Psychotherapy for bulimia nervosa on symptoms of depression: A meta-analysis of randomized controlled trials

Jake Linardon¹, Tracey Wade², Xochitl de la Piedad Garcia¹, Leah Brennan¹

¹School of Psychology, Australian Catholic University, 115 Victoria Parade/Locked Bag 4115, Melbourne, Victoria, Australia, 3065.
²School of Psychology, Flinders University, GPO Box 2100, Adelaide, South Australia, 5001.

Correspondence concerning this article should be addressed to Jake Linardon (Jake.Linardon@acu.edu.au), Faculty of Health Sciences, Australian Catholic University, 115 Victoria Parade/Locked Bag 4115, Melbourne, Victoria, 3065.

Phone: +613 9953 3662
Fax: +613 9953 3205

Word count

Abstract: 250
Manuscript: 4067
**Abstract**

**Objective:** Depressive symptoms are an important risk factor and consequence of binge eating and purging behavior in bulimia nervosa (BN). Although psychotherapy is effective in reducing symptoms of BN in the short- and long-term, it is unclear whether psychotherapy for BN is also effective in reducing depressive symptoms. This meta-analysis examined the efficacy of psychotherapy for BN on depressive symptoms in the short and long-term. **Method:** Randomized controlled trials (RCTs) on BN that assessed depressive symptoms as an outcome were identified. Twenty-six RCTs were included. **Results:** Psychotherapy was more efficacious at reducing symptoms of depression at post-treatment ($g=0.47$) than wait-lists. This effect was strongest when studies delivered therapist-led, rather than guided self-help, treatment. No significant differences were observed between psychotherapy and antidepressants. There was no significant post-treatment difference between CBT and other active psychological comparisons at reducing symptoms of depression. However, when only therapist-led CBT was analyzed, therapist-led CBT was significantly more efficacious ($g=0.25$) than active comparisons at reducing depressive symptoms. The magnitude of the improvement in depressive symptoms was predicted by the magnitude of the improvement in BN symptoms. **Discussion:** These findings suggest that psychotherapy is effective for reducing depressive symptoms in BN in the short-term. Whether these effects are sustained in the long-term is yet to be determined, as too few studies conducted follow-up assessments. Moreover, findings demonstrate that, in addition to being the front-running treatment for BN symptoms, CBT might also be the most effective psychotherapy for improving the symptoms of depression that commonly co-occur in BN.

**Keywords:** bulimia nervosa; cognitive-behavioural therapy; psychotherapy; depression; meta-analysis
Bulimia nervosa (BN) is a psychiatric disorder characterized by an over-evaluation of weight and shape and recurrent episodes of binge eating in combination with compensatory behavior (American Psychiatric Association, 2013). BN often runs a chronic course and is associated with serious medical complications, psychiatric comorbidity, and psychological impairment (Fairburn & Harrison, 2003). Randomized controlled trials (RCTs) have shown a range of psychological treatments, including cognitive-behavioral therapy (CBT), interpersonal psychotherapy (IPT) and dialectical behavior therapy (DBT), to be effective in reducing symptoms of BN in both the short and long-term (e.g., Agras, Walsh, Fairburn, Wilson, & Kraemer, 2000; Fairburn et al., 1991; Goldbloom et al., 1997; Poulsen et al., 2014; Safer, Telch, & Agras, 2001). Multiple systematic reviews of RCTs conclude that specific forms of CBT are the most effective treatment for BN (Hay, 2013; Hay, Bacaltchuk, Stefano, & Kashyap, 2009; Shapiro et al., 2007). This has prompted clinical guidelines to recommend CBT as the first line treatment for BN (Hay et al., 2014; National Institute of Clinical Excellence, 2017).

Many individuals with BN also suffer from depression, either meeting criteria for a comorbid depressive disorder or reporting elevated symptoms of depression. For instance, recent research has shown that 60% of individuals with BN met criteria for a lifetime major depressive disorder (Godart et al., 2015), and early research that compared individuals with BN to individuals with major affective disorder reported comparable levels of depressive symptoms across both conditions (Cooper & Fairburn, 1986). Depressive symptoms are also said to be a major risk factor, maintaining factor, and consequence of BN (Fairburn & Harrison, 2003; Puccio, Fuller-Tyszkiewicz, Ong, & Krug, 2016; Stice, 2001), and the two disorders share genetic risk factors (Slane, Burt, & Klump, 2011). Indeed, depressive symptoms have been shown to predict short and long-term BN persistence, poor treatment outcome, and relapse (Fahy
Depressive symptoms are often included as a secondary treatment outcome in RCTs evaluating BN treatment. Numerous RCTs of psychological treatments for BN have reported large improvements in depressive symptoms from pre-treatment to post-treatment and follow-up (e.g., Cooper & Steere, 1995; Fairburn et al., 1991; Poulsen et al., 2014; Safer et al., 2001; Wonderlich et al., 2014). There is also evidence that CBT is more effective than alternative psychotherapy approaches (e.g., psychodynamic therapy, short-term focal psychotherapy, and supportive suppressive therapy) at reducing symptoms of depression in BN (Fairburn, Kirk, O'Connor, & Cooper, 1986; Garner et al., 1993; Poulsen et al., 2014).

