



Published in final edited form as:

*Psychol Med.* 2017 March ; 47(4): 703–717. doi:10.1017/S0033291716002543.

## An adaptive randomized trial of dialectical behavior therapy and cognitive behavior therapy for binge-eating

E. Y. Chen<sup>1,\*</sup>, J. Cacioppo<sup>2</sup>, K. Fettich<sup>1</sup>, R. Gallop<sup>3,4</sup>, M. S. McCloskey<sup>1</sup>, T. Olino<sup>1</sup>, T. A. Zeffiro<sup>5</sup>

<sup>1</sup>Department of Psychology, Temple University, Philadelphia, PA, USA

<sup>2</sup>Department of Psychology, The University of Chicago, Chicago, IL, USA

<sup>3</sup>Department of Mathematics, West Chester University, West Chester, PA, USA

<sup>4</sup>Department of Psychiatry, Center for Psychotherapy Research, Perelman School of Medicine, University of Pennsylvania, USA

<sup>5</sup>Neurometrika, Potomac, MD, USA

### Abstract

**Background.**—Early weak treatment response is one of the few trans-diagnostic, treatment-agnostic predictors of poor outcome following a full treatment course. We sought to improve the outcome of clients with weak initial response to guided self-help cognitive behavior therapy (GSH).

**Method.**—One hundred and nine women with binge-eating disorder (BED) or bulimia nervosa (BN) (DSM-IV-TR) received 4 weeks of GSH. Based on their response, they were grouped into: (1) early strong responders who continued GSH (cGSH), and early weak responders randomized to (2) dialectical behavior therapy (DBT), or (3) individual and additional group cognitive behavior therapy (CBT+).

**Results.**—Baseline objective binge-eating-day (OBD) frequency was similar between DBT, CBT+ and cGSH. During treatment, OBD frequency reduction was significantly slower in DBT and CBT+ relative to cGSH. Relative to cGSH, OBD frequency was significantly greater at the end of DBT ( $d = 0.27$ ) and CBT+ ( $d = 0.31$ ) although these effects were small and within-treatment effects from baseline were large ( $d = 1.41, 0.95, 1.11$ , respectively). OBD improvements significantly diminished in all groups during 12 months follow-up but were significantly better sustained in DBT relative to cGSH ( $d = -0.43$ ). At 6- and 12-month follow-up assessments, DBT, CBT and cGSH did not differ in OBD.

\* Address for correspondence: E. Y. Chen, Ph.D., Department of Psychology, Temple University, 1701 N 13th Street, Philadelphia, PA, USA. Eunice.Chen@temple.edu.

#### Declaration of Interest

With regards to disclosure of financial relationships, the first author discloses annual royalties from Guilford Press and has consulted for Shire Pharmaceuticals. The second author discloses royalties from Norton Books and Cengage Publishing for academic books and chapters.

#### Supplementary material

The supplementary material for this article can be found at <http://dx.doi.org/10.1017/S0033291716002543>

**Conclusions.**—Early weak response to GSH may be overcome by additional intensive treatment. Evidence was insufficient to support superiority of either DBT or CBT+ for early weak responders relative to early strong responders in cGSH; both were helpful. Future studies using adaptive designs are needed to assess the use of early response to efficiently deliver care to large heterogeneous client groups.

### Keywords

Bulimia; binge-eating disorder; cognitive behavior therapy; dialectical behavior therapy; guided self-help; stepped care

---

## Introduction

How can we use stepped care to improve the lives of clients for whom standard treatments are not fully effective? One of the few trans-diagnostic and treatment-agnostic predictors of outcome after a full treatment course is ‘early weak response’ or ‘rapid response’ after 2–8 treatment weeks (Crits-Christoph *et al.* 2001; Szegedi *et al.* 2009; Aderka *et al.* 2012). This phenomenon holds across interventions including pharmacotherapy, psychotherapy and electroconvulsive therapy, and across major psychiatric disorders including depression, schizophrenia, generalized anxiety disorder, obsessive compulsive disorder, alcohol abuse and eating disorders (Crits-Christoph *et al.* 2001; Szegedi *et al.* 2009; Aderka *et al.* 2012).

Binge-eating disorder (BED) and bulimia nervosa (BN) affect up to 5% of individuals worldwide (Hudson *et al.* 2007). Both disorders are defined by repetitive binge-eating with loss of control. BN also involves compensatory behavior, like vomiting and increased risk for cardiovascular disease and other psychiatric disorders (Hudson *et al.* 2007). BED often co-occurs with obesity but is also associated with more disabling co-morbidities including major depression, diabetes, hypertension, and chronic pain (Kessler *et al.* 2014).

Guided self-help cognitive behavior therapy (GSH) and individually delivered cognitive behavior therapy (CBT) are the best-evidenced treatments for BED and BN, respectively (Hay *et al.* 2009; Wilson *et al.* 2010; Hay, 2013). GSH is a brief, cost-effective protocol delivered by non-specialists (Lynch *et al.* 2010), whereas individual CBT is typically administered by trained eating-disorders specialists (Wilson & Zandberg, 2012). Early weak response predicting later poor outcome is typically detected by the fourth outpatient CBT or GSH session for BED and BN (Agras *et al.* 2000; Wilson *et al.* 2002; Fairburn *et al.* 2004; Grilo *et al.* 2006, 2012, 2015; Masheb & Grilo, 2007; Grilo & Masheb, 2007; Hilbert *et al.* 2015; Thompson-Brenner *et al.* 2015). This suggests that not all individuals with binge-eating problems require the same type and intensity of therapy.

Stepped care adaptive treatment designs (Almirall *et al.* 2012; Nahum-Shani *et al.* 2012) tailor clients’ subsequent treatment given their initial response to a less intensive treatment. In a seminal study with BN, offering less intense GSH before more intensive psychotherapy, led to improved outcomes relative to offering the more intensive therapy first (Mitchell *et al.* 2011). However, it is unknown if early response can be used an adaptive algorithm to triage patients into more appropriate treatment.

Additional intensive psychotherapy may help early weak responders. One option for early weak responders is Dialectical Behavior Therapy (DBT) (Linehan, 1993a, b; Safer *et al.* 2009; Linehan *et al.* 2014). DBT has efficacy with treatment-resistant disorders, teaches a wide range of skills to manage emotions, and is highly intensive. DBT was developed for borderline personality disorder (Lieb *et al.* 2004), and has efficacy for substance use disorders (Linehan *et al.* 1999, 2002), depression (Lynch *et al.* 2007; Feldman *et al.* 2009), bipolar disorder (Van Dijk *et al.* 2013), BED (Telch *et al.* 2000, 2001; Safer & Jo, 2010), and BN (Safer *et al.* 2001). DBT broadly targets emotion dysregulation which is thought to maintain binge-eating and blends behavior learning principles with mindfulness practice and dialectical philosophy. This comprehensive multimodal treatment incorporates weekly group skills training, weekly individual psychotherapy organized by a target hierarchy, and 24-h phone coaching for clients. It also includes a weekly therapist consultation team and special protocols for suicidal and therapy-interfering behavior.

In contrast, CBT, is the best-evidenced treatment for eating disorders (Hay *et al.* 2009; Hay, 2013). For this study we used an adaptation of CBT (Fairburn, 2008) (denoted as CBT+). The CBT+ model posits that binge-eating is maintained by dietary restraint or vomiting, both resulting from overvaluation of weight and shape. Typically delivered individually, CBT+ is distinct in its behavioral specificity and efficiency of its delivery. Sessions are agenda-driven, and strategies are taught in a sequenced, parsimonious fashion.

We know that early strong response predicts better outcome after the full course of treatment; and early weak response predicts poorer outcome. Given this, the differential treatment hypothesis, posits that early strong responders to four sessions of GSH followed by continued GSH (cGSH) will have greater improvements in binge-eating than early weak responders administered four GSH sessions followed by intensive DBT or GSH then a CBT control (CBT+). The time-frame assessed for the differential treatment hypothesis, was from baseline to end of treatment. Our follow-up hypothesis, predicts that early strong responders in cGSH will have greater binge-eating reduction than early weak responders in DBT or CBT+ from the end of treatment to 6 and 12 months later. For these hypotheses, we examined if BN rather than BED diagnosis was associated with greater objective binge-eating day (OBD) frequency at baseline and after continued treatment.

