexual orientation, religious affiliation, living situation, reasons for degree non-completion, and economic support were collected as part of a standardized clinic-wide intake form used across multiple studies and are not presented.

Description of the treatments

Cognitive-Behavioral Therapy (CBT-BDD): Our CBT-BDD manual,[1] developed and pilot tested with an NIMH R34 grant,[2,3] consists of 22 individual 60-minute sessions provided over 24 weeks. If the patient responded to treatment (i.e., ≥ 30% reduction in BDD-YBOCS score), then booster sessions occurred one and three months after the end of treatment. The manual informs therapists about BDD, its assessment, and its treatment. It suggests helpful attitudes the and behaviors for the therapist and those to avoid (e.g., providing reassurance in response to reassurance seeking about appearance concerns). The three introductory sessions include psychoeducation, case formulation, motivational enhancement, and goal setting. Then core treatment components used with all patients are described (e.g., exposure and ritual prevention). Thereafter, optional treatment modules can be used with selected patients who have relevant symptoms requiring specific strategies (e.g., skin picking is treated with modified habit reversal). This ensures that treatment is personalized. Modules can be used for as long as necessary. All patients received relapse prevention (last two sessions). The manual includes didactic handouts, in-session forms, and homework forms.

CBT-BDD is a structured, skills-based treatment informed by the cognitive-behavioral model of BDD. Based on information from the evaluation, therapists and patients collaboratively derive an individualized cognitive-behavioral model of their BDD symptoms, including mechanisms (e.g., maladaptive beliefs and behaviors) that may cause and/or maintain the patient's BDD symptoms. This model is used to develop a treatment plan and select optional modules. Motivation Enhancement and Goal Setting: Motivational techniques engage patients who are ambivalent about starting therapy or to enhance motivation during treatment. When setting goals, therapist and patient discuss how treatment can enhance valued life activities (e.g., developing relationships, advancing in their job) that have been restricted due to BDD. Cognitive Interventions help patients evaluate and change maladaptive thoughts. They are integrated with mindfulness exercises (e.g., observing thoughts without expelling or clinging to them). These interventions promote a new perspective and a more flexible relationship with thoughts and feelings. Exposure and Ritual Prevention helps patients learn to gradually approach anxiety-provoking or avoided situations without ritualizing. Mindfulness/ Perceptual Retraining strategies help patients to develop a more holistic sense of their appearance and be more present in their day-to-day and social functioning. Advanced Cognitive Strategies help patients broaden the basis of their self-worth to include non-appearance factors (e.g., talents, personality). Relapse Prevention helps patients prevent, expect, and react effectively to future setbacks. Therapists can also select from optional treatment modules based on individual patients' symptoms: 1) Skin Picking and Hair Pulling Module uses habit reversal for these rituals aimed at improving appearance; 2) Muscularity and Shape/Weight Module is often used for patients (mostly males) with muscle dysmorphia, a malignant form of BDD associated with greater morbidity in which patients think they are insufficiently big and muscular. It is also used for other body shape/size concerns; 3) Cosmetic Treatment Module is for patients considering surgical, dermatologic, dental, or other

cosmetic treatment for BDD; and 4) Mood Management Module includes CBT approaches (e.g., activity scheduling, cognitive restructuring) for more severely depressed patients.

Supportive Psychotherapy (SPT): SPT matched CBT-BDD in length and number of sessions (22 individual, 60-minute sessions over 24 weeks with booster sessions for patients who respond to treatment one and three months after treatment ends). However, SPT differs from CBT-BDD in all core components. We adapted and enhanced Pinsker's published SPT manual[4] by adding three introductory sessions to provide participants with a treatment rationale and psychoeducation about BDD. These adaptations are similar to those used in prior NIMH-funded treatment studies[5] and were intended to enhance treatment engagement, quality, and credibility. Our manual delineates disallowed CBT-BDD techniques (e.g., cognitive restructuring, exposure and ritual prevention. To further differentiate CBT-BDD from SPT, we eliminated reframing, skills training, and role playing from SPT, which are not generally considered essential SPT elements. SPT focuses on therapeutic interactions and techniques that are nonspecific and common to psychotherapies. For example, SPT focuses on establishing and maintaining a therapeutic alliance, supporting adaptive coping skills, improving self-esteem, and expressing emotions[4,6]. Topics often include relationships, emotions, and current life issues. SPT techniques focus on conveying empathy, validation, and positive regard; comforting and supporting patients coping with distress; using a relaxed conversational style; providing praise and encouragement; clarification; and ventilation (allowing expression of intense feelings[4]). Sessions are non-directive with content driven by the patient.

