The American Psychiatric Association Practice Guidelines for the Treatment of Patients with Eating Disorders: Appendices

Table of Contents

Appendix A. Clinical Questions

Assessment Related Questions

As part of the initial assessment of adolescents and adults who present with a possible eating disorder:

- What questions related to food, eating patterns, and nutritional status are most important to ask?
- What questions related to motivation for treatment are most important to ask, if any?
- What questions related to co-occurring psychiatric disorders are most important to ask, if any?
- What questions related to co-occurring physical disorders or symptoms are most important to ask, if any?
- What other questions related to assessment are most important to ask, if any?
- What psychometric measures, clinician-administered rating scales, or self-report questionnaires are important to administer, if any?
- What laboratory tests or physiological measures are important to obtain, if any?
- Do any of the recommended assessments differ depending on the possible eating disorder diagnosis?

Determination of a Setting of Care

For adolescents and adults with a diagnosis of an eating disorder, what factors suggest the need for a higher level of care such as:

- Inpatient medical setting
- Acute inpatient psychiatric setting
- Longer-term inpatient psychiatric setting
- Intensive outpatient treatment program or day hospital program
- Other specialized eating disorders programs

Do any of the recommended factors that suggest a need for a higher level of care depend on the specific eating disorder diagnosis?

Refeeding Phase of Treatment

- What is the evidence for the effectiveness of treatments (including refeeding approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with AN who require refeeding?
- What is the evidence for harms associated with treatments (including refeeding approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with AN who require refeeding?
- Does the effectiveness of treatments for adolescents and adults with AN who require refeeding differ by age, sex, race, ethnicity, initial BMI, illness severity/chronicity (e.g., severe and enduring AN), coexisting conditions, or intensity of treatment setting?
- What is the appropriate target weight gain on a weekly basis for adolescents and adults with AN who require refeeding? Does the appropriate target weight gain vary by age, sex, race, ethnicity, initial BMI, illness severity/chronicity (e.g., severe and enduring AN), coexisting conditions, or intensity of treatment setting?
- What aspects of the physical examination are important to assess, if any, in adolescents and adults with AN who require refeeding? Do the necessary elements of physical examination vary

by age, sex, race, ethnicity, initial BMI, illness severity/chronicity (e.g., severe and enduring AN), coexisting conditions, or intensity of treatment setting?
- What laboratory tests or physiological measures are important to obtain, if any, in adolescents and adults with AN who require refeeding? Do the necessary laboratory tests or physiological measures vary by age, sex, race, ethnicity, initial BMI, illness severity/chronicity (e.g., severe and enduring AN), coexisting conditions, or intensity of treatment setting?

Treatment Once Malnutrition is Addressed

- What is the evidence for the effectiveness of treatments (including nutritional rehabilitation approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with AN once malnutrition has been addressed?
- What is the evidence for harms associated with treatments (including nutritional rehabilitation approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with AN once malnutrition has been addressed?
- Does the effectiveness of treatments for AN differ by age, sex, race, ethnicity, initial BMI, illness severity/chronicity (e.g., severe and enduring AN), coexisting conditions, or intensity of treatment setting?

Treatment to Address Bone Density Loss

- What is the evidence for the effectiveness of treatments (including hormonal therapy, bisphosphonates, and other interventions) alone or in combination to improve or prevent further deterioration in bone density for adolescents and adults with AN who have at least 6 months of amenorrhea?
- What is the evidence for the harms of treatments (including hormonal therapy, bisphosphonates, and other interventions) alone or in combination to improve or prevent further deterioration in bone density for adolescents and adults with AN who have at least 6 months of amenorrhea?
- Does the effectiveness of treatments for improving or preventing further deterioration in bone density differ by age, sex, race, ethnicity, initial BMI, illness severity/chronicity (e.g., severe and enduring AN), or coexisting conditions?

Bulimia Nervosa

- What is the evidence for the effectiveness of treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with BN?
- What is the evidence for harms associated with treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with BN?
- Does the effectiveness of treatments for BN differ by age, sex, race, ethnicity, initial BMI, illness severity/chronicity (e.g., multi-impulsive BN), or coexisting conditions?

Binge-Eating Disorder

- What is the evidence for the effectiveness of treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with BED?

- What is the evidence for harms associated with treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with BED?
- Does the effectiveness of treatments for BED differ by age, sex, race, ethnicity, initial BMI, illness severity/chronicity, or coexisting conditions?

Night Eating Syndrome

- What is the evidence for the effectiveness of treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with night eating syndrome?
- What is the evidence for harms associated with treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with night eating syndrome?
- Does the effectiveness of treatments for night eating syndrome differ by age, sex, race, ethnicity, initial BMI, illness severity/chronicity, or coexisting conditions?

Avoidant/Restrictive Food Intake Disorder

- What is the evidence for the effectiveness of treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with ARFID?
- What is the evidence for harms associated with treatments (including nutritional approaches, pharmacotherapy, psychotherapy, and other psychosocial interventions) alone or in combination for adolescents and adults with ARFID?
- Does the effectiveness of treatments for ARFID differ by age, sex, race, ethnicity, initial BMI, illness severity/chronicity, or coexisting conditions?

Appendix B. Search Strategies, Study Selection, Search Results, and Analytic Methods

This guideline is based on a systematic search of available research evidence conducted by APA staff, extraction of detailed information on included studies by Dr. Evidence (Santa Monica, CA) using the DOC Data 2.0 software platform, and network meta-analyses conducted by Heno Analytics (Vancouver, BC, Canada). The systematic search of available research evidence used MEDLINE (PubMed), Cochrane Library (Wiley), and PsycINFO (EBSCO) databases (see Tables B-1 through B-3). Results were limited to English-language, human-only studies that were clinical trials, observational studies, systematic reviews, or meta-analyses. Case reports, comments, editorials, and letters were excluded. Citations to registry links, abstracts, and proceedings were not included, unless also published in a peer-reviewed journal, because they did not include sufficient information to evaluate the risk of bias of the study. Searches covered the period from the start of each database to July 15, 2019. One search in each database was done for eating disorders in general, which also included AN, BN, and BED. Separate searches in each database were done for ARFID (see Tables B-4 through B-6). For the topic of bone density in AN, searches were done in MEDLINE (PubMed) and Cochrane Library (Wiley) (see Tables B-7 and B-8) because the topic was not specific to psychology literature. Updated searches were conducted using the same criteria for the period from January 1, 2019 to October 1, 2021 to assure that more recent evidence was incorporated into the guideline.

Table B-1. MEDLINE (PubMed) Search Strategy for AN, BN, and BED

Search ID#	Query	Search date: 07/15/2019	Search date: 10/01/2021
#1	Search (("anorexia"[MH] OR "anorexic"[TIAB] OR "anorexia"[TIAB]) AND "nervosa"[TIAB]) OR "anorexia nervosa"[TIAB] OR "anorexia nervosa"[MH] OR "bulimia"[MH] OR "bulimia nervosa"[MH] OR "bulimia"[TIAB] OR "bulimic"[TIAB] OR "binging"[TIAB] OR "purging"[TIAB] OR "binge eating"[TIAB] OR "binge eating disorder"[TIAB] OR "binge eating disorder"[MH] OR "eating disorder"[TIAB] OR "eating disorders"[TIAB]	35,365	40,241
#2	Search "randomized controlled trial"[PT] OR "randomisation"[TIAB] OR "randomised"[TIAB] OR "randomization"[TIAB] OR "randomized"[TIAB] OR "randomly"[TIAB] OR "placebo"[TIAB] OR "sham"[TIAB]	1,086,783	1,258,934
#3	Search "meta analysis as topic"[MeSH Major Topic] OR "meta analysis as topic"[MeSH Terms] OR "meta analysis"[TIAB] OR "meta analyses"[TIAB] OR "meta analytic"[TIAB] OR "metaanalysis"[TIAB] OR "metaanalyses"[TIAB] OR "systematic review"[TIAB] OR "systematic reviews"[TIAB] OR "meta analysis"[PT]	261,181	358,062
#4	Search "controlled clinical trial"[PT] OR "blinded" [TIAB] OR "case control" [TIAB] OR "clinical trial" [TIAB] OR "clinical trials" [TIAB] OR "Cohort Analysis" [TIAB] OR "cohort	2,312,245	2,734,741

	research" [TIAB] OR "cohort study" [TIAB] OR "cohort trial" [TIAB] OR "comparator group" [TIAB] OR "controlled studies" [TIAB] OR "controlled study" [TIAB] OR "controlled trial" [TIAB] OR "controlled trials" [TIAB] OR "double blind" [TIAB] OR "followup" [TIAB] OR "follow up" [TIAB] OR "longitudinal research" [TIAB] OR "longitudinal study" [TIAB] OR "longitudinal trial" [TIAB] OR "multicenter trial" [TIAB] OR "multicenter trials" [TIAB] OR "naturalistic research" [TIAB] OR "naturalistic study" [TIAB] OR "naturalistic trial" [TIAB] OR "prospective cohort" [TIAB] OR "prospective research" [TIAB] OR "prospective study" [TIAB] OR "prospective trial" [TIAB] OR "retrospective cohort" [TIAB] OR "retrospective research" [TIAB] OR "retrospective study" [TIAB] OR "retrospective trial" [TIAB] OR "single blind" [TIAB]		
#5	Search ("case reports"[PT] OR "comment"[PT] OR "editorial"[PT] OR "letter"[PT])	3,569,055	3,990,385
#6	Search (("animals"[MAJR] OR "animals"[MH] OR "animal"[TIAB] OR "animals"[TIAB] OR "rat"[TIAB] OR "mouse"[TIAB] OR "mice"[TIAB] OR "rodent"[TIAB] OR "rodents"[TIAB] OR "rats"[TIAB]) NOT ("humans"[MAJR] OR "humans"[MH] OR "human"[TIAB] OR "humans"[TIAB]))	4,562,314	4,858,349
#7	Search #1 AND (#2 OR #3 OR #4)	6,133	7,513
#8	Search #7 NOT #5	5,919	7,247
#9	Search #8 NOT #6	5,874	7,187
#10	Search (#9) AND "english"[Language]	5,588	6,891
#11	Search (("1960/01/01"[Date - Publication] : "2019/07/15"[Date - Publication])) AND #10	5,583	
#11	Search (("2019/01/01"[Date - Publication] : "2021/10/01"[Date - Publication])) AND #10		1,676

Table B-2: Cochrane Library (Wiley) Search Strategy for AN, BN, and BED

Search ID#	Search	Search date: 07/31/2019	Search date: 10/01/2021
#1	"anorexia nervosa":ti,ab,kw (Word variations have been searched)	1,000	1332
#2	"bulimia":ti,ab,kw (Word variations have been searched)	1,259	
#3	"binge eating":ti,ab,kw (Word variations have been searched)	998	
#4	"binging":ti,ab,kw (Word variations have been searched)	116	
#5	"purging":ti,ab,kw (Word variations have been searched)	368	
#6	"eating disorder":ti,ab,kw (Word variations have been searched)	1,741	1149
#7	"eating disorders":ti,ab,kw (Word variations have been searched)	1,524	
#8	MeSH descriptor: [Feeding and Eating Disorders] explode all trees	1,408	1737

#9	MeSH descriptor: [Anorexia Nervosa] explode all trees	475	557
#10	MeSH descriptor: [Bulimia] explode all trees	460	542
#11	MeSH descriptor: [Bulimia Nervosa] explode all trees	237	279
#12	MeSH descriptor: [Binge-Eating Disorder] explode all trees	206	310
#13	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12	3,837 (limited to trials and reviews)	1874 (limited to trials and reviews from Jan 2019 through Sept 2021)

Table B-3: PsycINFO (EBSCO) Search Strategy for AN, BN, and BED

Search ID#	Search	Search date: 07/31/2019	Search date: 10/01/2021
S1	DE "Eating Disorders" OR MH "Eating Disorders+" OR TI ("anorexia nervosa" OR "binge eating disorder" OR "binge eating disorders" OR "binge eating" OR "binging" OR "Bulimia Nervosa" OR "bulimia" OR "eating disorder" OR "eating disorders" OR "purging") OR AB ("anorexia nervosa" OR "binge eating disorder" OR "binge eating disorders" OR "binge eating" OR "binging" OR "Bulimia Nervosa" OR "bulimia" OR "eating disorder" OR "eating disorders" OR "purging") OR SU ("anorexia nervosa" OR "binge eating disorder" OR "binge eating disorders" OR "binge eating" OR "binging" OR "Bulimia Nervosa" OR "bulimia" OR "eating disorder" OR "eating disorders" OR "purging") OR KW ("anorexia nervosa" OR "binge eating disorder" OR "binge eating disorders" OR "binge eating" OR "binging" OR "Bulimia Nervosa" OR "bulimia" OR "eating disorder" OR "eating disorders" OR "purging") OR ((TI ("anorexic" OR "anorexia") OR AB ("anorexic" OR "anorexia") OR SU ("anorexic" OR "anorexia") OR KW ("anorexic" OR "anorexia")) AND (TI "nervosa" OR AB "nervosa" OR SU "nervosa" OR KW "nervosa"))	37,991	38,188
S2	DE "Between Groups Design" OR MH "Randomized Controlled Trials as Topic+" OR MM "Random Allocation" OR MM "Randomized Controlled Trial" OR TI ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR "randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham	228,168	238,681

	group" OR "sham") OR AB ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR "randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham group" OR "sham") OR SU ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR "randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham group" OR "sham") OR KW ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR "randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham group" OR "sham")		
S3	DE "Meta analysis"OR MM "Meta-Analysis as Topic" OR TI ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews") OR AB ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews") OR SU ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews") OR KW ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews")	52,978	65,026
S4	DE "Clinical Trials" OR DE "Cohort Analysis" OR DE "Followup Studies" OR DE "Longitudinal Studies" OR DE "Prospective Studies" OR DE "Retrospective Studies" OR	435,725	468,784

	MH "Case-Control Studies+" OR MH "Clinical Trials as Topic+" OR MH "Cohort Studies+" OR MM "Clinical Trials" OR MR "followup study" OR MR "longitudinal study" OR MR "retrospective study" OR MR "Treatment Outcome/Clinical Trial" OR TI ("blinded" OR "case control" OR "clinical trial" OR "clinical trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR "retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial") OR AB ("blinded" OR "case control" OR "clinical trial" OR "clinical trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR "retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial") OR SU ("blinded" OR "case control" OR "clinical trial" OR "clinical trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR "retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial") OR KW ("blinded" OR "case control" OR "clinical trial" OR "clinical		

	trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR "retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial")		
S5	TI ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter") OR AB ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter") OR SU ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter") OR KW ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter")	172,910	179,188
S6	(DE "Vertebrates" OR DE "Mammals" OR DE "Animals" OR DE "Rats" OR DE "Rodents" OR DE "Mice" OR TI "animals" OR TI "animal" OR TI "mouse" OR TI "mice" OR TI "rodent" OR TI "rodents" OR TI "rat" OR TI "rats" OR SU "animals" OR SU "animal" OR SU "mouse" OR SU "mice" OR SU "rodent" OR SU "rodents" OR SU "rat" OR SU "rats" OR KW "animals" OR KW "animal" OR KW "mouse" OR KW "mice" OR KW "rodent" OR KW "rodents" OR KW "rat" OR KW "rats" OR AB "animals" OR AB "animal" OR AB "mouse" OR AB "mice" OR AB "rodent" OR AB "rodents" OR AB "rat" OR AB "rats") NOT (PO "human" OR TI "humans" OR TI "human" OR AB "humans" OR AB "human" OR SU "humans" OR SU "human" OR KW "humans" OR KW "human")	300,133	299,102
S7	S1 AND (S2 OR S3 OR S4)	7,132	7,795
S8	S7 NOT S5	6,919	7,546
S9	S8 NOT S6	6,875	7,499
S10	S9 AND LA English	6,548	7,140
S11	S10 Limiters - Published Date: 19600101-20190731	6,544	
S11	S10 Limiters - Published Date: 20190101-20211001		999

Table B-4. MEDLINE (PubMed) Search Strategy for ARFID

Search ID#	Search	Search date: 07/15/2019	Search date: 10/01/2021

#1	Search (("avoidant restrictive food intake disorder") OR "selective eating") OR "arfid"	217	407
#2	Search "randomized controlled trial"[PT] OR "randomisation"[TIAB] OR "randomised"[TIAB] OR "randomization"[TIAB] OR "randomized"[TIAB] OR "randomly"[TIAB] OR "placebo"[TIAB] OR "sham"[TIAB]	1,086,783	1,258,934
#3	Search "meta analysis as topic"[MeSH Major Topic] OR "meta analysis as topic"[MeSH Terms] OR "meta analysis"[TIAB] OR "meta analyses"[TIAB] OR "meta analytic"[TIAB] OR "metaanalysis"[TIAB] OR "metaanalyses"[TIAB] OR "systematic review"[TIAB] OR "systematic reviews"[TIAB] OR "meta analysis"[PT]	261,181	358,062
#4	Search "controlled clinical trial"[PT] OR "blinded" [TIAB] OR "case control" [TIAB] OR "clinical trial" [TIAB] OR "clinical trials" [TIAB] OR "Cohort Analysis" [TIAB] OR "cohort research" [TIAB] OR "cohort study" [TIAB] OR "cohort trial" [TIAB] OR "comparator group" [TIAB] OR "controlled studies" [TIAB] OR "controlled study" [TIAB] OR "controlled trial" [TIAB] OR "controlled trials" [TIAB] OR "double blind" [TIAB] OR "followup" [TIAB] OR "follow up" [TIAB] OR "longitudinal research" [TIAB] OR "longitudinal study" [TIAB] OR "longitudinal trial" [TIAB] OR "multicenter trial" [TIAB] OR "multicenter trials" [TIAB] OR "naturalistic research" [TIAB] OR "naturalistic study" [TIAB] OR "naturalistic trial" [TIAB] OR "prospective cohort" [TIAB] OR "prospective research" [TIAB] OR "prospective study" [TIAB] OR "prospective trial" [TIAB] OR "retrospective cohort" [TIAB] OR "retrospective research" [TIAB] OR "retrospective study" [TIAB] OR "retrospective trial" [TIAB] OR "single blind" [TIAB]	2,312,245	2,734,741
#5	Search ("case reports"[PT] OR "comment"[PT] OR "editorial"[PT] OR "letter"[PT])	3,569,055	3,990,385
#6	Search (("animals"[MAJR] OR "animals"[MH] OR "animal"[TIAB] OR "animals"[TIAB] OR "rat"[TIAB] OR "mouse"[TIAB] OR "mice"[TIAB] OR "rodent"[TIAB] OR "rodents"[TIAB] OR "rats"[TIAB]) NOT ("humans"[MAJR] OR "humans"[MH] OR "human"[TIAB] OR "humans"[TIAB]))	4,562,314	4,858,349
#7	Search #1 AND (#2 OR #3 OR #4)	41	77
#8	Search #7 NOT #5	38	70
#9	Search #8 NOT #6	37	69
#10	Search (#9) AND "english"[Language]	37	69
#11	Search (("1960/01/01"[Date - Publication] : "2019/07/15"[Date - Publication])) AND #10	37	
#11	Search (("2019/01/01"[Date - Publication] : "2021/10/01"[Date - Publication])) AND #10		43

Table B-5: Cochrane Library (Wiley) Search Strategy for ARFID

Search ID#	Search	Search date: 07/31/2019	Search date: 10/01/2021
#1	(“ARFID”) ti, ab, kw	6	25
#2	(“avoidant restrictive”) ti, ab, kw	8	
#3	(“selective eating”) ti, ab, kw	5	
#4	#1 OR #2 OR #3	13 (limited to trials and reviews)	25 (limited to trials and reviews from Jan 2019 through Sept 2021)

Table B-6: PsycINFO (EBSCO) Search Strategy for ARFID

Search ID#	Search	Search date: 07/31/2019	Search date: 10/01/2021
S1	TI "selective eating" OR AB "selective eating" OR KW "selective eating" OR SU "selective eating" OR TI "ARFID" OR AB "ARFID" OR KW "ARFID" OR SU "ARFID" OR TI "avoidant restrictive" OR AB "avoidant restrictive" OR KW "avoidant restrictive" OR SU "avoidant restrictive"	182	314
S2	DE "Between Groups Design" OR MH "Randomized Controlled Trials as Topic+" OR MM "Random Allocation" OR MM "Randomized Controlled Trial" OR TI ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR "randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham group" OR "sham") OR AB ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR "randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham group" OR "sham") OR SU ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR	228,168	238,681

