Psychological Medicine (2014), **44**, 2241–2254. © Cambridge University Press 2013
doi:10.1017/S0033291713002791

REVIEW ARTICLE

Group therapy for people with bulimia nervosa: systematic review and meta-analysis

A. Polnay[1,2]*, V. A. W. James[3], L. Hodges[2], G. D. Murray[4], C. Munro[5] and S. M. Lawrie[2,3,6]

[1] *Edinburgh Psychotherapy Department, Royal Edinburgh Hospital, Edinburgh, UK*
[2] *Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK*
[3] *Scottish Mental Health Research Network, Kennedy Tower, Royal Edinburgh Hospital, Edinburgh, UK*
[4] *Centre for Population Health Sciences, University of Edinburgh Medical School, Teviot Place, Edinburgh, UK*
[5] *Anorexia Nervosa Intensive Treatment Team, Royal Edinburgh Hospital, Edinburgh, UK*
[6] *MRF/MRC CRTF Programme for Mental Health ('PsySTAR'), Kennedy Tower, Royal Edinburgh Hospital, Edinburgh, UK*

Background. Approximately 25% of people with bulimia nervosa (BN) who undertake therapy are treated in groups. National guidelines do not discriminate between group and individual therapy, yet each has potential advantages and disadvantages and it is unclear how their effects compare. We therefore evaluated how group therapy for BN compares with individual therapy, no treatment, or other therapies, in terms of remission from binges and binge frequency.

Method. We performed a systematic review and meta-analysis of randomized controlled trials of group therapies for BN, following standard guidelines.

Results. A total of 10 studies were included. Studies were generally small with unclear risk of bias. There was low-quality evidence of a clinically relevant advantage for group cognitive behavioural therapy (CBT) over no treatment at therapy end. Remission was more likely with group CBT *versus* no treatment [relative risk (RR) 0.77, 95% confidence interval (CI) 0.62–0.96]. Mean weekly binges were lower with group CBT *versus* no treatment (2.9 *v.* 6.9, standardized mean difference = −0.56, 95% CI −0.96 to −0.15). One study provided low-quality evidence that group CBT was inferior compared with individual CBT to a clinically relevant degree for remission at therapy end (RR 1.24, 95% CI 1.03–1.50); there was insufficient evidence regarding frequency of binges.

Conclusions. Conclusions could only be reached for CBT. Low-quality evidence suggests that group CBT is effective compared with no treatment, but there was insufficient or very limited evidence about how group and individual CBT compared. The risk of bias and imprecise estimates of effect invite further research to refine and increase confidence in these findings.

Received 10 June 2013; Revised 15 October 2013; Accepted 15 October 2013; First published online 15 November 2013

Key words: Bulimia nervosa, cognitive behavioural therapy, group therapy, systematic review.

Background

Bulimia nervosa (BN) is characterized by binge eating, behaviour to prevent weight gain, and usually shame and secrecy about behaviour (Fairburn & Cooper, 1982; APA, 2000). The lifetime prevalence of BN is estimated at 2–4% (Favaro *et al.* 2003; Keski-Rahkonen *et al.* 2009). It is commonly associated with mental and physical problems, such as depression, hypokalaemia from vomiting, and gastrointestinal consequences.

Cognitive behavioural therapy (CBT) is the most established treatment (National Institute for Health and Care Excellence, 2004; Hay *et al.* 2009). Interpersonal therapy (IPT) is effective but has been reported to take longer to achieve the same effect (National Institute for Health and Care Excellence, 2004). Most people are treated individually, but a substantial proportion (estimated at 15–37%) are treated in groups (Newton *et al.* 1993; Rosenvinge & Klusmeier, 2000; Von Ranson & Robinson, 2006). An appealing aspect of group therapy for BN is that its direct cost is lower than individual therapy (Mitchell *et al.* 1999). It is not clear, however, how individual therapy and group therapy compare in terms of effectiveness. There are differences in processes and dynamics between group and individual therapy, even if the therapy modality is the same (Yalom & Leszcz, 2005). Indeed, in depression, individual CBT has been found to be more effective than group CBT at the end of therapy, but not at follow-up (Huntley *et al.* 2012).

* Address for correspondence: A. Polnay, Division of Psychiatry, University of Edinburgh, Kennedy Tower, Royal Edinburgh Hospital, Edinburgh EH10 5HF, UK.
(Email: adam.polnay@nhs.net)

One hypothesis is that an individualized approach for BN would be superior to group in terms of treatment effect, as individual therapy may allow for more specific formulation of the person's difficulties and for the therapy to be tailored to these (Morrison, 2001). On the other hand, group treatment for BN might be more effective than individual because of its potential to normalize shameful experiences (Lavender *et al.* 2012), and shorter waiting times (Morrison, 2001) which may improve motivation for treatment (Schmidt *et al.* 2008). One of the contentious issues in the literature is whether attrition rate is higher in the group setting (Mahon, 2000).

To our knowledge, no recent systematic review addresses how individual and group therapies compare, or how group therapies compare with no treatment. Neither the NICE guideline (National Institute for Health and Care Excellence, 2004) nor the Cochrane review for BN (Hay *et al.* 2009) discriminate between the individual and group settings. The Cochrane review excluded studies that compared individual and group therapy, noting that this comparison would entail a new review. NICE recommends CBT, or IPT as an alternative. CBT targets abnormal eating patterns and underlying thinking processes, and attempts to introduce more adaptive cognitions and behaviour. IPT focuses on interpersonal problems that trigger binge-eating. An older systematic review did discriminate between settings, but did not include any studies that directly compared individual with group therapies (Thompson-Brenner *et al.* 2003).

The main objective of this review is to assess the effects of group therapies for BN for out-patients with the condition. We compare the effects of group therapies with the same therapy delivered individually, no treatment, and with other therapies. Our primary measures of therapy effect are remission from binges and binge frequency. Given the uncertainty about whether attrition is higher from group therapy, we included drop-out from therapy as a secondary outcome measure.

Method

We conducted a systematic review and meta-analysis of randomized controlled trials (RCTs) of group psychological therapies for BN, following standard guidelines (Centre for Reviews and Dissemination, 2009; Higgins & Green, 2011).

Search strategy

Electronic searches were made of Medline, PsycINFO and EMBASE on 6 May 2011 and updated on 5 April 2013. We used the following search algorithm: (exp Bulimia Nervosa/ OR bulimia OR eating dis$) AND (exp Randomised Controlled Trials as Topic/ OR random$) AND (exp Psychotherapy, Group/ OR group tr$ OR group ses$ OR group therapy). We sourced additional papers from prior systematic reviews, reference lists of included papers, and the Cochrane Central Register of Controlled Trials.