Training and fidelity of therapists

Doctoral- or masters-level therapists (2-3 per treatment per site) received training on their respective treatment manual, passed a knowledge test, and participated in weekly cross-site supervision teleconferences with a CBT (SW) or SPT (SMO) expert at MGH. A higher proportion of SPT therapists had doctoral degrees at MGH (2 of 2) than RIH (1 of 3), and SPT therapists at MGH had treated more patients prior to study involvement (weighted mean ± SD: 81 ± 84) than at RIH (12 ± 32). To avoid allegiance effects, therapists provided only the treatment for which they were qualified and for which they received study-specific training. We prevented contamination across therapy conditions via: 1) training on the treatments' distinctiveness, 2) ongoing supervision, 3) explication in the SPT manual of disallowed CBT techniques, 4) therapist self-rated checklists of any disallowed therapy technique during a session, and 5) fidelity monitoring, including 6) a checklist of items in adherence measures on disallowed techniques.

All sessions were audiotaped, and 16% (n=334) were randomly selected to be rated by an independent, doctoral-level reviewer for fidelity and quality using measures modified from prior NIMH trials[3,5,7]. Adherence raters were experienced in CBT and SPT and further trained and supervised by the CBT and SPT supervisors. Core elements of each treatment session (4-20 items) were rated for adherence on a 7-point scale (1=not at all to 7=completely adherent) and then a global rating of adherence was assigned. Competence was rated on 18 aspects (39 items) for CBT-BDD and 13 aspects (27 items) for SPT. Each item was scored on a 5-point scale (1=not at all to 5=completely competent) and then a global rating of competence was assigned.

For each SPT session as part of the adherence rating form, the rater was asked to indicate on a checklist whether the therapist used any of the following CBT techniques, and if so, to rate the number of minutes per session focused on these CBT techniques: self-monitoring, cognitive restructuring, identifying cognitive errors, exposure, response prevention, mindfulness/perceptual retraining, activity scheduling, and other noncritical SPT techniques that were disallowed in the current protocol to prevent potential crossover with CBT (reframing, skills training, role playing). If the therapist used or suggested any disallowed techniques, including CBT strategies or SPT strategies that may potentially overlap with CBT, the rater was instructed to give an overall adherence rating for the session < 5 (i.e., less than moderately adherent). Procedures were in place to recertify, and replace if necessary, therapists who for 2 consecutive sessions received an adherence or competence rating below the certification standard (overall scores of >6 [mostly or completely] on adherence and >4 [mostly or completely] on competence); however, this never occurred. Among the specific CBT-BDD adherence ratings, six criteria evaluated specific aspects of protocol adherence (e.g., review of the previous session, agenda setting, homework review) at sessions after the first one, followed by 1-12 adherence questions that were content-specific (e.g., psychoeducation using the CBT model, identifying rituals, addressing negative thinking before doing an exposure). For SPT, sessions were rated for adherence to either specific "enhanced" SPT therapy content (e.g., assessment of and feedback about BDD symptoms, providing an overview of BDD) or core SPT strategies (e.g., discussion of current concerns, facilitation of self-esteem, support of adaptive skills). With regard to specific competence criteria, therapists from both treatment arms were evaluated on aspects of conversational communication style, knowledge, empathy, dealing with treatment interfering behaviors and other common therapy skills. CBT-BDD therapist competence was further evaluated by assessing CBT-related skills such as properly designing behavioral experiments/exposures, reinforcing

patients' change talk, and verifying working hypotheses. By contrast, SPT therapist competence was further examined by assessing skills such as using praise, encouragement, and discussions about self-esteem in appropriate contexts.

We examined treatment and site differences in global treatment adherence and competence ratings with Fisher's exact tests. To explore potential site differences in specific SPT adherence and competence criteria, we used Wilcoxon's rank sum test for non-parametric data, because most items (>75%) were rated as completely adherent or competent. For competence or adherence aspects that had multiple specific item ratings, we calculated the mean of specific items prior to analysis. To make reports of subtle site differences more informative, we also calculated the percentage of ratings that were not rated as perfectly adherent/competent, as inter quartile ranges often did not capture this information. To be counted as 'perfectly' adherent/competent, the mean of all component items had to be equal to the highest competence (=5) or adherence (=7) rating. (See results below.)