We employed a Phase II adaptive clinical trial design that allows treatment individualization based on a client's initial response (e.g. Nahum-Shani *et al.* 2012). Women with BED or BN received four GSH sessions. If response to four GSH sessions was weak, clients were randomized into either DBT or CBT+. Clients with strong treatment responses continued with GSH (cGSH). The primary outcomes were OBDs assessed at baseline, after GSH, and after cGSH, DBT or CBT+, and 6 and 12 months later. Secondary outcomes were OBD abstinence, vomiting frequency and abstinence, eating disorder psychopathology, body mass index (BMI), number of co-occurring Axis I disorders, and global assessment of functioning.

## Method

This study was conducted from June 2009 to June 2013, and treatment occurred within a hospital outpatient adult eating disorders program. The protocol was approved by the local institutional review board. Written informed consent was obtained after the study was described to the participants. Participants were paid for completing assessments. The CONSORT chart in Fig. 1 outlines the study design. The Supplementary material lists a glossary of acronyms.

*Participants* were 109 women aged  $\geq 18$  years who met Diagnostic Statistical Manual IV-TR (DSM-IVTR) (APA, 2000) criteria for BED or BN in the past 3 months and resided within commuting distance. Participants were medically stable, allowing outpatient treatment. All participants met DSM-5 (APA, 2013) criteria for BN or BED. Exclusion criteria included current bipolar disorder or psychosis, use of appetite suppressants, past bariatric surgery, current eating-disorder treatment, and pregnancy. Participants taking psycho-tropic medication were eligible if stable on the medication for at least 1 month. (Table 1).

## Procedure

After a phone screen and medical clearance, eligible participants were screened at baseline using the Eating Disorders Examination-16 (EDE; Fairburn, 2008), the Structured Clinical Interview for DSM-IV, Axis I (SCID-I; First & Gibbon, 2004).

*Stage 1.* Participants were given a GSH CBT manual (Fairburn, 1995) and administered four weekly 20- to 30-min sessions with a Masters-level trainee therapist. GSH utilizes six additive sequential steps: orientation to a model of binge-eating as maintained by dietary restraint arising from overvaluation of weight and shape, goal-setting, psycho-education, food monitoring, behavioral strategies and relapse prevention. A generic protocol of suicide risk assessment and management was utilized (Linehan, 1999). Treatment dropout was defined per hospital policy as missing two scheduled consecutive sessions without notice or if no EDE was conducted after GSH.

After the fourth GSH session, clients were administered an EDE to assess early response magnitude. Given previous studies (Grilo *et al.* 2006, 2012, 2015; Grilo & Masheb, 2007; Masheb & Grilo, 2007), early strong response was defined as  $\geq 65\%$  reduction in OBD or vomiting frequency (whichever was more frequent), from baseline to after GSH. Early weak response was defined as  $< 65\%$  reduction in these behaviors. This rule accounted for high-risk BN clients with high frequency vomiting, but lower frequency binge-eating, and utilizes the finding that vomiting is a better metric of early weak response in BN (Agras *et al.* 2000; Fairburn *et al.* 2004).

*Stage 1: Results for N = 109.* From baseline to after 4 weeks of GSH, the average OBD frequency was reduced by 89% for early strong responders (baseline: mean = 14.60, S.D. = 5.96; after GSH: mean = 1.64, S.D. = 1.97) and 15% for early weak responders (baseline: mean = 13.42, S.D. = 7.19; after GSH: mean = 11.35, S.D. = 6.93). Early strong responders comprised 38.5% (42/109) of the sample.

*Stage 2: Early strong responders* could continue cGSH for up to 24 weekly 20–30 min sessions, but could end earlier by mutual client and therapist agreement. cGSH used the same model of binge-eating, self-help strategies and suicide risk and dropout protocol as GSH.

*Stage 2: Early weak responders* were randomized to either 6 months of group and individual DBT or CBT+. An *a priori* power analysis showed that a sample of 67 early weak responders would have >85% power to detect a medium to large OBD frequency reduction. The power analysis used a Poisson regression model, assumed one predictor, a base rate risk ratio of 0.29 over 28 days, and a 20% dropout rate using a two-tailed test where  $\alpha = 0.05$ . We assumed a rate ratio of 0.71 comparing DBT and waitlist for BED (Telch *et al.* 2001), and 1.75 comparing CBT and placebo for BED (Grilo *et al.* 2006).

A block randomization scheme was used, stratifying by BN or BED diagnosis, which was generated and concealed in envelopes by a biostatistician. The research coordinator who assigned clients to treatment did not conduct assessments or treatment. A different therapist than the one who administered the first four GSH sessions administered the second stage of cGSH for early strong responders, and DBT or CBT+ for early weak responders. Assessors were blind to treatment assignment and did not treat randomized clients. Clients could choose to discontinue treatment but continue to be assessed for the length of the study.

*DBT for early weak responders* involved 6 months of weekly sessions of: 2 h of skills group, 1 h of individual therapy, 2 h with a therapist consultation team, and 24-h phone coaching. As in a previous trial (Safer & Jo, 2010), these DBT manuals were used (Linehan, 1993a, b; Safer *et al.* 2009). Binge-eating is viewed as an attempt to escape or to modulate aversive emotions and over time reinforced as escape behavior, becoming an over-learned dysfunctional response to dysregulated emotions. As a principle-driven treatment, individual DBT sessions target the highest ranking behavior from a target hierarchy of: life-threatening behaviors including suicide attempts, therapy-interfering behaviors like missing sessions, and quality-of-life interfering behaviors, such as substance abuse, as well as increasing behavioral skills. DBT group skills teach emotion regulation, mindfulness, distress tolerance, and interpersonal effectiveness skills. Therapeutic strategies attempt to balance change and acceptance. DBT specifies detailed protocols for suicidal and therapy-interfering behavior. DBT defines dropout as missing four consecutive scheduled sessions of either group or individual DBT.

*CBT+ for early weak responders* involved 6 months of weekly sessions of: 2 h of group therapy, 1 h of individual therapy, 2 h of therapist case conference, and a 24-h psychiatry resident on-call. CBT+ was modified to control for treatment dosage and content in DBT. A group therapy protocol (Chen *et al.* 2003) was added to standard individual treatment delivery (Fairburn, 2008). The overlapping emotional eating module (Fairburn, 2008) was removed given the focus of DBT on emotion regulation (Linehan *et al.* 1991; Linehan, 1993a, b). Individual CBT+ sessions have four stages: identifying behavioral strategies, identifying barriers to change, reviewing eating disorder maintenance mechanisms, and relapse prevention. The structure of the group sessions involved review of homework followed by didactics. CBT+ utilized the same protocol for suicide risk and dropout as GSH.

### Therapist qualifications/training

There were five DBT, six CBT+, and 10 cGSH therapists. All therapists were Masters-level clinicians trained and supervised by E.Y.C.. Separate 5-day workshops for CBT+ and DBT and a 1-day workshop for GSH/cGSH were conducted annually. Therapists in DBT and CBT+ had two supervised training clients prior to working with study clients. Training and weekly supervision were conducted separately for each therapy. Treatment adherence was assessed (see Supplementary material).

### Assessments

Clients were assessed after cGSH, DBT, or CBT+ and at 6- and 12-month follow-up assessments using the EDE. The EDE assesses frequency of, and abstinence from, eating disorder behavior and psychopathology over the last month. Assessors were 10 Masters-level doctoral students in clinical psychology. High assessor reliability was obtained for BED and BN diagnoses on the EDE using a standard set of tapes, with intra-class correlations ranging from 0.79 to 0.91. The 12-month follow-up assessment focused on binge-eating, vomiting and EDE outcomes. BMI was assessed with weight and height scales. If the participant could not be seen in person, height and weight were obtained in a physician's office and results were emailed to the study team. Co-occurring Axis I disorders were assessed up to the 6-month follow-up with the Longitudinal Interview Follow-up Evaluation Psychiatric Status Ratings (Keller *et al.* 1987). Up to the 6-month follow-up, assessor ratings of global assessment of functioning from DSM-IV-TR (APA, 2000) were used where scores of 51–60 referred to moderate impairment; 61–70, mild impairment; and 71–80, slight impairment.