Inclusion and exclusion criteria

These were specified *a priori*.

Types of studies

This review included studies if they were RCTs. We had no language restrictions.

Types of participants

We included studies if participants were aged 18 years or over and had BN [Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV; APA, 2000)]. For studies that did not use DSM-IV, we examined the participant characteristics and study inclusion criteria, and if participants appeared to correspond to DSM-IV BN the study was included.

If studies contained participants with other diagnoses, where possible we extracted data for participants with BN. If this was not possible then studies were excluded if >20% of the participants had other diagnoses. There were no restrictions on gender, comorbidity or medication use. Participants with binge-eating disorder or eating disorder not otherwise specified (EDNOS) were excluded.

Types of intervention

We included face-to-face psychological therapies for BN delivered in a group setting. In line with other reviews, we defined group treatment as consisting of three or more participants (Huntley *et al.* 2012). Studies of out-patient or community treatment were included; those of in-patients were excluded. As this review focused on face-to-face therapy, guided self-help was excluded.

Types of comparator arms

The comparator arms were the same therapy delivered individually, no treatment (including waiting list), or another therapy either in individual or group format.

Outcome measures

To allow comparison with NICE and Cochrane, we used similar measures.

https://doi.org/10.1017/S0033291713002791 Published online by Cambridge University Press

Main outcomes

(1) Remission from binge eating, defined as 100% cessation of binge eating. If studies only reported remission from binge eating and vomiting combined, this was used.
(2) Mean frequency of binges per week.

Secondary outcomes

(1) Mean score on a scale measuring depressive symptoms. If more than one measure was reported we used clinician-rated scales preferentially.
(2) Drop-out, defined as the ending of therapy by the patient. (Note this outcome is not measured for the second comparison, as treatment drop-out is not applicable to 'no treatment' arms.)

Time points

(1) End of therapy (primary time point).
(2) Longest follow-up reported (secondary time point).

Selection of studies and data extraction

One author (A.P.) read all the abstracts of papers found in the literature search. We then obtained all potentially eligible papers and two authors (A.P. and V.J.) reviewed them collaboratively against the inclusion and exclusion criteria. Any disagreements were resolved by discussion and a consensus reached. Reasons for exclusion were noted. We extracted data into Excel and later RevMan (Review Manager version 5.2; Nordic Cochrane Centre, Denmark), with two authors checking for accuracy of transcription. We attempted to contact the first author of studies where appropriate to ask for further outcome data. As well as outcomes, information extracted included: study author; year of publication; nature of sample; type of intervention and comparisons; sample size; number of sessions; and therapist level.

Risk of bias

We applied the Cochrane Collaboration's tool for assessing risk of bias (Higgins & Green, 2011) to each included study. This consists of six domains: sequence generation; allocation concealment; blinding; incomplete outcome data; selective outcome reporting; and other sources of bias. Each domain was rated as high, low, or unclear risk of bias. For clinical trials of therapies, blinding of therapists and participants to treatment allocation is clearly not possible, and in line with a previous meta-analysis (Leichsenring & Rabung, 2008) we did not judge a study as high risk of bias from inadequate blinding for this reason alone.

Dealing with missing data

We preferentially extracted intention-to-treat outcome data, where available, over 'per-protocol' or 'treatment-received' analyses. We described data as intention-to-treat when participants remained in the groups to which they had been randomized regardless of adherence to intervention or control; outcome measures were sought on all participants (and authors may have attempted to impute missing data). In accordance with the Cochrane guidelines, we did not impute data on missing individuals ourselves, but assessed the potential impact of incomplete outcome data in the risk of bias section.

Measures of treatment effect

For dichotomous outcomes relative risk (RR) was calculated, which is the proportion of adverse events in the intervention group divided by the proportion of adverse events in the comparison group. A RR of <1 indicates a favourable outcome for the intervention group. For continuous outcomes standardized mean difference (SMD) was calculated.

Assessment of heterogeneity

A meta-analysis was considered if there were two or more studies for a particular comparison. We assessed heterogeneity using the I^2 test, with a value <30% considered to indicate mild heterogeneity, 30–50% moderate heterogeneity and >50% substantial heterogeneity (Higgins & Green, 2011). If heterogeneity was mild, studies were combined in a meta-analysis. If moderate or substantial heterogeneity was detected, we examined the studies for possible differences in study participants, intervention (intensity and level of training of therapist) and study quality – on the basis of this a decision was made whether studies were similar enough to combine in a meta-analysis.

Studies with multiple arms

If within a particular meta-analysis there was a trial with multiple relevant arms, to prevent double-counting of participants in a common arm, if appropriate we collapsed groups to create, effectively, a two-arm trial. If this was not possible we planned to split the common group into two or more smaller groups, and entered two or more comparisons into that meta-analysis. For studies with multiple relevant arms within a particular meta-analysis, we assessed heterogeneity according to the Cochrane Handbook, section 16.5.5 (Higgins & Green, 2011).

https://doi.org/10.1017/S0033291713002791 Published online by Cambridge University Press

Pooling of data

Where meta-analyses were considered appropriate, a fixed-effects model was used if heterogeneity was mild, and a random-effects model used where heterogeneity was moderate or substantial.

Quality of evidence

Flowchart 2 from the NICE guideline (National Institute for Health and Care Excellence, 2004) was used to guide the interpretation of size and precision of effects. Where there was sufficient evidence to reach conclusions the quality of evidence was assessed using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system (Atkins *et al.* 2004) and tables created using GRADEpro (version 3.6).

Results

The literature search yielded 143 articles [for PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) study flow diagram, see Fig. 1]. After screening the abstracts 39 full-text articles were reviewed for eligibility. In all, 11 articles were included describing 10 studies involving 484 participants in total (Yates & Sambrailo, 1984; Freeman *et al.* 1985, 1988; Kirkley *et al.* 1985; Lee & Rush, 1986; Laessle *et al.* 1987; Leitenberg *et al.* 1988; Wolf & Crowther, 1992; Sundgot-Borgen *et al.* 2002; Chen *et al.* 2003; Nevonen & Broberg, 2006). Online Supplementary Table S1 lists excluded studies with rationale for exclusion.

Characteristics of included studies

Table 1 shows a summary of study characteristics.