	"randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham group" OR "sham") OR KW ("controlled clinical trial" OR "multicenter trial" OR "multicenter trials" OR "placebo" OR "random allocation" OR "random assignment" OR "random" OR "randomisation" OR "randomised assignment" OR "randomised controlled" OR "randomised" OR "randomization" OR "randomized assignment" OR "randomized controlled" OR "randomized" OR "randomly allocated" OR "randomly assigned" OR "randomly" OR "sham control" OR "sham group" OR "sham")		
S3	DE "Meta analysis"OR MM "Meta-Analysis as Topic" OR TI ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews") OR AB ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews") OR SU ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews") OR KW ("meta analyses" OR "meta analysis" OR "metaanalyses" OR "meta-analyses" OR "metaanalysis" OR "meta-analysis" OR "metaanalytic" OR "meta-analytic" OR "systematic review" OR "systematic reviews")	52,978	65,026
S4	DE "Clinical Trials" OR DE "Cohort Analysis" OR DE "Followup Studies" OR DE "Longitudinal Studies" OR DE "Prospective Studies" OR DE "Retrospective Studies" OR MH "Case-Control Studies+" OR MH "Clinical Trials as Topic+" OR MH "Cohort Studies+" OR MM "Clinical Trials" OR MR "followup study" OR MR "longitudinal study" OR MR "retrospective study" OR MR "Treatment Outcome/Clinical Trial" OR TI ("blinded" OR "case control" OR "clinical trial" OR "clinical trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR	435,725	468,784

	"retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial") OR AB ("blinded" OR "case control" OR "clinical trial" OR "clinical trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR "retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial") OR SU ("blinded" OR "case control" OR "clinical trial" OR "clinical trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR "retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial") OR KW ("blinded" OR "case control" OR "clinical trial" OR "clinical trials" OR "Cohort Analysis" OR "cohort research" OR "Cohort Studies" OR "cohort study" OR "cohort trial" OR "comparator group" OR "controlled clinical trial" OR "controlled studies" OR "controlled study" OR "controlled trial" OR "controlled trials" OR "double blind" OR "followup research" OR "followup study" OR "followup trial" OR "longitudinal research" OR "longitudinal study" OR "longitudinal trial" OR "naturalistic research" OR "naturalistic study" OR "naturalistic trial" OR "prospective cohort" OR "prospective research" OR "prospective study" OR "prospective trial" OR "retrospective cohort" OR "retrospective research" OR "retrospective study" OR "retrospective trial" OR "single blind" OR "trial")		
S5	TI ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter") OR AB ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter")	172,910	179,188

	OR SU ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter") OR KW ("case report" OR "case reports" OR "case series" OR "comment" OR "commentary" OR "editorial" OR "letter")		
S6	(DE "Vertebrates" OR DE "Mammals" OR DE "Animals" OR DE "Rats" OR DE "Rodents" OR DE "Mice" OR TI "animals" OR TI "animal" OR TI "mouse" OR TI "mice" OR TI "rodent" OR TI "rodents" OR TI "rat" OR TI "rats" OR SU "animals" OR SU "animal" OR SU "mouse" OR SU "mice" OR SU "rodent" OR SU "rodents" OR SU "rat" OR SU "rats" OR KW "animals" OR KW "animal" OR KW "mouse" OR KW "mice" OR KW "rodent" OR KW "rodents" OR KW "rat" OR KW "rats" OR AB "animals" OR AB "animal" OR AB "mouse" OR AB "mice" OR AB "rodent" OR AB "rodents" OR AB "rat" OR AB "rats") NOT (PO "human" OR TI "humans" OR TI "human" OR AB "humans" OR AB "human" OR SU "humans" OR SU "human" OR KW "humans" OR KW "human")	300,133	299,102
S7	S1 AND (S2 OR S3 OR S4)	38	83
S8	S7 NOT S5	36	72
S9	S8 NOT S6	36	72
S10	S9 AND LA English	36	72
S11	S10 Limiters - Published Date: 19600101-20190731	36	
S11	S10 Limiters - Published Date: 20190101-20211001		32

Table B-7. MEDLINE (PubMed) Search Strategy for Bone Density in AN

Search	Query	Search date: 07/31/2019	Search date: 10/01/2021
#1	Search (("anorexia"[MH] OR "anorexic"[TIAB] OR "anorexia"[TIAB]) AND "nervosa"[TIAB]) OR "anorexia nervosa"[TIAB] OR "anorexia nervosa"[MH] OR "bulimia"[MH] OR "bulimia nervosa"[MH] OR "bulimia"[TIAB] OR "bulimic"[TIAB] OR "binging"[TIAB] OR "purging"[TIAB] OR "binge eating"[TIAB] OR "binge eating disorder"[TIAB] OR "binge eating disorder"[MH] OR "eating disorder"[TIAB] OR "eating disorders"[TIAB]	35,365	40,241
#2	Search "randomized controlled trial"[PT] OR "randomisation"[TIAB] OR "randomised"[TIAB] OR "randomization"[TIAB] OR "randomized"[TIAB] OR "randomly"[TIAB] OR "placebo"[TIAB] OR "sham"[TIAB]	1,086,783	1,258,934
#3	Search "meta analysis as topic"[MeSH Major Topic] OR "meta analysis as topic"[MeSH Terms] OR "meta analysis"[TIAB] OR "meta analyses"[TIAB] OR "meta analytic"[TIAB] OR "metaanalysis"[TIAB] OR	261,181	358,062

	"metaanalyses"[TIAB] OR "systematic review"[TIAB] OR "systematic reviews"[TIAB] OR "meta analysis"[PT]		
#4	Search "controlled clinical trial"[PT] OR "blinded" [TIAB] OR "case control" [TIAB] OR "clinical trial" [TIAB] OR "clinical trials" [TIAB] OR "Cohort Analysis" [TIAB] OR "cohort research" [TIAB] OR "cohort study" [TIAB] OR "cohort trial" [TIAB] OR "comparator group" [TIAB] OR "controlled studies" [TIAB] OR "controlled study" [TIAB] OR "controlled trial" [TIAB] OR "controlled trials" [TIAB] OR "double blind" [TIAB] OR "followup" [TIAB] OR "follow up" [TIAB] OR "longitudinal research" [TIAB] OR "longitudinal study" [TIAB] OR "longitudinal trial" [TIAB] OR "multicenter trial" [TIAB] OR "multicenter trials" [TIAB] OR "naturalistic research" [TIAB] OR "naturalistic study" [TIAB] OR "naturalistic trial" [TIAB] OR "prospective cohort" [TIAB] OR "prospective research" [TIAB] OR "prospective study" [TIAB] OR "prospective trial" [TIAB] OR "retrospective cohort" [TIAB] OR "retrospective research" [TIAB] OR "retrospective study" [TIAB] OR "retrospective trial" [TIAB] OR "single blind" [TIAB]	2,312,245	2,734,741
#5	Search ("case reports"[PT] OR "comment"[PT] OR "editorial"[PT] OR "letter"[PT])	3,569,055	3,990,385
#6	Search (("animals"[MAJR] OR "animals"[MH] OR "animal"[TIAB] OR "animals"[TIAB] OR "rat"[TIAB] OR "mouse"[TIAB] OR "mice"[TIAB] OR "rodent"[TIAB] OR "rodents"[TIAB] OR "rats"[TIAB]) NOT ("humans"[MAJR] OR "humans"[MH] OR "human"[TIAB] OR "humans"[TIAB]))	4,562,314	4,858,349
#7	Search "estrogen" OR "estrogens" OR "dehydroepiandrosterone" OR "hormone replacement" OR "gonadal steroid" OR "gonadal steroids" OR "teriparatide" OR "alendronate" OR "risedronate" OR "risedronic acid" OR "oral contraception" OR "oral contraceptive" OR "estradiol" OR "vitamin d" OR "growth hormone" OR "estrogen replacement therapy"[MeSH Terms] OR "estrogenic steroids, alkylated"[MeSH Terms] OR "estrogens"[MeSH Terms] OR "dehydroepiandrosterone"[MeSH Terms] OR "hormone replacement therapy"[MeSH Terms] OR "gonadal steroid hormones"[MeSH Terms] OR "teriparatide"[MeSH Terms] OR "alendronate"[MeSH Terms] OR "risedronic acid"[MeSH Terms] OR "contraceptive agents"[MeSH Terms] OR "estradiol"[MeSH Terms] OR "vitamin d"[MeSH Terms] OR "growth hormone"[MeSH Terms]	558,581	598,484
#8	Search #1 AND #7	1,267	1,349
#9	Search #8 AND (#2 OR #3 OR #4)	187	211
#10	Search #9 NOT #5	183	206
#11	Search #10 NOT #6	179	200
#12	Search #11 AND "english"[Language]	172	193

#13	Search #12 AND (("1960/01/01"[Date - Publication] : "2019/07/15"[Date - Publication]))	172	
#13	Search #12 AND (("2019/01/01"[Date - Publication] : "2021/10/01"[Date - Publication]))		26

Table B-8: Cochrane Library (Wiley) Search Strategy for Bone Density in AN

Search ID#	Search	Search date: 07/31/2019	Search date: 10/01/2021 (from Jan 2019 through Sept 2021)
#1	(bone density):ti,ab,kw OR ("estrogen"):ti,ab,kw OR (estrogens):ti,ab,kw OR ("dehydroepiandrosterone"):ti,ab,kw OR ("hormone replacement therapy"):ti,ab,kw	25,268	8427
#2	("gonadal steroid"):ti,ab,kw OR ("teriparatide acetate"):ti,ab,kw OR ("alendronate sodium"):ti,ab,kw OR ("risedronate sodium"):ti,ab,kw OR ("oral contraceptive agent"):ti,ab,kw	1,146	139
#3	("vitamin d"):ti,ab,kw OR ("growth hormone"):ti,ab,kw with Cochrane Library publication date	15,508	6276
#4	("anorexia nervosa"):ti,ab,kw (Word variations have been searched)	956	388
#5	#1 OR #2 OR #3	38,591	13872
#6	#4 AND #5	112	30

Four reviewers (L.J.F., S.-H.H., J.Y., and T.C.) screened the results of the initial search, with each abstract and title screened by two reviewers according to APA's general screening criteria: RCT, systematic review or meta-analysis, or observational study with a sample of at least 50 individuals; human; and study of the effects of a specific intervention or psychiatric disorder or symptoms. If discrepancies were noted among reviewers' ratings, an additional opinion was given by a third individual and consensus was achieved among the reviewers. For the updated search, abstracts were screened in the same fashion by two reviewers (L.J.F. and S.-H.H.) with discrepancies resolved by discussion and consensus among the reviewers. Abstracts identified using this approach were then reviewed by one individual (S.-H. H.), with verification by a second reviewer (L. J. F.) to determine whether they met eligibility criteria as defined by

the PICOT elements (see Table B-9). If the publication characteristics were not clear from the initial title and abstract review, full text review occurred.

Table B-9: PICOT elements for eating disorders systematic review

Participants/population

- Age ≥10
- Diagnosed with an eating disorder (anorexia nervosa, bulimia nervosa, binge-eating disorder, night eating syndrome, avoidant restrictive food intake disorder) with diagnosis as defined by DSM-III, DSM-III-R, DSM-IV, DSM-IV-TR, DSM-5 (Section II or Section III), or ICD-10, as applicable.
- For mixed population studies, the eating disorder of interest had to account for ≥75% of the total population

Interventions

- Psychotherapies
    - Individual cognitive-behavioral therapy
    - Individual dialectical behavioral therapy
    - Individual interpersonal therapy
    - Individual supportive psychotherapy
    - Psychodynamically informed individual therapy
    - Maudsley Anorexia Treatment for Adults (MANTRA) – for AN only
    - Family-based therapy
    - Other approaches to family or couples' therapy
    - Group therapy
    - Psychoeducation
    - Other psychotherapies
- Pharmacotherapies
    - SSRI: citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline
    - SNRI: desvenlafaxine, duloxetine, venlafaxine, levomilnacipran
    - Other antidepressants (e.g., mirtazapine, bupropion)
    - Second generation antipsychotic agents: aripiprazole, asenapine, brexpiprazole, cariprazine, iloperidone, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, ziprasidone
    - Anticonvulsants (e.g., topiramate)
    - Other medications: benzodiazepine, metoclopramide
- Other interventions
    - Nutritional rehabilitation
    - Inpatient management
    - Specialist supportive clinical management (i.e., including support, education, advice, praise)
    - Neurostimulation therapies (e.g., ECT, TMS, tDCS)
    - Self-help/12 step programs

- Refeeding approaches – for AN only
    - Intravenous tube feeding
    - Nasogastric (NG) continuous tube feeding
    - NG bolus tube feeding
    - Supplemental overnight tube feeding
- Treatments to improve or prevent deterioration of bone density – for AN only
    - Calcium and vitamin D supplementation
    - Hormone replacement therapy
    - Bisphosphonates
    - Moderate exercise (if no history of compulsive exercising)

Comparators

- Placebo
- Treatment as usual
- Wait list control
- General psychiatric management
- Interventions listed for inclusion

Outcomes

- BMI
- Percent ideal body weight (IBW)
- Other measures related to body weight
- Behavioral events (e.g., binges, purging)
- Other eating related outcomes (e.g., rating scale metrics) – if primary outcome or pre-specified secondary outcome measure
- Partial or complete response/remission
- Bone density changes – for AN
- Return of menses – for AN
- Prevention/reduction of co-occurring psychiatric conditions
- Quality of life
- Functioning
- Treatment adherence rates
- Study withdrawal rates, all cause
- Study withdrawal rates due to adverse events or serious adverse events
- Treatment emergent side effects (e.g., sedation, gastrointestinal disturbances, lightheadedness, cardiovascular changes, sleep disturbance, headache, sexual dysfunction) – for pharmacotherapy studies
- Rates of rehospitalization
- Suicidal behaviors including suicide deaths and attempts
- All-cause mortality

Treatment duration

- ≥10 days mean/median treatment duration for refeeding studies
- ≥8 weeks for studies of other interventions

Setting of care

- Any

Study design

- RCTs with N≥20
- Non-randomized clinical trials with N≥50
- Observational studies, comparative, with N≥50
    - Cross-sectional
    - Prospective cohort
    - Retrospective cohort
    - Non-concurrent cohort
    - Case-control
- Pooled analyses of the above study designs

For the literature searches described above, PRISMA Diagrams were generated (Page et al. 2021) as shown in Figures B-1 through B-4.

Figure B-1. PRISMA diagram for general search for studies on eating disorders, including AN, BN, BED, and night eating syndrome

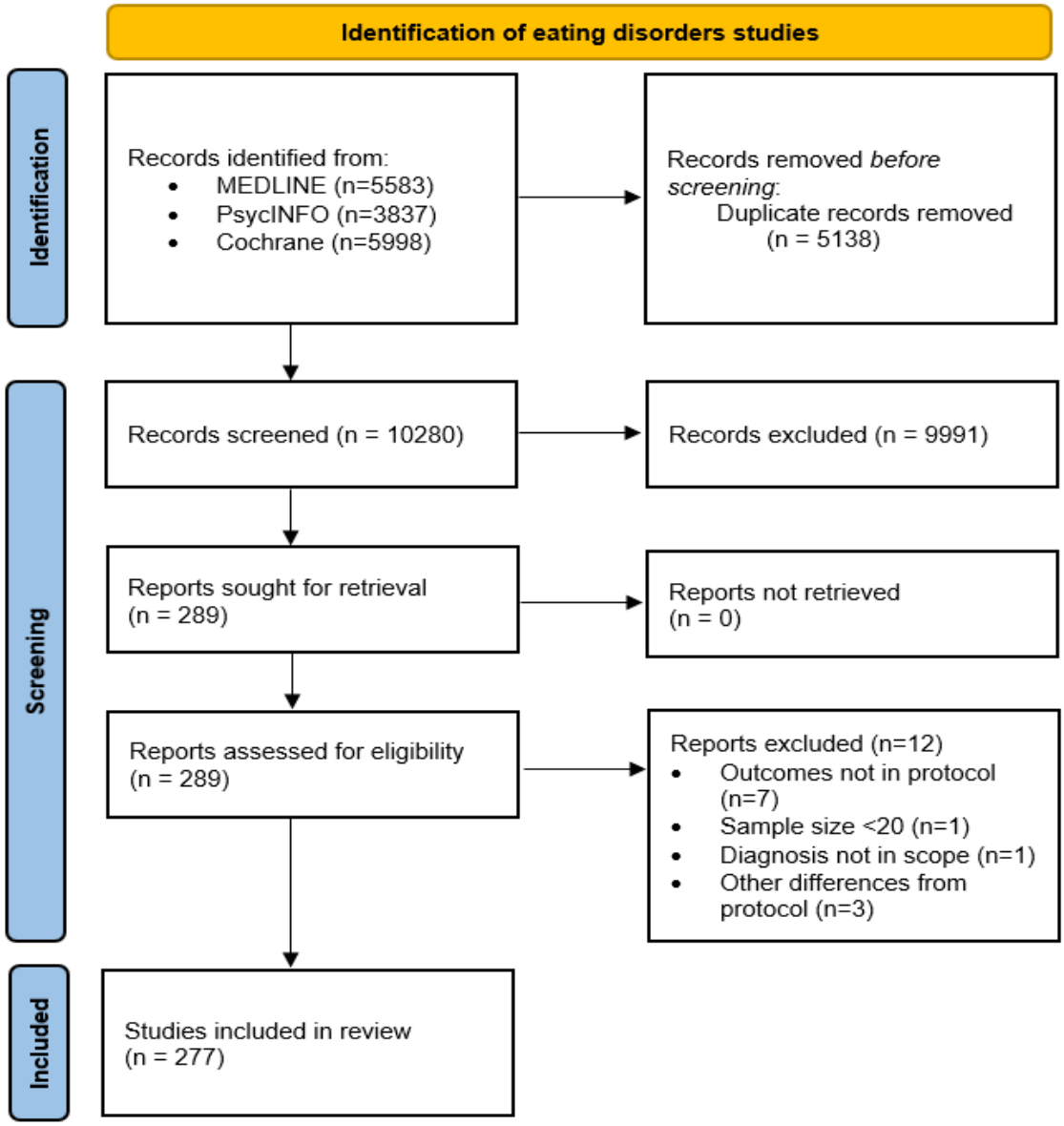


Figure B-2. PRISMA diagram for treatments to address bone density reductions in individuals with AN