For quality assurance of the adherence-competence ratings, 16% (53/334) of the rated sessions were randomly selected to be double-rated by a second, independent doctoral-level clinician rater. Inter-rater reliability was evaluated using Gwet's AC1 criterion.[8] Percent agreement between two independent adherence-competence raters on global session ratings was very high (80% agreement for adherence, 93% for competence), and chance-adjusted inter-rater reliability was high for adherence (AC1=0.80) and very high for competence ratings (AC1=0.94).

Training and fidelity of independent evaluators

Clinician-administered measures were administered by masters- or doctoral-level independent evaluators (IEs) blinded to treatment. IEs received rigorous training on study instruments, had to demonstrate high inter-rater reliability with a gold standard before administering assessments, and participated in regular supervision (by KAP). All interviews were audio-taped, and 15% of randomly selected tapes were rated to prevent drift. Inter-rater reliability (intraclass correlations)[9] was excellent for the BDD-YBOCS (ICC=0.95) and BABS (ICC=0.97).

Randomization procedures

At the baseline visit, subjects were randomized in a 1:1 ratio to CBT-BDD or SPT. The MGH Biostatistics Center randomized all subjects (stratified by site) using the secure web-based RS2 system.

Blinding procedures

Independent evaluators remained blind until the last participant completed all study procedures. To ensure blindness, IEs: 1) were not told treatment assignments; 2) were trained to focus on outcome measurement only; and 3) did not attend treatment supervision, weekly study meetings, or cross-site conference calls. Patients, therapists, and study staff were reminded regularly not to discuss treatment conditions with the IE, and treatment and assessment recordings were kept separately. In the event the blind was inadvertently broken, the independent evaluator completed a Treatment Unblinding Form, and the participant was assigned a new, blinded independent evaluator. There were eight incidents of reported unblinding: five among patients assigned to CBT (two at MGH, three at RIH), and three among patients assigned to SPT (two at MGH, one at RIH); all occurred within the first 16 weeks of treatment, and each patient was immediately assigned a new, blinded independent evaluator. The IE was asked to guess the treatment condition of each subject after completion of follow-up assessments or, for dropouts, after the last treatment session. IEs were first asked: "Which therapy do you believe the patient was randomized to?" with the answer options of CBT and SPT (there was no 'don't know' option), followed by: "How certain are you in your knowledge of which therapy the patient was randomized to?" with the answer options 0 = Not at all, 1 = A little certain, 2 = Somewhat certain, 3 = Moderately certain, 4 = Fairly certain, and 5 = Completely certain. The answers from the first question were re-coded to 'don't know' if the IE answered 'not at all [certain]'. The blinding assessment, originally only assessed in cases when IEs thought that they had been unblinded, was added belatedly to the trial for all participants after 07/29/13, so that the collection of blinding data for 23% (27/120) participants' IEs was missed. We used Bang's 2x3 blinding index[10] to calculate the blinding index with confidence interval for each treatment. Bang's BI can range from -1 to +1, where 0 indicates the best possible blinding situation, positive values indicate failure in masking above random chance (i.e., more people guessing their treatment assignment correctly than would be expected by random chance), and negative values indicate either successful masking or rather a failure of masking in the other direction (i.e., more people guess the incorrect treatment than would be expected by random chance).

Adverse events monitoring

All adverse events were documented using a standardized Adverse Events Form administered by a research assistant and reviewed by the therapist every 4 weeks at the beginning of therapy sessions to obtain information about new adverse events that occurred over the past 4 weeks. It was also administered at the post-treatment and 3- and 6-month follow-up visits. An adverse event was defined as any new psychological or physical symptom. The severity (i.e. grade) of all adverse events was determined using the Common Terminology Criteria for Adverse Events (CTCAE) v4.0 criteria. Serious adverse events used the standard FDA definition and included any of the following: death, life-threatening events (including suicide attempts), inpatient hospitalization or prolongation of existing hospitalization, disability/incapacitation, congenital anomalies/birth defects, other serious events that may jeopardize the health or well-being of the subject or require medical or surgical intervention to prevent one of the above outcomes.