Studies and participants

All studies were carried out in developed countries. Studies were published between 1984 and 2006, with six out of 10 published in the 1980s. Participants were recruited from out-patient clinics (Freeman *et al.* 1985, 1988; Laessle *et al.* 1987; Sundgot-Borgen *et al.* 2002; Nevonen & Broberg, 2006), via advertisements (Kirkley *et al.* 1985; Lee & Rush, 1986), out-patients and advertisements (Yates & Sambrailo, 1984; Leitenberg *et al.* 1988; Wolf & Crowther, 1992) and from out-patients and general practitioners (Chen *et al.* 2003). All participants were female. The mean age was 24.3 years (S.D.=5.8). Of 270 participants assigned to group therapy, 50 (18.5%) dropped out of treatment (range of drop-out was 0% to 37%). Of the studies, two studies did not report exclusion criteria (Lee & Rush, 1986; Laessle *et al.* 1987). For the remainder, typically patients were excluded with psychiatric co-morbidities (including drug/alcohol problems), high suicide risk, or were medically compromised. Also, three studies excluded participants if patients were taking medication (Sundgot-Borgen *et al.* 2002; Chen *et al.* 2003; Nevonen & Broberg, 2006).

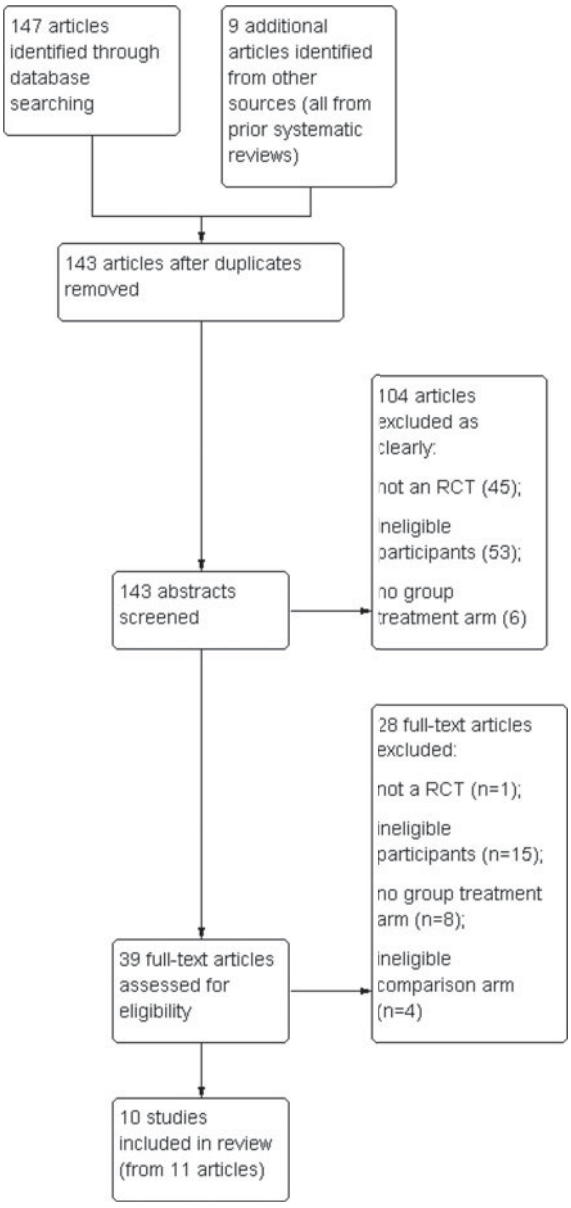


Fig. 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) study flow diagram. RCT, Randomized controlled trial.

Interventions

The mean number of sessions was 16.5 (S.D.=6.1, range 6–24), two studies were brief therapy (less than 12 sessions) (Yates & Sambrailo, 1984; Wolf & Crowther, 1992) and the remainder were medium term (12–24

https://doi.org/10.1017/S0033291713002791 Published online by Cambridge University Press

Table 1. *Characteristics of included studies*

Study (year)	Country, how recruited and mean age (S.D.)	Mean duration of eating disorder symptoms, years (S.D.)	Exclusion criteria	Interventions (*n*)	Therapist level and no. of sessions (duration)
Yates & Sambrailo (1984)	Australia, adverts and referrals, 27.9 years (5.8)	7.8 (5.4)	'Gross psychopathology'	Group CBT (12) Group CBT + ERP (12)	Not reported, six (1.5 h)
Kirkley *et al.* (1985)	USA, adverts, 27.9 years (S.D. not reported)	Age of onset of BN 17.9 years (S.D. not reported)	Drug/alcohol use, psychosis, suicidal	Group CBT (14) Group eclectic therapy – behavioural and non-directive component (14)	Doctoral-level clinical psychologists, 16 (1.5 h)
Lee & Rush (1986)	USA, adverts, 27.7 years (5.3)	Not reported	Not reported	Group CBT (15) Waiting list (15)	Doctoral student in psychology with 4 years of clinical experience, 12 (2 h)
Laessle *et al.* (1987)	Germany, out-patient clinics, 23.4 years (2.5)	6.2 (2.1)	Not reported	Group CBT (8) Waiting list (9)	2 years of clinical experience, 24 (duration not reported)
Leitenberg *et al.* (1988)	USA, adverts and referrals, 26.6 years (6.7)	6.94 (S.D. not reported)	Laxative use, alcohol problems, psychosis, high suicide risk	Group CBT (12) Group CBT + ERP (30) Waiting list (12)	Qualified therapists with >6 years of experience, or graduate students in psychology with >2 years of experience, 24 (2 h)
Freeman *et al.* (1985, 1988)	UK, out-patient clinics, 24.2 years (5.6)	6.0 (4.9)	Psychosis, current treatment	Group eclectic therapy – supportive with behavioural component (30) Individual CBT (32) Individual BT (30)	Clinical psychologist or nurse practitioner, 15 (1 h)
Wolf & Crowther (1992)	USA, out-patient clinic and adverts, 26.3 years (8.2)	5.7 (4.0)	Previous CBT, current treatment	Group CBT (15) Group BT (15) Waiting list (12)	Post-masters graduate student in clinical psychology with 4 years of clinical experience, 10 (2 h)
Sundgot-Borgen *et al.* (2002) Review includes unpublished data	Norway, out-patient clinics, 22.5 years (2.8)	5.7 (3.0)	Co-morbidity, use of medication	Group CBT (16) Group nutritional counselling (17) Waiting list (16)	CBT therapist for therapy, registered dietitian for nutritional counselling, 16 (2 h)
Chen *et al.* (2003)	USA, out-patient clinics and general practitioners, 25.8 years (7.2)	9.6 (7.3)	Co-morbid mental illness, medically compromised, use of medication	Group CBT (30) Individual CBT (30)	Clinical psychology graduate student with 2 years of CBT, 19 (1.5 h for group, 50 min for individual)
Nevonen & Broberg (2006)	Sweden, out-patients, 20.7 years (2.0)	4.8 (2.9)	Psychotic illness, drug or alcohol abuse, suicidal behaviour, use of medication	Group CBT then group IPT (44) Individual CBT then individual IPT (42)	Four senior therapists with long experience in eating disorders, 10 CBT then 13 IPT (group 2 h, individual 1 h)