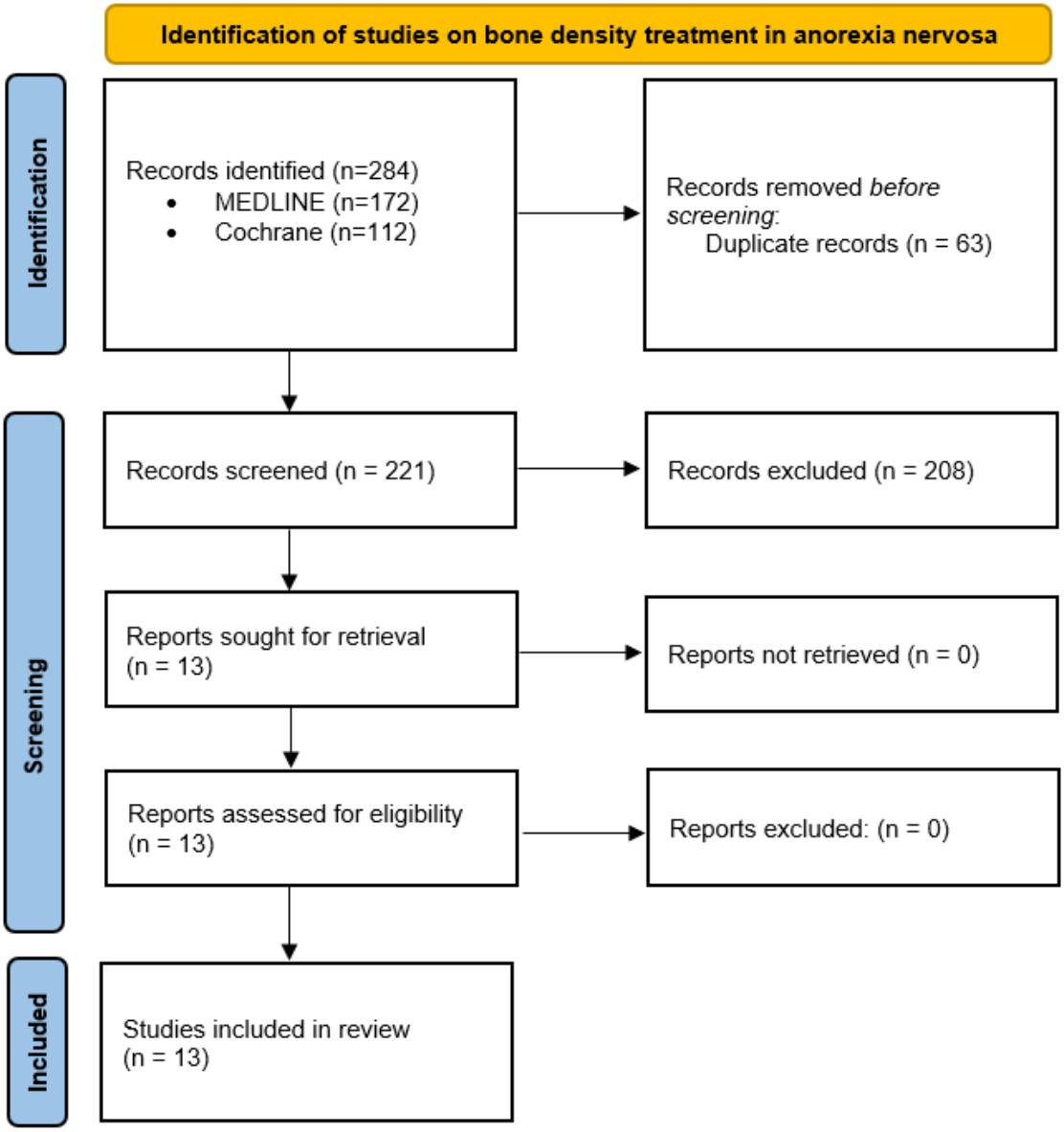


Figure B-3. PRISMA diagram for studies on ARFID

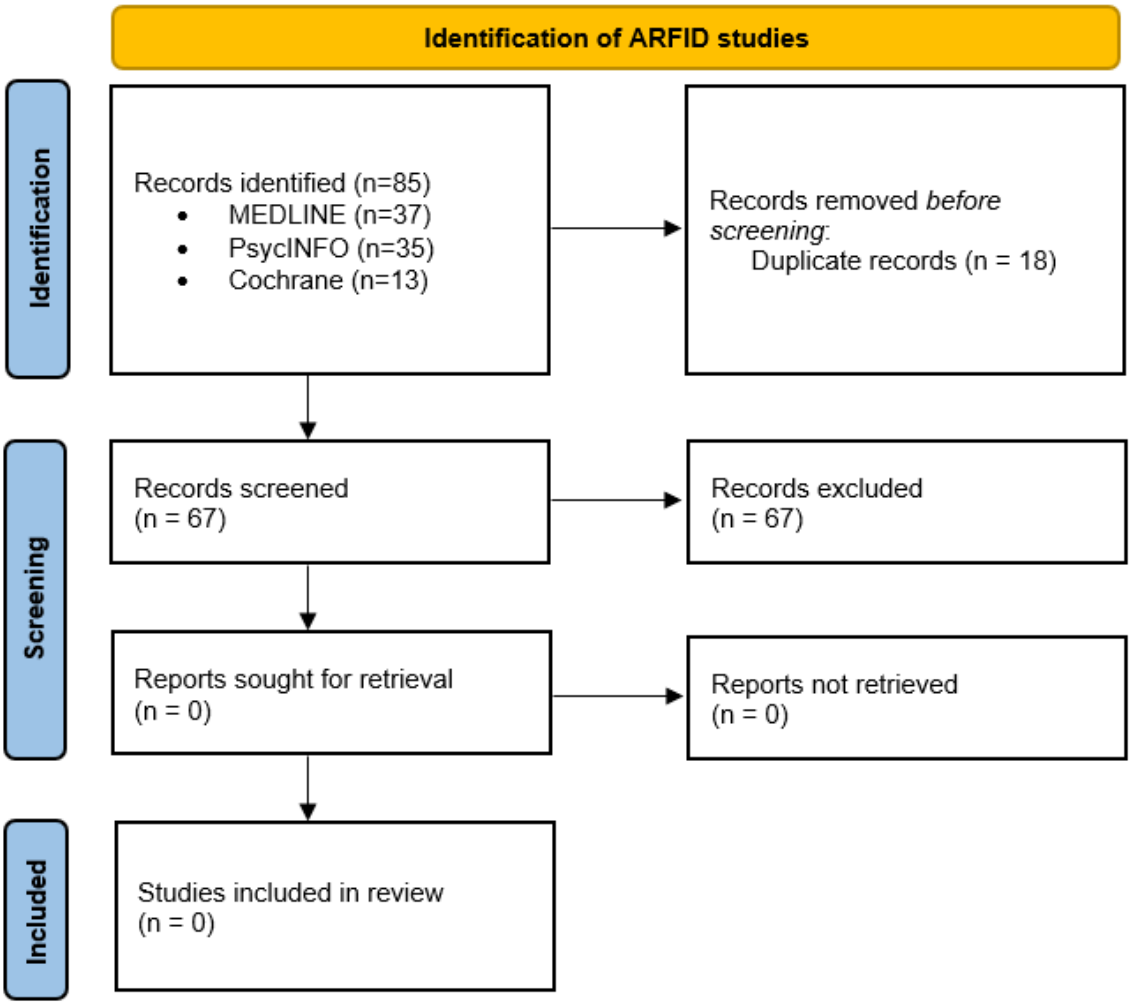
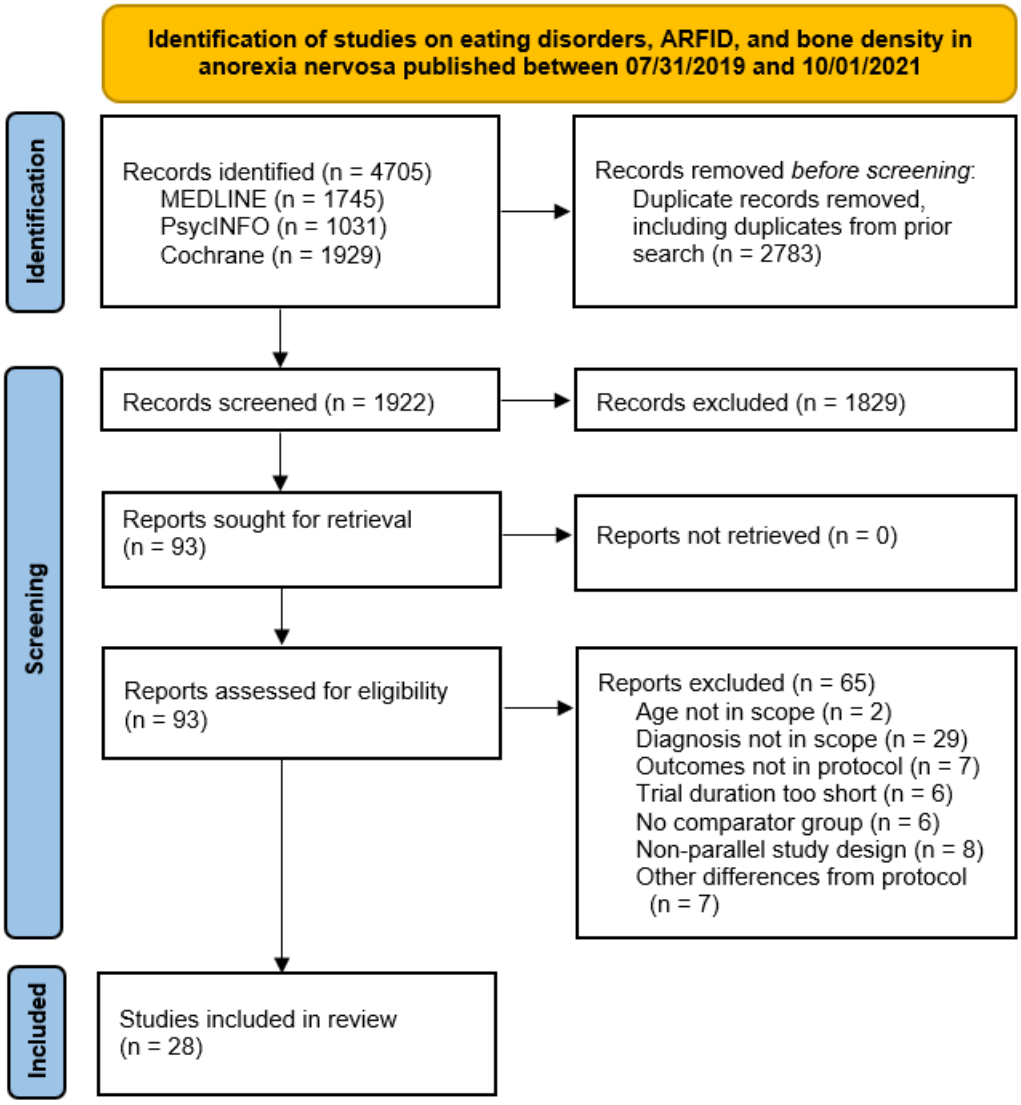


Figure B-4. PRISMA diagram for studies in updated search



For each trial identified for inclusion from the initial search, detailed information was extracted by Dr. Evidence (Santa Monica, CA) using the DOC Data 2.0 software platform. A small number of studies were excluded during this extraction phase, as noted in Figure B-1. Dr. Evidence processes included verifications and quality checks on data extraction. In addition to specific information about each reported outcome, extracted information included citation; study design; treatment arms (including doses, sample sizes); co-intervention, if applicable; trial duration and follow-up duration, if applicable; country; setting; funding source; sample characteristics (e.g., mean age, percent nonwhite, percent female, percent with co-occurring condition); and rates of attrition, among other data elements. For the updated search, information was extracted and verified by APA staff (S.-H.H., J.M.).

Summary tables (see Appendices E and H) include specific details for each study identified for inclusion from the literature search. Factors relevant to risk of bias were also identified for each RCT that contributed to a guideline statement. Risk of bias was determined using the Cochrane Risk of Bias 2.0

tool (Sterne et al. 2019) by one reviewer (J.M.) and verified by an additional reviewer (S.-H.H. or L.J.F.). Risk of bias ratings are included in summary tables (see Appendix E) with specific factors contributing to the risk of bias for each study shown in Appendix F. Ratings of the strength of supporting evidence were determined, in accordance with the AHRQ's Methods Guide for Effectiveness and Comparative Effectiveness Reviews (Agency for Healthcare Research and Quality 2014), by the methodologist (L.J.F.) and reviewed by members of the SRG and GWG.

Network meta-analyses were conducted by Heno Analytics (Vancouver, BC, Canada) using the extracted outcome data, the DOC Data 2.0 software platform (Dr. Evidence; Santa Monica, CA), R software (R Core Team 2020), and Just another Gibbs sampler (JAGS), a program for simulation from Bayesian hierarchical models using Markov chain Monte Carlo methods (Plummer 2021). The feasibility of conducting each NMA was assessed using recommended approaches based on subject characteristics, network connectivity, and definitions and availability of outcomes (Cope et al. 2014). Outcomes were grouped, where possible as shown in Tables B-10 through B-12.

Table B-10. Outcomes and combined outcomes of AN studies

Outcome	Outcome combined
BMI	BMI
Weight	Weight
BDI	Depression (calculated as SMD)
DASS, depression	
BSI, depression	
HADS, depression	
MFQ	
BAI	Anxiety (calculated as SMD)
DASS, anxiety	
BSI, anxiety	
STAI	
HADS, anxiety	
Study withdrawal	Study withdrawal
Study withdrawal, all cause	
Mortality, all cause	Mortality
Mortality, anorexia nervosa	
Treatment adherence	Treatment adherence
Treatment adherence > 50%	
Compliance	
Disease response, recovery	Disease response, remission/recovery
Disease response, remission	
Disease response, remission, full	

Disease response, complete response	
Treatment discontinuation	Treatment discontinuation
EDE	Eating disorder (calculated as SMD)
EDE, global	
EDE-Q	
EDI	
EDI-2	
SEED ANTSI	
YBC-EDS	
CIA	Social functioning (calculated as SMD)
SIAB-EX, general psychopathology and social integration	
WSAS	
Q-LES-Q	Quality of life (calculated as SMD)
WHO-QoL	
SF-Physical component	
Hospitalization	Hospitalization
Re-hospitalization	Re-hospitalization
Percent IBW	Percent IBW
Percent EBW	

Abbreviations: BAI=Beck Anxiety Inventory; BDI=Beck Depression Inventory; BMI=body mass index; BSI=Brief Symptom Inventory; CIA=Clinical Impairment Assessment; DASS=Depression Anxiety Stress Scales; EBW=Expected Body Weight; EDE=Eating Disorder Examination; EDE-Q=Eating Disorder Examination Questionnaire; EDI=Eating Disorder Inventory; HADS= Hospital Anxiety and Depression Scale; IBW=Ideal Body Weight; MFQ=Mood and Feelings Questionnaire; Q-LES-Q=Quality of Life Enjoyment and Satisfaction Questionnaire; SEED ANTSI=Short Evaluation of Eating Disorders Anorexia Nervosa Total Severity Index; SF=Short Form; SIAB-EX=Structured Interview for Anorexic and Bulimic Syndromes; SMD=standardized mean difference; STAI=State-Trait Anxiety Inventory; WHO-QoL=World Health Organization Quality of Life; WSAS=Work and Social Adjustment Scale; YBC-EDS=Yale-Brown-Cornell Eating Disorders Scale

Table B-11. Outcomes and combined outcomes of BN studies

Outcome	Outcomes Combined
BMI	BMI
Weight	Weight
BDI	Depression
HDRS	
HDRS 21 items	
HDRS 17 items	
Depression scale	
Study withdrawal	Study withdrawal

Treatment adherence, completed	Treatment adherence
Treatment adherence ≥4 sessions	
Disease response, remission	Disease response, remission
Disease response, complete remission	
Binge eating	Binge eating
Binge eating, objective	
Binge eating, abstinence	Binge eating, abstinence
Binge eating, abstinence OR remission and compensatory behaviors, abstinence OR remission	
Binge eating, marked response, reduction 75%-99%	
Treatment discontinuation	Treatment discontinuation
EDE	Eating disorder scale
EAT	
EDI	
EDI-2	
Purging	Purging
Purging, abstinence	Purging, abstinence
Bulimic episodes, abstinence and purging, abstinence	
Binge eating, marked improvement OR remission ≥ 75%-100% and/or purging, marked improvement OR remission ≥ 75%-100%	
RSES	Self esteem

Abbreviations: BDI=Beck Depression Inventory; BMI=body mass index; EAT=Eating Attitude Test; EDE=Eating Disorder Examination; EDI=Eating Disorder Inventory; HDRS=Hamilton Depression Rating Scale; RSES=Rosenberg Self-Esteem Scale

Table B-12. Outcomes and combined outcomes of BED studies

Outcome	Outcome combined
Adherence, completed treatment (Binary)	Adherence, completed treatment
Adverse event, serious (Binary)	Adverse event, serious
BDI	Depression
BMI	BMI
Binge eating	Binge eating
EDE-Q, binge eating	
EDE-I, binge eating	
BES	Binge eating scale
Binge-eating disorder (Binary)	Binge-eating disorder (Binary)
Binge eating, abstinence (Binary)	Binge eating, abstinence
CGI-S	CGI-S
CGI-I, very much improved (Binary)	CGI-I, very much improved
Constipation (Binary)	Constipation
Disease response, remission (Binary)	Disease response, remission

Diarrhea (Binary)	Diarrhea
Dizziness (Binary)	Dizziness
Drowsiness (Binary)	Drowsiness
Fatigue (Binary)	Fatigue
Insomnia (Binary)	Insomnia
Study withdrawal (Binary)	Study withdrawal
Y-BOC-BE	Y-BOC-BE
Weight	Weight
Xerostomia (Binary)	Xerostomia

Abbreviations: BDI=Beck Depression Inventory; BES=Binge Eating Scale; BMI=body mass index; CGI-I=Clinical Global Impression-Improvement; CGI-S=Clinical Global Impressions-Severity; EDE-I=Eating Disorder Examination Interview; EDE-Q=Eating Disorder Examination Questionnaire; Y-BOC-BE=Yale-Brown Obsessive Compulsive Scale Modified for Binge Eating

Outcomes were also grouped and analyzed at specified time points, where feasible, as well as at all time points. The selection of a random-effects model versus a fixed effects model for the NMAs was based on the deviation information criterion (DIC). Typically, the random-effects model was chosen based on anticipated heterogeneity and DIC criteria but where the DIC for the two models was comparable or when the network contained only two studies, a fixed effects model was used. When all treatment arms reported a value of zero for an outcome (such as number of adverse events), the study was not included in the NMA. Endonodal studies were also excluded. When the network included closed loops, the consistency of relative treatment effects was assessed based on direct as well as indirect evidence. For BN and BED, the NMA included pharmacotherapies as well as psychotherapies; however, because of the relatively small number of pharmacotherapy studies in AN, these studies were reviewed qualitatively rather than quantitatively.

Appendix C. Review of Research Evidence Supporting Guideline Statements

Assessment and Determination of Treatment Plan

Statement 1 – Screening for Presence of an Eating Disorder

APA recommends (1C) screening for the presence of an eating disorder as part of an initial psychiatric evaluation.

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice, from epidemiologic data on the prevalence of eating disorders in clinical populations, and from data on the validation of eating disorder screening tools. Together, the strength of research evidence is rated as low. This recommendation is also consistent with the recommendations of Guideline I, "Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History" in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016).

A detailed systematic review to support this statement is outside the scope of this guideline; however, a less comprehensive search of the literature identified two meta-analyses on the diagnostic test characteristics of the SCOFF. Botella and colleagues (Botella et al. 2013) examined the diagnostic accuracy of the SCOFF as compared to a diagnostic reference group among 15 studies, which included 882 cases and 4,350 controls. The diagnostic reference conditions included structured diagnostic interviews in some studies (e.g., CIDI, DSM-IV, MINI, SCAN) and specified thresholds in psychometric tools (e.g., EAT-40, EAT-26, EDI-2) in other studies. Languages in which the SCOFF was studied included Catalan, English, Finnish, French, German, Italian, and Spanish. Taken together, the pooled estimates for sensitivity and specificity of the SCOFF were 0.80 and 0.93 respectively with corresponding estimates of 0.882 and 0.925 when studies were limited to those that used an interview format for a diagnostic reference. A subsequent meta-analysis by Kutz and colleagues (Kutz et al. 2020) included 10 additional validation studies of the SCOFF and found pooled estimates for sensitivity and specificity of 0.86 and 0.83, respectively. However, they also noted significant heterogeneity in these estimates that the authors attributed to differences in study methodology or sample characteristics among the studies. Higher values for diagnostic accuracy were noted when samples included more women than men as well as in case-control studies, which primarily included individuals at risk for AN or BN. As in the study of Botella and colleagues, use of an interview as a diagnostic standard was also associated with higher diagnostic accuracy. Overall, Kutz and colleagues found pooled positive and negative likelihood ratios of 5.0 and 0.17 suggesting that the SCOFF is moderately helpful in detecting the presence of an eating disorder or in ruling out the presence of an eating disorder. However, the available evidence on the utility of the SCOFF has not included validation samples in diverse community settings, primary care practices, or general psychiatric settings. In addition, many of the studies of the SCOFF were rated as having some or high levels of bias, with only two studies described as having a low risk of bias in all study quality domains.

Other screening measures for eating disorders have also been proposed but have been less well studied. The Eating Disorder Screen for Primary Care (Cotton et al. 2003) was developed for use in primary care and tested in a London sample of 129 university students and 104 primary care patients. A threshold score of 2 on the scale yielded a sensitivity and specificity of 1.0 and 0.71 respectively. The Screen for

Disordered Eating (Maguen et al. 2018) was also developed for use in primary care settings and can screen for BED as well as AN and BN. Using a score of 2 as a threshold, this screening measure had a sensitivity and specificity of 0.91 and 0.58 respectively. However, the generalizability of the findings is not clear as there were few individuals who had an eating disorder diagnosis. In addition, the study was conducted through a Veterans Hospital Administration medical center and included 407 female veterans, so it is unlikely to be representative of general community or psychiatric samples. Both screening measures would benefit from additional study in larger and more diverse samples.

Grading of the Overall Supporting Body of Research Evidence for Screening for Presence of an Eating Disorder

On the basis of the limitations of the evidence for screening for the presence of an eating disorder, no grading of the body of research evidence is possible.

Statement 2 – Initial Evaluation of Eating History

APA recommends (1C) that the initial evaluation of a patient with a possible eating disorder include assessment of

- **the patient's height and weight history (e.g., maximum and minimum weight, recent weight changes);**
- **presence of, patterns in, and changes in restrictive eating, food avoidance, binge eating, and other eating-related behaviors (e.g., rumination, regurgitation, chewing and spitting);**
- **patterns and changes in food repertoire (e.g., breadth of food variety, narrowing or elimination of food groups);**
- **presence of, patterns in, and changes in compensatory and other weight control behaviors, including dietary restriction, compulsive or driven exercise, purging behaviors (e.g., laxative use, self-induced vomiting), and use of medication to manipulate weight;**
- **percentage of time preoccupied with food, weight, and body shape;**
- **prior treatment and response to treatment for an eating disorder;**
- **psychosocial impairment secondary to eating or body image concerns or behaviors; and**
- **family history of eating disorders, other psychiatric illnesses, and other medical conditions (e.g., obesity, inflammatory bowel disease, diabetes mellitus).**

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice. Data from the expert survey specifically confirms the importance of obtaining information about the patient's height and weight as part of the initial evaluation (see Appendix D). Expert opinion also suggests that conducting assessments of height, weight, eating history, and family history of eating disorders as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and treatment safety. For additional details of the initial psychiatric evaluation, see Guideline I, "Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History" in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016). A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield RCTs related to this

recommendation in the context of eating disorder treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Assessment of an Eating History

On the basis of the limitations of the evidence for assessment of an eating history in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Statement 3 – Quantitative Measures

APA recommends (1C) that the initial psychiatric evaluation of a patient with a possible eating disorder include weighing the patient and quantifying eating and weight control behaviors (e.g., frequency, intensity, or time spent on dietary restriction, binge eating, purging, exercise, and other compensatory behaviors).