Survival analysis of drop-out

For the survival analysis, time to dropout from therapy was defined as: (1) for drop-outs, the number of days from the first to the last treatment plus 7 to account for the fact that people were expected weekly for therapy; (2) for therapy completers, the number of days from the first to the last treatment session that was not a booster session, censored at the last therapy session. Two participants dropped out of therapy, but completed IE assessments until the end of follow-ups and were counted as drop-outs for this analysis. Time to drop-out was left-truncated (i.e., the first therapy visit was also the last) for six participants (1 MGH CBT-BDD, 0 MGH SPT, 4 RIH CBT-BDD, 1 RIH SPT). Time to drop-out was modeled using a Kaplan-Meier estimator of the survival function with censoring. Tests of equality of the survival function across strata (here grouped into treatment-by-site strata) were evaluated with one Chisquare-based log-rank test for overall equivalency, and two follow-up tests of CBT-BDD vs. SPT at each site.

Assessment of missingness methods

In exploratory follow-up analyses, we ran a series of proportional hazard models to assess whether likely demographic and clinical characteristics were associated with time to drop-out. In each proportional hazards model, we kept treatment, site, site by treatment, and baseline BDD-YBOCS scores as predictors, because they were part of the original LGM model. We then added one additional baseline predictor at a time to see their relative contribution to the risk of therapy drop-out. The baseline predictors tested were gender (male vs. female), age, racial/ethnic minority status (vs. white, non-Hispanic), education (more than high school education vs. high school graduate or less), marital status (single, never married vs. all other), any axis 1 comorbidity (vs. none), avoidant personality disorder (yes vs. no), any psychotropic medication use (yes vs. no), and treatment credibility scores.

1b. eResults

Treatment adherence and competence

We were not able to detect any significant differences in the global adherence or competence ratings between the two treatment arms at either site (adherence Fisher's exact test [FET] p=0.47 for MGH, p=0.07 for RIH; competence FET p>0.99 for MGH, p=0.25 for RIH), but global adherence ratings were slightly higher for CBT therapists than for SPT therapists at RIH. Considering differences between the sites in the treatment effect of CBT versus SPT, we examined whether therapist adherence and competence differed between sites. Again, we were not able to detect any significant site differences in CBT global adherence (FET p=0.85) or global competence (FET p=0.51) ratings between sites. However, SPT adherence global ratings were more frequently rated 'completely' or 'mostly' adherent at MGH (98%) than at RIH (93%; FET p=0.03), whereas we found no difference in the SPT competence global ratings (FET p=0.32). Some differences were found for some specific aspects of SPT competence, however (see next paragraph). Overall, adherence and competence were high in both treatments and at both sites, with 93-99% of all rated sessions evaluated as 'mostly' or 'completely' adherent, and 96-100% of all rated sessions evaluated as 'completely competent.'

Exploratory analyses of potential site differences in specific SPT adherence and competence ratings revealed that MGH therapists were rated as slightly more consistently excellent in 3 out of 14 specific competence criteria: their collaborative, conversational communication style (7 items; Median [IQR]: 4.9 [4.9 – 5.0], 47% perfectly competent at MGH vs. 4.9 [4.9-4.9], 13% perfectly competent at RIH, p<0.001), clear communication (3 items; 5.0 [5.0 – 5.0], 99% perfectly competent at MGH vs. 5.0 [5.0-5.0], 89% perfectly competent at RIH; p=0.005), and discussions of how the patient's daily concerns related to self-esteem (1 item; 5.0 [5.0 – 5.0], 97% perfectly competent at MGH vs. 5.0 [5.0-5.0], 85% perfectly competent at RIH; p<0.05). The slight site difference in SPT global adherence ratings was not reflected in any specific adherence criterion (all p>0.08).

Therapist self-reports of disallowed treatment techniques were completely absent or rare in both treatment arms. Of the 1,105 SPT sessions provided in total, 0% (0/618) of SPT sessions at MGH had therapist self-reports of CBT or other techniques, and only 2% (10/487) of SPT sessions at RIH had self-reports of CBT techniques (or, more generally, only 5% (22/487) used CBT or other non-SPT techniques). Of the 1,091 total CBT-BDD sessions provided, 0% (0/624) CBT-BDD sessions at MGH had therapist self-reports of non-CBT techniques, and 1% (4/467) CBT-BDD sessions at RIH had self-reports of non-CBT techniques. Based on the independently evaluated adherence/competence ratings, only two of the n=187 rated SPT sessions used any disallowed strategies: one session at MGH included 21-25 minutes of "Exposure, Activity scheduling," and another included 0-5 minutes of "Response prevention". There were no reports of SPT strategies used in the n=200 CBT-BDD sessions rated.