S.D., Standard deviation; CBT, cognitive behavioural therapy; ERP, exposure with response prevention; BN, bulimia nervosa; BT, behavioural therapy; IPT, interpersonal therapy.

sessions). The therapists were qualified clinicians (Freeman *et al.* 1985, 1988; Kirkley *et al.* 1985; Sundgot-Borgen *et al.* 2002; Nevonen & Broberg, 2006), psychologists in training (Lee & Rush, 1986; Wolf & Crowther, 1992; Chen *et al.* 2003), a mixture of qualified clinicians and those in training (Leitenberg *et al.* 1988), or unclear (Yates & Sambrailo, 1984; Laessle *et al.* 1987).

Of the studies, three (Leitenberg *et al.* 1988; Wolf & Crowther, 1992; Sundgot-Borgen *et al.* 2002) contained three treatment arms and contributed to more than one comparison.

Risk of bias

Only two studies provided sufficient information to assess the adequacy of random sequence generation or allocation concealment (Kirkley *et al.* 1985; Chen *et al.* 2003). In one study it appeared that the outcome assessor for the primary outcome was not blind to treatment allocation (Chen *et al.* 2003). Three studies were judged as high risk of bias from incomplete outcome data (Yates & Sambrailo, 1984; Kirkley *et al.* 1985; Leitenberg *et al.* 1988). In terms of selective reporting, 80% of studies reported on all outcomes as specified in the methods, although we note that no study protocols were available. In one study the behavioural therapy (BT) arm had a higher frequency of binges at baseline compared with the other study arms (Wolf & Crowther, 1992). In another study the no-treatment arm had a lower frequency of binges at baseline compared with the other arms (Sundgot-Borgen *et al.* 2002). Overall, the level of bias across studies was unclear, introducing some doubt into the results. (See online Supplementary Fig. S1 for assessment of methodological quality table.)

Sample size

The sample size in the studies was low; one study had eight participants in the intervention arm, six studies had 10–19 participants per arm, two studies had 20–39 participants per arm and one study had 40–50 participants per arm. Freeman *et al.* (1988) was the only study to report a power calculation to inform sample size. Chen *et al.* (2003) based their sample size on guidelines but did not do a formal power calculation.

Comparison 1 – group therapy versus the same therapy delivered individually

There was one study of CBT (Chen *et al.* 2003), and one of sequenced CBT then IPT (CBT/IPT) (Nevonen & Broberg, 2006). Table 2 (section 2.1) shows the results and analyses.

Remission

For CBT there is limited evidence suggesting that there is a clinically relevant difference in remission from binge eating and vomiting at the end of therapy, with group CBT (30/30 non-remission) inferior to individual CBT (24/30 non-remission) [N=1, n=60; RR 1.24, 95% confidence interval (CI) 1.03–1.50]. At follow-up there is insufficient evidence to allow conclusions (N=1, n=60; RR 1.04, 95% CI 0.86–1.25).

For CBT/IPT there is insufficient evidence to determine if there is a clinically relevant difference in terms of remission from binge eating and vomiting between group (26/44 non-remission) and individual (29/42 non-remission) CBT/IPT at end of therapy (N=1, n=86; RR 0.86, 95% CI 0.62–1.18) or at follow-up.

Frequency of binges, depression symptoms and drop-out from treatment

There is insufficient evidence to determine if there is a clinically relevant difference in these outcomes between group and individual CBT, or between group and individual CBT/IPT (see Table 2, section 2.1).

Comparison 2 – group therapy versus no treatment

There were five studies of CBT (Lee & Rush, 1986; Laessle *et al.* 1987; Leitenberg *et al.* 1988; Wolf & Crowther, 1992; Sundgot-Borgen *et al.* 2002); one of BT (Wolf & Crowther, 1992); and one of nutritional counselling (Sundgot-Borgen *et al.* 2002). See Table 2 (section 2.2) for a summary of results, and Fig. 2 for forest plots of primary outcomes.

CBT

Remission. There is limited evidence suggesting that there is a clinically relevant difference in remission at the end of therapy, with group CBT (23/31 non-remission) superior to no treatment (30/31 non-remission) (N=2, n=62; RR 0.77, 95% CI 0.62–0.96).

Heterogeneity as measured by I^2 was 0%. At follow-up there is insufficient evidence to allow conclusions.

Frequency of binges. The baseline mean frequency of binges per week across both arms was 8.0 (S.D.=10.2; N=4, n=100). There is limited evidence suggesting that there is a clinically relevant difference in mean frequency of binges per week at the end of therapy, with group CBT (mean=2.9, S.D.=4.8) superior to no treatment (mean=6.9, S.D.=10.4) (N=4, n=98; SMD =−0.56, 95% CI −0.96 to −0.15, I^2=0%)[1]†. At follow-up there is insufficient evidence to allow conclusions.

† The notes appear after the main text.

https://doi.org/10.1017/S0033291713002791 Published online by Cambridge University Press