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice. Expert opinion suggests that conducting such assessments as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and treatment safety. Diagnostic criteria for specific eating disorders include consideration of the patient's weight and the frequency of weight control behaviors; clinical trials typically include measurement of these parameters for assessing treatment outcomes. Quantitative approaches to assessment are also suggested as part of the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016) in Guideline VII, "Quantitative Assessment."

A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any RCTs related to this recommendation in the context of eating disorder evaluation or treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Use of Quantitative Measures in a Patient With a Possible Eating Disorder

On the basis of the limitations of the evidence for use of quantitative measures in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Statement 4 – Identification of Co-occurring Conditions

APA recommends (1C) that the initial psychiatric evaluation of a patient with a possible eating disorder identify co-occurring health conditions, including co-occurring psychiatric disorders.

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice. Expert opinion suggests that conducting such assessments as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and treatment safety. For additional details, see Guideline I, "Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History", Guideline II. Substance Use Assessment, and Guideline VI, "Assessment of Medical Health," in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016). A detailed systematic review to support this statement is outside the

scope of this guideline; however, less comprehensive searches of the literature did not yield any RCTs related to this recommendation in the context of eating disorder evaluation or treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Assessment of Co-occurring Conditions in a Patient With a Possible Eating Disorder

On the basis of the limitations of the evidence for assessment of co-occurring conditions in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Statement 5 – Initial Review of Systems

APA recommends (1C) that the initial psychiatric evaluation of a patient with a possible eating disorder include a comprehensive review of systems.

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice as well as expert opinion on commonly observed or clinically important abnormalities in individuals with an eating disorder. Expert opinion also suggests that conducting such assessments as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and treatment safety. For additional details, see Guideline I, "Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History" and Guideline VI, "Assessment of Medical Health," in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016). A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any RCTs related to this recommendation in the context of eating disorder evaluation or treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Review of Systems in a Patient With a Possible Eating Disorder

On the basis of the limitations of the evidence for conducting a review of systems in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Statement 6 – Initial Physical Examination

APA recommends (1C) that the initial physical examination of a patient with a possible eating disorder include assessment of vital signs, including temperature, resting heart rate, blood pressure, orthostatic pulse, and orthostatic blood pressure; height, weight, and BMI (or percent median BMI, BMI percentile, or BMI Z-score for children and adolescents); and physical appearance, including signs of malnutrition or purging behaviors.

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice as well as data from the expert survey (see Appendix D). Expert opinion also suggests that abnormal findings on the physical examination are commonly observed or clinically important in individuals with an eating disorder and that conducting such assessments as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and treatment safety. For additional details, see Guideline VI, "Assessment of Medical Health," in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016). A detailed

systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any RCTs related to this recommendation in the context of eating disorder evaluation or treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Conducting an Initial Physical Examination in a Patient With a Possible Eating Disorder

On the basis of the limitations of the evidence for conducting an initial physical examination in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Statement 7 – Initial Laboratory Assessment

APA recommends (1C) that the laboratory assessment of a patient with a possible eating disorder include a complete blood count and a comprehensive metabolic panel, including electrolytes, liver enzymes, and renal function tests.

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice as well as well as data from the expert survey (see Appendix D). Expert opinion also suggests that laboratory test abnormalities are commonly observed or clinically important in individuals with an eating disorder and that conducting such assessments as part of the initial psychiatric evaluation improves diagnostic accuracy, appropriateness of treatment selection, and treatment safety. For additional details, see Guideline VI, "Assessment of Medical Health," in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016). A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any RCTs related to this recommendation in the context of eating disorder evaluation or treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Conducting an Initial Laboratory Assessment in a Patient With a Possible Eating Disorder

On the basis of the limitations of the evidence for conducting an initial laboratory assessment in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Statement 8 – Initial Electrocardiogram

APA recommends (1C) that an electrocardiogram be done in patients with a restrictive eating disorder, patients with severe purging behavior, and patients who are taking medications that are known to prolong QTc intervals.

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice as well as data from the expert survey (see Appendix D). Expert opinion and literature reports also suggest that a number of clinically important cardiac and electrocardiographic abnormalities can occur in individuals with an eating disorder (Frederiksen et al. 2018a, 2018b, 2021; Giovinazzo et al. 2019; Hanachi et al. 2020; Krantz et al. 2020; Sachs et al. 2016) As a result, conducting an ECG under specified circumstances as part of the initial evaluation may improve diagnostic accuracy, appropriateness of treatment selection, and treatment safety. For additional details, see Guideline VI,

"Assessment of Medical Health," in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016). A detailed systematic review to support this statement is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any RCTs related to this recommendation in the context of eating disorder evaluation or treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Conducting an Initial Electrocardiogram in a Patient With a Possible Eating Disorder

On the basis of the limitations of the evidence for conducting an initial ECG in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Statement 9 – Treatment Plan, Including Level of Care

APA recommends (1C) that patients with an eating disorder have a documented, comprehensive, culturally appropriate, and person-centered treatment plan that incorporates medical, psychiatric, psychological, and nutritional expertise, commonly via a coordinated multidisciplinary team.

Support for this statement comes from general principles of assessment and clinical care in psychiatric practice. For additional details, see the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd Edition (American Psychiatric Association 2016). Information from the expert survey (see Appendix D) also informs portions of the implementation section on determining a level of care. A detailed systematic review to support the importance of treatment planning is outside the scope of this guideline; however, less comprehensive searches of the literature did not yield any RCTs related to this recommendation in the context of eating disorders evaluation or treatment. Consequently, the strength of research evidence is rated as low.

Grading of the Overall Supporting Body of Research Evidence for Treatment Planning in a Patient With a Possible Eating Disorder

On the basis of the limitations of the evidence for developing and documenting a comprehensive, culturally appropriate, and person-centered treatment plan in a patient with a possible eating disorder, no grading of the body of research evidence is possible.

Anorexia Nervosa

Statement 10 – Medical Stabilization, Nutritional Rehabilitation, and Weight Restoration for Patients With Anorexia Nervosa

APA recommends (1C) that patients with anorexia nervosa who require nutritional rehabilitation and weight restoration have individualized goals set for weekly weight gain and target weight.

Setting Individualized Goals for Weekly Weight Gain and Target Weight

No RCTs or other clinical trials were found that specifically assessed whether setting individualized goals for weekly weight gain and target weight improved outcomes as compared to not setting such goals. Instead, evidence for setting individualized goals for weekly weight gain and for target weight comes from expert opinion (see Appendix D) and indirect inferences from studies of weight gain with differing levels of initial caloric intake (see Appendix H) as well as studies of AN prognosis (Garber et al. 2013,

2021; Golden et al. 2013, 2021; Imbierowicz et al. 2002; O'Connor et al. 2016). Consequently, the strength of research evidence is rated as low.

There has previously been concern about the occurrence of physiological abnormalities including refeeding syndrome with high calorie nutritional rehabilitation; however, evidence from RCTs (Garber et al. 2021; Golden et al. 2021; O'Connor et al. 2016), retrospective studies (Golden et al. 2013; Imbierowicz et al. 2002), and clinical experience (Haas et al. 2021) suggest that use of higher caloric intake is associated with greater weekly weight gain and shorter hospital stays, and does not result in significant differences in physiological abnormalities (Garber et al. 2016). Evidence also suggests that weight gain during hospitalization and more rapid normalization of weight are associated with improved long-term outcomes (Glasofer et al. 2020; Redgrave et al. 2021) as is normalization of eating behaviors (Cooper et al. 2021). Thus, setting weekly weight gain targets with a focus on a relatively rapid return to the target weight is likely to be associated with enhanced short- and long-term outcomes. The evidence also suggests that using higher values for initial daily caloric intake (e.g., 1,500 to 2,000 kcal/day increasing by 200 kcal/day) is likely to be more effective in accomplishing weight gain goals than lower values of initial caloric intake or slower increases in caloric intake.

In the expert survey, there was concurrence that goals for kcal/day should be based on initial and target weights and anticipated/recommended rate of weight gain. In addition, most of the experts agreed that lower target goals for weight gain or caloric intake were appropriate for outpatients as compared to inpatients. There was substantial variability in expert opinion for the appropriateness of targets for daily caloric intake, but 40 to 60 kcal/kg/day had higher mean and median ratings than other caloric ranges for inpatient, intensive outpatient, or partial hospital settings and a range of 30 to 60 kcal/kg/day was rated as most appropriate for patients treated in an office-based outpatient setting. Target weight gains of 2 to 3 lbs/week (0.9 to 1.36 kg/week) were viewed as most appropriate for adolescents and adults in inpatient, intensive outpatient, or partial hospital settings with 1 to 2 lbs/week (0.45 to 0.9 kg/week) receiving the highest mean ratings for appropriate weekly weight gain targets in office-based outpatient settings. However, the expert survey was conducted prior to the publication of recent evidence suggesting that higher rates of weight gain are associated with better outcomes.

Grading of the Overall Supporting Body of Research Evidence for Medical Stabilization, Nutritional Rehabilitation, and Weight Restoration for Patients With Anorexia Nervosa

On the basis of the limitations of the evidence for setting individualized goals set for weekly weight gain and target weight, no grading of the body of research evidence is possible.

Use of Supplementary Feeding Approaches to Promote Adequate Caloric Intake in Individuals With Anorexia Nervosa

No specific recommendation was made about the use of supplementary feeding approaches such as NGT feeding. In the expert survey, supplemental overnight tube feeding and NGT feeding, either continuous or bolus, were rated as moderately appropriate for adolescents and for adults whereas intravenous feeding was rated as less appropriate (see Appendix D). Available evidence from research studies is limited. One retrospective study (Robb et al. 2002) compared oral refeeding supplemented by nocturnal NGT feeding (N=52) to oral refeeding alone (N=48) in female inpatients with AN. The hospital

length of stay was comparable in the two groups, but individuals who received nocturnal NGT feeding had greater absolute weight gains. Nevertheless, the starting daily caloric intake in this study was only 1,200 kcal/day and final values for daily caloric intake were 3,255 kcal/day with supplemental NGT feeding as compared to 2,508 kcal/day with oral refeeding alone, suggesting that comparable weight gains could be achieved without NGT feeding with more aggressive oral refeeding. Another retrospective study (Agostino et al. 2013) used data from patients who had been admitted to an academic inpatient unit over an 8-year period and who had either received "bolus" feeding, with calories divided into oral meals and snacks (N=134), or initial continuous NGT feeding for approximately 10 days followed by transition to oral refeeding (N=31). Individuals who had received NGT feeding had greater initial rates of weight gain (2.3 lbs/week vs. 1.5 lbs/week) and a shorter length of stay than those who received oral feeding only; however, the mean weight gain during the admission was comparable for the two treatment groups. An additional prospective study (Rigaud et al. 2007a) used a randomized design in 81 subjects but staggered admissions to their inpatient nutrition unit based on the patients' assigned treatment. After 2 months of treatment, individuals who received NGT feeding took longer to relapse than those who did not (34.3+/-8.2 weeks vs. 26.8+/-7.5 weeks; $p<0.05$) and also had significantly greater increases in weight, fat-free mass, and caloric intake. Weekly weight gain averaged 2 lbs per week with oral refeeding alone as compared to approximately 3 lbs per week in the NGT feeding group, again suggesting that comparable weight gain could be achieved with more aggressive oral refeeding alone. These findings are consistent with the conclusions of a systematic review that used somewhat different inclusion criteria (Garber et al. 2016) and showed that supplementary NGT feeding could increase caloric intake and weight gain but that comparable rates of weight gain can also be achieved without NGT feeding. Adverse effects of NGT feeding were not well studied, but NGT feeding was noted to cause some psychological distress, particularly early in treatment (Rigaud et al. 2007a). Expert opinion also suggests that some patients have difficulty transitioning to oral intake for caloric needs once NGT feeding is initiated. Excessive reliance on NGT feeding can also be an impediment to developing a broader repertoire of food choices with a balance of macronutrient content, each of which can affect outcomes (Schebendach et al. 2008, 2011, 2012).

Use of Medication to Support Weight Gain During Nutritional Rehabilitation

No specific recommendation was made about the use of medication to support weight gain during nutritional rehabilitation. In the expert survey (see Appendix D), the use of a second-generation antipsychotic medication had a median rating of moderately appropriate but other medications including antidepressants, metoclopramide, benzodiazepines, and anticonvulsants were rated as less appropriate or inappropriate.

The most robust evidence for the use of a medication in AN is for olanzapine. Attia and colleagues (Attia et al. 2019) conducted a randomized, double-blind, placebo-controlled, multicenter, outpatient trial of olanzapine in individuals who had not consistently gained weight over the preceding 4 weeks. Participants received olanzapine (N=75) or a comparable number of placebo pills (N=77) with treatment begun at 2.5 mg/day for 2 weeks, increased to 5 mg/day for 2 weeks, and then increased to the maximum dose of 10 mg/day for the remainder of the 16-week trial. Attrition was significant with only 55% of the sample completing the trial but did not differ between olanzapine and placebo groups.

Despite this, the intent-to-treat analysis showed no differences in adverse effects between the olanzapine and placebo groups, modest benefits of olanzapine on weight gain, and a non-significant tendency for participants who received olanzapine to be described as much or very much improved. A smaller (N=34) randomized double-blind, placebo-controlled trial in day hospital patients (Bissada et al. 2008) also showed greater weight gain with flexibly dosed olanzapine (2.5 to 10 mg/day) during the 10-week trial. Participants treated with olanzapine also showed a greater diminution in obsessive symptoms than participants who received placebo, and rates of adverse effects did not differ between the groups. An additional randomized, double-blind, placebo-controlled, outpatient trial (Brambilla et al. 2007) assigned participants to placebo or to olanzapine 2.5 mg daily for 1 month followed by olanzapine 5 mg daily for 2 months. All 30 participants received CBT for the duration of the trial. At study endpoints, both groups had experienced an increase in BMI but there was no additional benefit of olanzapine over placebo.

SSRIs and other antidepressant medications are sometimes used in individuals with AN to address co-occurring disorders. However, it was difficult to draw conclusions from studies of fluoxetine and citalopram during nutritional rehabilitation because small sample sizes, large attrition rates, and observational study designs contributed to significant study biases (Barbarich et al. 2004; Fassino et al. 2002; Halmi et al. 2005; Kaye et al. 2001; Ruggiero et al. 2003). An additional randomized, controlled, double-blind study (N=93) examined the effects of flexibly dosed fluoxetine (up to 80 mg/daily) as compared to placebo in weight stabilized patients (Walsh et al. 2006). No differences between the groups were found at 52 weeks in terms of maintenance of body weight, although only 42% of the initial sample completed the study.

Interventions to Promote Optimal Bone Density in Individuals With Anorexia Nervosa and Amenorrhea

No specific recommendation was made about the use of specific interventions to improve bone density or prevent further deterioration in bone density for individuals with AN who have had at least 6 months of amenorrhea.

The experts in the GWG and comments received in the expert survey suggested that the initial focus of treatment should be weight restoration (see Appendix D). In addition, in the expert survey, calcium and vitamin D supplementation were rated as highly appropriate and moderate exercise was rated as moderately appropriate in individuals without a history of compulsive exercising. Use of hormone replacement therapy or bisphosphonates were rated as minimally appropriate or inappropriate.

The research evidence on interventions to maintain or improve bone density in individuals with AN is relatively limited. In terms of the effects of bisphosphonates, a 1-year RCT of alendronate 10 mg daily as compared to placebo in 32 adolescents with AN and amenorrhea did not produce statistically significant effects on femoral neck or lumbar spine BMDs as measured by DXA (Golden et al. 2005). All participants also received vitamin D 400 IU and calcium 1,200 mg daily. The authors did note that an increase in body weight was the most important contributor to improvement in BMDs. Another 1-year RCT of 77 outpatients with AN compared risedronate 35 mg weekly, low-dose transdermal testosterone replacement therapy, combination therapy, and placebo (Miller et al. 2011). Although testosterone

therapy increased lean body mass by 1.9%, it did not affect BMD and there was no change in overall body mass in either group. Risedronate did increase BMD in the hip and in the posteroanterior and lateral spine, but changes were modest (2% to 4%) and of unclear clinical significance. An additional RCT (N=41) compared etidronate 200 mg daily to the combination of calcium 600 mg daily and alfacalcidol 1 mg daily, and to placebo (Nakahara et al. 2006) and found benefits for both active treatments at 3 months, but the primary outcome measure was tibial speed of sound making the findings difficult to generalize.

Dehydroepiandrosterone (DHEA) has also been studied in a small double-blind, randomized, controlled trial of outpatients with AN (Bloch et al. 2012). Participants (N=26) were treated for 6 months with DHEA 100 mg daily or placebo and both treatment groups also received calcium carbonate 600 mg and vitamin D3 200 IU daily; however, no effects on bone density were noted. Other studies of DHEA have used combination interventions (DHEA 50 mg and concomitant 20 µg ethinyl estradiol/0.1 mg levonorgestrel combined oral contraceptive; Divasta et al. 2012, 2014) or compared DHEA to other active comparators (Gordon et al. 2002).

Several trials have examined the effects of estrogen and progestin on BMD. Klibanski and colleagues (Klibanski et al. 1995) followed subjects with AN and amenorrhea for an average of 1.5 years after random assignment to replacement therapy (Premarin 0.625 mg on days 1 to 25 with Provera 5 mg on days 16 to 25 or oral contraceptive containing 35 µg ethinyl estradiol, based on patient preference) or no replacement therapy. Estrogen therapy was not associated with significant changes in BMD although increased body weight was associated with improved BMD. In another double-blind study, 22 participants with AN were randomly assigned to transdermal estradiol (plus cyclic medroxyprogesterone) or placebo in a study for 12 months (Faje et al. 2012). Although a statistically significant difference was noted in the estradiol treatment group, the clinical impact of the treatment was unclear, and the primary research focus was on the role of sclerostin as a mediator of treatment effects. Based on patient preference, Golden and colleagues (Golden et al. 2002) assigned subjects to oral contraceptive containing 20 to 35 µg ethinyl estradiol (N=22) or standard treatment, which included calcium supplementation (N=28). After a mean length of follow-up of 23.1 months, there was no difference in bone density between the treatment groups although bone density improvements were noted in participants whose body mass increased during the trial. Misra and colleagues (Misra et al. 2011) randomly assigned mature girls with AN (N=96) to placebo or to transdermal beta-estradiol 100 mg patch applied twice weekly with medroxyprogesterone 2.5 mg prescribed daily for 10 days each month. Although participants in the active treatment group had a greater increase in BMD in spine as compared to placebo, changes in BMD were associated with change in weight and inversely associated with height, baseline age, and years since menarche.

An additional approach to osteopenia or osteoporosis that has been studied in individuals with AN is teriparatide, which is the biologically active N-terminus portion of parathyroid hormone. In a small study (N=21), participants were randomly assigned to teriparatide 20 µg subcutaneously or placebo (Fazeli et al. 2014). At 6 months, there was a significant increase in posteroanterior and lateral spine BMD but no difference in BMD in the hip or femoral neck. The treatment was well tolerated but additional replication of these findings is needed.

Statement 11 – Psychotherapy in Adults With Anorexia Nervosa

APA recommends (1B) that adults with anorexia nervosa be treated with an eating disorder-focused psychotherapy, which should include normalizing eating and weight control behaviors, restoring weight, and addressing psychological aspects of the disorder (e.g., fear of weight gain, body image disturbance).