Blinding effectiveness

Overall, IE's guessed treatment assignment correctly in 24% (15/46 in CBT-BDD, 7/47 in SPT) of the patients they evaluated, incorrectly in 17% (5/46 in CBT-BDD, 11/47 in SPT) of cases, and were not at all certain in most cases (59% overall; 26/46 in CBT-BDD, 29/47 in SPT). Bang's blinding index was 0.22 (95%CI: 0.04, 0.40) for CBT-BDD treatment and -0.09 (95%CI: -0.26, 0.09) for SPT treatment. This indicates that the blinding was largely successful, due to the large number of unknown treatment assignments. In cases of incorrect guesses, IEs tended to guess that patients had received CBT-BDD more often than that patients had received SPT.

Primary and secondary outcomes during follow-up

As indicated in the manuscript text, we were not able to detect any changes in primary (BDD-YBOCS) and secondary (BABS, BDI-II, SDS, QLESQ-SF) symptoms during follow-up (from week 24 to week 50). The results of these analyses are presented in eTable 1.

eTable 1. Tests of intercept differences, overall slope effects, and slope differences in random intercept random coefficient models of primary and secondary treatment outcomes during follow-up (weeks 24-50).

Outcome, Effect	Effect Type [a]	Estimate [95% CI] [b]	*p*	
Primary outcome				
BDD symptom severity (BDD-YBOCS)				
Treatment	intercept difference	-6.64 [-12.35, -0.93]	0.02	*
Site	intercept difference	0.20 [-5.51, 5.91]	0.94	
Treatment*Site	intercept difference	2.35 [-3.36, 8.06]	0.42	
Time	overall change [c]	-1.04 [2.78, 0.71]	0.24	
Time*Treatment	effect of treatment overall [c]	0.86 [-2.62, 4.35]	0.62	
Time*Site	effect of site overall [c]	-1.52 [-5.00, 1.97]	0.39	
Time*Treatment*Site	site diff. in treatment effect [c]	1.12 [-2.36, 4.60]	0.52	
Secondary outcomes				
Insight (BABS)				
Treatment	intercept difference	0.64 [-3.47, 4.74]	0.84	
Site	intercept difference	0.41 [-3.70, 4.52]	0.76	
Treatment*Site	intercept difference	-1.27 [-6.97, 4.43]	0.76	
Time	overall change [c]	-0.83 [-1.98, 0.32]	0.15	
Time*Treatment	effect of treatment overall [c]	-0.87 [-3.17, 1.43]	0.45	
Time*Site	effect of site overall [c]	-0.16 [-2.45, 2.14]	0.89	
Time*Treatment*Site	site diff. in treatment effect [c]	1.36 [-0.93, 3.66]	0.24	
Depression severity (BDI-II)				
Treatment	intercept difference	-7.52 [-14.04, -1.00]	0.02	*
Site	intercept difference	-3.20 [-9.72, 3.32]	0.33	
Treatment*Site	intercept difference	-0.37 [-6.89, 6.15]	0.91	
Time	overall change [c]	1.11 [-0.65, 2.88]	0.21	
Time*Treatment	effect of treatment overall [c]	2.02 [-1.51, 5.54]	0.26	
Time*Site	effect of site overall [c]	-2.96 [-6.48, 0.57]	0.10	
Time*Treatment*Site	site diff. in treatment effect [c]	0.96 [-2.56, 4.49]	0.59	
Functional impairment (SDS)				
Treatment	intercept difference	-1.83 [-7.43, 3.78]	0.52	
Site	intercept difference	-2.88 [-8.48, 2.73]	0.31	
Treatment*Site	intercept difference	2.13 [-3.48, 7.73]	0.45	
Time	overall change [c]	-0.50 [-2.26, 1.26]	0.58	