### Statistical analysis

Given the high correlation between OBD frequency and objective binge episode (OBE) frequency ( $r = 0.91$  to  $r = 1.00$ ,  $n = 109$ ), we only analyzed OBD frequency. Participants with any outcome assessments were included in the analysis. Pattern mixture modeling (Hedeker & Gibbons, 2006), and a sensitivity analysis using a Markov-Chain Monte Carlo imputation method were used to assess the potential informative nature of missing data (Yuan, 2011). These analyses suggest that data was missing at random (see Supplementary Table and Results). Generalized linear mixed model (GLMM) analyses provides robust estimates when data is missing at random (Little & Rubin, 2014). GLMM was conducted using SAS/STAT9.4 (SAS Institute Inc., 2015) and R version 3.1.2 (R Core Team, 2014). We modeled the OBD frequency trajectories of cGSH, DBT, and CBT+ groups from baseline through the end of treatment (differential treatment hypothesis), and from end of treatment to 12-month follow-up (follow-up hypothesis) using longitudinal Poisson GLMM (Atkins *et al.* 2013) that included random intercepts and slopes. Intercept estimates from the longitudinal GLMM were used to test group differences at the first time-point of each analysis. Slope estimates comparing the temporal trajectories of the DBT and CBT+ groups relative to cGSH were conducted to test group differences in OBD reduction over time. Model based estimates over the relevant timeframe were made to compare interventions at the end of treatment, and 6- and 12-month follow-up assessments. Given potential outcome differences due to BED or BN diagnosis (Fairburn *et al.* 2000; Kessler *et al.* 2014), we first

fit the models without controlling for covariates and then reanalyzed with diagnosis as a covariate. Two-tailed tests were used, where  $p < 0.05$ .

Rate ratios (RR) and their upper and lower 95% confidence intervals (CIs) for comparing groups over the time-frame for OBD frequency were estimated. To clarify effect sizes, RRs comparing groups over time, over time-frames, or for a covariate were converted to Cohen's  $d$ . Descriptive findings and Cohen's  $d$  and upper and lower 95% CIs within treatment group were reported for secondary outcomes for both hypotheses. We denote effect sizes as large when  $d \geq 0.8$  as large, and medium when  $0.5 \leq d < 0.8$  (Cohen, 1988).

For OBD or vomiting, a frequency  $\geq 8$  days/month was regarded as clinically significant given DSM-IV-TR (APA, 2000) BED or BN diagnoses. The non-clinical range for the EDE total scores was  $< 2.45$ , which is 2 S.D. above the mean EDE total scores for a healthy non-eating disorders group ( $N = 344$ ) (Jacobson & Truax, 1991; Fairburn & Cooper, 1993). Because BED is particularly associated with increased risk for obesity (Hudson *et al.* 2007), we separated BED and BN subgroups in reporting this outcome. DSM-IV-TR (APA, 2000) global assessment of functioning categories were used to assess clinically significant improvement in general functioning.

## Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

## Results

There was 7% study or assessment dropout at the end of treatment, a further 6% at the 6-month follow-up, and an additional 10% at the 12-month follow-up with no differences between the cGSH, DBT, and CBT+ groups. Among those randomized to DBT, 31% were treatment dropouts and 45% were CBT+ treatment dropouts. Of the 42 early strong responders in cGSH, 26% dropped out (see Supplementary Table S1).

## Differential treatment hypothesis

**Primary outcome**—BN or BED diagnosis at baseline was not associated with greater OBD frequency. There were no differences in baseline OBD frequency between cGSH compared to DBT or CBT+. All groups were in the clinical range at baseline; however, after four sessions of GSH, the cGSH group fell in the non-clinical range while DBT and CBT+ fell in the non-clinical range only after intensive treatment. Fig. 2 and Table 2 show that from baseline to the end of treatment, the decline in OBD frequency was significantly faster in cGSH relative to DBT or CBT+ (respectively,  $RR = 1.36$ ,  $d = 0.82$ ;  $RR = 1.36$ ,  $d = 0.83$ ). Within-treatment Cohen's  $d$  from baseline to end of treatment was 1.41 for cGSH, 0.95 for DBT and 1.11 for CBT+. Model-based estimates covering the treatment period showed that at the end of treatment OBD frequency was significantly greater in DBT relative to cGSH ( $t_{99} = 2.38$ ,  $p = 0.019$ ,  $RR = 1.55$ ,  $d = 0.27$ ) and CBT+ relative to cGSH ( $t_{99} = 2.69$ ,  $p = 0.008$ ,  $RR = 1.65$ ,  $d = 0.31$ ).

**Secondary outcomes**—From baseline to the end of treatment, cGSH, DBT and CBT+ had large effects on increasing OBD abstinence ( $d = 1.12, 0.92, 0.89$ , respectively) (see Table 3).

At both baseline and the end of treatment, cGSH was in the non-clinical range for vomiting frequency. DBT and CBT+ were in the clinical range at baseline. From baseline to the end of treatment, vomiting frequency reduced and vomiting abstinence increased in DBT and CBT+. DBT had a medium effect ( $d = 0.50$ ) on vomiting frequency reduction, which fell in the non-clinical range at the end of treatment. However, for CBT+, vomiting frequency remained in the clinical range at both baseline and at the end of treatment.

All treatments led to large reductions in EDE total scores from baseline to end of treatment. EDE scores for all groups were in the clinical range at baseline but moved to the non-clinical range after GSH for early strong responders in cGSH and to the non-clinical range only after intensive DBT and CBT+ for early weak responders.

From baseline to end of treatment, cGSH and DBT had medium effects ( $d = 0.65, 0.58$ , respectively) in reducing the number of co-occurring Axis I disorders whereas CBT+ led to a large reduction where  $d = 0.84$ .

Global assessment of functioning scores were in the ‘mild’ impairment range at both baseline and end of cGSH but moved from ‘moderate’ at baseline to ‘mild’ at the end of DBT and CBT+.

Treatments did not result in notable changes in BMI from baseline to end of treatment for BED but CBT+ was associated with a medium increase of  $d = -0.48$  in BMI for BN.

### Follow-up hypothesis

**Primary outcome**—From the end of treatment to 12-month follow-up, OBD frequency significantly increased in all groups (RR = 1.17, 95% CI 1.03–1.25,  $B = 0.16$ ,  $p = 0.02$ ,  $d = 0.46$ ) although all remained in the non-clinical range at the end of treatment and at 6- and 12-month follow-up after being in the clinical range at baseline (see Fig. 2). Within-group effect size changes in OBD frequency from end of treatment to 12-month follow-up were small:  $d = -0.28, -0.30$ , and  $0.10$  for cGSH, CBT+ and DBT, respectively.

Deterioration of improvement in OBD frequency during follow-up only differed in DBT: gradual increases in OBD frequency during follow-up were significantly less in DBT compared to cGSH (RR = 0.82,  $p = 0.03$ ,  $d = -0.43$ ) (see Fig. 2 and Table 2). There was no significant difference in the rate of change of OBD frequency in CBT+ compared to cGSH from the end of treatment through follow-up. Model-based estimates at 6- and 12-month follow-ups, showed that OBD frequency did not differ between DBT relative to cGSH (6 months:  $t_{173} = 1.59$ ,  $p = 0.11$ , RR = 1.80,  $d = 0.27$ ; 12 months:  $t_{173} = 1.01$ ,  $p = 0.31$ , RR = 1.47,  $d = 0.11$ ) or CBT+ relative to cGSH (6 months:  $t_{173} = 1.55$ ,  $p = 0.12$ , RR = 1.79,  $d = 0.24$ ; 12 months:  $t_{173} = 1.55$ ,  $p = 0.12$ , RR = 1.82,  $d = 0.41$ ).

Participants diagnosed with BN compared to BED had marginally greater OBD frequency after treatment (RR = 1.94, 95% CI 0.99–3.79,  $B = 0.66$ ,  $p = 0.052$ ,  $d = 0.38$ ).



**Secondary outcomes**—From the end of treatment to 12-month follow-up, participants in cGSH, DBT and CBT+ showed medium increases in OBD abstinence, respectively  $d = 0.55$ ,  $0.62$ ,  $0.46$ .

From the end of treatment to 12-month follow-up, vomiting frequency increased slightly after DBT and cGSH, and increased then decreased after CBT+. Vomiting frequencies after cGSH and DBT and throughout follow-up fell in the non-clinical range while CBT+ was in the clinical range until the 12-month follow-up. Vomiting abstinence rates declined in all groups over follow-up.

During follow-up, there was a medium-sized loss of improvement in EDE total scores for cGSH, DBT and CBT+,  $d = 0.53$ ,  $0.58$ ,  $0.50$ , respectively. However, at the 12-month follow-up, scores remained in the non-clinical range.