To date, one early meta-analysis has examined the effects of CBT for BN on symptoms of depression (Hay et al., 2009). The authors concluded that CBT significantly outperformed wait-list controls ($k=7$, $d = 0.69$, 95% CI = -1.09, -0.30) but not active psychological comparisons ($k=13$, $d= -0.28$, 95% CI = -0.57, 0.00) in reducing depressive symptoms at post-treatment in individuals with BN and binge eating disorder (Hay et al., 2009). However, key questions still remain. First, evidence suggests that treatment non-response tends to be higher in BN than in BED samples (Castellini et al., 2011), suggesting we need to know if the beneficial effects of CBT or psychotherapy on depressive symptoms apply to individuals with BN specifically. Second, it is unknown whether these improvements are CBT-specific or whether other distinct psychotherapy approaches also show similar improvements in depressive symptoms. Third, moderators of the effects of CBT for BN on depressive symptoms have not been tested. For instance, it is not known whether different psychotherapy modalities that vary in duration have similar effects on depressive symptoms (i.e., a dose-response relationship), or
whether the improvements in depressive symptoms following psychotherapy are simply a consequence of BN symptom improvement. The latter question is crucial for understanding the mechanisms through which psychotherapy for BN achieves its effects on reducing depressive symptoms (Kazdin, 2007). Finally, since 2009, seven additional RCTs of psychotherapy for BN that have assessed symptoms of depression have been published, allowing for more adequately powered analyses.

The current meta-analysis therefore aims to examine the efficacy of psychotherapy for BN on symptoms of depression. In particular, we aim to compare (a) psychotherapy to inactive control groups; (b) psychotherapy to pharmacotherapy; and (c) CBT to any alternative psychological treatment on symptoms of depression at post-treatment and follow-up. In addition, we also aim to test whether these effects are associated with or moderated by the type of treatment modality, the duration of treatment, the degree of BN symptom improvement, whether use of antidepressant was an exclusion criterion, and the quality of included studies.

Method

This review was conducted in accordance to the latest Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009).

Search strategy

The primary search strategy involved searching the Medline and PsycInfo database on the 8th of May 2017. The following terms were searched and combined using the “AND” Boolean operator: “bulimia nervosa”, bulimi* AND psychotherap*, therap*, treat*, intervention*, “self help”, self-help, “self care”, self-care, counsel*, bibliotherap*, AND random*, trial*, RCT,
controlled, allocat*, assign*. The secondary search strategy involved searching the reference list of included studies and relevant reviews.

**Inclusion criteria**

We included (a) published RCTs (b) that compared any form of psychotherapy to either an inactive control (defined as either a wait-list or placebo control) or pharmacotherapy condition, or compared CBT to an alternative psychotherapy condition, (c) in individuals with BN, and (d) reported symptoms of depression at post-treatment or follow-up. Consistent with previous reviews, psychotherapy was defined as an intervention in which the central element was verbal communication between a client and a therapist, or as a psychological treatment in the form of a website or book which the participant worked through somewhat independently but with guided support from a therapist (e.g., Kolovos, Kleiboer, & Cuijpers, 2016). We therefore excluded studies that (a) compared two different variants of CBT (e.g., group vs individual), (b) compared two or more psychotherapies for BN of which one was not cognitive-behavioral based, and (c) did not or could not provide any data allowing for the calculation of an effect size. We excluded four studies because an effect size could not be calculated and the relevant data necessary for calculating an effect size could not be provided (Freeman, Barry, Dunkeld-Turnbull, & Henderson, 1988; Freeman, Sinclair, Turnbull, & Annandale, 1985; Mitchell et al., 2001; Ordman & Kirschenbaum, 1985).

**Study selection**

The first author (JL) conducted the search. Once the outputs from the databases were combined, all duplicate records were removed, and titles and abstracts were screened (by JL). To maximize identification of relevant articles, all RCTs that compared either a psychological treatment to a control condition or CBT to an alternative psychotherapy condition in individuals
with BN were read by the first author in their entirety. This was because measures of depressive
symptoms are typically a secondary outcome reported, and are often not mentioned in the title or
abstract. Authors JL and LB discussed trials in which their inclusion was uncertain. A decision
was made for these trials after consensus was reached. In total, 27 articles (reporting 26 RCTs)
met full inclusion criteria. A flowchart of the search strategy is presented in Figure 1.

Quality assessment

The validity of trials was assessed using the four criteria of the Cochrane Collaboration
Risk of Bias tool (Higgins & Green, 2011). This risk of bias tool assesses potential sources of
biases in RCTs, such as the adequate generation of allocation sequence, the concealment of
allocation to treatment conditions, blinding of outcome assessors, and dealing with incomplete
data. Dealing with incomplete data was assessed as low risk when ITT analyses were conducted.
The two other criteria of the Cochrane Collaboration tool were not used, as there was no
indication that there were selective outcome reporting or other potential sources of bias. The first
author (JL) and an independent research assistant conducted the assessments. Ratings were
cross-checked, and any discrepancy was discussed and resolved.