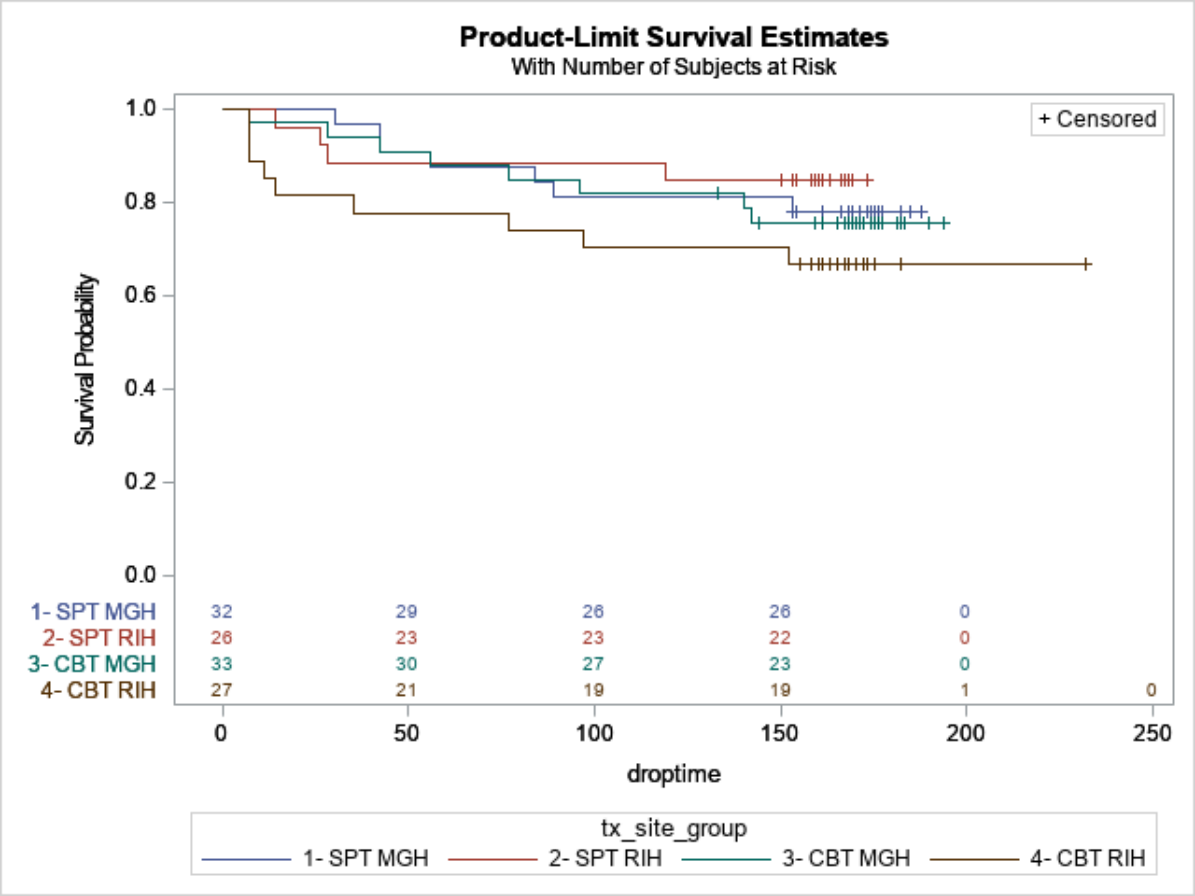
		Time*Treatment	effect of treatment overall [c]	-0.32 [-3.85, 3.20]	0.85	
		Time*Site	effect of site overall [c]	-2.55 [-6.07, 0.97]	0.15	
		Time*Treatment*Site	site diff. in treatment effect [c]	-1.40 [-4.93, 2.12]	0.43	
	Quality of life (QLESQ-SF)					
		Treatment	intercept difference	5.60 [-7.09, 18.29]	0.38	
		Site	intercept difference	-10.36 [-23.05, 2.33]	0.11	
		Treatment*Site	intercept difference	2.37 [-10.32, 15.06]	0.71	
		Time	overall change [c]	-1.12 [-5.14, 2.91]	0.58	
		Time*Treatment	effect of treatment overall [c]	-0.90 [-8.96, 7.15]	0.82	
		Time*Site	effect of site overall [c]	-4.25 [-12.31, 3.80]	0.30	
		Time*Treatment*Site	site diff. in treatment effect [c]	-3.18 [-11.24, 4.87]	0.43	

Note: Significance is indicated by * for $p<0.05$. [a] effect type refers to the model parameters tested in the latent growth curve models, Treatment, Site, and Treatment by Site tested the mean differences between groups at week 24 (based on people who attended the week 24 assessment, only, not including early drop-outs). Time tested the overall mean slope during follow-up, and Time*Treatment, Time*site, and Time*Treatment*Site tested mean slope differences between treatments, sites, or treatments within each site during follow-up; [b] units are scale-points on each scale except for the QLESQ-SF, where the units are percentage score points; [c] change is over 26 weeks (weeks 24 to 50) of assessment during follow-up. BDD-YBOCS = Yale-Brown Obsessive Compulsive Scale Modified for BDD; BABS = Brown Assessment of Beliefs Scale (higher total scores indicating less insight); BDI-II = Beck Depression Inventory-II (higher scores indicating greater depression severity); SDS = Sheehan Disability Scale global functioning score (higher scores indicating greater functional impairment); QLESQ-SF = Quality of Life Enjoyment and Satisfaction Questionnaire-Short form (higher percent scores indicating greater quality of life).