Number of co-occurring Axis I disorders and global assessment of functioning scores were assessed after treatment and 6 months later. At 6-month follow-up, cGSH had a medium effect,  $d = 0.68$ , in reducing the number of co-occurring Axis I disorders. Global assessment of functioning scores were ‘mild’ category at the end of DBT and CBT+ and at 6-month follow-up while the cGSH group improved from ‘mild’ to ‘slightly’ impaired at the end of treatment to 6-month follow-up. These improvements were large for cGSH and CBT+ and medium for DBT,  $d = -0.88$ ,  $-0.89$ ,  $-0.54$ , respectively.

DBT, CBT+ and cGSH did not result in changes in BMI over follow-up for BED or BN.

## Discussion

Early response magnitude may be useful for tailoring treatment delivery in a stepped care model. After four GSH sessions, the OBD frequency of early strong responders no longer met clinical levels and this was sustained through follow-up. Our results parallel those found in BN showing that stepped care is superior to standard treatment (Mitchell *et al.* 2011) and extend these findings to BED.

Statistically, early strong responders in cGSH had faster and greater reductions in OBD frequency by the end of treatment than early weak responders in DBT or CBT+ although these differences were small (relative to cGSH, DBT  $d = 0.27$  and CBT+  $d = 0.31$ ). DBT, CBT+ and cGSH had large effects on OBD reduction at the end of treatment ( $d = 0.95$ ,  $1.11$ ,  $1.41$ , respectively) and by then, on average, all groups no longer met the frequency criteria for BN or BED diagnoses (APA, 2000). While intensive DBT and CBT+ for early weak responders resulted in similar clinical improvement in OBD frequency to early strong responders in cGSH at the end of treatment, this was evident statistically at 6- and 12-month follow-ups where there were no group differences. OBD abstinence, eating disorders psychopathology, and number of co-occurring Axis I disorders similarly improved in early weak compared to early strong responders by the end of treatment. At the end of treatment, BN diagnosis was weakly associated with greater OBD frequency, regardless of early response status or intervention.

During follow-up, improvements in OBD frequency, and OBD and vomiting abstinence and eating disorder psychopathology diminished in both early weak and strong responders but not to clinically significant baseline levels with the exception of vomiting abstinence. At 6-month follow-up, only early strong responders in cGSH continued to have reductions in number of co-occurring Axis I disorders. While global functioning improved in all groups, for early weak responders, this effect was greatest in early strong responders in cGSH.

Results did not strongly favor either DBT or CBT+ relative to cGSH. Both DBT and CBT+ led to large improvements in OBD frequency by the end of treatment although OBD frequency was still significantly greater in both groups compared to cGSH. While all groups showed deterioration in OBD improvements made during treatment, at follow-up, relative to cGSH, DBT showed significantly less deterioration ( $d = -0.43$ ). Nonetheless, both DBT and CBT+ remained in the non-clinical range for OBD frequency during follow-up. For vomiting frequency, after DBT participants moved from the clinical to non-clinical range, while those in CBT+ made this transition 1 year later. Although vomiting abstinence rates in DBT and CBT+ were always less than in cGSH at all time-points, these were greater after DBT than CBT+ although these rates were comparable by 1-year follow-up and both were similar to baseline. Unlike DBT, CBT+ had a large effect in reducing co-occurring Axis I disorders after treatment ( $d = 0.84$  v.  $0.58$ ) and improving global functioning ( $d = -0.93$  v.  $-0.44$ ). However, both DBT and CBT+ led to similar improvements in global functioning, though still less than cGSH, at 6-month follow-up.

It is important to note the limitations of our study design. We did not compare early weak responders in cGSH directly with DBT or CBT+, as we did not randomize early weak responders to a third cGSH arm. The efficacy of CBT+ was possibly compromised in the attempt to make a controlled comparison by removing the emotional eating module and adding a group to individual CBT. We were only sufficiently powered to conduct statistical analysis on OBD frequency and could not draw definitive conclusions about BN diagnosis as a predictor or treatment effects on vomiting. We did not conduct 12-month follow-up assessments for non-eating disorder variables.

Given these limitations, and the inconclusive differences between DBT and CBT+, we can only surmise that there is insufficient evidence to reject the null hypothesis in comparing DBT or CBT+ with cGSH for early weak responders with binge-eating.

Although DBT and CBT+ may not have had strong differential treatment effects, their mechanisms of action may differ, accounting, for example why CBT+ and DBT acted differently on OBD and vomiting frequency. DBT skills mediate changes in Borderline Personality Disorder behavior (Neacsiu *et al.* 2010), while in CBT, reduction in weight concerns mediate binge abstinence for binge-eating (Dingemans *et al.* 2007). Future studies of DBT and CBT mediators may establish what intermediate targets may yield changes in outcome and examine BN as a moderator of outcome. Adding CBT to DBT may further help early weak responders (Federici & Wisniewski, 2013; Harned *et al.* 2014). Finally, there may be a ceiling effect for skills-based interventions (Johnsen & Friborg, 2015) for binge eating, suggesting the need to further understand the underlying neurobiology across disorders to inform a new generation of treatments (Val-Laillet *et al.* 2015).

This is the first randomized trial comparing a broad emotion regulation treatment, DBT, with an eating disorders-focused form of CBT for early weak responders with BED and BN. Strengths of the study include the use of an adaptive clinical trial design, strong retention despite treatment and therapist transitions, inclusion of both BED and BN, and controlling for treatment dosage and assessor blinding in the randomized arms. Future studies are needed that use a sequential multiple assignment randomized trial design (Almirall *et al.* 2014), allowing randomization of early strong and early weak responders to both more and less intensive arms. Future Phase III or equivalence studies may test different adaptations of CBT and early strong response guidelines. Future studies may employ early strong response as a guideline for stepped care for other eating disorders including anorexia nervosa. Evaluation of the cost-effectiveness of stepped care, and the dissemination of stepped care models in diverse settings are needed.

Unfortunately, less than half of individuals with BED or BN receive any treatment (Kessler *et al.* 2014). This study suggests that an early strong response to brief treatment may be one way of tailoring individualized effective treatment to a large, heterogeneous client group, although enhancing the maintenance of effects warrant further investigation.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

This study received support from the National Institute of Mental Health (1K23 MH081030-01 to the first author). Role of the Funder: The sponsor had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The authors thank the study participants, and students: Jessica Weissman, MA, Matthew Southward, MA, Kara Christensen, Johnny Berona, MA, Brooke Slawinski, Suzanna So, Jaleesa Akuoko, MS, Peggy Chau, Kay Segal, PsyD, Joy Harrington, PsyD, Dawn Eichen, Ph.D., Lindsey Ohler, PsyD, Taylor Dryman, MA, Richa Aggawal, MA, Kalina Eneva, MA., Angelina Yiu, MA Jean Arlt, MPhil., and Susan Murray.

([Clinicaltrials.gov#](https://clinicaltrials.gov/ct2/show/study/NCT00965705): Stepped Care Treatment for Binge-eating. NCT00965705.)