Meta-analysis

Meta-analyses were performed for the three main comparisons: (1) psychotherapy versus
inactive controls (2) psychotherapy versus pharmacotherapy; and (3) CBT versus any alternative
psychotherapy approach. When studies compared two different psychotherapy conditions to a
control group, the sample size of the control group was halved to avoid double counting in the
meta-analysis (Higgins & Green, 2011). Where feasible, analyses were also performed at follow-
up (i.e., the last reported follow-up). Intention to treat (ITT) data were analyzed and prioritized over completer data.

Effect sizes (standardized mean difference; SMD) were calculated by dividing the difference between the post-treatment psychotherapy group mean and the post-treatment control group mean by the pooled standard deviation (Lipsey & Wilson, 2001). If means and standard deviations were not reported, SMD was calculated using conversion equations from significance tests (Borenstein, Hedges, Higgins, & Rothstein, 2009). We then converted SMD to Hedges $g$ to correct for sample size. To compute a pooled effect size, each study’s effect size was weighted by its inverse variance. Effect sizes were coded such that positive $g$ values indicate that the psychotherapy (or CBT) condition had lower depressive symptoms than the comparison condition. Small (0.2), medium (0.5) and large (0.8) effects were specified. Comprehensive Meta-Analysis was used to calculate effect sizes (Borenstein et al., 2009).

Since heterogeneity was expected among the studies, a random effects model was used. Heterogeneity was assessed through the $I^2$ statistic. The $I^2$ statistic assesses the degree of heterogeneity, where a value of 0% indicates no observed heterogeneity, 25% low heterogeneity, 50% moderate heterogeneity, and 75% as high heterogeneity (Higgins & Thompson, 2002).

We conducted subgroup analyses to identify potential moderators. For the subgroup analyses, a pooled effect size was calculated for each subgroup, and a test was conducted to determine whether the effect sizes for subgroups differ significantly from each other (Borenstein et al., 2009). A mixed effects model was used, which pools studies within a subgroup using a random effects model, but tests for differences between subgroups using a fixed effects model (Borenstein et al., 2009). Significant differences between subgroups are tested by the $Q_{between}$ statistic. Subgroup analyses were conducted for the following categorical moderators:
psychotherapy modality (therapist-led versus guided self-help); psychotherapy type (CBT versus other); whether studies excluded participants who were concurrently taking medication (yes excluded versus not excluded); and study quality (Jadad score ≥ 3 versus Jadad score ≤ 2).

Continuous moderators were tested using meta-regression. We examined whether there was a relationship between depressive symptoms and (a) the number of treatment sessions provided, (b) changes in binge eating, and (c) changes in purging. For the meta-regression, the effect size for changes in depressive symptoms was set as the dependent variable, while the number of treatment sessions, the effect size for binge eating, and the effect size for purging was set as the independent variable. Unstandardized coefficients were computed for meta-regression analyses.

The Fail-Safe $N$ was calculated to address potential publication bias (Rosenthal, 1991). The Fail-Safe $N$ estimates how many missing studies would need to be included in the meta-analysis for the effect size to become not significantly different from zero. An effect that is considered robust to publication bias is where the number of studies needed to reduce the effect size to zero is greater than the number of studies included in the meta-analysis (Rosenthal, 1991).

Results

Characteristics of included studies

Twenty-six RCTs (27 papers) met full inclusion criteria. Table 1 presents the characteristics of included studies. All studies sampled adults, with the exception of one study that sampled adolescents (Le Grange, Lock, Agras, Bryson, & Jo, 2015). The majority of included studied ($k=19$) used the Beck Depression Inventory (BDI) as a measure of depressive symptoms. Ten studies compared psychotherapy for BN to a wait-list control and one compared
psychotherapy for BN to a pill-placebo. Of these 11 trials, eight delivered a cognitive-behavioral intervention, with five studies delivering CBT in a therapist-led format, two studies delivering CBT in a guided self-help format, and one study delivering CBT in a pure self-help format. The other psychotherapy conditions for this comparison were dialectical behavior therapy ($N_{comp}=2$), non-specific supportive therapy ($N_{comp}=2$), and behavior therapy ($N_{comp}=1$).

Five studies compared psychotherapy for BN to pharmacotherapy. All pharmacotherapy trials administered an antidepressant; four administered fluoxetine and one administered imipramine. Four trials administered CBT as the psychotherapy condition, and one trial administered CBT and a supportive psychotherapy condition.

Fifteen trials compared CBT to an alternative psychotherapy. Of these, 13 delivered therapist-led CBT, one delivered guided self-help CBT, and one delivered pure self-help CBT. Comparison conditions varied, and included behavior therapy, supportive therapy, psychodynamic therapy, emotion and social mind training, integrative cognitive-affective therapy, interpersonal psychotherapy, family therapy, and a mindfulness-based intervention (See Table 1).

Of the 26 RCTs, 14 trials excluded participants who were concurrently taking medication, whereas 12 trials did not exclude concurrent medication use. Critically, all of these trials ensured that all participants had been on a stable dose of medication for a minimum of six weeks, and five of these trials reported how many participants were on medication (<10% of the total sample were concurrently taking medication).

The quality of included trials varied. Fifteen trials reported an adequate sequence generation, five trial reported adequate allocation concealment, all trials used a self-report measure of depressive symptoms, and 13 trials conducted ITT analyses. Four trials met all four
quality criteria, seven trials met three criteria, six trials met two criteria, and nine trials met just one criteria.