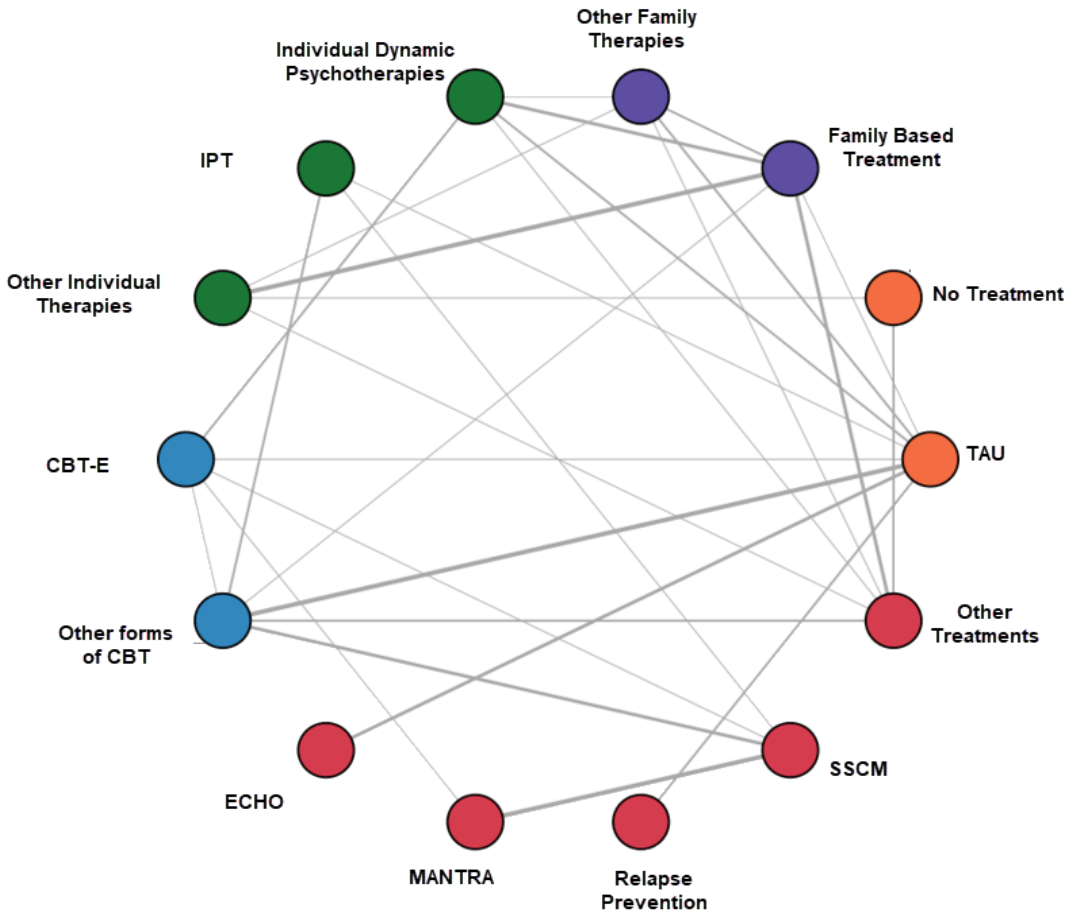
Support for this statement comes from the expert survey (Appendix D) and from an NMA of studies of psychotherapies in AN. In the expert survey, psychotherapy alone was rated as highly appropriate as an initial intervention in all age groups. In terms of specific psychotherapies, individual CBT was rated as highly appropriate as was FBT for adolescents. Other approaches to family or couples therapy and group therapies were rated as moderately appropriate as were individual IPT, supportive therapy, and SSCM. Psychodynamically-informed psychotherapy was rated as less appropriate.

In the NMA, as compared to no treatment, weight-related outcomes (i.e., BMI change, weight change, % IBW) were improved with family therapies, CBT-E, other forms of CBT, MANTRA, individual dynamic psychotherapies (e.g., focal psychoanalytic psychotherapy, FPT, ego-oriented individual therapy, or cognitive analytical therapy), SSCM, and treatment as usual (TAU). Relapse prevention therapy and ECHO also showed benefits as compared to no treatment. Benefits were also observed with CBT-E, individual dynamic psychotherapy, and FBTs that place parents in charge of their child's eating as compared to TAU. Furthermore, as compared to no treatment or TAU, weight-related outcomes improved with a heterogenous set of other therapies (e.g., group outpatient therapy with dietary counselling, family group psychoeducation, multidisciplinary outpatient psychotherapy, dietary advice, cognitive remediation therapy followed by CBT, art therapy in combination with FBT, nutritional counseling, educational behavioral therapy, body awareness therapy in combination with family therapy) that were grouped together for analysis. Consequently, the strength of research evidence is rated as moderate.

Network Meta-Analysis of Psychotherapies for Anorexia Nervosa

An NMA of psychotherapies for AN was conducted using the studies identified as described in Appendix B. Overall, the NMA contained 36 trials (42 publications) with 90 treatment arms and 3,560 subjects. The network was well-connected, with most treatments connected to multiple other treatments. A number of studies compared treatments that were similar enough to have been grouped together for purposes of the NMA (Allan et al. 2018; Eisler et al. 2000, 2007; Herscovici et al. 2017; Le Grange et al. 2016; Lock et al. 2005, 2006a, 2015b; Madden et al. 2015). These endonodal studies were not included in the NMA.

Figure C-1. Network graph of psychotherapies for AN



Note: Nodes represent a treatment. Node colors indicate broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: CBT=cognitive-behavioral therapy; CBT-E=enhanced cognitive-behavioral therapy for eating disorders; ECHO=Experienced Caregivers Helping Others; IPT=interpersonal psychotherapy; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; SSCM=Specialist Supportive Clinical Management; TAU=treatment as usual

In terms of the baseline characteristics of participants in studies of AN, the majority of participants were female (86-100% among 33 trials), with a mean age of about 20 years (range 13 to 34 years among 30 trials). Baseline mean BMI values ranged from 14 to 18 kg/m^2 (24 trials) with a mean weight of 82 to 103 lbs in the 9 trials that reported weight instead of BMI.

There was some heterogeneity in subject ages, but no concerns about heterogeneity in terms of sex, BMI, or weight. The greatest variability among studies related to the duration of follow-up assessments. All time points were used for analyses, but sensitivity analyses of age and follow-up time were conducted to identify possible effects of follow-up time on outcomes. In general, these did not show relevant differences as compared to the original analysis.

Table C-1. AN NMA feasibility and network characteristics

Outcome	Interventions: Total (NMA)	Studies: Total (NMA)	Trials per direct comparison	Subjects per comparison	Total Subjects in NMA	Network connected? (Number of networks)	NMA feasible?
BMI change from baseline	12	17	1-3	12-134	1,426	Yes (1)	Yes
Weight change from baseline	8 (6)	6 (4)	1	14-82	307	No	Yes
Percent IBW	5	4	1	12-82	168	Yes (1)	Yes
Depression change from baseline	8	5	1	34-128	N/A	No (3)	Yes
Anxiety change from baseline	9	6	1	12-128	N/A	No (3)	Yes
Study withdrawal	12	12	1-2	12-130	950	Yes	Yes
Mortality	11	6		20-134	N/A	No (3)	No
Treatment adherence rate	8 (6)	5 (4)	1	16-72	459	No (2)	Yes
Treatment discontinuation	9 (6)	5 (4)	1	16-82	354	Yes	Yes
Disease response, recovery/remission	10 (7)	7 (6)	1-2	15-82	688	No (2)	Yes
Hospitalization	7	7	1	11-82	568	Yes	Yes
Eating disorder scale change	7	8	1-3	30-134	874	Yes	Yes
Social Functioning change	6	5	1-2	31-130	N/A	No (2)	Yes
Quality of Life	3	3			N/A	No (3)	No

Abbreviations: AN=anorexia nervosa; BMI=body mass index; IBW=ideal body weight IndividualDyn=individual dynamic psychotherapies; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; N/A=not available; NMA=network meta-analysis; SSCM=Specialist Supportive Clinical Management; TAU=treatment as usual

Table C-2. Statistically favored comparisons from the AN NMA

Intervention	Comparison	Outcome	Statistical values
FBT	No treatment	BMI change	RMD 2.81 (0.95, 4.64)
	TAU	Percent IBW	RMD 4.24 (0.32, 8.04)
	Other individual therapies	Recovery/remission	RR 1.92 (1.12, 3.56)
Other family therapies	No treatment	BMI change	RMD 2.53 (0.50, 4.58)
	Other individual therapies	Hospitalization	RR 0.14 (0.03, 0.53)
CBT-E	No Treatment	BMI change	RMD 2.35 (0.51, 4.23)
		Weight change	RMD 15.24 (1.74, 28.79)
	TAU	Weight change	RMD 6.11 (0.48, 11.95)
		Eating disorder scale change	RMD -0.12 (-0.21, -0.03)
		Recovery/remission	RR 3.63 (1.27, 13.90)
	Individual dynamic therapies	Eating disorder scale change	RMD -0.19 (-0.28, -0.10)
Other forms of CBT	No treatment	BMI change	RMD 2.24 (0.48, 4.06)
		Weight change	RMD 16.27 (1.69, 30.87)
	TAU	Recovery/remission	RR 2.00 (1.30, 3.36)
	Individual dynamic therapies	Eating disorder scale change	RMD -0.66 (-1.32, -0.00)
		Hospitalization	RR 0.43 (0.21, 0.81)
Individual dynamic therapies	No treatment	BMI change	RMD 2.60 (0.74, 4.51)
		Weight change	RMD 15.72 (3.11, 28.46)
	TAU	BMI change	RMD 0.37 (0.03, 0.73)
		Weight change	RMD 6.65 (1.34, 12.10)
		Recovery/remission	RR 3.58 (1.32, 13.38)
	Other individual therapies	Hospitalization	RR 0.19 (0.04, 0.61)

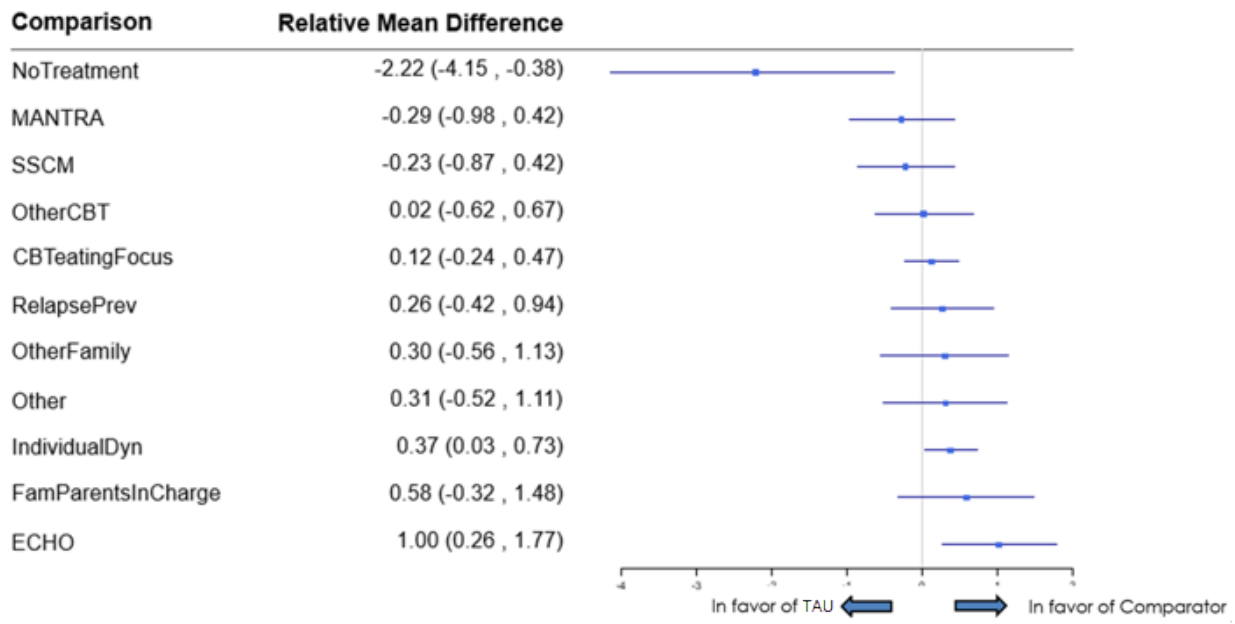
MANTRA	No treatment	BMI change	RMD 1.96 (0.09, 3.91)
	SSCM	Hospitalization	RR 1.28 (1.02, 1.66)
SSCM	No treatment	BMI change	RMD 2.01 (0.14, 3.88)
Other individual therapies	FBT	Study withdrawal	RR 0.29 (0.07, 0.89)
	Other therapies	Study withdrawal	RR 0.20 (0.03, 0.96)
Other therapies	No treatment	BMI change	RMD 2.52 (0.84, 4.23)
		Weight change	RMD 15.19 (5.02, 25.55)
	TAU	Percent IBW	RMD 9.28 (1.83, 16.36)
TAU	No treatment	BMI change	RMD 2.22 (0.38, 4.15)
	Relapse prevention therapy	Study withdrawal	RR 0.24 (0.11, 0.48)
Relapse prevention therapy	No treatment	BMI change	RMD 2.47 (0.53, 4.51)
ECHO	No treatment	BMI change	RMD 3.26 (1.24, 5.30)
	CBT-E	BMI change	RMD 0.89 (0.04, 1.74)
	MANTRA	BMI change	RMD 1.28 (0.27, 2.33)
	SSCM	BMI change	RMD 1.23 (0.24, 2.28)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; CBT-E=enhanced cognitive-behavioral therapy for eating disorders; ECHO=Experienced Carers Helping Others; FBT=family-based treatment; IBW=ideal body weight; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; NMA=network meta-analysis; RMD=relative mean difference; RR=relative risk; SSCM=Specialist Supportive Clinical Management; TAU=treatment as usual

Subgroup analyses were conducted based on age and follow-up duration. For BMI change from baseline among adolescents, FBT and the heterogenous group of other therapies were superior to individual dynamic psychotherapies based on 4 interventions in 3 studies including 63 participants. For BMI change from baseline among adults, all treatments were superior to no treatment and individual dynamic psychotherapies were superior to SSCM, MANTRA, and TAU (11 interventions in 15 studies including 1,312 participants). With a follow-up duration of 10 months, all treatments were superior to no treatment (11 interventions, 15 studies, 1,312 participants for BMI change from baseline; 5 interventions, 4 studies, 307 participants for weight change from baseline). With a follow-up duration of 20 months, CBT-E was superior to MANTRA, individual dynamic psychotherapies were superior to MANTRA and TAU, and FBT was superior to CBT-E, other forms of CBT, individual dynamic psychotherapies, MANTRA, and relapse prevention therapy for BMI change from baseline (10 interventions, 10 studies, 1,163 participants). For weight change from baseline at 20 months, there were no differences among the treatments (4 interventions, 2 studies, 237 participants). Taken together, these subgroup analyses should be interpreted with caution, because almost all of the treatment comparisons included only one study.

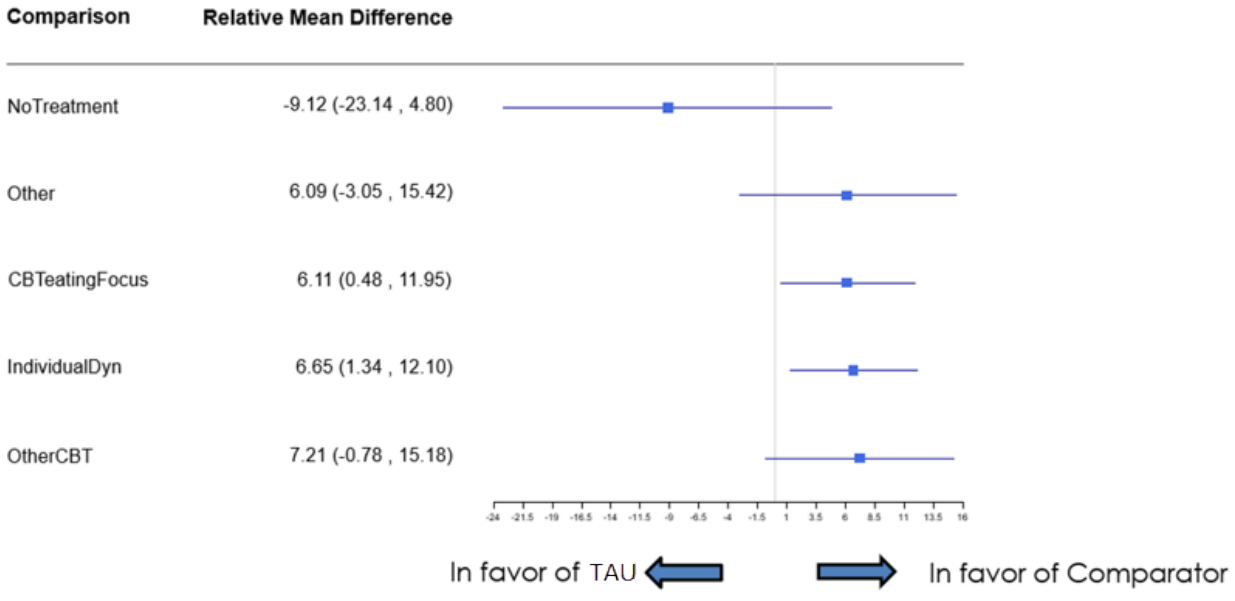
Forest Plots of Psychotherapies for Anorexia Nervosa

Figure C-2. BMI change from baseline at all time points



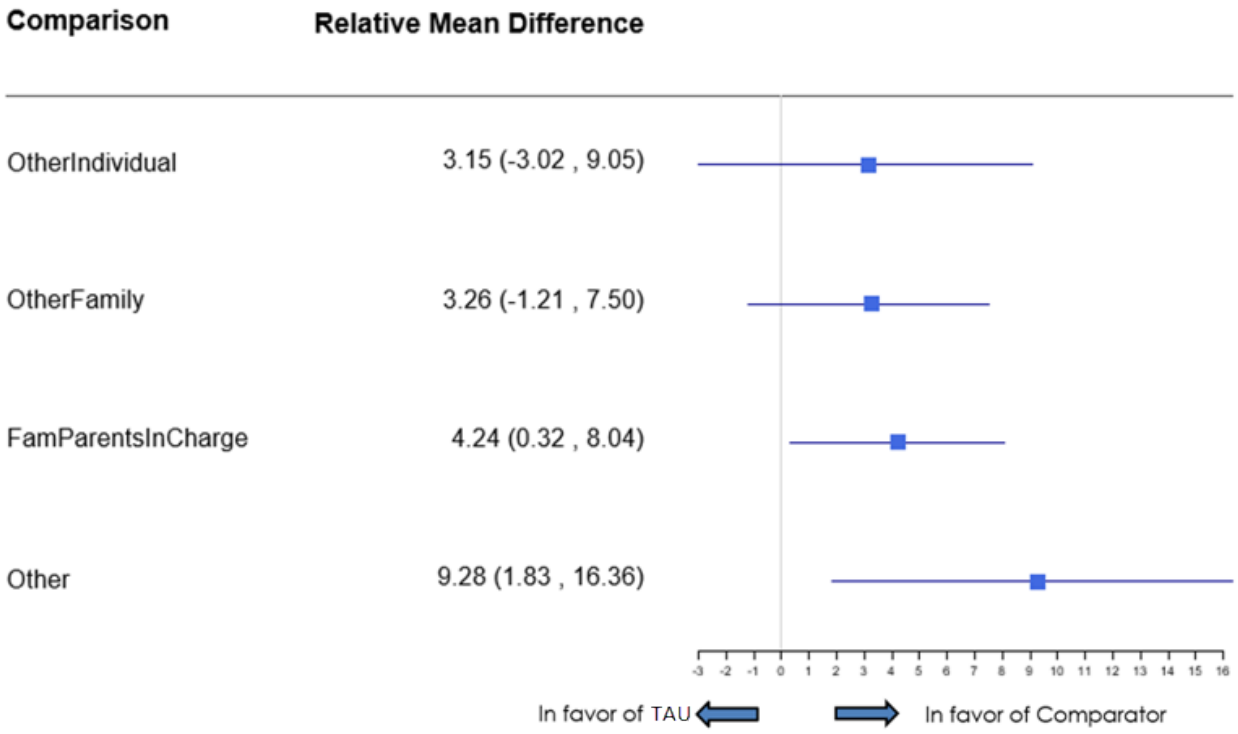
Abbreviations: BMI=body mass index; CBT=cognitive-behavioral therapy; ECHO=Experienced Carers Helping Others; FamParentsInCharge=family treatment with parents with in charge; IndividualDyn=individual dynamic psychotherapies; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; OtherFamily=other family therapies; RelapsePrev=relapse prevention therapy; SSCM=Specialist Supportive Clinical Management; TAU=treatment as usual

Figure C-3. Weight change from baseline at all time points



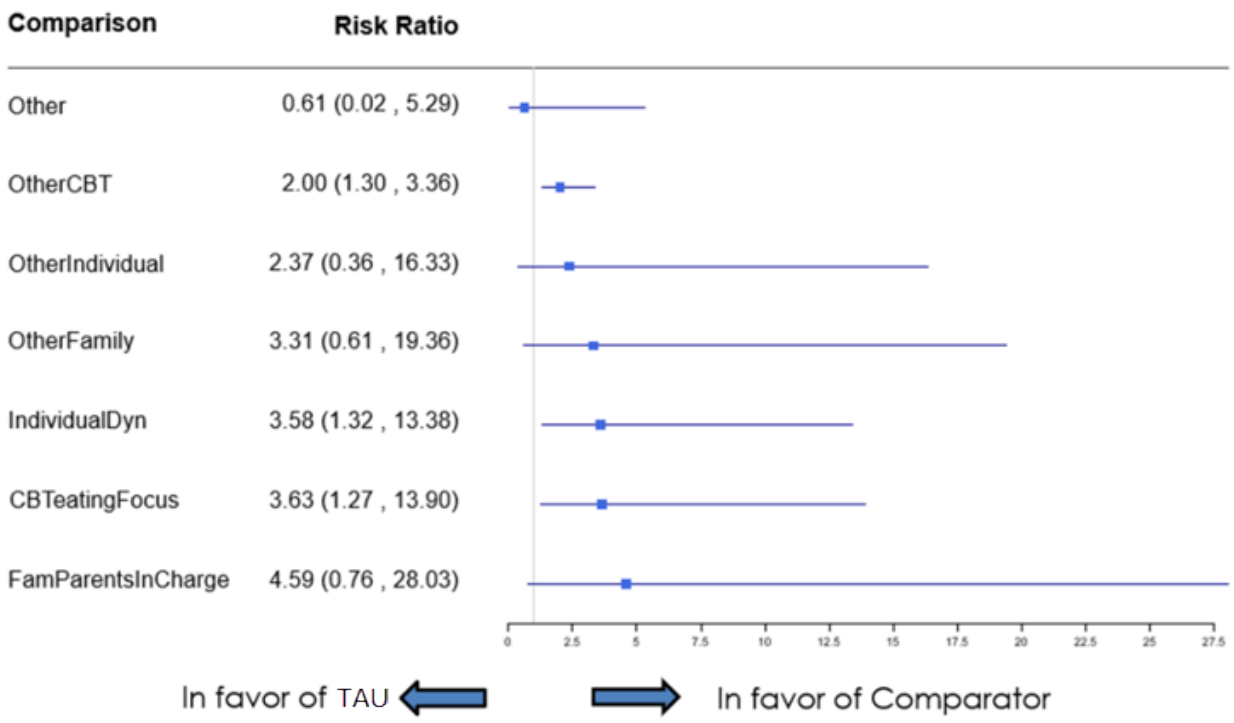
Abbreviation: CBT=cognitive-behavioral therapy; IndividualDyn=individual dynamic psychotherapies; TAU=treatment as usual