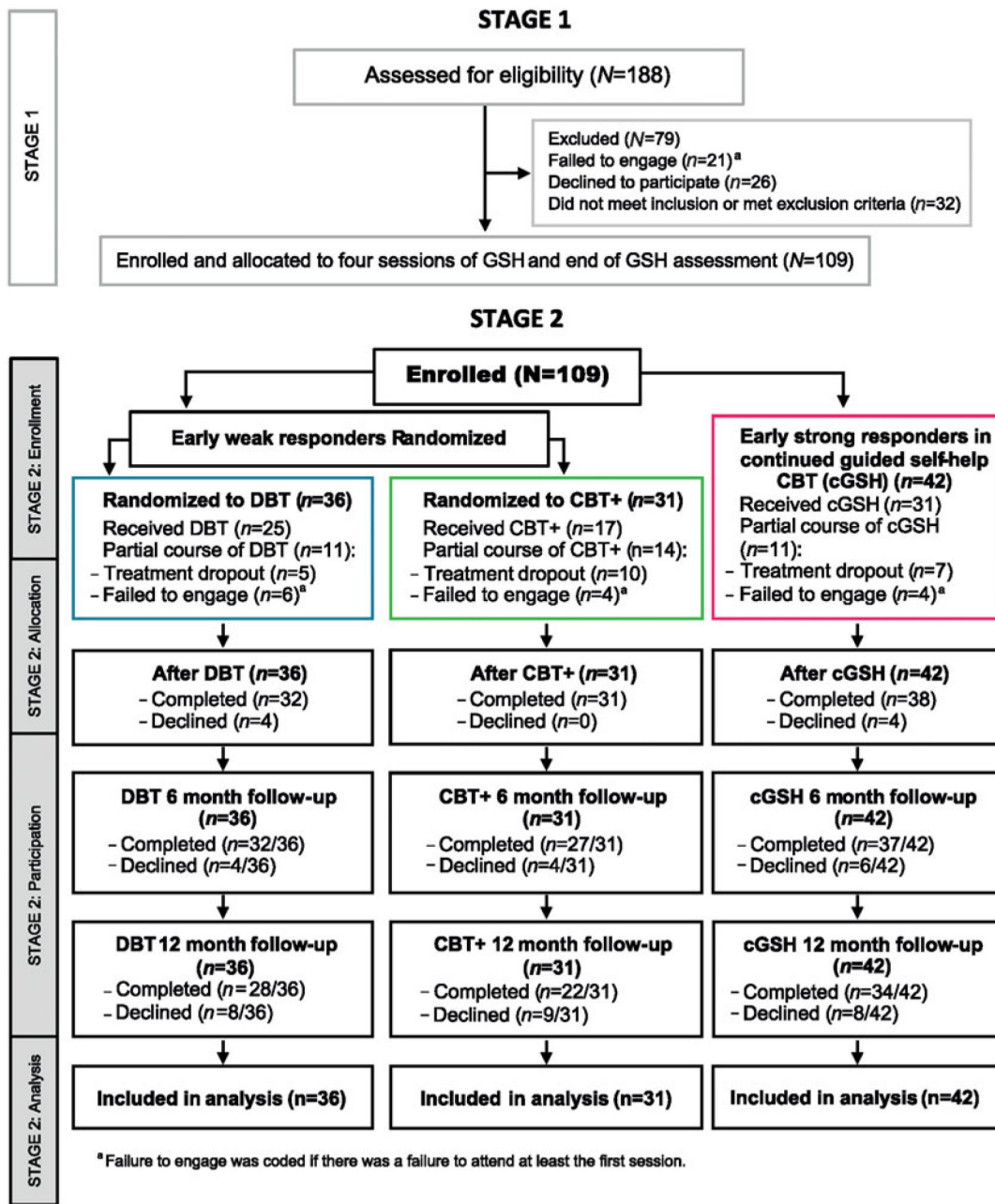
## References

- Aderka IM, Nickerson A, Bøe HJ, Hofmann SG (2012). Sudden gains during psychological treatments of anxiety and depression: a meta-analysis. *Journal of Consulting and Clinical Psychology* 80, 93. [PubMed: 22122290]
- Agras W, Crow S, Halmi K, Mitchell JE, Wilson G, Kraemer H (2000). Outcome predictors for the cognitive behavior treatment of bulimia nervosa: data from a multisite study. *American Journal of Psychiatry* 157, 1302–1308. [PubMed: 10910795]
- Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA (2012). Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. *Statistics in Medicine* 31, 1887–1902. [PubMed: 22438190]
- Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA (2014). Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. *Translational Behavioral Medicine* 4, 260–274. [PubMed: 25264466]
- APA (2000). *Diagnostic and Statistical Manual of Mental Disorders, 4th edn Training Revision*. American Psychiatric Association: Washington, DC.
- APA (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th edn Text Revision*. American Psychiatric Association: Washington, DC.

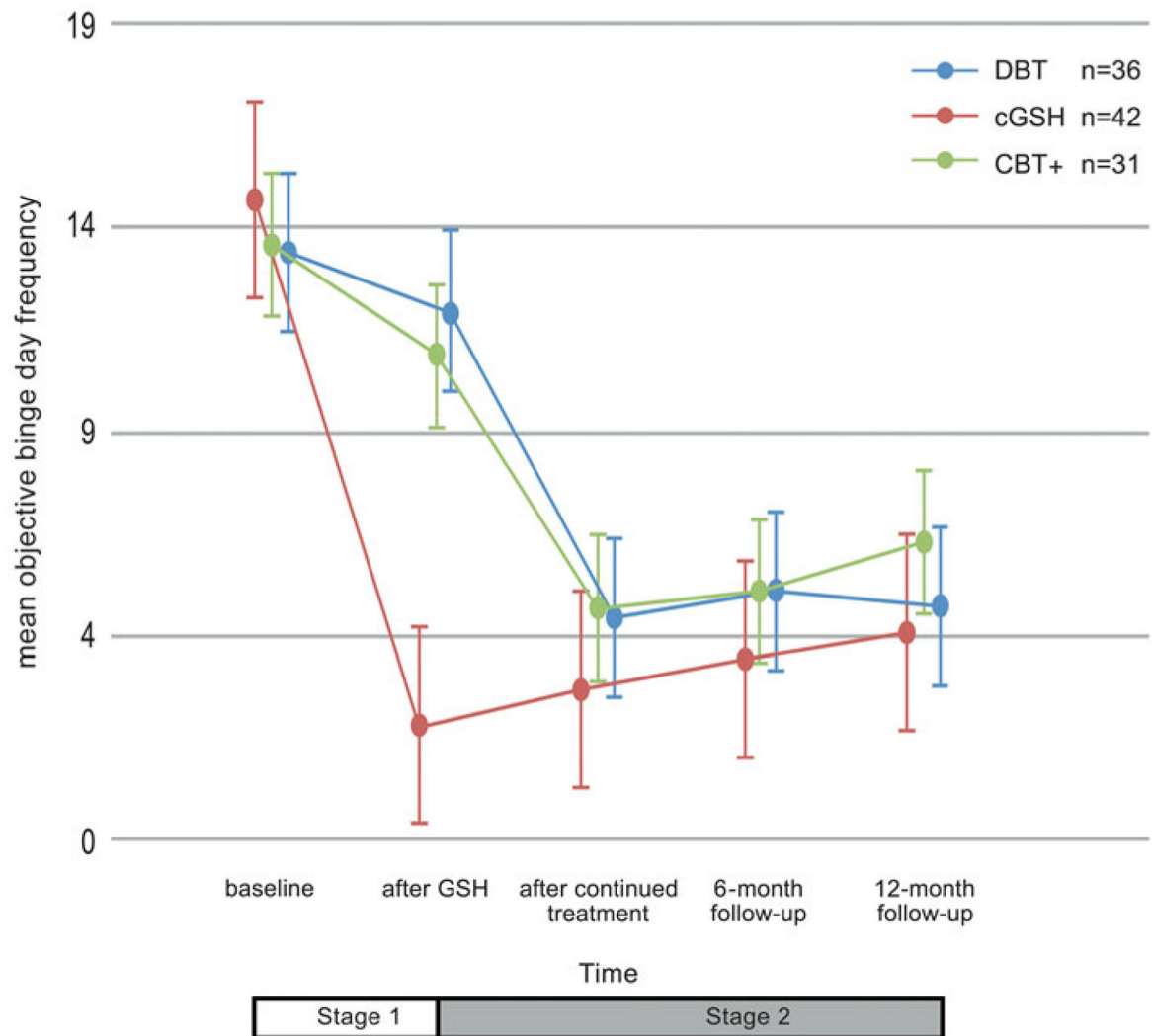
- Atkins DC, Baldwin SA, Zheng C, Gallop RJ, Neighbors C (2013). A tutorial on count regression and zero-altered count models for longitudinal substance use data. *Psychology of Addictive Behaviors* 27, 166. [PubMed: 22905895]
- Chen E, Touyz SW, Beumont PJV, Fairburn CG, Griffiths R, Butow P, Russell J, Schotte DE, Gertler R, Basten C (2003). Comparison of group and individual cognitive-behavioral therapy for patients with bulimia nervosa. *International Journal of Eating Disorders* 33, 241–254. [PubMed: 12655619]
- Cohen J (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates: Hillsdale, NJ.
- Crits-Christoph P, Connolly MB, Gallop R, Barber JP, Tu X, Gladis M, Siqueland L (2001). Early improvement during manual-guided cognitive and dynamic psychotherapies predicts 16-week remission status. *Journal of Psychotherapy Practice and Research* 10, 145. [PubMed: 11402077]
- Dingemans AE, Spinhoven P, van Furth EF (2007). Predictors and mediators of treatment outcome in patients with binge eating disorder. *Behaviour Research and Therapy* 45, 2551–2562. [PubMed: 17643390]
- Fairburn C (1995). *Overcoming Binge Eating*. Guilford Press: New York.
- Fairburn C, Cooper Z (1993). The eating disorder examination In *Binge Eating: Nature, Assessment and Treatment* (ed. Fairburn C and Wilson G), pp. 317–366. Guilford Press: New York.
- Fairburn CG (2008). *Cognitive Behavior Therapy and Eating Disorders*. Guilford Press: New York, NY, USA.
- Fairburn CG, Agras W, Walsh B, Wilson G, Stice E (2004). Prediction of outcome in Bulimia Nervosa by early change in treatment. *American Journal of Psychiatry* 161, 2322–2324. [PubMed: 15569910]
- Fairburn CG, Cooper Z, Doll HA, Norman P, O'Connor M (2000). The natural course of bulimia nervosa and binge eating disorder in young women. *Archives of General Psychiatry* 57, 659–665. [PubMed: 10891036]
- Federici A, Wisniewski L (2013). An intensive DBT program for patients with multidagnostic eating disorder presentations: a case series analysis. *International Journal of Eating Disorders* 46, 322–331. [PubMed: 23381784]
- Feldman G, Harley R, Kerrigan M, Jacobo M, Fava M (2009). Change in emotional processing during a dialectical behavior therapy-based skills group for major depressive disorder. *Behaviour Research and Therapy* 47, 316–321. [PubMed: 19232571]
- First MB, Gibbon M (2004). The structured clinical interview for DSM-IV axis I disorders (SCID-I) and the structured clinical interview for DSM-IV Axis II disorders (SCID-II) In *Comprehensive Handbook of Psychological Assessment* (ed. Hilsenroth MJ & Segal DL), pp. 134–143. John Wiley & Sons, Inc: Hoboken, NJ.
- Grilo C, White M, Wilson G, Gueorguieva R, Masheb R (2012). Rapid response predicts 12-month post-treatment outcomes in binge-eating disorder: theoretical and clinical implications. *Psychological Medicine* 42, 807–817. [PubMed: 21923964]
- Grilo CM, Masheb RM (2007). Rapid response predicts binge eating and weight loss in binge eating disorder: findings from a controlled trial of orlistat with guided self-help cognitive behavioral therapy. *Behaviour Research and Therapy* 45, 2537–2550. [PubMed: 17659254]
- Grilo CM, Masheb RM, Wilson G (2006). Rapid response to treatment for Binge eating disorder. *Journal of Consulting and Clinical Psychology* 74, 602–613. [PubMed: 16822116]
- Grilo CM, White MA, Masheb RM, Gueorguieva R (2015). Predicting meaningful outcomes to medication and self-help treatments for binge-eating disorder in primary care: the significance of early rapid response. *Journal of Consulting and Clinical Psychology* 83, 387. [PubMed: 25622201]
- Harned MS, Korslund KE, Linehan MM (2014). A pilot randomized controlled trial of Dialectical Behavior Therapy with and without the Dialectical Behavior Therapy Prolonged Exposure protocol for suicidal and self-injuring women with borderline personality disorder and PTSD. *Behaviour Research and Therapy* 55, 7–17. [PubMed: 24562087]
- Hay P (2013). A systematic review of evidence for psychological treatments in eating disorders: 2005–2012. *International Journal of Eating Disorders* 46, 462–469. [PubMed: 23658093]