**Psychotherapy vs inactive controls**

Fourteen comparisons of psychotherapy to inactive controls (i.e., wait-list, placebo) on depressive symptoms were included. The mean effect size for depressive symptoms was moderate, statistically significant ($g = 0.50; 95\% CI = 0.22, 0.79$), and in favor of the psychotherapy at post-treatment. There was moderate to high heterogeneity ($I^2 = 71.06$). The fail-safe $N$ was 120, indicating no publication bias. When the BDI was used only as the measure of depressive symptoms ($N_{comp} = 12$), the mean effect size was $g = 0.47$ ($95\% CI = 0.15, 0.81$).

Analyses were not performed at follow-up as only two studies reported follow-up data.

The mean effect sizes for binge eating ($N_{comp} = 7, g = 0.61, 95\% CI = 0.21, 1.01$) and purging ($N_{comp} = 11, g = 0.64, 95\% CI = 0.32, 0.95$) were moderate, statistically significant, and in favor of psychotherapy interventions. Heterogeneity was large ($I^2 = 70.87$, binge eating, $I^2 = 64.16$ for purging), and there was no indication of publication bias (fail-safe $N = 46$ and 110, respectively).

Results from the subgroup analyses across the three outcomes at post-treatment can be seen in Table 2. The effect sizes for depressive symptoms and purging were significantly larger for studies that delivered therapist-led psychological treatments compared to studies that delivered self-help treatments. For binge eating, the effect size was significantly larger for studies that delivered alternative psychological treatments ($N_{comp} = 2$) than for studies that delivered any mode of CBT ($N_{comp} = 6$). The two studies that contributed to this former subgroup (alternative interventions) delivered therapist-led DBT. When we performed an analysis in which the effect size was compared for studies that delivered therapist-led DBT ($N_{comp} = 2$) to studies
that delivered therapist-led CBT ($N_{comp}=3$), the effect size was still significantly larger (for binge eating) for DBT studies. The effect size of depressive symptoms for the subgroup of studies that excluded concurrent medication use was also moderate and statistically significant.

For all studies included in the psychotherapy versus wait-list comparison, meta-regression analyses revealed that there was a positive and statistically significant relationship between the number of treatment sessions ($B=0.05$, 95% $CI=0.01$, 0.10, $p=0.010$) and the effect size for depressive symptoms. There was no statistically significant relationship between the effect size for depressive symptoms and the effect size for binge eating ($B=0.14$, 95% $CI=-0.38$, 0.67, $p=0.595$) and purging ($B=0.51$, 95% $CI=-0.03$, 1.06, $p=0.068$).

**Psychotherapy vs antidepressants**

Only five studies (6 comparisons) were included in this meta-analysis comparing psychotherapy to antidepressants on depressive symptoms. The mean effect size was small, non-significant, and in favor of antidepressant medication ($g=-0.11$, 95% $CI=-0.34$, 0.12). No heterogeneity was present ($I^2=0.68$). The mean effect size for binge eating ($g=-0.23$, 95% $CI=-0.61$, 0.15) and purging ($g=-0.36$, 95% $CI=-0.77$, 0.04) was small and non-significant. Given the small number of studies included in this analysis, subgroup analyses were not performed. No follow-up data were available.

**Cognitive-behavioral therapy vs alternative psychotherapies**

Eighteen comparisons were included in this meta-analysis comparison of CBT versus alternative psychotherapies on depressive symptoms. The mean effect size was small, non-significant ($g=0.18$, 95% $CI=-0.03$, 0.38) and in favor of CBT. There was a moderate amount of heterogeneity present ($I^2=54.27$). A similar effect size was observed when the BDI was only used as the measure of depressive symptoms ($N_{comp}=13$, $g=0.15$, 95% $CI=-0.09$, 0.40). At
follow-up, the mean effect size was also small, non-significant, and in favor of CBT ($N_{comp}=11$, \(g=0.14, 95\% \ CI = -0.06, 0.35\)).

The mean effect size for binge eating was small to moderate, statistically significant, and in favor of CBT ($N_{comp}=13, g=0.30, 95\% \ CI = 0.07, 0.53$). There was moderate heterogeneity ($I^2 = 53.56$) and no indications of publication bias (Fail-safe $N=33$). The mean effect size for purging was small, non-significant, and in favor of CBT ($N_{comp}=17, g=0.18, 95\% \ CI = -0.05, 0.40$). Heterogeneity was moderate ($I^2 = 58.19$).

Results from the subgroup analyses of CBT vs active comparisons across the three outcomes at post-treatment can be seen in Table 2. CBT was significantly more efficacious than alternative psychotherapies on reducing symptoms of depression and binge eating only when CBT was led by a therapist (as opposed to guided self-help CBT). Meta-regression analyses reveal that there was no relationship between the effect size for depressive symptoms and the number of treatment sessions administered ($B=0.02, 95\% \ CI = -0.01, 0.06, p=.260$). There was a statistically significant relationship between the effect size for depressive symptoms and the effect size for binge eating ($B=0.72, 95\% \ CI = 0.36, 1.09, p<.001$) and purging ($B=0.68, 95\% \ CI = 0.34, 1.02, p<.001$), indicating that greater changes in binge eating and purging during CBT, relative to alternative psychotherapies, were associated with greater changes in depressive symptoms (See Figure 2 and 3).