Table 2. *Summary of effects for comparisons 1 and 2*

Comparison (study reference)	No of studies	Intervention				Comparison				SMD	RR	(95% CI)
		Mean	(S.D.)[a]	Event	Subjects, *n*	Mean	(S.D.)[a]	Event	Subjects, *n*			
2.1 Group *versus* same therapy individually												
CBT (Chen *et al.* 2003)												
Remission end of therapy	1			30	30			24	30		1.24	(1.03 to 1.50)[b]
Remission follow-up	1			27	30			26	30		1.04	(0.86 to 1.25)
Frequency of binges end of therapy[c]	1	2.64	(4.46)		30	1.92	(3.22)		30	0.18		(−0.33 to 0.68)
Frequency of binges follow-up	1	2.4	(3.54)		30	2.62	(3.56)		30	−0.06		(−0.57 to 0.45)
Depression score end of therapy[d]	1	14.33	(10.6)		30	15.39	(11.91)		30	−0.09		(−0.60 to 0.41)
Depression score follow-up	1	13.37	(10.68)		30	16.7	(12.74)		30	−0.28		(−0.79 to 0.23)
Drop-out	1			8	30			8	30		1	(0.43 to 2.31)
CBT/IPT (Nevonen & Broberg, 2006)												
Remission end of therapy	1			26	44			29	42		0.86	(0.62 to 1.18)
Remission follow-up	1			32	44			26	42		1.17	(0.87 to 1.58)
Frequency of binges end of therapy	1	1.6	(2.2)		44	1.2	(1.5)		42	0.21		(−0.21 to 0.63)
Frequency of binges follow-up	1	2.1	(2.3)		44	1.3	(2.1)		42	0.36		(−0.07 to 0.79)
Depression score end of therapy	1	17	(14.5)		44	13	(11.6)		42	0.3		(−0.12 to 0.73)
Depression score follow-up	1	15	(14)		44	13	(10.5)		42	0.16		(−0.26 to 0.58)
Drop-out	1			5	44			3	42		1.59	(0.41 to 6.25)
2.2 Group therapy *versus* no treatment												
CBT												
Remission end of therapy (Lee & Rush, 1986; Sundgot-Borgen *et al.* 2002)	2			23	31			30	31		0.77	(0.62 to 0.96)[b]
Remission follow-up (Sundgot-Borgen *et al.* 2002)	1			14	16			16	16		0.88	(0.71 to 1.09)
Frequency of binges end of therapy (Lee & Rush, 1986; Laessle *et al.* 1987; Wolf & Crowther, 1992; Sundgot-Borgen *et al.* 2002)	4	2.9	(4.8)		51	6.9	(10.4)		47	−0.56		(−0.96 to −0.15)[b]
Frequency of binges follow-up (Sundgot-Borgen *et al.* 2002)	1	4.36	(3.37)		14	4.5	(2.33)	13		−0.05		(−0.80 to 0.71)
Depression score end of therapy (Lee & Rush, 1986; Leitenberg *et al.* 1988)	2	14.5	(8.53)		26	20.23	(13.39)		26	−0.52		(−1.08 to 0.03)
Behavioural therapy (Wolf & Crowther, 1992)												
Frequency of binges end of therapy	1	4.4	(6.75)		15	3.55	(2.3)		11	0.15		(−0.63 to 0.93)

Table 2 (*cont.*)

Comparison (study reference)	No of studies	Intervention				Comparison				SMD	RR	(95% CI)
		Mean	(S.D.)[a]	Event	Subjects, *n*	Mean	(S.D.)[a]	Event	Subjects, *n*			
Nutritional counselling (Sundgot-Borgen *et al.* 2002)												
Remission end of therapy	1			16	17			16	16		0.94	(0.80 to 1.11)
Remission follow-up	1			17	17			16	16		1	(0.89 to 1.12)
Frequency of binges end of therapy	1	4.9	(2.6)		17	4.9	(2.72)		13	0		(−0.72 to 0.72)
Frequency of binges follow-up	1	6.8	(3.67)		17	4.5	(2.33)		13	0.71		(−0.04 to 1.45)

S.D., Standard deviation; SMD, standardized mean difference; RR, relative risk; CI, confidence interval; CBT, cognitive behavioural therapy; IPT, interpersonal therapy.

[a] If more than one study per comparison, the combined mean and S.D. are reported.

[b] Limited evidence of a clinically relevant difference.

[c] Frequency of binges is per week.

[d] Depression symptoms: all studies used Beck Depression Inventory (BDI-I).

Depression symptoms. There is insufficient evidence to determine if there is a clinically relevant difference at the end of therapy between group CBT and no treatment in terms of depressive symptoms (N=2, n=52; SMD=−0.52, 95% CI −1.08 to 0.03, I^2=0%). This outcome was not reported for follow-up.

BT

There was insufficient evidence to allow conclusions about effect on frequency of binges at end of therapy. The one study did not report on other outcomes.

Nutritional counselling

There was insufficient evidence to allow conclusions regarding effect on remission from binge-eating or frequency of binges. The one study did not report depression symptoms.

Comparison 3 – group therapy versus another therapy

Group CBT was compared with the following group therapies: CBT+exposure with response prevention (ERP) (Yates & Sambrailo, 1984; Leitenberg *et al.* 1988); BT (Wolf & Crowther, 1992); eclectic therapy (involved behavioural and non-directive components) (Kirkley *et al.* 1985); nutritional counselling (Sundgot-Borgen *et al.* 2002). One study compared group eclectic therapy with individual CBT (Freeman *et al.* 1985, 1988).

For all comparisons and outcomes there was either insufficient evidence to reach conclusions, or the outcome was not reported. Online Supplementary Fig. S2 shows details of these results and forest plots of effects.

Quality of evidence

There was sufficient evidence to reach conclusions for group CBT *versus* individual CBT (remission), and group CBT *versus* no treatment (remission and frequency of binges).

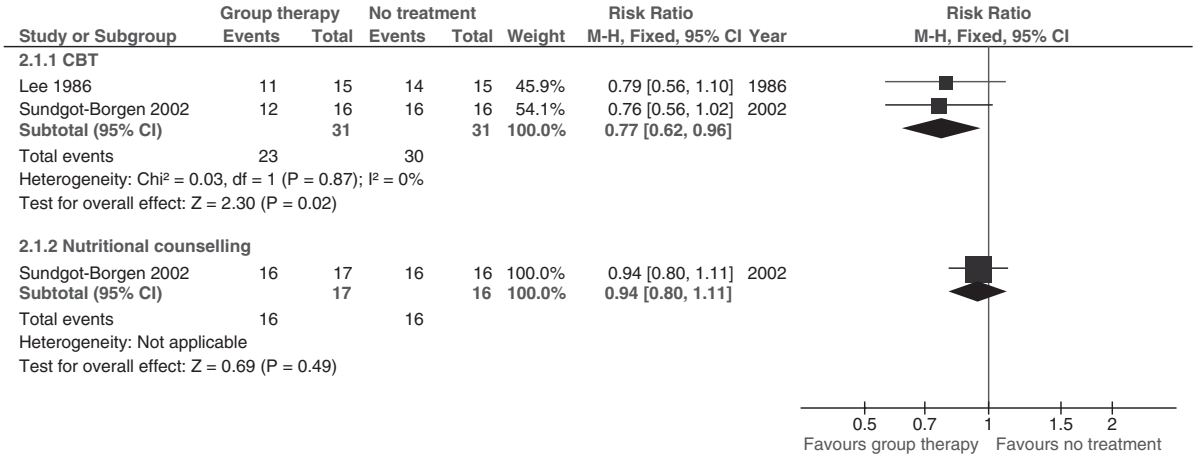
A strength of the evidence was that the participant characteristics were relevant to the research question. We were unable to assess risk of publication bias, as there were too few studies to carry out a funnel plot.