- Hay PPI, Bacaltchuk J, Stefano S, Kashyap P (2009). Psychological treatments for bulimia nervosa and bingeing. *Cochrane Database of Systematic Reviews* 4, DOI: 10.1002/14651858.CD000562.pub3.
- Hedeker D, Gibbons RD (2006). *Longitudinal Data Analysis*. John Wiley & Sons, Hoboken, New Jersey.
- Hilbert A, Hildebrandt T, Agras WS, Wilfley DE, Wilson GT (2015). Rapid response in psychological treatments for binge eating disorder. *Journal of Consulting and Clinical Psychology* 83, 649. [PubMed: 25867446]
- Hudson JI, Hiripi E, Pope HG Jr., Kessler RC (2007). The prevalence and correlates of eating disorders in the national comorbidity survey replication. *Biological Psychiatry* 61, 348–358. [PubMed: 16815322]
- Jacobson NS, Truax P (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 59, 12–19. [PubMed: 2002127]
- Johnsen TJ, Friberg O (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: a meta-analysis. *Psychological Bulletin* 141, 747. [PubMed: 25961373]
- Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P, Andreasen NC (1987). The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* 44, 540–548. [PubMed: 3579500]
- Kessler RC, Shahly V, Hudson JI, Supina D, Berglund PA, Chiu WT, Gruber M, Aguilar-Gaxiola S, Alonso J, Andrade LH (2014). A comparative analysis of role attainment and impairment in binge-eating disorder and bulimia nervosa: results from the WHO World Mental Health Surveys. *Epidemiology and Psychiatric Sciences* 23, 27–41. [PubMed: 24054053]
- Lieb K, Zanarini MC, Schmahl C, Linehan MM, Bohus M (2004). Borderline personality disorder. *Lancet* 364, 453–461. [PubMed: 15288745]
- Little RJ, Rubin DB (2014). *Statistical Analysis with Missing Data*. John Wiley & Sons, Hoboken, New Jersey.
- Linehan MM (1993a). *Cognitive-Behavioral Treatment of Borderline Personality Disorder*. The Guilford Press: New York, NY.
- Linehan MM (1993b). *Skills Training Manual for Treating Borderline Personality Disorder*. The Guilford Press: New York, NY.
- Linehan MM (1999). Standard protocol for assessing and treating suicidal behaviors for patients in treatment In *The Harvard Medical School guide to suicide assessment and intervention* (ed. Jacobs DG), pp. 146–187. Jossey-Bass Inc, Publishers: San Francisco, CA.
- Linehan MM (2014). *DBT® Skills Training Manual*. Guilford Publications: New York.
- Linehan MM, Armstrong HE, Suarez A, Allmon D, Heard HL (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry* 48, 1060–1064. [PubMed: 1845222]
- Linehan MM, Dimeff LA, Reynolds SK, Comtois KA, Welch SS, Heagerty P, Kivlahan DR (2002). Dialectical behavior therapy versus comprehensive validation therapy plus 12-step for the treatment of opioid dependent women meeting criteria for borderline personality disorder. *Drug and Alcohol Dependence* 67, 13–26. [PubMed: 12062776]
- Linehan MM, Schmidt III H, Dimeff LA, Craft JC, Kanter J, Comtois KA (1999). Dialectical behavior therapy for patients with borderline personality disorder and drug-dependence. *American Journal of Addictions* 8, 279–292.
- Lynch FL, Striegel-Moore RH, Dickerson JF, Perrin N, DeBar L, Wilson GT, Kraemer HC (2010). Cost-effectiveness of guided self-help treatment for recurrent binge eating. *Journal of Consulting and Clinical Psychology* 78, 322–333. [PubMed: 20515208]
- Lynch TR, Cheavens JS, Cukrowicz KC, Thorp SR, Bronner L, Beyer J (2007). Treatment of older adults with co-morbid personality disorder and depression: a dialectical behavior therapy approach. *International Journal of Geriatric Psychiatry* 22, 131–143. [PubMed: 17096462]