**Discussion**

This meta-analysis examined the efficacy of psychotherapy for BN on symptoms of depression. Psychotherapy was found to be more effective than inactive controls (i.e., wait-lists) at post-treatment in reducing depressive and bulimic symptoms in individuals with BN. The
The effect size observed in the current study ($g = 0.47$) was slightly lower than the effect size observed in a recent meta-analysis that examined the efficacy of psychotherapy in depressed populations ($g = 0.71$) (Cuijpers et al., 2013). On the other hand, we found no clear benefit of psychotherapy over antidepressant medication for individuals with BN in terms of depressive symptoms. This was not unexpected, as antidepressants have been shown to be just as effective as psychotherapy in the short-term at reducing depressive symptoms across a range of clinical conditions (Spielmans, Berman, & Usitalo, 2011). In addition, CBT was equally efficacious to active psychological controls at reducing depressive and bulimic symptoms at post-treatment. However, when only therapist-led CBT was analyzed (i.e., self-help studies were omitted), CBT was significantly more efficacious than active psychological controls at reducing depressive and bulimic symptoms.

The type of psychotherapy modality delivered emerged as a consistent moderating variable. In particular, therapist-led psychological treatments, including CBT, was typically associated with larger effect sizes in depressive symptom improvement than abbreviated guided self-help interventions. The size of the effect comparing therapist-led psychological interventions to wait-list controls ($g = 0.47$) and therapist-led CBT to active comparisons ($g = 0.25$) was quite similar to the effect size reported in Hay and colleagues’ meta-analysis ($d = 0.69$, and $d = 0.28$, respectively). The fact that therapist-led psychological treatments are typically longer in duration than guided self-help treatments suggests that the amount of therapist contact might, at least in part, play a role in effectively reducing co-occurring symptoms of depression in BN. Indeed, this dose-response relationship was further supported in our meta-regression analysis which demonstrated a positive relationship between the number of treatment sessions and the effect size in depressive symptoms.
When CBT was compared to active controls, a strong association between the effect size for depressive symptoms and the effect size for BN symptoms was observed. Specifically, greater improvements in binge eating and purging were associated with greater improvements in depressive symptoms during CBT. Although the data precludes causal inferences, these findings might suggest that improvements in depressive symptoms could be a byproduct of BN symptom improvement, particularly since a primary goal of CBT is to eliminate BN symptoms via a collection of treatment strategies designed to directly target the maintaining mechanisms of BN (Fairburn, Marcus, & Wilson, 1993). CBT for eating disorders has traditionally devoted minimal attention to directly targeting depressive symptoms, though enhanced versions of CBT (CBT-E) now incorporate mood regulation strategies designed to address negative affective states (Fairburn, 2008). Importantly, only two trials delivered this enhanced version of CBT.

There has been a recent call for prioritizing the delivery of treatments that can target more than one type of problem, as these “best buy” interventions are purported to be cost-effective and appropriate to implement within the constraints of a local health system (Kazdin, Fitzsimmons-Craft, & Wilfley, 2017). Interpersonal psychotherapy has been touted as a best buy intervention for eating disorders, as IPT can produce improvements in BN symptoms and interpersonal functioning and general psychiatric symptoms (Wilfley et al., 1993; Wilfley et al., 2002). In relation to the current study, although therapist-led DBT was associated with greater improvements in binge eating than therapist-led CBT relative to wait-lists, we found therapist-led CBT to directly outperform any other psychotherapy approach on depressive symptom improvement. Therefore, our findings indicate that therapist-led CBT for BN can be categorized as a best buy intervention, as CBT can not only reduce BN symptoms and improve quality of life
(Linardon & Brennan, 2017), but it can also effectively reduce any co-occurring symptoms of depression.

There are limitations to the current study that must be considered. First, although an association between improvements in BN and depressive symptoms was observed, the nature of the analyses did not allow us to determine whether improvements in depressive symptoms were a cause or a consequence of the improvements observed in bulimic symptoms. While there is reason to suspect that improvements in depressive symptoms are a consequence of BN symptom improvement (Puccio et al., 2016), particularly since a rapid reduction in BN symptoms independent of depressive symptoms are often observed within the first few weeks of CBT (Linardon, Brennan, & de la Piedad Garcia, 2016; Linardon, de la Piedad Garcia, & Brennan, 2016; Thompson-Brenner, Shingleton, Sauer-Zavala, Richards, & Pratt, 2015; Vall & Wade, 2015), the possibility of reverse causality cannot be ruled out until further research examines the temporal nature of these relationships, as has been conducted with changes in therapeutic alliance (Graves et al., 2017). Second, the number of comparisons contributing to some of the subgroup analyses was small. Thus, we may have lacked adequate statistical power to detect some moderation effects (Borenstein et al., 2009). Third, the current findings only apply to outcomes at post-treatment assessment. Too few studies assessed the efficacy of psychotherapy for BN on depressive symptoms at follow-up. Thus, it is unclear whether the beneficial effects of psychotherapy for BN on depressive symptoms are sustained well after treatment ends. Follow-up assessments are therefore important. Finally, only English studies that were published in peer-review journals were included. Publication bias was not evident in the current meta-analysis. However, because unpublished studies are more likely to report non-significant findings, our effect sizes might have been inflated.
In sum, the current study demonstrated that psychotherapy for BN is efficacious for not only reducing symptoms of BN, but also for reducing symptoms of depression. The greatest improvements in depressive symptoms are made when psychotherapy is led by a therapist rather than when delivered in a guided self-help format. The current findings also indicate that greater improvements in depressive symptoms during CBT might be explained by greater improvements in BN symptoms. The fact that therapist-led CBT for BN was more effective at reducing depressive symptoms than alternative psychotherapies suggests that CBT for BN has a powerful therapeutic effect for more than one type of problem, and can therefore be categorized as a “best buy” psychological treatment for BN.