Group CBT versus individual CBT

A limitation of the evidence for this comparison lies in the imprecise estimate of effect. Furthermore, there is risk of bias, as it appeared that the outcome assessor was not blind to treatment allocation.

https://doi.org/10.1017/S0033291713002791 Published online by Cambridge University Press

(*a*) Non-remission from binge eating



(*b*) Frequency of binges per week

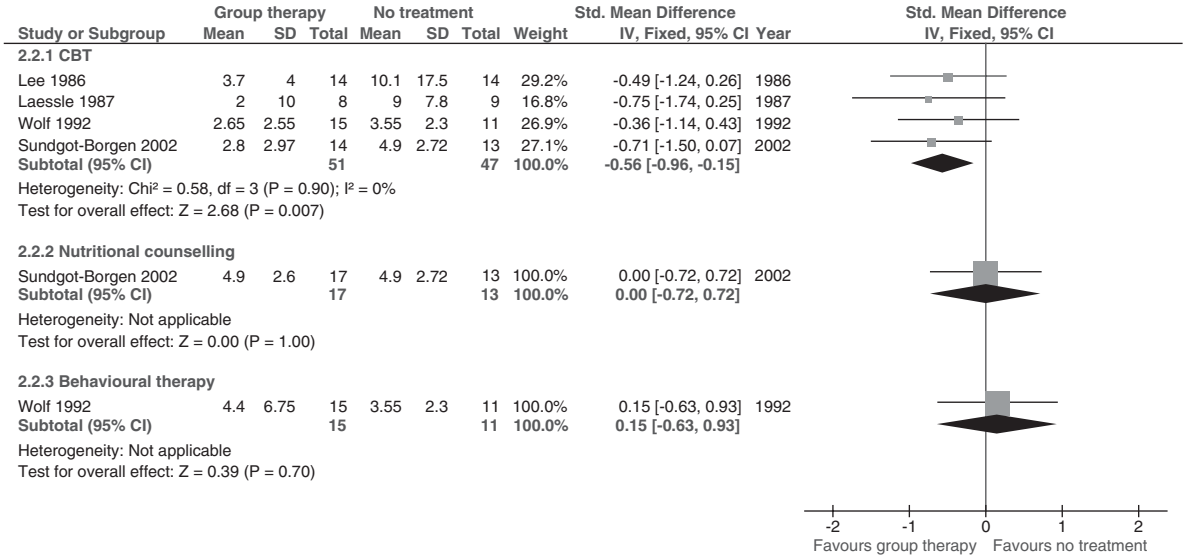


Fig. 2. Forest plots of primary outcomes for group therapy *versus* no treatment at end of therapy. (*a*) Non-remission from binge eating. (*b*) Frequency of binges per week. M-H, Mantel–Haenszel; CI, confidence interval; CBT, cognitive–behavioural therapy; df, degrees of freedom; SD, standard deviation; IV, inverse variance.

Group CBT versus no treatment

The results appeared consistent across studies. However, the randomization and allocation were inadequately described, and the estimate of effect was imprecise – these factors lower the quality of evidence.

Overall quality of evidence

Tables 3 and 4 show overall quality of evidence using the GRADE system, which was assessed as low for both comparisons.

Discussion

Main findings

This review addresses two questions in the treatment of BN: the efficacy of group therapy; and how individual and group therapies compare. The findings are summarized in Tables 3 and 4.

We found five studies comparing group CBT with no treatment. Low-quality evidence suggests a clinically relevant difference in favour of group CBT for both remission (23/30 did not remit with group CBT, 30/31 with no treatment; RR 0.77, 95% CI 0.62–0.96) and frequency of binges (mean binges per week

https://doi.org/10.1017/S0033291713002791 Published online by Cambridge University Press

Table 3. *GRADE profile: should group CBT versus individual CBT be used for bulimia nervosa?*

Quality assessment							Summary of findings				
							Study event rates (%)			Anticipated absolute effects	
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	With individual CBT	With group CBT	Relative effect (95% CI)	Risk with individual CBT	Risk difference with group CBT (95% CI)
Non-remission from binge eating											
60 (one study)	Serious[a]	No serious inconsistency	No serious indirectness	Serious[b]	Undetected	⊕⊕○○ Low[a,b] Due to risk of bias, imprecision	24/30 (80%)	30/30 (100%)	RR 1.24 (1.03 to 1.5)	800 per 1000	192 more per 1000 (24 more to 400 more)

GRADE, Grading of Recommendations, Assessment, Development and Evaluation; CBT, cognitive behavioural therapy; CI, confidence interval; RR, relative risk.

[a] It appears that the outcome assessor was not blind to treatment allocation.

[b] The CI includes an area of small clinical effect.

2.9 with group CBT, 6.9 with no treatment; SMD = −0.56, 95% CI −0.96 to −0.15) at the end of therapy. There was insufficient evidence at follow-up for further conclusions.

The situation for group CBT compared with individual CBT is less clear, with a single study providing low-quality evidence that individual CBT may be superior at end of therapy in terms of remission (30/30 did not remit with group CBT, 24/30 with individual therapy; RR 1.24, 95% CI 1.03–1.50). There was insufficient evidence to reach conclusions at follow-up, or for frequency of binges at any time point.

There was insufficient evidence to conclude how drop-out compared between individual and group therapy. However, we note that the drop-out rate of participants assigned to group therapy arms across all studies was 18.5%, which is fairly low.

For other therapies and comparisons there was either insufficient or no evidence on which to draw conclusions.

Strengths and limitations

The studies described here are relevant to clinical practice, as the patient characteristics are similar to those treated in groups in out-patient clinics, i.e. tending to be patients without significant suicide or medical risk, or significant co-morbidity. In addition, the experience of the therapist and the duration and intensity of the interventions correspond to clinical practice.

One limitation, however, lies in the inclusion of female patients only, given that 8–10% of people with BN are male (Bushnell *et al.* 1994; Garfinkel *et al.* 1995).

The main limitation is that studies of group therapy are small and CIs large, making it difficult to draw precise and, often, any conclusions. By contrast, when the individual and group settings are pooled, conclusions can be reached for CBT, IPT, focal supportive psychotherapy and CBT+ERP (National Institute for Health and Care Excellence, 2004).

The strengths of the review process are that we followed standard guidelines and applied rigorous methods for assessing bias and quality of evidence. Furthermore, there was no language restriction in the search and we approached authors directly for further outcome data. In terms of possible limitations, we did not assess the effect of the number of sessions on outcome. Second, we could not take into account any tendency for individuals treated within a single therapy group to have similar outcomes to each other. To overcome this it would be necessary to compare the variance between and within actual therapy groups.