- Masheb RM, Grilo CM (2007). Rapid response predicts treatment outcomes in binge eating disorder: implications for stepped care. *Journal of Consulting and Clinical Psychology* 75, 639–644. [PubMed: 17663617]
- Mitchell JE, Agras S, Crow S, Halmi K, Fairburn CG, Bryson S, Kraemer H (2011). Stepped care and cognitive-behavioural therapy for bulimia nervosa: randomised trial. *British Journal of Psychiatry* 198, 391. [PubMed: 21415046]
- Nahum-Shani I, Qian M, Almirall D, Pelham WE, Gnagy B, Fabiano GA, Waxmonsky JG, Yu J, Murphy SA (2012). Experimental design and primary data analysis methods for comparing adaptive interventions. *Psychological Methods* 17, 457. [PubMed: 23025433]
- Neacsiu AD, Rizvi SL, Linehan MM (2010). Dialectical behavior therapy skills use as a mediator and outcome of treatment for borderline personality disorder. *Behaviour Research and Therapy* 48, 832–839. [PubMed: 20579633]
- R Core & Team (2014). R: A language and environment for statistical computing. In R Foundation for Statistical Computing: Vienna, Austria.
- Safer DL, Jo B (2010). Outcome from a randomized controlled trial of group therapy for binge eating disorder: comparing dialectical behavior therapy adapted for binge eating to an active comparison group therapy. *Behavior Therapy* 41, 106–120. [PubMed: 20171332]
- Safer DL, Telch CF, Agras WS (2001). Dialectical behavior therapy for bulimia nervosa. *American Journal of Psychiatry* 158, 632–634. [PubMed: 11282700]
- Safer DL, Telch CF, Chen EY (2009). *Dialectical Behavior Therapy as adapted for Binge Eating Disorder and Bulimia*. The Guilford Press: New York.
- SAS Institute Inc. (2015). *Language Reference: Concepts In SAS® 9.4 Language Reference: Concepts*: Cary, NC.
- Szegedi A, Jansen WT, van Willigenburg APP, van der Meulen E, Stassen HH, Thase ME (2009). Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder: a meta-analysis including 6562 patients. *Journal of Clinical Psychiatry* 70, 344–353. [PubMed: 19254516]
- Telch CF, Agras WS, Linehan MM (2000). Group dialectical behavior therapy for binge-eating disorder: a preliminary, uncontrolled trial. *Behavior Therapy* 31, 569–582.
- Telch CF, Agras WS, Linehan MM (2001). Dialectical behavior therapy for binge eating disorder. *Journal of Consulting and Clinical Psychology* 69, 1061–1065. [PubMed: 11777110]
- Thompson-Brenner H, Shingleton RM, Sauer-Zavala S, Richards LK, Pratt EM (2015). Multiple measures of rapid response as predictors of remission in cognitive behavior therapy for bulimia nervosa. *Behaviour Research and Therapy* 64, 9–14. [PubMed: 25462877]
- Val-Laillet D, Aarts E, Weber B, Ferrari M, Quaresima V, Stoeckel L, Alonso-Alonso M, Audette M, Malbert C, Stice E (2015). Neuroimaging and neuromodulation approaches to study eating behavior and prevent and treat eating disorders and obesity. *NeuroImage: Clinical* 8, 1–31. [PubMed: 26110109]
- Van Dijk S, Jeffrey J, Katz MR (2013). A randomized, controlled, pilot study of dialectical behavior therapy skills in a psychoeducational group for individuals with bipolar disorder. *Journal of Affective Disorders* 145, 386–393. [PubMed: 22858264]
- Wilson G, Fairburn CC, Agras W, Walsh B, Kraemer H (2002). Cognitive-behavioral therapy for bulimia nervosa: time course and mechanisms of change. *Journal of Consulting and Clinical Psychology* 70, 267–274. [PubMed: 11952185]
- Wilson GT, Wilfley DE, Agras WS, Bryson SW (2010). Psychological treatments of binge eating disorder. *Archives of General Psychiatry* 67, 94–101. [PubMed: 20048227]
- Wilson GT, Zandberg LJ (2012). Cognitive-behavioral guided self-help for eating disorders: effectiveness and scalability. *Clinical Psychology Review* 32, 343–357. [PubMed: 22504491]
- Yuan Y (2011). Multiple imputation using SAS software. *Journal of Statistical Software* 45, 1–25.



**Fig. 1.** Consort flow diagram of Stage 1 (GSH) and Stage 2 (cGSH, DBT, and CBT+).



Note: Error bars reflect one standard error of the mean;

Objective binge day frequency over the last 28 days as assessed by the Eating Disorders Examination -16, Fairburn, 2008;

GSH = first four sessions of Guided self-help; cGSH = continued Guided self-help for early strong responders; DBT = Dialectical Behavior Therapy; CBT + = group and individual Cognitive-Behavior Therapy.

**Fig. 2.**

Objective binge-day frequency changes in DBT and CBT+ groups relative to early strong responders in cGSH at the end of treatment and at follow-up.



**Table 1.**

Demographic and clinical characteristics for the sample at baseline (N = 109)

Variable	Early strong responders		Early weak responders			
	cGSH (n = 42)		DBT (n = 36)		CBT+ (n = 31)	
	n	%	n	%	n	%
Racial background <sup>a</sup>						
Caucasian	31	73.8	27	75	22	71
African-American	7	16.7	6	16.7	6	19.4
American-Indian/Alaska native	0	0	0	0	0	0
Asian	0	0	1	2.8	0	0
Native Hawaiian/Other Pacific Islander	0	0	0	0	0	0
Other	4	9.5	2	5.6	3	9.7
Hispanic <sup>a</sup>	4	9.5	1	2.8	6	19.4
Current marital status						
Single	24	57.1	19	52.8	17	54.8
Married	13	31	12	33.3	7	22.6
Divorced	4	9.5	3	8.3	6	19.4
Separated	1	2.4	2	5.6	0	0
Widowed	0	0	0	0	1	3.2
Highest level of education						
Graduated high school	2	4.8	3	8.3	1	3.2
Some college	5	11.9	10	27.8	6	19.4
Graduated 2 years college	6	14.3	0	0	4	12.9
Graduated 4 years college	6	14.3	7	19.4	8	25.8
Some graduate/professional school	8	19	4	11.1	3	9.7
Currently employed	33	78.6	27	75	21	67.6
No. with lifetime Axis I disorders <sup>b</sup>	21	50	20	55.6	18	58.1
No. with Axis II disorders	9	22	18	50	8	26
Binge-eating disorder <sup>c</sup>	36	85.7	23	61.3	19	61.3
Bulimia nervosa <sup>c</sup>	6	14.3	13	38.7	12	38.7
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Age, years	38.6	12.2	38.2	13.1	37.8	13.9
Gross annual income, dollars	38 762	22 390	30 472	23 135	31 806	25 112

cGSH, Continued guided self-help cognitive behavior therapy; DBT, dialectical behavior therapy; CBT+, group and individual cognitive behavior therapy control.

<sup>a</sup>Self-identified classification and options defined by the investigator.

<sup>b</sup>DSM-IV-TR diagnoses made using the Structured Clinical Interview for DSM-IV, Axis I (SCID-I; First & Gibbon, 2004). This did not count eating disorders. Current = last year; lifetime = before the last year.

<sup>c</sup>In the last month from the Eating Disorders Examination-16 (Fairburn, 2008).

**Table 2.**

Descriptive statistics and main results for objective binge-day frequency

	Baseline		After GSH		End of treatment		6-month follow-up		12-month follow-up	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
<b>Objective binge-day frequency<sup>a</sup></b>										
<i>Early strong responders</i>										
cGSH ( <i>n</i> = 42)	14.60	5.96	1.64	1.97	2.62	5.70	3.31	6.47	3.97	4.48
<i>Early weak responders</i>										
DBT ( <i>n</i> = 36)	13.31	8.05	11.86	7.40	4.31	7.00	4.97	5.89	4.61	7.21
CBT+ ( <i>n</i> = 31)	13.52	6.32	10.84	6.46	4.55	7.06	4.96	7.38	6.18	6.59
<b>Objective binge-episode frequency<sup>a</sup></b>										
<i>Early strong responders</i>										
cGSH ( <i>n</i> = 42)	17.36	9.63	1.79	2.48	3.13	7.81	3.86	8.49	3.97	4.48
<i>Early weak responders</i>										
DBT ( <i>n</i> = 36)	22.22	27.30	15.31	11.71	6.53	16.42	6.44	9.21	6.00	12.09
CBT+ ( <i>n</i> = 31)	21.65	17.00	14.81	12.24	5.00	8.09	7.78	15.67	6.18	6.59
<b>Between group contrasts for objective binge-day frequency<sup>b,c</sup></b>										
<i>Differential treatment hypothesis (baseline to end of treatment)</i>										
	<b>RR</b>	<b>B</b>	<b>p</b>	<b>95% CI<sup>c</sup> for RR</b>						
DBT v. cGSH										
Intercept	0.84	-0.18	0.22	0.63–1.11						
Time effect	1.36	0.31	<0.001	1.18–1.57						
CBT+ v. cGSH										
Intercept	0.89	-0.12	0.44	0.67–1.19						
Time effect	1.36	0.31	<0.001	1.18–1.57						
<i>Follow-up hypothesis (end of treatment to 12-month follow-up)</i>										
DBT v. cGSH										
Intercept	2.20	0.79	0.04	1.04–4.66						
Time effect	0.82	-0.20	0.03	0.68–0.97						
CBT+ v. cGSH										
Intercept	1.76	0.56	0.15	0.81–3.78						
Time effect	1.02	0.02	0.84	0.85–1.22						

GSH, Guided self-help; cGSH, continued guided self-help; DBT, dialectical behavior therapy; CBT+, group and individual cognitive behavior therapy; RR, rate ratio; CI, confidence interval.