or related disorders in a student population. *Psychological Medicine, 41*(2), 407-417. doi:10.1017/S0033291710000711


Figure 1
1. PRISMA flowchart of literature search

2. Records identified through database searching (n=1258)

3. Additional records identified through other sources (n = 0)

   Records after duplicates removed (n = 797)

   Records screened (n=797)

   Full-text articles assessed for eligibility (n 119)

   Studies included in qualitative synthesis (n=31)

   Studies included in quantitative synthesis (n=27 papers, 26 RCTs)

   Records excluded (n=678)

   Full-text articles excluded, with reasons (n=88)
   - No measure of depression (n=29)
   - Not a psychological treatment (n=7)
   - Not one of the three main comparisons analyzed (n=29)
   - Not a BN sample (n=17)
   - Not a RCT (n=6)
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample age (M SD)</th>
<th>CBT intervention (n)</th>
<th>Alternative psychotherapy (n)</th>
<th>Inactive control or pharmacotherapy condition (n)</th>
<th>Depression measure</th>
<th>Concurrent medication use an exclusion criteria?</th>
<th>Quality assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Agras, Schneider, Arnow, Raeburn, &amp; Telch, 1989)</td>
<td>29.2 (8.6)</td>
<td>Therapist-led (n=17) 14 individual sessions</td>
<td>Self-monitoring group (n=16) 14 individual sessions</td>
<td>Wait-list (n=18)</td>
<td>BDI</td>
<td>Yes</td>
<td>+ ? SR -</td>
</tr>
<tr>
<td>(Banasiak, Paxton, &amp; Hay, 2005)</td>
<td>29.5 (8.72)</td>
<td>GSH (n=54) 9 Sessions</td>
<td>-</td>
<td>Wait-list (n=55)</td>
<td>BDI</td>
<td>Yes</td>
<td>+ + SR +</td>
</tr>
<tr>
<td>(Carter et al., 2003)</td>
<td>27.0 (8.00)</td>
<td>PSH (n=28) 8 sessions</td>
<td>Supportive PSH (n=28) 8 sessions</td>
<td>Wait-list (n=29)</td>
<td>BDI</td>
<td>No. n= not specified</td>
<td>+ + SR +</td>
</tr>
<tr>
<td>(Cooper &amp; Steere, 1995)</td>
<td>23.8 (NA)</td>
<td>Therapist-led (n=13) 19 sessions</td>
<td>Behavior therapy (n=13) 19 sessions</td>
<td>-</td>
<td>BDI</td>
<td>No. n = not specified</td>
<td>? ? SR -</td>
</tr>
<tr>
<td>(Davis, McVey, Heinmaa, Rockert, &amp; Kennedy, 1999)</td>
<td>27.1 (5.3)</td>
<td>Therapist-led (n=37) 20 sessions</td>
<td>-</td>
<td>Wait-list (n=19)</td>
<td>BDI</td>
<td>Yes</td>
<td>? ? SR -</td>
</tr>
<tr>
<td>(Fairburn et al., 1991)</td>
<td>24.2 (NA)</td>
<td>Therapist-led (n=21) 19 sessions</td>
<td>IPT (n=22)</td>
<td>-</td>
<td>BDI</td>
<td>No. n = not specified</td>
<td>? ? SR -</td>
</tr>
<tr>
<td>(Fairburn et al., 1986)</td>
<td>22.9 (4.4)</td>
<td>Therapist-led (n=11) 19 sessions</td>
<td>BT (n=19) Short-term focal psychotherapy (n=11)</td>
<td>-</td>
<td>MADRS</td>
<td>Yes</td>
<td>+ ? SR -</td>
</tr>
<tr>
<td>(Garner et al., 1993)</td>
<td>23.7 (4.4)</td>
<td>Therapist-led (n=25) 19 sessions</td>
<td>Supportive expressive therapy (n=25) 19 sessions</td>
<td>-</td>
<td>BDI</td>
<td>No. n = not specified</td>
<td>+ - SR -</td>
</tr>
<tr>
<td>Study</td>
<td>Mean (SD)</td>
<td>Intervention Details</td>
<td>Comparator Details</td>
<td>Outcome Measures</td>
<td>Follow-up</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------------------</td>
<td>-----------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>(Goldbloom et al., 1997)</td>
<td>25.8 (5.5)</td>
<td>Therapist-led (n=14) 16 sessions</td>
<td>Fluoxetine (n=12) BD</td>
<td>Yes (for CBT group)</td>
<td>? ? SR -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Gulec et al., 2014)</td>
<td>28.2 (7.8)</td>
<td>Online non-specific internet based intervention (n=44) 6 sessions</td>
<td>Wait-list (n=51) DASS</td>
<td>No. n = not specified</td>
<td>? ? SR -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Hill, Craighead, &amp; Safer, 2011)</td>
<td>22.6 (5.6)</td>
<td>-</td>
<td>DBT-AF (n=18) 12 sessions</td>