Figure C-4. Percent ideal body weight at all time points



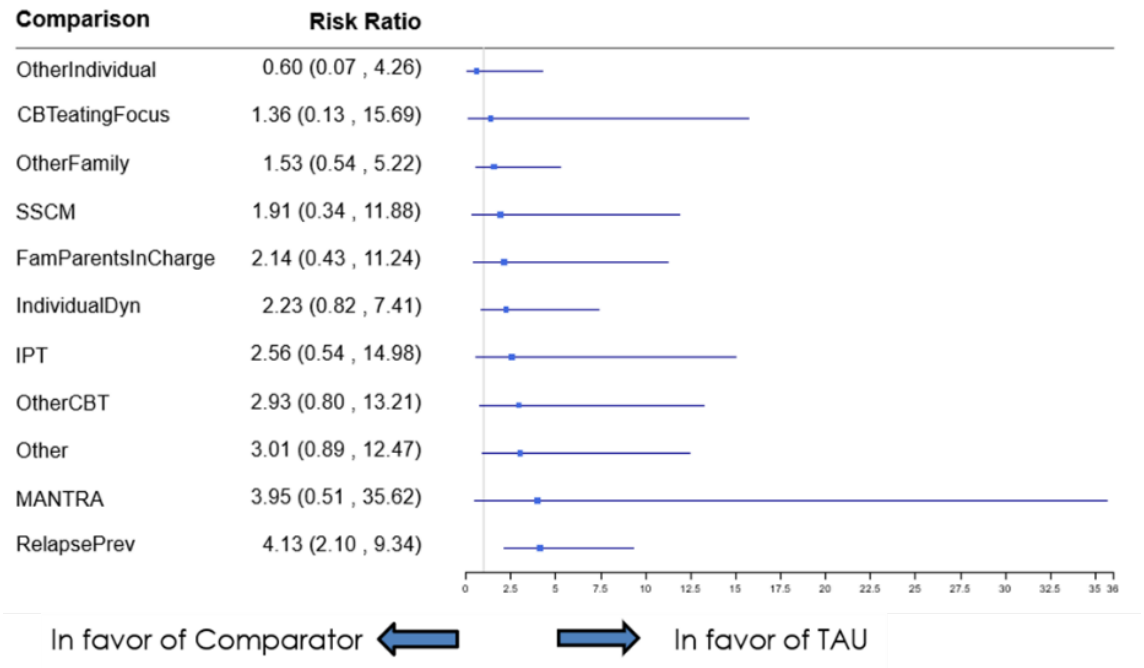
Abbreviation: FamParentsInCharge=family treatment with parents with in charge; OtherFamily=other family therapies; OtherIndividual=other individual therapies; TAU=treatment as usual

Figure C-5. Disease response recovery/remission at all time points



Abbreviation: CBT=cognitive-behavioral therapy; FamParentsInCharge=family treatment with parents with in charge; IndividualDyn=individual dynamic psychotherapies; OtherFamily=other family therapies; OtherIndividual=other individual therapies; TAU=treatment as usual

Figure C-6. Study withdrawal at all time points



Abbreviations: CBT=cognitive-behavioral therapy; FamParentsInCharge=family treatment with parents with in charge; IndividualDyn=individual dynamic psychotherapies; IPT=interpersonal psychotherapy; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; OtherFamily=other family therapies; OtherIndividual=other individual therapies; RelapsePrev=relapse prevention therapy; SSCM=Specialist Supportive Clinical Management; TAU=treatment as usual

Detailed Review of Evidence: Cognitive-Behavioral Therapy

In adults with AN, the largest body of evidence is related to CBT. In the NMA, as compared to no treatment, CBT was associated with greater changes in BMI (CBT-E RMD 2.35; 95% CI 0.51, 4.23; other forms of CBT RMD 2.24; 95% 0.48, 4.06) and in weight (CBT-E RMD 15.24; 95% CI 1.74, 28.79; other forms of CBT RMD 16.27; 95% CI 1.69, 30.87). Rates of recovery or remission were also higher for CBT as compared to TAU (CBT-E RR 3.63; 95% CI 1.27, 13.90; other forms of CBT RR 2.00; 95% CI 1.30, 3.36).

The Anorexia Nervosa Treatment of Outpatients (ANTOP) study was conducted in Germany and randomly assigned 242 adults with AN to 10 months of treatment with CBT-E, FPT, or optimized TAU (Zeeck et al. 2018; Zipfel et al. 2014). All of the groups showed modest improvements in weight-related outcomes at the end of treatment and 3 months and 12 months after treatment discontinuation. For example, mean BMI values at the end of treatment were 18.2 with FPT, 18.1 with CBT-E, and 17.95 with optimized TAU. Mean BMI values were slightly higher in the CBT-E group at the end of treatment, but this difference did not persist. Rates of attrition with CBT were less than with TAU. Dalle Grave and colleagues (Dalle Grave et al. 2013a) compared focused CBT-E to complex broad CBT, which also addressed features such as mood intolerance, clinical perfectionism, low self-esteem, or interpersonal difficulties. In this Italian inpatient sample, which included 80 participants with AN who were randomly assigned to 20 weeks of treatment, both groups showed increased weight and BMI. However, no differences were found between the two CBT approaches at the end of treatment or at 44-week or 72-week follow-up assessments. Byrne and colleagues (Byrne et al. 2017) in an Australian multi-center trial randomly assigned 120 participants with AN to 10 months of treatment with CBT-E, MANTRA, or SSCM. Although treatment was only completed by 60% of participants, all three treatments were associated with improvements in weight and associated features of eating disorder psychopathology. At 12-month follow-up, outcomes continued to be comparable. In a small Australian RCT of adolescents and young adults (Ball and Mitchell 2004), 25 participants were randomly assigned to 12 months of outpatient treatment with either CBT or behavioral family therapy (based on behavioral family systems therapy). Both groups showed improvements in weight, with 60% of participants in each group achieving a good outcome at the end of treatment and at 18-month follow-up. Lock and colleagues (Lock et al. 2013) conducted a feasibility study in which 46 participants with AN were randomly assigned to 8 weeks of initial treatment with either CBT or cognitive remediation therapy, with both groups subsequently receiving an additional 16 weeks of CBT. Although attrition during the initial phase of the trial was greater for individuals who received CBT, overall attrition and overall weight-related outcomes did not differ for the two intervention approaches. Another RCT (Carter et al. 2011; McIntosh et al. 2005) compared 20 weeks of CBT (N=19) to 20 weeks of IPT (N=21) or SSCM (N=16) and found comparable weight related outcomes at the end of the study and at the end of long-term follow-up (mean 6.7 years). Although attrition at the end of long-term follow-up was substantial, attrition rates were not significantly different among the treatment groups. Touyz and colleagues (Stiles-Shields et al. 2013;

Touyz et al. 2013) randomly assigned 63 participants with longstanding AN (mean duration 16.6 years) to 8 months of treatment with either CBT or SSCM. At baseline, the mean BMI of the sample was 16.2 kg/m² and at 20 months both groups had had a mean change in BMI of 0.7 kg/m² (SD ± 1.22) vs. 0.7 kg/m² (SD ± 1.29). The clinical implications of this study are unclear. On one hand, neither group had a worsening of average BMI during the study, yet the impact of treatment was relatively small. Also, despite the chronicity of illness in the participants, their prior treatment history was unclear. Together with the findings of the NMA, these studies suggest that CBT is beneficial in individuals with AN but that differing forms of CBT have comparable efficacies as do other psychotherapies in comparison with CBT.

Two studies have examined the effects of CBT in individuals who have been stabilized in an inpatient eating disorders treatment program. Pike and colleagues (Pike et al. 2003) randomly assigned 33 participants to 10 sessions of either CBT or nutritional counseling at the time of discharge at which point their weight had reached at least 90% of their IBW. Each group received 50 sessions of treatment during the year after discharge. Treatment with CBT was more likely to yield a good response (44% vs. 7%) or a complete response (17% vs. 0%) at 1 year as compared to nutritional counseling and participants in the nutritional counseling group were more likely to withdraw from the study (0% vs. 20%) during the treatment phase. An additional study examined whether CBT (N=46) either alone or in combination with fluoxetine differed from TAU (N=42) in participants who had been stabilized in a hospital-based program and maintained a minimum BMI of 19.5 kg/m^2 for several weeks (Carter et al. 2009). Participants in the TAU group were not randomly assigned but instead chose not to participate in the fluoxetine vs. placebo portion of the trial, did not meet other inclusion criteria, or completed the hospital-based program after the CBT trial had ended. In the CBT arm, outcomes for fluoxetine and placebo did not differ (Walsh et al. 2006). Relapse was defined as BMI ≤ 17.5 for 3 months or resumption of regular binge-eating or purging behaviors for 3 months; with either definition, time to relapse was significantly greater in the CBT group. In addition, a substantially greater proportion of participants who received CBT remained in remission from AN at 1 year (65% with CBT vs. 34% with TAU). These studies suggest that CBT may reduce the risk of relapse or prolong the time to relapse for AN after discharge from inpatient eating disorders treatment.

Detailed Review of Evidence: Specialist Supportive Clinical Management;

As described in the discussion of CBT, SSCM has been used as an active comparator group in a number of trials and has shown comparable weight related outcomes to CBT (Byrne et al. 2017; Carter et al. 2011; McIntosh et al. 2005; Stiles-Shields et al. 2013; Touyz et al. 2013), IPT (Carter et al. 2011; McIntosh et al. 2005), and MANTRA (Byrne et al. 2017; Schmidt et al. 2012, 2015, 2016). In the NMA, SSCM showed a greater change in BMI than no treatment (RMD 2.01; 95% CI 0.14, 3.88)

Detailed Review of Evidence: Maudsley Model of Anorexia Nervosa Treatment for Adults

MANTRA has also been studied as a treatment for AN. The NMA shows benefit for MANTRA in the change in BMI (RMD 1.96; 95% CI 0.09, 3.91) as compared to no treatment.

Schmidt and colleagues conducted two studies of MANTRA in the United Kingdom. In the initial study (Schmidt et al. 2012), 71 participants with a BMI of <18.5 kg/m^2 and a diagnosis of either AN or eating disorder not otherwise specified were randomly assigned to 20 weekly sessions of either MANTRA or

SSCM. Both groups showed improvement with treatment but there were no significant differences between the treatments in weight or BMI change at the end of treatment or at 1 year. Those treated with MANTRA were more likely to have received day treatment or been hospitalized (RR 1.28; 95% CI 1.02, 1.66), but absolute numbers were small. In the subsequent Maudsley Outpatient Study of Treatments for Anorexia Nervosa and Related Conditions (MOSAIC) study (Schmidt et al. 2015, 2016), 142 participants were randomly assigned to MANTRA or to SSCM. Participants received 20 to 30 sessions of treatment depending on clinical severity and individuals in either treatment group could also receive additional dietician or carer sessions. As in the initial study, both treatment groups showed improvements in weight and BMI and these improvements were maintained at 12 months. A follow-up study that included 73% of the initial sample also showed no difference between outcomes in the two groups although both groups showed higher proportions of individuals who had recovered at 24 months as compared to rates of recovery at the end of treatment. This finding of comparable effects of MANTRA and SSCM is also consistent with the conclusions of Byrne and colleagues' comparison of CBT-E, MANTRA, or SSCM (Byrne et al. 2017).

Detailed Review of Evidence: Family Therapies in Adults

The majority of studies of family therapy have been conducted in adolescents and emerging adults (see Appendix C, Statement 12). Dare and colleagues (Dare et al. 2001) studied adults with AN (mean age 26.3 years) and compared 1 year of family therapy (N=22) to 7 months of cognitive analytic therapy (N=22), 1 year of focal psychoanalytic psychotherapy (N=21), or 1 year of low contact routine treatment (N=19). A greater proportion of those treated with family therapy or focal psychoanalytic psychotherapy showed significant improvement as compared to TAU; however, the number of participants in each group was relatively small, attrition rates varied among treatment groups, and the differences in treatment duration introduced confounding effects. Russell and colleagues (Russell et al. 1987) also studied family therapy in comparison with individual supportive therapy in adolescents and adults with AN and in individuals with BN. Among the participants with AN who were assessed at a 5-year follow-up (Eisler et al. 1997), family therapy appeared to be more beneficial in those with an onset of illness before age 18 and a short illness duration whereas individual supportive therapy appeared to be more effective in those with a later onset of illness. However, the numbers in each subgroup were small and features of the study design were associated with a high risk of bias. Based on these findings, the evidence for family therapy in adults is contradictory and limited.

Detailed Review of Evidence: Experienced Carers Helping Others

Three studies examined the effects of ECHO, a family psychoeducation program that includes a book and 5 DVDs. Hodsoll and colleagues (Hodsoll et al. 2017) identified individuals with AN or atypical AN who were 13 to 20 years of age and randomly assigned a participating family member to receive TAU (N=50), unguided ECHO (N=49), or ECHO paired with telephone guidance (N=50). The two ECHO groups had comparable effects and seemed to have more carer skills and fewer accommodating and enabling behaviors. Despite these differences, the effect on body mass of the adolescent with AN was minimal and adherence to the ECHO intervention was limited. Salerno and colleagues (Salerno et al. 2016) included 149 participants aged 12 to 21 years with AN or atypical AN and 225 of their caregivers. Families were randomly assigned to TAU, unguided ECHO, or ECHO paired with 5 sessions of telephone

guidance. The study suggested that the ECHO intervention may reduce the association of participant and caregiver distress at 1-year follow-up but there was no predictive ability in terms of changes in body mass with ECHO as compared to TAU. Magill and colleagues (Magill et al. 2016) in the United Kingdom conducted a follow-up assessment of 178 participants with AN and 268 caregivers who were randomly assigned at discharge from an inpatient program to receive TAU or ECHO, with the ECHO intervention accompanied by a maximum of 10 telephone guidance sessions per family. At 24 months, the differences between the groups did not reach statistical significance but approximately one-third of the sample had been lost to attrition. Together, these studies suggest a potential for modest effects in caregiver experience and the NMA suggests a benefit of ECHO as compared to no treatment in change in BMI (RMD 3.26; 95% CI 1.24, 5.30). Nevertheless, there is a significant risk of bias in these conclusions due to study attrition and low adherence to the ECHO intervention.

Detailed Review of Evidence: Individual Dynamically-Oriented Psychotherapy

A number of studies compared individual dynamically-oriented psychotherapies to other psychotherapeutic approaches or to TAU. In the NMA, the individual dynamically-oriented psychotherapies were associated with a significant change in BMI as compared to no treatment (RMD 2.60; 95% CI 0.74, 4.51) or TAU (RMD 0.37; 95% CI 0.03, 0.73), significant change in weight as compared to no treatment (RMD 15.72; 95% CI 3.11, 28.46) or TAU (RMD 6.65; 95% CI 1.34, 12.10), and a significant increase in the likelihood of recovery or remission (RR 3.58; 95% CI 1.32, 13.38) as compared to TAU. As described in the discussion of CBT, the German ANTOP study randomly assigned 242 adults with AN to 10 months of treatment with CBT-E, FPT, or optimized TAU (Zeeck et al. 2018; Zipfel et al. 2014). All of the groups showed modest improvements in weight-related outcomes during the study, with additional improvements noted 3 months and 12 months after the end of treatment. Although mean BMI values were slightly higher in the CBT-E group at the end of treatment, weight restoration was incomplete and mean BMI was less than 20. Despite these modest effects, there was a statistically significant increase in the rates of AN remission in the FPT group at 12 months after the end of treatment. Rates of attrition were also lower with FPT as compared to the TAU group. Dare and colleagues at the Maudsley Hospital (Dare et al. 2001) randomly assigned 84 participants with AN to 7 months of cognitive analytic therapy or to 1 year of family therapy, focal psychoanalytic psychotherapy, or low contact routine treatment. Although a greater proportion of those treated with family therapy or focal psychoanalytic psychotherapy showed significant improvement as compared to TAU, the number of participants in each group was relatively small, attrition rates varied among treatment groups, and the differences in treatment duration introduced confounding effects. Treasure and colleagues at the Maudsley Hospital (Treasure et al. 1995) randomly assigned 30 participants with AN to 20 weeks of outpatient treatment with cognitive analytical therapy or educational behavioral therapy. Both groups showed improvements in weight-related outcomes but there was significant attrition, and the sample size was small. A series of studies by Robin and colleagues (Robin et al. 1994, 1995, 1999) in the United States randomly assigned adolescents with AN to behavioral family systems therapy or ego-oriented individual therapy. An initial study which included 24 participants found a statistically greater increase in BMI with behavioral family systems therapy than with ego-oriented individual therapy, although other comparisons were limited by the small sample size (Robin et al. 1994, 1995). A subsequent study of 37 participants (Robin et al. 1999) did not find significant differences in weight-related outcomes between

the two treatments although a greater proportion of individuals treated with behavioral family systems therapy had a resumption of menses as compared to those treated with ego-oriented individual therapy. Overall, the individual dynamically-oriented psychotherapies show evidence of benefit, but studies have used different treatment methods and most have small sample sizes. In addition, there does not appear to be an advantage to these therapies as compared to CBT, behavioral approaches, or TAU in adults; findings in adolescents in comparison to behavioral family systems therapy were inconsistent.

Grading of the Overall Supporting Body of Research Evidence for Psychotherapy in Adults With Anorexia Nervosa

o Magnitude of effect: The magnitude of effect is low to moderate. In the results of the NMA, values of the RMD for weight change were approximately 15 lbs, relative to no treatment although CIs were wide. Similarly, values of the RMD for BMI change were 2.2 to 2.6. Values compared to TAU were less, but relative risk values for recovery or remission for effective psychotherapies compared to TAU were substantial and ranged from 2 to 3.63, albeit with wide CIs. With longer follow-up times, the magnitude of the effect was less pronounced but still clinically significant in many studies.

o Risk of bias: The risk of bias for the supporting body of research evidence is moderate. Of the RCTs on psychotherapy other than FBT in AN in adults, 7 studies had a low risk of bias, 5 had a moderate risk of bias, and 2 had a high risk of bias. A high risk of bias was most often associated with an inappropriate method of random assignment or a lack of specification of the method that was used.

o Applicability: The included studies all involve individuals with AN diagnosed using DSM criteria. The majority of the psychotherapy studies were in outpatient settings although some studies enrolled participants during an inpatient hospitalization or at the time of hospital discharge. Almost all of the studies were conducted in the US, the UK, Europe, or Australia. Although health system policies differ among these countries, the findings are expected to be generally applicable to US and Canadian patients. Study participants are primarily young, white, and female, with many studies only including women participants. Applicability of the evidence to older individuals and individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low rates of inclusion, limiting applicability of the findings.

o Directness: Direct. Although the majority of studies included a large number of outcome variables, almost all included a weight related outcome as a primary or secondary outcome measure. Rates of disease response or recovery were also included in a number of studies.

o Consistency: Consistency was variable, when it could be determined. However, in the NMA, for active interventions or active comparators, there was a consistent benefit of treatment demonstrated as compared to TAU or to a wait list control groups.

o Precision: Imprecise. For comparisons in the NMA, CIs were wide and overlapped each other.

o Dose-response relationship: There is insufficient information to determine whether there is a relationship between treatment response and treatment frequency or duration.

o Confounding factors (including likely direction of effect): For all psychotherapy studies, the participant and the therapist are aware of the treatment that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention.

o Publication bias: Due to the small number of studies for each treatment comparison, funnel plots were not able to be done to assess for publication bias. Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

o Overall strength of research evidence: The overall strength of research evidence is moderate, based on the results of the NMA, which included studies that typically had a low to moderate risk of bias. Studies that compared active interventions to TAU or to wait list control groups, showed consistent benefits of psychotherapy in AN. Studies that included an intervention and an active comparator constituted the bulk of the research, and these typically showed benefits of each treatment but no consistent superiority of one treatment as compared to another. Nevertheless, these treatment by time effects provide additional supportive evidence of treatment effects.