<sup>a</sup>Assessed over the last month. For mean objective binge-eating day or episode frequency to be within the non-clinical range for binge-eating disorder or bulimia nervosa <8 mean objective binges per month (DSM-IV-TR) (APA, 2000).

<sup>b</sup>Reference group is early strong responders or the GSH group for these comparisons.

<sup>c</sup>Rate ratios: effect size comparisons between groups.

**Table 3.**

Descriptive statistics and results for the secondary outcome variables

	Early strong responders		Early weak responders		cGSH	Hypotheses	d <sup>a</sup>	95% CI	d	95% CI			
	Mean	S.D.	DBT (n = 42)	S.D.							CBT+ (n = 36)	S.D.	CBT+ (n = 31)
<b>(a) Continuous secondary outcome measures</b>													
Eating Disorders Examination total score <sup>b,c</sup>													
Baseline	3.07	0.87	3.16	0.93	3.10	0.98							
After GSH	1.95	0.85	2.94	1.18	2.62	1.20							
End of tx	1.83	1.11	1.77	1.08	1.86	1.04	Different tx	1.01	0.81 to 1.20	1.08	0.86 to 1.30	1.40	1.07 to 1.72
6-mo. follow-up	1.77	1.14	1.98	1.06	1.86	1.24							
12-mo. follow-up	2.26	1.43	2.15	0.92	2.36	1.46	Follow-up	0.53	-0.01 to 1.08	0.58	0.03 to 1.13	0.50	-0.12 to 1.13
Body mass index for BED <sup>b</sup>													
Baseline	35.89	8.42	38.33	10.57	36.18	10.76							
After GSH	35.80	8.39	38.57	10.77	36.57	10.81							
End of tx	34.68	8.03	38.30	10.12	36.32	10.70	Different tx	0.33	-0.37 to 1.03	-0.34	-1.44 to 0.77	-0.06	-1.24 to 1.12
6-mo. follow-up	34.49	8.16	37.81	9.83	34.95	10.50							
12-mo. follow-up	34.83	9.13	37.14	10.97	33.92	7.96	Follow-up	0.29	-1.43 to 2.02	-0.15	-0.28 to -0.03	-0.24	-2.21 to 1.73
Body mass index for BN <sup>b</sup>													
Baseline	30.66	7.95	24.60	6.92	23.50	4.37							
After GSH	31.18	8.33	24.89	7.66	24.10	5.58							
End of tx	31.98	9.90	25.74	8.43	24.34	5.79	Different tx	-0.32	-4.70 to 4.06	-0.38	-1.40 to 0.65	-0.48	-1.60 to 0.65
6-mo. follow-up	34.51	9.58	25.41	7.96	25.48	7.70							
12-mo. follow-up	34.76	12.91	25.03	5.12	26.06	7.09	Follow-up	-0.36	-9.96 to 9.24	0.12	-2.42 to 2.65	-0.27	-3.15 to 2.61
Global assessment of functioning <sup>b,d</sup>													
Baseline	61.45	10.60	56.64	11.92	57.39	11.04							
After GSH	65.19	9.56	56.00	9.46	58.55	11.34							
End of tx	69.69	11.43	60.53	8.68	64.94	11.05	Different tx	-0.66	-2.66 to 1.34	-0.44	-2.09 to 1.21	-0.93	-3.90 to 2.03
6-mo. follow-up	72.47	9.10	62.16	8.33	65.85	12.46	Follow-up	-0.88	-5.07 to 3.31	-0.54	-4.37 to 3.29	-0.89	-4.91 to 3.14

(b) Count secondary outcome measures (negative binomial)

Vomiting episode frequency <sup>b,e</sup>		DBT (n = 36)		CBT+ (n = 31)		cGSH		DBT		CBT+			
	n	%	n	%	n	%	d	95% CI	d	95% CI	d	95% CI	
Baseline	1.40	4.92	17.39	35.44	17.58	41.21							
After GSH	0.07	0.34	8.44	18.98	9.87	21.59							
End of tx	0.03	0.16	2.44	9.73	7.90	25.84	Different tx	0.30	-0.50 to 1.10	0.50	-3.48 to 4.47	0.23	-15.32 to 15.77
6-mo. follow-up	0.08	0.37	3.19	8.65	10.81	30.19							
12-mo. follow-up	2.32	3.57	6.46	15.48	4.36	8.82	Follow-up	-0.17	-2.13 to 1.80	0.34	-7.09 to 7.77	0.27	-18.79 to 19.32
Number of co-occurring Axis I disorders <sup>b</sup>													
Baseline	1.12	1.47	0.94	1.31	0.94	0.93							
After GSH	0.31	0.64	0.67	0.99	0.68	0.91							
End of tx	0.28	0.72	0.25	0.51	0.16	0.37	Different tx	0.65	0.45 to 0.86	0.58	0.35 to 0.81	0.84	0.50 to 1.18
6-mo. follow-up	0.14	0.42	0.44	0.88	0.44	0.89	Follow-up	0.68	0.18 to 1.18	0.40	-0.11 to 0.91	0.42	0.03 to 0.81

(c) Binomial secondary outcome measures

Objective binge-day abstinence<sup>b</sup>

Baseline	0	0	0	0	0	0							
After GSH	17	41	1	3	1	3							
End of tx	22	56	15	47	14	45	Different tx	1.12	1.04 to 1.20	0.92	0.84 to 1.01	0.89	0.71 to 1.08
6-mo. follow-up	18	50	10	31	12	44							
12-mo. follow-up	8	24	8	29	4	18	Follow-up	0.55	0.40 to 0.70	0.62	0.44 to 0.80	0.46	0.29 to 0.64

Objective binge episode abstinence<sup>b,f</sup>

Baseline	0	0	0	0	0	0						
After GSH	17	41	1	3	1	3						
End of tx	22	56	15	47	14	45						
6-mo. follow-up	18	50	10	31	12	44						
12-mo. follow-up	8	24	8	29	4	18						

Vomiting episode abstinence<sup>b,f</sup>

Baseline	37 <sup>g</sup>	88	22 <sup>g</sup>	61	19 <sup>g</sup>	61						
After GSH	40	95	22	61	20	65						
End of tx	38	97	28	88	23	74						

6-mo. follow-up	36	95	25	76	21	72
12-mo. follow-up	23	68	17	61	13	59

GSH, Guided self-help; cGSH, continued guided self-help; DBT, dialectical behavior therapy; CBT+, group and individual cognitive behavior therapy; BED, binge-eating disorder; BN, bulimia nervosa; CI, confidence interval; mo., month; tx, treatment.

<sup>a</sup>  $d$  = Cohen's  $d$ . These effect sizes are used for within treatment group comparisons. For the differential treatment (Different tx) hypothesis the effect size was calculated from baseline to end of treatment. For the follow-up hypothesis, the effect size is calculated from end of treatment to 12-months follow-up. A rough rubric for effect sizes are that small effects are around  $d = 0.2$ , medium  $d = 0.5$  and large  $d = 0.8$  (Cohen, 1988).

<sup>b</sup> Assessed over the last month.

<sup>c</sup> For total Eating Disorders Examination (EDE) scores to be within the non-clinical range mean < 2.45. Mean = 2.45 is the 2 S.D. cut-off for a healthy sample of  $N = 337$  from Fairburn & Wilson (1993).

<sup>d</sup> Global assessment of functioning scores of 51–60 refer to moderate impairment; 61–70, mild impairment; and 71–80, slight impairment.

<sup>e</sup> For mean vomiting frequency to be within the non-clinical range for BN (DSM-IV-TR) (APA, 2000)  $m < 8$  vomiting episodes/month.

<sup>f</sup> Objective binge episode abstinence and vomiting episode abstinence were not analyzed because they were respectively highly correlated with objective binge day abstinence and vomiting frequency.

<sup>g</sup> One participant in cGSH who met BN criteria over-exercised (28 days/last month); and one participant in DBT predominantly over-exercised (28 days/last month).