<td>Wait-list (n=14) BDI</td>
<td>No. n = 2 participants on medication</td>
<td>+ ? SR +</td>
<td></td>
</tr>
<tr>
<td>(Hsu et al., 2001)</td>
<td>24.5 (6.4)</td>
<td>Therapist-led (n=24) 12 sessions</td>
<td>Support group (n=24) 12 sessions</td>
<td>HADRS</td>
<td>Yes</td>
<td>+ ? SR +</td>
<td></td>
</tr>
<tr>
<td>(Jacobi, Dahme, &amp; Dittmann, 2002)</td>
<td>26.0 (5.8)</td>
<td>Therapist-led (n=19) 20 sessions</td>
<td>-</td>
<td>Fluoxetine (n=16) BDI</td>
<td>Yes (for CBT group)</td>
<td>? ? SR +</td>
<td></td>
</tr>
<tr>
<td>(Lavender et al., 2012)</td>
<td>27.7 (7.6)</td>
<td>Therapist-led (n=21) 17 sessions</td>
<td>Emotion social mind training (n=23)</td>
<td>-</td>
<td>No. n = not specified</td>
<td>+ ? SR +</td>
<td></td>
</tr>
<tr>
<td>(Le Grange et al., 2015)</td>
<td>15.7 (1.5)</td>
<td>Therapist-led (n=58) 18 sessions</td>
<td>FBT-BN (n=51) -</td>
<td>BDI</td>
<td>No. n = 10 participants</td>
<td>+ + SR +</td>
<td></td>
</tr>
<tr>
<td>(Lee &amp; Rush, 1986)</td>
<td>27.7 (5.3)</td>
<td>Therapist-led (n=15) 12 sessions</td>
<td>-</td>
<td>Wait-list (n=15) BDI</td>
<td>Yes</td>
<td>+ ? SR +</td>
<td></td>
</tr>
<tr>
<td>(Leitenberg, Rosen, Gross, Nudelman, &amp; Vara, 1988)</td>
<td>25.0 (5.4)</td>
<td>Therapist-led (n=22) 24 sessions</td>
<td>Exposure response prevention (n=11) 24 sessions</td>
<td>-</td>
<td>BDI</td>
<td>Yes</td>
<td>? ? SR -</td>
</tr>
<tr>
<td>Study Reference</td>
<td>Baseline</td>
<td>Treatment Details</td>
<td>Control Details</td>
<td>Outcome Measure</td>
<td>Additional Details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>----------</td>
<td>------------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>---------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| (Mitchell et al., 1990)                             | 22.8 (4.3)| Therapist-led (n=33)
10 sessions                                               | -               | Imipramine (n=45) | HADRS Yes (for CBT group) | ? - SR - |
| (Poulson et al., 2014)                              | 25.8 (4.9)| Therapist-led (n=36)
20 sessions                                               | Psychoanalysis (n=36) | -               | BDI No. n = 10 participants | ++ SR + |
| (Safer et al., 2001)                                | 34.0 (11.0)| -                                                                                | Therapist-led DBT (n=14) 20 sessions | Wait-list (n=15) | BDI Yes. | ?? SR - |
| (Sánchez-Ortiz et al., 2011)                        | 23.9 (5.9)| GSH (n=31) Sessions = NS                                                         | -               | Wait-list (n=36) | HADS No. n = 6 participants | ++ SR - |
| (Steele & Wade, 2008)                               | 25.7 (5.6)| GSH (n=15); 8 sessions
CBT GSH for perfectionism (n= 17) 8 sessions             | Mindfulness-based CT (n=15) | -               | DASS No. n = 10 participants | + - SR + |
| (Thackwray, Smith, Bodfish, & Meyers, 1993)         | 31.3 (10.4)| Therapist-led (n=13) 8 sessions                                                  | BT (n=13)       | -               | BDI Yes. | ?? SR - |
| (Walsh, Fairburn, Mickley, Sysko, & Parides, 2004)  | 30.6 (7.8)| GSH (n=25) 8 sessions                                                           | -               | Fluoxetine (n=20) | BDI Yes. | ?? SR + |
| (Walsh et al., 1997)                                | 25.8 (4.4)| Therapist-led (n=25) 20 sessions                                                | Supportive expressive therapy (n=22) 20 sessions | Fluoxetine (n=28) | BDI Yes. | ?? SR + |
| (Wonderlich et al., 2014)                           | 27.3 (9.6)| Therapist-led (n=36) 21 sessions                                                | Integrative cognitive affective therapy (n=36) 20 sessions | -               | BDI No. n = not specified | + ? SR + |