Statement 12 – Family-Based Treatment in Adolescents and Emerging Adults With Anorexia Nervosa

APA recommends (1B) that adolescents and emerging adults with anorexia nervosa who have an involved caregiver be treated with eating disorder-focused family-based treatment, which should include caregiver education aimed at normalizing eating and weight control behaviors and restoring weight.

Support for this statement comes from the expert survey (Appendix D) and from an NMA of studies of psychotherapies in AN (Appendix C, Statement 11). In the expert survey, FBT was rated as highly appropriate as an initial intervention in adolescents. The expert survey did not specifically assess the appropriateness of interventions for emerging adults, ages 18-26 years of age. The NMA also did not address treatment of emerging adults but, in adolescents, found that FBTs that included placing the family in charge of the patients' eating led to greater changes in BMI than no treatment (RMD 2.81; 95% CI 0.95, 4.64) and greater changes in %IBW than TAU (RMD 4.24; 95% CI 0.32, 8.04). Consequently, the strength of research evidence is rated as moderate.

Detailed Review of Evidence: Family-Based Treatment

Dare and colleagues (Dare et al. 1990) compared family therapy (10.5 +/- 8.9 sessions in 1 year) to individual supportive therapy (15.9 +/- 8.5 in 1 year) in individuals with AN. Random assignment to one of these treatment groups occurred after the participants' weight was restored in an inpatient setting and after they were divided into subgroups that were felt likely to predict prognosis. For the subgroup who had an illness onset ≤ 18 years of age and an illness duration of <3 years, there was a significant difference in the proportion of individuals with a good or intermediate outcome (9 of 10 participants

[90%] treated with family therapy as compared to 2 of 11 [18%] treated with individual therapy). Individuals treated with family therapy had a percent average change in body weight of 25.5% as compared to a 15.5% average change in those who received individual supportive therapy. For the subgroup who had an illness onset ≤ 18 years of age and an illness duration of >3 years, the treatments were comparable; 4 of 10 participants (40%) treated with family therapy had a good or intermediate outcome as compared to 3 of 9 (33%) treated with individual therapy. In contrast, for those with an age of onset > 18 years of age, treatment with individual therapy yielded a good or intermediate outcome in 3 of 7 participants (43%) as compared to 1 of 7 (14%) treated with family therapy. Individuals treated with individual supportive therapy had a percent average change in body weight of 19.9% as compared to a 5.4% change in those who received family therapy. The authors did note that, with the older participants, the emphasis of family therapy was less on placing the family in charge of eating and more on eliminating use of symptoms as a form of communication; this shift in approach may affect interpretation of their findings. By the end of the study, one-third of participants had dropped out, but attrition was comparable in the two treatment groups. At 5-year follow-up, for the subgroup who had an illness onset ≤ 18 years of age and an illness duration of <3 years, there was a significant difference in the proportion of individuals with a good or intermediate outcome, 9 of 10 participants (90%) treated with family therapy as compared to 6 of 11 (55%) treated with individual therapy. Individuals treated with family therapy had an average percent body weight of 103.4% as compared to 94.4% in those who received individual supportive therapy. For the subgroup who had an illness onset ≤ 18 years of age and an illness duration of >3 years, the treatments were comparable; 4 of 10 participants (40%) treated with family therapy had a good or intermediate outcome as compared to 5 of 9 (55%) treated with individual therapy. The average percent body weight was 86.9% versus 95.7%, respectively. In contrast, for those with an age of onset > 18 years of age treatment with individual therapy yielded a good or intermediate outcome in 6 of 7 participants (86%) as compared to 4 of 7 (57%) with family therapy with corresponding values for average percent body weight of 97.5% and 93.7%, respectively.

Robin and colleagues randomly assigned adolescents with AN to behavioral family systems therapy or ego-oriented individual therapy. With 24 initial participants (Robin et al. 1994, 1995) found a statistically greater increase in BMI with behavioral family systems therapy than with ego-oriented individual therapy, but this was not confirmed when all 37 participants were included in the analysis (Robin et al. 1999). Lock and colleagues (Lock et al. 2010) compared twenty-four 1-hour sessions of FBT (N=61) to thirty-two 45-minute sessions of adolescent-focused individual therapy (which they note is comparable to the ego-oriented individual therapy of Robins and colleagues). For participants in the individual therapy group (N=60), up to 8 sessions could be held with parents. At the end of 1-year of treatment, the difference in full remission between the groups did not reach statistical significance (41.8% vs. 22.6%, p=0.055); however, the proportion of participants in remission was greater for FBT as compared to adolescent-focused individual therapy at 18 months (40% vs. 18%, p=0.029) and at 24 months (49% vs. 23%, p=0.024). Participants treated with FBT also were less likely to have been hospitalized by 24 months (15% vs. 37%, p=0.02). For those who achieved remission, outcomes were generally stable, regardless of the treatment that had been received (Le Grange et al. 2014b). Aspects of family functioning, such as communication, behavioral control, and affective involvement, also seemed to show more improvement with FBT than with adolescent-focused individual therapy (Ciao et al. 2015).

Agras and colleagues (Agras et al. 2014) randomly assigned 164 adolescents with AN to FBT or to systemic family therapy, which focused on the family system and its communication. Both interventions consisted of sixteen 1-hour sessions during 9 months of therapy. There were no significant differences between the groups in weight-related outcomes at the end of treatment or at 12-month follow-up (%IBW at 36 weeks 92.1% vs. 91.1% and at 88 weeks 94.6% vs. 93.3%; remission rates at 36 weeks 33.1% vs. 25.3% and at 88 weeks 40.7% vs. 39.0%). Nevertheless, participants who received FBT gained weight more quickly at the beginning of treatment, had fewer days of hospitalization, and had a lower cost of treatment than individuals who were treated with systemic family therapy.

In individuals aged 17 to 24 with AN, Nyman-Carlsson and colleagues (Nyman-Carlsson et al. 2020) compared a combination of family and individual therapy (N=37) with a maximum of 40 90-minute sessions to individual CBT (N=37) with a maximum of 60 1-hour sessions. Both treatments were associated with improvements at 18 months as compared to baseline, with no statistically significant differences between the treatments on measured outcomes. For example, BMI increased from 16.49 at baseline to 19.61 at the end of treatment in the CBT group and from 16.54 to 19.33, respectively, in the group receiving combined family and individual therapy. At follow-up, rates of recovery were 89% with CBT and 81% for combined family and individual therapy. Although data was able to be analyzed for all participants, attrition from treatment was substantial; only 32% of CBT treated participants and 51% of family/individual therapy participants attended at least 75% of the maximum number of sessions.

Additional studies of FBT have examined different delivery approaches or different durations of treatment. Eisler and colleagues (Eisler et al. 2000, 2007) provided FBT using a conjoint format with one-hour sessions (16.4 +/- 8.9 sessions in 12 months) or a separate format with distinct 45-minute sessions for the adolescent and for other family members (15.5 +/- 6.8 sessions in 12 months). For the sample as a whole, both groups showed improvement but there were no statistically significant differences in weight-related outcomes between the treatments. When the group was split based on the level of maternal expressed emotion (EE), there was no difference in treatment outcomes when maternal EE was low. When maternal EE was high, 8 of 10 participants had good or intermediate outcome with separated family therapy as compared to 2 of 7 participants with conjoint family therapy. At 5-year follow-up, more participants in the separated family therapy had normal menstrual function (95% vs. 72%). In other respects, outcomes were comparable with a good outcome in 72.2% of participants who received conjoint family therapy and 80% of those who received separated family therapy.

Le Grange and colleagues (Le Grange et al. 2016) compared FBT (N=55) to parent focused treatment (N=51) in which treatment sessions consisted of a 15-minute nurse visit with the adolescent and a 50-minute therapy session with the parents. With FBT, a 50-minute conjoint session followed weighing of the adolescent by the primary therapist. At the end of 18 sessions delivered over 6 months, remission rates with parent-focused treatment were greater than with FBT (43.1% vs. 21.8%, p=0.16). However, this difference dissipated by the 6-month and 12-month follow-up assessments and % median BMI did not differ between the groups at any of the assessment times. Using 63 adolescents, 89 mothers, and 64 fathers in the same sample, Allan and colleagues (Allan et al. 2018) analyzed audio recordings at the beginning and end of treatment to assess relationships between outcome and parental EE. Parent focused treatment was associated with a reduction in maternal criticism whereas an increase in

maternal criticism was more likely to occur with conjoint FBT. For the sample as a whole, remission was more likely to occur when EE was persistently low or decreased with treatment as compared to persistently high or increased EE during treatment.

Gabel and colleagues (Gabel et al. 2014) conducted a retrospective analysis of using a multiple family group format to conduct FBT with hospitalized adolescents. Addition of multiple family group treatment to TAU (N=25) as compared to TAU alone was associated with a significantly greater %IBW at 1-year follow-up (99.6% vs. 95.4%, $p<0.05$). Another small study in hospitalized adolescents (Geist et al. 2000) randomly assigned families to FBT (N=12) or family group psychoeducation (N=13) and found no differences in outcome at 4 months of treatment or upon hospital discharge.

A small pilot study (Lock et al. 2015b) attempted to enhance FBT by adding intensive parental coaching if participants had not begun to respond by 4 sessions of treatment. Addition of the parental coaching seemed to improve outcomes in the poorly responsive group and their outcomes became comparable to those in the rapidly responsive group. Nevertheless, only 12 families received intensive parental coaching in this adaptive design, making the data difficult to interpret. In another small study (Lock et al. 2018), treatment with FBT was augmented with either art therapy (N=15) or cognitive remediation (N=15) in an effort to improve response to FBT in adolescents with obsessive-compulsive features. Both groups showed improvements with treatment but there were no significant differences in weight-related outcomes. Although fewer individuals in the cognitive remediation group dropped out of treatment, both augmentation approaches were feasible.

Lock and colleagues (Lock et al. 2005) also assessed the impact of treatment duration on outcomes with FBT, randomly assigning 86 adolescents with AN to 10 sessions of FBT over 6 months or 20 sessions of FBT over 12 months. At 12 months, there were no differences in weight-related outcomes or EDE scores between the short-term and long-term treatment groups. However, those who were assigned to the long-term treatment had greater rates of dropping out of treatment than those assigned to short-term treatment (Lock et al. 2006b) At a later follow-up assessment (mean follow-up duration 3.96 years), the treatment outcomes remained comparable (Lock et al. 2006a).

Only one small study (Herscovici et al. 2017) has attempted to look at the role of different components of FBT in contributing to clinical response. In this study, participants were randomly assigned to have a family meal intervention (N=11) or no family meal intervention (N=12) as part of FBT. Both groups improved with 6 months of treatment but there were no significant differences between the treatment groups at the end of treatment or at 6-month follow-up.

Lock and colleagues (Lock et al. 2021) randomly assigned 40 adolescents aged 12 to 18 years to 15 60-minute FBT video sessions or 12 20-minute FBT-GSH sessions (N=40). The primary outcomes of recruitment, retention, and acceptability showed no differences between treatments and the findings were comparable to those of prior studies using face-to-face interventions. Both the FBT video and FBT-GSH groups improved in terms of the percent of expected body weight (%EBW; 84.47 at baseline, 92.97 at end of treatment, and 94.11 at 3-month follow-up for FBT video; 80.55, 90.80, and 93.10, respectively

for FBT-GSH). Parallel improvements were seen in global EDE scores with no significant differences found between the treatment conditions.

A retrospective study conducted in Sweden (Wallin and Holmer 2021) compared inpatient care to a family treatment apartment program in which the family assumes responsibility for meal support. Patients who had stayed in either program for at least 10 days between 1990 and 2009 were contacted about participating in follow-up assessments, which occurred an average of 14.2 years after treatment. 43 of 86 (50%) of patients admitted to the family treatment apartment program agreed to participate as compared to 25 of 63 (40%) of patients admitted to the inpatient program. The two groups were comparable in their %EBW on admission, but inpatient treatment was associated with a longer length of stay, higher weight gain per week, and higher %EBW at discharge, all of which were statistically significant. Readmission during follow-up was comparable for the two groups as was readmission within the first year, although more individuals who received inpatient treatment were readmitted in the first 6 months as compared to family treatment apartment participants. At follow-up assessments, 32% of the participants continued to meet criteria for AN. Although there was no difference in remission rate between the treatment groups, participants who were in the family treatment apartment program had better outcomes on several quality of life and symptom outcomes (e.g., RAND 36, SCL 90, Eating Disorders General Inventory General Psychological Maladjustment Component, Morgan Russell Outcome Assessment Schedule).

In addition to FBT, other approaches to family therapy in adolescents with AN have been studied (Ball and Mitchell 2004; Godart et al. 2012), but samples sizes have been small making it difficult to draw conclusions about these interventions.

For emerging adults, research on specific psychotherapies is limited; however, several small open-label studies show support for the use of FBT in individuals who have involved parents, guardians, or other care partners (Chen et al. 2016; Dimitropoulos et al. 2018).

Grading of the Overall Supporting Body of Research Evidence for Family-Based Treatment in Adolescents and Emerging Adults With Anorexia Nervosa

- Magnitude of effect: The magnitude of effect is moderate. In the NMA, the mean change in BMI as compared to no treatment was 2.81, the mean change in %IBW as compared to TAU was 4.24, and the relative risk value for recovery or remission as compared to other individual therapies was 1.92.

- Risk of bias: Of the RCTs on FBT in AN, 3 studies had a low risk of bias, 5 had a moderate risk of bias, and 4 had a high risk of bias. A moderate or high risk of bias was most often associated with an inappropriate method of random assignment, a lack of specification of the method that was used for random assignment, or bias due to outcome measurement, including missing outcome data.

- Applicability: The included studies all involve adolescents and emerging adults with AN, diagnosed using DSM criteria. The majority of the studies were in outpatient settings although some studies enrolled participants during an inpatient hospitalization or at the time of hospital discharge. Almost all of the studies were conducted in the US, the UK, Europe, or Australia. Although health system policies differ among these countries, the findings are expected to be generally applicable to US and

Canadian patients. Study participants are typically white and female, with many studies only including women participants. Applicability of the evidence to individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low rates of inclusion, limiting applicability of the findings.

o Directness: Direct. Although the majority of studies included a large number of outcome variables, almost all included a weight related outcome as a primary or secondary outcome measure.

o Consistency: In the small number of studies that included a TAU comparator group, there was a consistent benefit of FBT. In studies that compared FBT to other active interventions, there was a consistent benefit for FBT, even when the two treatments being compared showed no difference in their effects.

o Precision: Imprecise. For comparisons in the NMA, CIs were wide and overlapped each other.

o Dose-response relationship: There is insufficient information to determine whether there is a relationship between treatment response and treatment frequency or duration. A single study examined effects of 6 months of FBT as compared to 12 months of FBT and found no difference for the participants as a whole, although there was a suggestion that some patient subgroups may do better with longer treatment.

o Confounding factors (including likely direction of effect): For all psychotherapy studies, the participant and the therapist are aware of the treatment that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention.

o Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

o Overall strength of research evidence: The overall strength of research evidence is moderate, based on the results of the NMA, which included studies with a mix of low, moderate, and high risks of bias. In addition, several studies compared active interventions to TAU and showed consistent benefits of FBT in adolescents and emerging adults with AN. Studies that included intervention and active comparator groups constituted the bulk of the research. Results typically showed treatment by time effects that are consistent with benefits of psychotherapy, but no consistent superiority of one treatment as compared to another.

Bulimia Nervosa

Statement 13 – Cognitive-Behavioral Therapy and Serotonin Reuptake Inhibitor Treatment for Adults With Bulimia Nervosa

APA recommends (1C) that adults with bulimia nervosa be treated with eating disorder-focused cognitive-behavioral therapy and that a serotonin reuptake inhibitor (e.g., 60 mg fluoxetine daily) also

be prescribed, either initially or if there is minimal or no response to psychotherapy alone by 6 weeks of treatment.

Support for this statement comes from the expert survey (Appendix D) and from an NMA of studies of interventions in BN; however, the strength of research evidence is rated as low because of the high risk of bias of most of the studies. In the expert survey, individual CBT, nutritional rehabilitation, and psychoeducation were rated as highly appropriate for adolescents as well as adults. SSRIs, DBT, and group therapy were rated as moderately to highly appropriate for both adolescents and adults. For adolescents, FBT was also rated as moderately to highly appropriate whereas IPT and couples therapy were rated as moderately to highly appropriate for adults. In terms of initial interventions, psychotherapy alone was rated as highly appropriate and a combination of medications and psychotherapy was rated as moderately to highly appropriate for both adolescents and adults. CBT was recommended as the most appropriate initial psychotherapy and an SSRI was recommended as the most appropriate medication. Notably, use of self-help approaches or medication alone was rated as minimally appropriate.

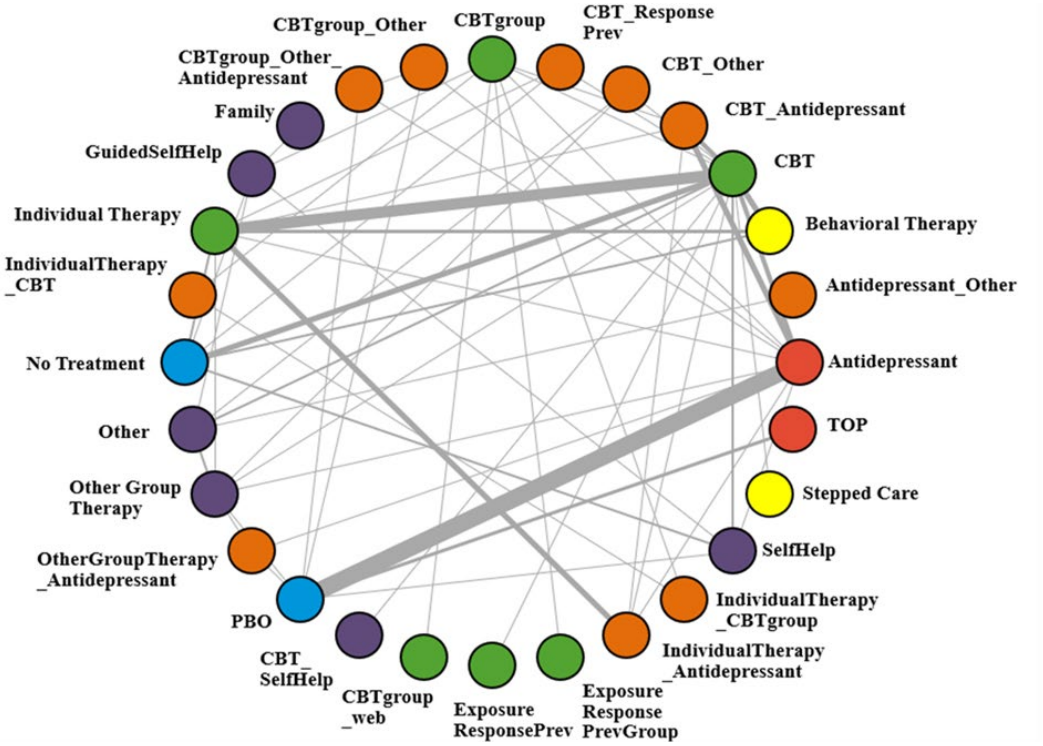
The NMA found that, in comparison to no treatment, individual CBT was associated with reductions in purging frequency (RMD -7.01; 95% CI -15.27, -0.76) and an increased likelihood of abstinence from binge eating (RR 4.97; 95% CI 1.76, 15.29) and purging (RR 11.15; 95% CI 1.87, 132.66). Individual CBT was also associated with an increased likelihood of abstinence from binge eating (RR 3.25; 95% CI 1.37, 10.86 as compared to placebo) and reductions in binge-eating frequency (RMD -7.90; 95% CI -15.42, -0.76 as compared to placebo and RMD -4.24; 95% CI -8.13, -0.30 as compared to antidepressants). Individuals who received CBT were also less likely to withdraw from treatment than those who had been randomly assigned to placebo (RR 0.29; 95% CI 0.08, 0.94) or antidepressants (RR 0.25; 95% CI 0.07, 0.73). CBT in combination with other therapies also reduced vomiting frequency as compared to placebo (RMD -3.06; 95% CI -6.56, -0.03). Group CBT was similarly effective in increasing the likelihood of binge-eating abstinence (RR 5.36; 95% CI 1.02, 26.05 as compared to no treatment; RR 3.61; 95% CI 1.02, 14.84 as compared to placebo) and reducing vomiting frequency (RMD -3.06; 95% CI -6.04, -0.29 as compared to placebo). Antidepressant medications as a group, including SSRIs, reduced binge-eating frequency (RMD -4.29; 95% CI -7.60, -1.07) and increased the likelihood of binge-eating abstinence (RR 2.23; 95% CI 1.47, 4.25) relative to placebo. The combination of CBT plus an antidepressant also reduced binge-eating frequency RMD -9.88; 95% CI -18.68, -1.75) and increased the likelihood of binge-eating abstinence (RR 2.70; 95% CI 1.01, 7.09) relative to placebo. In addition, concomitant depressive symptoms were reduced by combination treatment with CBT and an antidepressant (RMD -11.74; 95% CI -21.90, -1.84, RMD -5.52; 95% CI -10.58, -0.46, and RMD -12.91; 95% CI -25.29, -0.30, as compared to no treatment, placebo, and self-help, respectively). Placebo treatment was associated with less weight increase than antidepressant medications, but none of the other treatments were associated with changes in weight-related outcomes.