Finally, the approach of comparing effects of therapy settings may be a simplification – there may be patients

https://doi.org/10.1017/S0033291713002791 Published online by Cambridge University Press

Table 4. *GRADE profile: should group CBT versus no treatment be used for bulimia nervosa?*

Quality assessment							Summary of findings				
							Study event rates (%)			Anticipated absolute effects	
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	With no treatment	With group CBT	Relative effect (95% CI)	Risk with no treatment	Risk difference with group CBT (95% CI)
Non-remission from binge eating											
63 (two studies)	Serious[a]	No serious inconsistency	No serious indirectness	Serious[b]	Undetected[c]	⊕⊕○○ Low[a,b,c] Due to risk of bias, imprecision	30/31 (96.8%)	23/31 (74.2%)	RR 0.77 (0.62 to 0.96)	968 per 1000	223 fewer per 1000 (39 fewer to 368 fewer)
Frequency of binges (better indicated by lower values)											
98 (four studies)	Serious[a,d]	No serious inconsistency	No serious indirectness	Serious[b]	Undetected[c]	⊕⊕○○ Low[a,b,c,d] Due to risk of bias, imprecision	N.A.	N.A.	–	The mean frequency of binges in the control groups was 6.9 binges per week	The mean frequency of binges in the intervention groups was 0.56 S.D. lower (0.96 to 0.15 lower)

GRADE, Grading of Recommendations, Assessment, Development and Evaluation; CBT, cognitive–behavioural therapy; CI, confidence interval; RR, relative risk; N.A., not applicable.

[a] The potential limitations across studies are likely to lower confidence in the estimate of effect.

[b] The CIs include an area of small clinical effect.

[c] There were insufficient studies to use a funnel plot, and publication bias cannot be ruled out.

[d] In Sundgot-Borgen *et al.* (2002) the control group had less severe frequency of binges at baseline – the effect of group CBT may have been underestimated in this study.

with BN for whom group treatment is likely to be more successful than individual, and vice versa (Polivy & Federoff, 1997).

Implications

Clinicians who offer or wish to offer group CBT for BN may feel a degree of reassurance from its superiority over no treatment. However, the uncertainty of its effect invites further research (see below).

The lack of clear evidence in BN about how group and individual CBT compare is shared across conditions as a whole (Tucker & Oei, 2007). The direct cost of group CBT for BN is one-fifth to two-fifths of that for individual CBT (Mitchell *et al.* 1999) – while this may appeal to those planning services, the relative efficacy needs to be clarified to establish overall how the two settings compare in terms of cost-effectiveness.

The Cochrane review (Hay *et al.* 2009), pooling individual and group settings, reported the effect of CBT compared with no treatment to be RR 0.69 (95% CI 0.61–0.79) for remission, and SMD = −0.94 (95% CI −1.19 to −0.70) for frequency of binges. These effects are somewhat larger than in our review. Interpreting indirect comparisons is problematic, as two sets of trials may differ in relevant aspects such as characteristics of participants and the intervention. Studies of individual therapy for BN tend to be more recent than those of group, and it has been suggested than improvements in therapy techniques could partly account for the observed higher effect in individual therapy studies (Keel, 2004). As it stands, however, the quality of evidence is stronger for individual CBT.

To resolve the question of whether the treatment setting of CBT matters would be a matter of conducting a further RCT of group *versus* individual CBT, which is adequately powered and has robust methodology to minimize risk of bias. Such a trial would measure outcomes at follow-up as well as end of therapy, to test whether group CBT might have a delayed effect relative to individual CBT. To widen its relevance men should be included, documenting reasons if they decline to participate or drop-out. Given the finding of a comparable effect for individual IPT and CBT at follow-up (National Institute for Health and Care Excellence, 2004), there would also be interest for future trials of group IPT compared with an established treatment.

Another avenue for research would be to evaluate if tailoring the treatment setting to the patient adds benefit. A RCT could be designed where in one of the treatment arms patients are assigned, through a process of assessment, to the setting that appears to best suit the patient (i.e. individual or group CBT); the comparison arm would assign all patients to the most established treatment, currently individual CBT.

Finally, the question of the efficacy of group CBT for BN needs to be viewed in context of developments in BN treatment as a whole. In particular, a large RCT (n = 293) found that a stepped-care sequence was more successful and cost-effective than individual CBT and fluoxetine at follow-up (Mitchell *et al.* 2011; Crow *et al.* 2013). In this stepped-care sequence, individual CBT was offered as a third stage if patients had ongoing bulimic symptoms after guided self-help and then fluoxetine. For services interested in delivering feasible and cost-effective treatment for BN, the strong evidence reported in support of this stepped-care approach puts the limited evidence that our review found for group CBT in perspective. This stepped-care research also raises the question of when might be the most beneficial stage to offer patients group CBT.

Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291713002791.

Acknowledgements

We would like to thank Professor Andrew Mcintosh, Professor of Biological Psychiatry, University of Edinburgh, for his contributions to the design of the article and helpful comments. The present review received no financial support.

Declaration of Interest

None.

Note

[1] One study (Laessle *et al.* 1987) provided data on baseline mean and S.D. of frequency of binges, and the mean and S.D. of change, but not end of therapy mean and S.D. End of therapy means were calculated from the data. We decided to conservatively assume that the S.D.s had not changed from baseline. This study was the smallest in this comparison and had the lowest weight assigned. Experimentally imputing a lower S.D. for the intervention group into the meta-analysis did not result in very much change at all to the pooled SMDs.

References

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, text revision. American Psychiatric Publishing: Washington, DC.

Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schünemann HJ, Edejer T, Varonen H, Vist GE, Williams JW Jr, Zaza S; GRADE Working Group (2004). Grading quality of evidence and strength of recommendations. *British Medical Journal (Clinical Research Ed.)* **328**, 1490.

Bushnell JA, Wells JE, McKenzie JM, Hornblow AR, Oakley-Browne MA, Joyce PR (1994). Bulimia comorbidity in the general population and in the clinic. *Psychological Medicine* **24**, 605–611.

Centre for Reviews and Dissemination (2009). *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. Centre for Reviews and Dissemination, University of York: York.

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; discussion 255–256.

Crow SJ, Agras WS, Halmi KA, Fairburn CG, Mitchell JE, Nyman JA (2013). A cost effectiveness analysis of stepped care treatment for bulimia nervosa. *International Journal of Eating Disorders* **46**, 302–307.