Survival analysis of drop-out

We were not able to detect significant differences in survival times across treatment-by-site strata overall (X^2(df=3)=2.75, p=0.43) or at either site (MGH: X^2=0.06, p=0.81; RIH: X^2=2.55, p=0.11). Survival curves of time to drop-out from therapy during the treatment period are shown in the eFigure.

eFigure. Survival curves of time to drop-out from therapy during the treatment period.



All participants who completed treatment have observations censored on the day of their last non-booster therapy session. Numbers underneath the survival curves indicate the number of people still in therapy at that time point in each treatment by site arm.

Assessment of missingness results

We were unable to detect any significant effects of the examined baseline risk factors for therapy dropout after adjusting for multiple test. Before adjusting for multiple tests, being single, never married was associated with a 3-fold increase of risk of earlier dropout compared to other marital statuses. The estimated hazard ratios for all risk factors evaluated are presented in eTable 2.

eTable 2. Baseline characteristics evaluated as risk factors for early drop-out from treatment.

Baseline predictor	Hazard ratio	95% LCL	95% UCL	*p*	adj. *p**
Male (vs. female)	1.10	0.46	2.62	0.83	0.83
Age (in years)	0.97	0.94	1.01	0.12	0.39
Racial minority (vs. white, non-Hispanic)	0.54	0.16	1.83	0.32	0.58
More than high school education (vs. not)	0.47	0.18	1.24	0.13	0.39
Single, never married (vs. other)	3.34	1.30	8.58	0.01	0.09
Axis 1 comorbidity (vs. not)	0.74	0.31	1.74	0.49	0.74
Avoidant personality disorder (vs. not)	1.12	0.51	2.47	0.78	0.83
Any psychotropic medication use (vs. not)	1.14	0.53	2.45	0.73	0.83
Treatment credibility (scale score)	0.85	0.65	1.10	0.22	0.50

* adjusted for false discovery rate

Literature cited

1. Wilhelm S, Phillips KA, Steketee G. *A cognitive-behavioral treatment manual for body dysmorphic disorder.* New York, NY: Guilford Press; 2013.
2. Wilhelm S, Phillips KA, Fama JM, Greenberg JL, Steketee G. Modular cognitive-behavioral therapy for body dysmorphic disorder. *Behav The*r. 2011;42:624–633.
3. Wilhelm S, Phillips KA, Didie E, Buhlmann U, Greenberg JL, Fama JM, Keshaviah A, Steketee G: Modular cognitive-behavioral therapy for body dysmorphic disorder: a randomized controlled trial. *Behav Ther*. 2014; 45(3):314-27.
4. Pinsker H. *A Primer of Supportive Psychotherapy*. Hillsdale, NJ: The Analytic Press; 1997.
5. Wilhelm S, Peterson AL, Piacentini J, Woods DW, Deckersbach T, Sukhodolsky DG, Chang S, Liu H, Dziura J, Walkup JT, Scahill L: Randomized trial of behavior therapy for adults with Tourette syndrome. *Arch Gen Psychiatry*. 2012; 69(8):795-803.
6. Winston A, Rosenthal RN, Pinsker H, eds. *Introduction to Supportive Psychotherapy*. Washington, DC: American Psychiatric Pub; 2004.
7. Tobin K, Davey-Rothwell MA, Nonyane BA, Knowlton A, Wissow L, Latkin CA: RCT of an integrated CBT-HIV intervention on depressive symptoms and HIV risk. *PloS One*. 2017; 12(12):e0187180.
8. Gwet, K. *Handbook of Inter-Rater Reliability: How to measure the level of agreement between two or multiple raters*. Gaithersburg, MD: Stataxis Publishing Company; 2001.
9. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979; 86(2):420.
10. Bang H, Ni L, Davis CE. Assessment of Blinding in Clinical Trials. *Control Clin Trials.* 2004; 25(2):143–156.