Note: GSH= guided self-help; DBT-AF= dialectical behavior therapy – appetite focused; CT= cognitive therapy; FBT = family-based therapy; BDI= beck depression inventory; HADS = hospital depression and anxiety scale; DASS= Depression, anxiety and stress scale; MADRS= Montgomery Asberg depression rating scale; In the
quality assessment column a + sign (low risk), a ‘?’ (unclear), or a - sign (high risk) is given for the four items of risk of bias: allocation sequence; concealment of allocation to conditions; blinding of assessors; and intention-to treat analyses. For Blinding of assessor we reported “SR” when only self-report outcome measures were used.
### Table 2
Subgroup analyses for the psychotherapy vs inactive comparison and the CBT vs active comparisons

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Subgroups</th>
<th>Depressive symptoms</th>
<th>Binge eating</th>
<th>Purging</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ncomp g (95% CI)</td>
<td>Qbp</td>
<td>Ncomp g (95% CI)</td>
<td>Qbp</td>
</tr>
<tr>
<td>Psychotherapy vs. inactive control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychotherapy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBT</td>
<td>8</td>
<td>0.52 (0.12, 0.92)</td>
<td>5</td>
<td>0.37 (0.02, 0.71)</td>
</tr>
<tr>
<td>Other psychotherapy</td>
<td>6</td>
<td>0.48 (0.03, 0.95)</td>
<td>2</td>
<td>1.41 (0.67, 2.15)</td>
</tr>
<tr>
<td>TL CBT only</td>
<td>5</td>
<td>0.92 (0.45, 1.37)</td>
<td>3</td>
<td>0.49 (0.12, 0.86)</td>
</tr>
<tr>
<td>TL other psychotherapy</td>
<td>4</td>
<td>0.70 (0.16, 1.24)</td>
<td>2</td>
<td>1.41 (0.67, 2.15)</td>
</tr>
<tr>
<td><strong>Modality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Therapist-led</td>
<td>8</td>
<td>0.78 (0.42, 1.15)</td>
<td>4</td>
<td>0.94 (0.41, 1.48)</td>
</tr>
<tr>
<td>Guided self-help</td>
<td>4</td>
<td>0.41 (-0.01, 0.83)</td>
<td>3</td>
<td>0.26 (-0.27, 0.79)</td>
</tr>
<tr>
<td>Pure self-help</td>
<td>2</td>
<td>-0.25 (-0.85, 0.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Concurrent med use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excluded</td>
<td>9</td>
<td>0.61 (0.23, 0.99)</td>
<td>5</td>
<td>0.61 (0.77, 1.14)</td>
</tr>
<tr>
<td>Not Excluded</td>
<td>5</td>
<td>0.33 (-0.14, 0.81)</td>
<td>2</td>
<td>0.64 (-0.19, 1.48)</td>
</tr>
<tr>
<td><strong>CBT vs. active control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Modality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Therapist-led</td>
<td>15</td>
<td>0.25 (0.03, 0.46)</td>
<td>15</td>
<td>0.33 (0.08, 0.58)</td>
</tr>
<tr>
<td>Guided self-help</td>
<td>2</td>
<td>0.05 (-0.57, 0.68)</td>
<td>3</td>
<td>0.07 (-0.58, 0.72)</td>
</tr>
<tr>
<td>Pure self-help</td>
<td>1</td>
<td>-0.52 (-1.30, 0.26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Concurrent med use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excluded</td>
<td>7</td>
<td>0.39 (0.04, 0.73)</td>
<td>10</td>
<td>0.25 (-0.01, 0.51)</td>
</tr>
</tbody>
</table>

Q: What is the Qp value for the comparison of Psychotherapy vs. inactive control, CBT only, and TL CBT only? 
A: 0.907, 0.008, 0.909

Q: What is the Qp value for the comparison of Psychotherapy vs. inactive control, Other psychotherapy, and TL other psychotherapy? 
A: 0.548, 0.012, 0.144

Q: What is the Qp value for the comparison of Psychotherapy vs. inactive control, Therapist-led, and Guided self-help? 
A: 0.016, 0.075, 0.003

Q: What is the Qp value for the comparison of Psychotherapy vs. inactive control, Excluded, and Not Excluded? 
A: 0.372, 0.955, 0.795

Q: What is the Qp value for the comparison of CBT vs. active control, Therapist-led, and Guided self-help? 
A: 0.171, 0.455, 0.542

Q: What is the Qp value for the comparison of CBT vs. active control, Excluded, and Not Excluded? 
A: 0.39 (0.04, 0.73), 10, 0.25 (-0.01, 0.51), 7, 0.23 (-0.14, 0.61)
| Not excluded | 11 | 0.07 (-0.17, 0.31) | 3 | 0.47 (-0.03, 0.97) | 10 | 0.14 (-0.14, 0.43) | .134 | .458 | .820 |

Note: \(N_{\text{comp}}\) = number of comparisons; \(Q_{bp}\) = p value test of whether the effect size for subgroups are significant; CBT = cognitive-behavioral therapy; TL = therapist-led; med = medication; bolded numbers signify statistical significance.
Figure 2:

Meta-regression of the relationship between binge eating and depressive symptom severity for the CBT vs active comparison

Note: Studies are represented by circles. The size of the circle is proportional to each study’s weight; larger circles are studies that have a smaller standard error and a smaller variance component.
Figure 3:

Meta-regression of the relationship between purging and depressive symptom severity for the CBT vs active comparison

Note: Studies are represented by circles. The size of the circle is proportional to each study’s weight; larger circles are studies that have a smaller standard error and a smaller variance component.