Fairburn CG, Cooper PJ (1982). Self-induced vomiting and bulimia nervosa: an undetected problem. *British Medical Journal (Clinical Research Edition)* **284**, 1153–1155.

Favaro A, Ferrara S, Santonastaso P (2003). The spectrum of eating disorders in young women: a prevalence study in a general population sample. *Psychosomatic Medicine* **65**, 701–708.

Freeman C, Sinclair F, Turnbull J, Annandale A (1985). Psychotherapy for bulimia: a controlled study. *Journal of Psychiatric Research* **19**, 373–378.

Freeman CP, Barry F, Dunkeld-Turnbull J, Henderson A (1988). Controlled trial of psychotherapy for bulimia nervosa. *British Medical Journal (Clinical Research Ed.)* **296**, 521–525.

Garfinkel PE, Lin E, Goering P, Spegg C, Goldbloom DS, Kennedy S, Kaplan AS, Woodside DB (1995). Bulimia nervosa in a Canadian community sample: prevalence and comparison of subgroups. *American Journal of Psychiatry* **152**, 1052–1058.

Hay PP, Bacaltchuk J, Stefano S, Kashyap P (2009). Psychological treatments for bulimia nervosa and binging. *Cochrane Database of Systematic Reviews*, Issue **4**. Art no.: CD000562. doi:10.1002/14651858.CD000562.pub3.

Higgins J, Green S (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011] (http://www.cochrane-handbook.org). Accessed October 2013.

Huntley AL, Araya R, Salisbury C (2012). Group psychological therapies for depression in the community: systematic review and meta-analysis. *British Journal of Psychiatry* **200**, 184–190.

Keel PK (2004). Review: individual psychotherapy may be an effective treatment for people with bulimia nervosa. *Evidence Based Mental Health* **7**, 11.

Keski-Rahkonen A, Hoek HW, Linna MS, Raevuori A, Sihvola E, Bulik CM, Rissanen A, Kaprio J (2009). Incidence and outcomes of bulimia nervosa: a nationwide population-based study. *Psychological Medicine* **39**, 823–831.

Kirkley BG, Schneider JA, Agras WS, Bachman JA (1985). Comparison of two group treatments for bulimia. *Journal of Consulting and Clinical Psychology* **53**, 43–48.

Laessle RG, Waadt S, Pirke KM (1987). A structured behaviorally oriented group treatment for bulimia nervosa. *Psychotherapy and Psychosomatics* **48**, 141–145.

Lavender A, Startup H, Naumann U, Samarawickrema N, Dejong H, Kenyon M, van den Eynde F, Schmidt U (2012). Emotional and social mind training: a randomised controlled trial of a new group-based treatment for bulimia nervosa. *PLoS ONE* **7**, e46047.

Lee NF, Rush AJ (1986). Cognitive–behavioral group therapy for bulimia. *International Journal of Eating Disorders* **5**, 599–615.

Leichsenring F, Rabung S (2008). Effectiveness of long-term psychodynamic psychotherapy: a meta-analysis. *Journal of the American Medical Association* **300**, 1551–1565.

Leitenberg H, Rosen JC, Gross J, Nudelman S, Vara LS (1988). Exposure plus response-prevention treatment of bulimia nervosa. *Journal of Consulting and Clinical Psychology* **56**, 535–541.

Mahon J (2000). Dropping out from psychological treatment for eating disorders: what are the issues? *European Eating Disorders Review* **8**, 198–216.

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**, 391–397.

Mitchell JE, Peterson CB, Agras S (1999). Cost-effectiveness of psychotherapy for eating disorders. In *Cost-Effectiveness of Psychotherapy: A Guide for Practitioners, Researchers and Policy Makers* (ed. N. E. Miller), pp. 270–278. Oxford University Press: New York.

Morrison N (2001). Group cognitive therapy: treatment of choice or sub-optimal option? *Behavioural and Cognitive Psychotherapy* **29**, 311–332.

National Institute for Health and Care Excellence (2004). *CG9 Eating Disorders: NICE Guideline*. NICE: London.

Nevonen L, Broberg AG (2006). A comparison of sequenced individual and group psychotherapy for patients with bulimia nervosa. *International Journal of Eating Disorders* **39**, 117–127.

Newton JT, Robinson PH, Hartley P (1993). Treatment for eating disorders in the United Kingdom. Part II. Experiences of treatment: a survey of members of the Eating Disorders Association. *European Eating Disorders Review* **1**, 10–21.

Polivy J, Federoff I (1997). Group psychotherapy. In *Handbook of Treatment for Eating Disorders*, 2nd edn (ed D. Garner and P. Garfinkel), pp. 462–475. Guilford Press: New York.

Rosenvinge JH, Klusmeier AK (2000). Treatment for eating disorders from a patient satisfaction perspective: a Norwegian replication of a British study. *European Eating Disorders Review* **8**, 293–300.

Schmidt U, Andiappan M, Grover M, Robinson S, Perkins S, Dugmore O, Treasure J, Landau S, Eisler I, Williams C (2008). Randomised controlled trial of CD-ROM-based cognitive–behavioural self-care for bulimia nervosa. *British Journal of Psychiatry* **193**, 493–500.

Sundgot-Borgen J, Rosenvinge JH, Bahr R, Schneider LS (2002). The effect of exercise, cognitive therapy, and nutritional counseling in treating bulimia nervosa. *Medicine and Science in Sports and Exercise* **34**, 190–195.

Thompson-Brenner H, Glass S, Westen D (2003). A multidimensional meta-analysis of psychotherapy for bulimia nervosa. *Clinical Psychology: Science and Practice* **10**, 269–287.

Tucker M, Oei TPS (2007). Is group more cost effective than individual cognitive behaviour therapy? The evidence is not solid yet. *Behavioural and Cognitive Psychotherapy* **35**, 77–91.

Von Ranson KM, Robinson KE (2006). Who is providing what type of psychotherapy to eating disorder clients? A survey. *International Journal of Eating Disorders* **39**, 27–34.

Wolf EM, Crowther JH (1992). An evaluation of behavioral and cognitive–behavioral group interventions for the treatment of bulimia nervosa in women. *International Journal of Eating Disorders* **11**, 3–15.

Yalom ID, Leszcz M (2005). The therapist: working in the here-and-now. In *Theory and Practice of Group Psychotherapy*, 5th revised edn (ed. I. D. Yalom and M. Leszcz), pp. 172–173. Basic Books: New York.

Yates AJ, Sambrailo F (1984). Bulimia nervosa: a descriptive and therapeutic study. *Behaviour Research and Therapy* **22**, 503–517.