Network Meta-Analysis of Treatments for Bulimia Nervosa

Overall, the network contains 60 trials with 28 treatment arms and 5,202 subjects. In addition, the network is well-connected, with most treatments connected to multiple other treatments. Where possible, outcomes were grouped, although differences in definitions for binge-eating and purging

outcomes often made this challenging. A number of studies were endonodal and not included in the network (Habibzadeh et al. 2010; Ghaderi 2006; Crosby et al. 1993; Mitchell et al. 1993; Thompson-Brenner et al. 2016; Zeeck et al. 2009a, 2009b). An additional study (Fernández-Aranda et al. 2009) was not an RCT and not included in the network.

Figure C-7. Network graph of treatments for BN



Note: Nodes represent a treatment. Node colors indicat broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; PBO=placebo; TOP=topiramate

Table C-3. BN NMA feasibility and network characteristics

Outcome	Interventions: Total	Studies: Total	Trials per direct comparison	Total Subjects in NMA
BMI change from baseline	14	14	1-4	1,226
Weight change from baseline	7	8	1-4	695
Eating disorder scale change from baseline	17	20	1-5	2,245
Binge-eating abstinence	17	20	1-6	1,541
Binge-eating frequency change from baseline (all units)	17	31	1-8	2,863
Binge-eating frequency change from baseline (per week)	13	16	1-11	1,420
Binge-eating frequency change from baseline (per month)	10	12	1-3	1,207
Purging frequency change from baseline (all units)	11	11	1-5	790
Purging frequency change from baseline (per week)	11	6	1	418
Purging frequency change from baseline (per month)	3	3	1-2	247

Purging abstinence	6	5	1-2	272
Vomiting frequency change (all units)	17	23	1-3	2,241
Vomiting frequency change (per week): Network 1	9	6	1-2	432
Vomiting frequency change (per week): Network 2	4	4	1-3	376
Vomiting frequency change (per month)	9	10	1-3	653
Vomiting abstinence	12	10	1-2	1,003
Study withdrawal	26	21	1-7	2,528
Treatment discontinuation: Network 1	12	8	1-4	727
Treatment discontinuation: Network 2	3	2	1	277
Disease response, remission/recovery	11	11	1-2	1,424
Depression scale change from baseline	16	27	1-4	2,370
Self-esteem change from baseline	10	9	1-3	558
Treatment adherence rate	6	2	1	160

Abbreviations: BMI=body mass index; BN=bulimia nervosa; NMA=network meta-analysis

In terms of baseline characteristics of subjects in BN studies, the majority of subjects were female (90-100% among 64 trials), with a mean age of about 20 to 25 years (range 17.4 to 41 years among 41 trials). Baseline mean BMI ranged from 20.6 to 31.7 kg/m^2 (25 trials) with a mean weight of 123 to 202 lbs in the 14 trials that reported weight instead of BMI. There were no concerns about heterogeneity in these variables and they are consistent with the characteristics of patients seen in clinical practice. Baseline vomiting frequency was reported in 18 trials and ranged from 3 to 18 episodes per week and 21 to 53 episodes per month. In terms of baseline rates of laxative abuse, 9 trials reported this information with rates that range from 0 to 18.7%, leading to concerns about possible heterogeneity. There also appeared to be heterogeneity in the proportion of subjects with prior AN (range of 7 to 45.6% among 15 trials), although many of the trials did not report this information. Variability was also present in baseline scores on the BDI (range of 12 to 26.5 among 20 trials).

Table C-4. Statistically favored comparisons from the BN NMA

Intervention	Comparison	Outcomes	Statistical values
Antidepressants	Placebo	Binge-eating abstinence	RR 2.23 (1.47, 4.25)
	Placebo	Binge-eating frequency, all units	RMD -4.29 (-7.60, -1.07)
	Placebo	Binge-eating frequency, per week	RMD -3.54 (-5.51, -1.72)
Behavioral therapy	No treatment	Binge-eating abstinence	RR 5.02 (1.43, 21.49)
CBT	No Treatment	Binge-eating abstinence	RR 4.97 (1.76, 15.29)
		Purging frequency, all units	RMD -7.01 (-15.27, -0.76)
		Purging abstinence	RR 11.15 (1.87, 132.66)
	Placebo	Binge-eating abstinence	RR 3.25 (1.37, 10.86)
		Binge-eating frequency, all units	RMD -7.90 (-15.42, -0.76)
		Binge-eating frequency, per week	RMD -7.77 (-12.30, -3.59)
		Study withdrawal	RR 0.29 (0.08, 0.94)
	Antidepressants	Binge-eating frequency, per week	RMD -4.24 (-8.13, -0.30)
		Study withdrawal	RR 0.25 (0.07, 0.73)
	Other therapy	Treatment adherence	RR 1.41 (1.01, 2.20)
	Other group therapy	Treatment adherence	RR 1.58 (1.10, 2.54)
CBT + antidepressant	No treatment	Depression scales	RMD -11.74 (-21.90, -1.84)
	Placebo	Binge-eating abstinence	RR 2.70 (1.01, 7.09)
		Depression scales	RMD -5.52 (-10.58, -0.46)

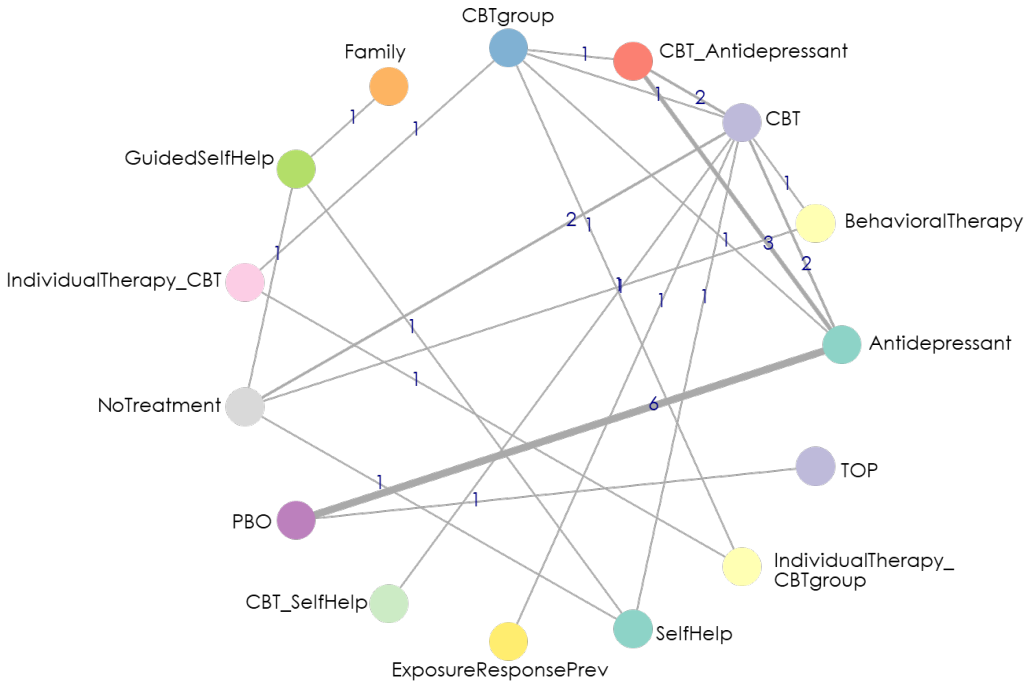
		Binge-eating frequency, all units	RMD -9.88 (-18.68, -1.75)
		Binge-eating frequency, per week	RMD -8.37 (-14.04, -2.82)
	Self-Help	Depression scales	RMD -12.91 (-25.29, -0.30)
CBT Group	No treatment	Eating disorder scale	RR -17.47 (-33.02, -2.14)
		Binge-eating abstinence	RR 5.36 (1.02, 26.05)
		Depression scales	RMD -9.96 (-18.74, -1.57)
	Placebo	Binge-eating abstinence	RR 3.61 (1.02, 14.84)
		Vomiting frequency, all units	RMD -3.06 (-6.04, -0.29)
		Vomiting frequency, per week	RMD -7.02 (-13.85, -0.67)
CBT + self-help	No treatment	Binge-eating abstinence	RR 4.25 (1.04, 18.93)
CBT + other	No treatment	Binge-eating frequency, per week	RMD -6.97 (-13.02, -0.96)
	Placebo	Vomiting frequency, all units	RMD -3.06 (-6.56, -0.03)
		Binge-eating frequency, per week	RMD -11.68 (-18.48, -5.63)
	Antidepressants	Binge-eating frequency, per week	RMD -8.18 (-14.55, -2.12)
	Other therapy	Treatment adherence	RR 1.46 (1.06, 2.31)
	Other group therapy	Treatment adherence	RR 1.65 (1.17, 2.63)
Individual therapy	No treatment	Depression scales	RMD -10.67 (-20.78, -0.51)
	Placebo	Study withdrawal	RR 0.20 (0.06, 0.68)
		Binge-eating frequency, per week	RMD -7.79 (-12.96, -2.86)
	Antidepressants	Study withdrawal	RR 0.17 (0.05, 0.53)
	Other group therapy + antidepressants	Study withdrawal	RR 0.24 (0.06, 0.88)
	Other group therapy	Study withdrawal	RR 0.49 (0.23, 1.00)
Individual therapy + antidepressant	No treatment	Depression scales	RMD -17.16 (-29.63, -4.51)
		Binge-eating frequency, all units	RMD -12.86 (-22.12, -4.43)
		Binge-eating frequency, per week	RMD -25.64 (-33.88, -17.67)
	Placebo	Depression scales	RMD -10.74 (-20.39, -1.57)
		Binge-eating frequency, all units	RMD -16.07 (-27.09, -5.71)
		Binge-eating frequency, per week	RMD -30.48 (-39.59, -21.99)
	Antidepressants	Binge-eating frequency, all units	RMD -11.77 (-22.50, -1.68)
		Binge-eating frequency, per week	RMD -26.85 (-35.47, -18.45)
	CBT	Depression scales	RMD -8.37 (-16.51, -0.37)
		Binge-eating frequency, all units	RMD -8.14 (-16.12, -0.41)
		Binge-eating frequency, per week	RMD -22.68 (-30.69, -14.56)
	CBT + antidepressants	Binge-eating frequency, per week	RMD -22.00 (-31.51, -12.68)
	CBT + other	Binge-eating frequency, per week	RMD -18.80 (-27.88, -9.51)
	CBT group	Binge-eating frequency, all units	RMD -12.64 (-23.47, -3.06)
		Binge-eating frequency, per week	RMD -23.27 (-32.24, -14.54)
	CBT group + other	Binge-eating frequency, all units	RMD -16.17 (-29.98, -2.65)
	CBT group + other + antidepressants	Binge-eating frequency, all units	RMD -15.68 (-29.89, -2.35)
	Exposure response prevention	Depression scales	RMD -12.53 (-24.03, -0.26)
	Individual therapy	Binge-eating frequency, all units	RMD -11.25 (-18.40, -5.14)
		Binge-eating frequency, per week	RMD -22.65 (-29.68, -15.60)
	Individual therapy + antidepressant	Study withdrawal	RR 11.40 (1.24, 121.11)
	Other therapy	Binge-eating frequency, per week	RMD -21.69 (-29.74, -13.46)
	Other group therapy	Binge-eating frequency, all units	RMD -8.99 (-17.69, -1.10)
		Binge-eating frequency, per week	RMD -21.89 (-29.59, -14.05)
	Other group therapy + antidepressants	Binge-eating frequency, per week	RMD -22.64 (-31.80, -13.71)

	Self-help	Depression scales	RMD -18.12 (-32.65, -3.16)
		Binge-eating frequency, per week	RMD -26.00 (-36.10, -16.36)
	Stepped care	Binge-eating frequency, all units	RMD -12.22 (-24.53, -0.37)
Exposure response prevention	No treatment	Eating disorder scale	RR -19.25 (-34.31, -3.39)
	Placebo	Vomiting frequency, all units	RMD -3.72 (-7.27, -0.54)
	Individual therapy	Vomiting frequency, all units	RMD -2.25 (-5.48, -0.06)
	Self-help	Vomiting frequency, all units	RMD -2.67 (-5.53, -0.03)
Guided self-help	No treatment	Binge-eating Abstinence	RR 3.49 (1.35, 9.13)
		Purging frequency, all units	RMD -23.99 (-41.04, -5.20)
		Vomiting frequency, all units	RMD -7.53 (-14.17, -0.53)
	Placebo	Purging frequency, all units	RMD -24.02 (-45.19, -0.40)
		Vomiting frequency, all units	RMD -9.37 (-16.22, -2.05)
		Study withdrawal	RR 0.04 (0.00, 0.40)
		Binge-eating frequency, all units	RMD -14.19 (-27.40, -0.39)
	Antidepressants	Vomiting frequency, all units	RMD -8.75 (-15.73, -1.36)
		Study withdrawal	RR 0.04 (0.00, 0.33)
	Antidepressants + self-help	Vomiting frequency, all units	RMD -8.00 (-15.12, -0.73)
	CBT	Vomiting frequency, all units	RMD -7.35 (-14.16, -0.12)
	CBT + antidepressants	Study withdrawal	RR 0.08 (0.01, 0.93)
	CBT + self-help	Study withdrawal	RR 0.08 (0.01, 0.84)
	Individual therapy	Vomiting frequency, all units	RMD -7.91 (-15.07, -0.77)
	Individual therapy + antidepressants	Study withdrawal	RR 0.09 (0.01, 0.81)
	Lithium	Study withdrawal	RR 0.07 (0.01, 0.78)
	Other group therapy	Study withdrawal	RR 0.11 (0.01, 0.77)
	Other group therapy + antidepressants	Study withdrawal	RR 0.05 (0.00, 0.52)
	Self-help	Vomiting frequency, all units	RMD -8.27 (-15.05, -1.25)
	Topiramate	Study withdrawal	RR 0.06 (0.01, 0.60)
Other therapies	No treatment	Binge-eating frequency, per week	RMD -4.00 (-7.62, -0.35)
	Placebo	Binge-eating frequency, per week	RMD -8.68 (-14.09, -3.57)
	Antidepressants	Binge-eating frequency, per week	RMD -5.17 (-10.21, -0.17)
Other group therapy	No treatment	Binge-eating frequency, per week	RMD -3.76 (-7.47, -0.11)
	Placebo	Binge-eating frequency, per week	RMD -8.51 (-12.82, -4.16)
	Antidepressants	Binge-eating frequency, per week	RMD -4.97 (-8.70, -1.04)
		Study withdrawal	RR 0.36 (0.11, 0.92)
Other group therapy + antidepressant	Placebo	Binge-eating frequency, per week	RMD -7.78 (-12.95, -2.84)
Self-help	No treatment	Binge-eating Abstinence	RR 3.68 (1.29, 10.27)
	Placebo	Study withdrawal	RR 0.08 (0.01, 0.59)
	Antidepressants	Study withdrawal	RR 0.07 (0.01, 0.48)
	Other group therapy + antidepressants	Study withdrawal	RR 0.10 (0.01, 0.79)
	Topiramate	Study withdrawal	RR 0.11 (0.01, 0.91)
Placebo	Antidepressants	Weight change	RR 5.50 (0.96, 10.95)
No treatment	Placebo	Study withdrawal	RR 0.15 (0.03, 0.67)
	Antidepressant	Study withdrawal	RR 0.13 (0.02, 0.55)
	Other group therapy + antidepressants	Study withdrawal	RR 0.18 (0.03, 0.86)

Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; NMA=network meta-analysis; RMD=relative mean difference; RR=relative risk

Network Meta-Analysis Networks and Forest Plots by Outcome Measure

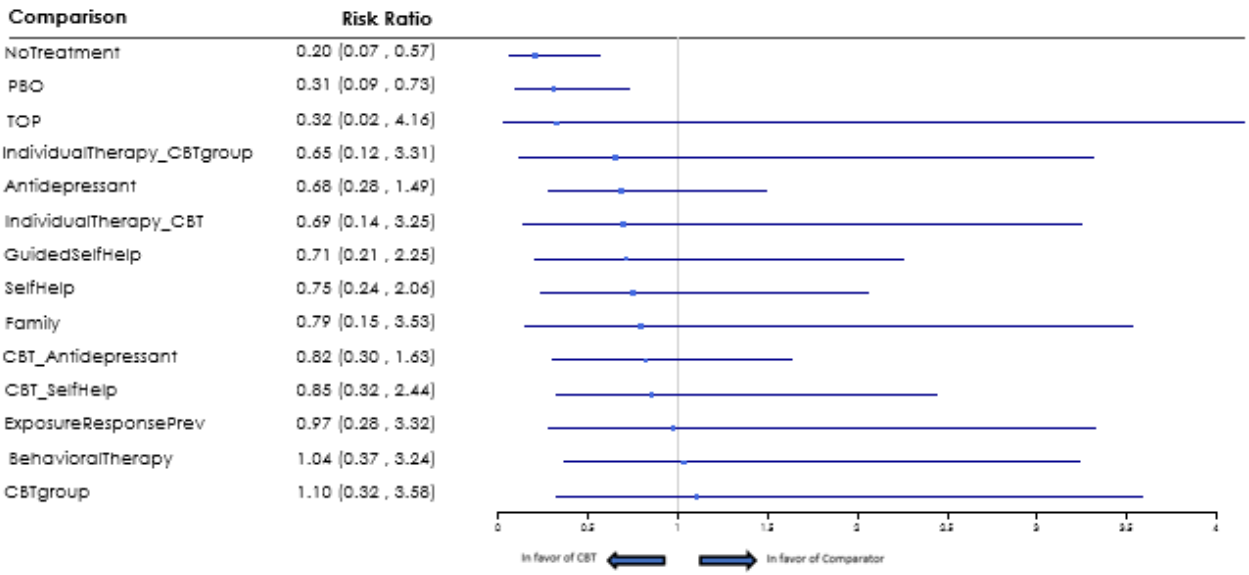
Figure C-8. Network graph of treatments for BN as compared to CBT for the outcome of binge-eating abstinence.



Note: Nodes represent a treatment. Node colors indicate broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; ExposureResponsePrev=Exposure and response prevention; PBO=placebo; TOP=topiramate

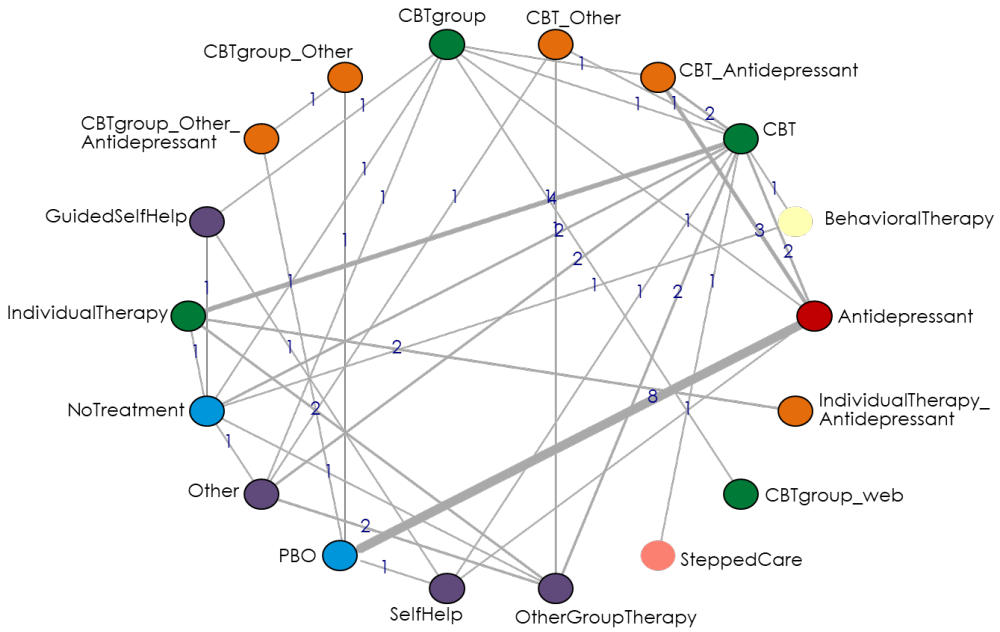
Figure C-9. Forest plot of binge-eating abstinence at all time points as compared to CBT.

Statistically significant differences are present for antidepressants, CBT, group CBT, and CBT in combination with antidepressants as compared to placebo and for CBT, behavioral therapy, group CBT and CBT plus self-help as compared to no treatment.



Abbreviations: CBT=cognitive-behavioral therapy; PBO=placebo; ExposureResponsePrev=Exposure and response prevention; TOP=topiramate

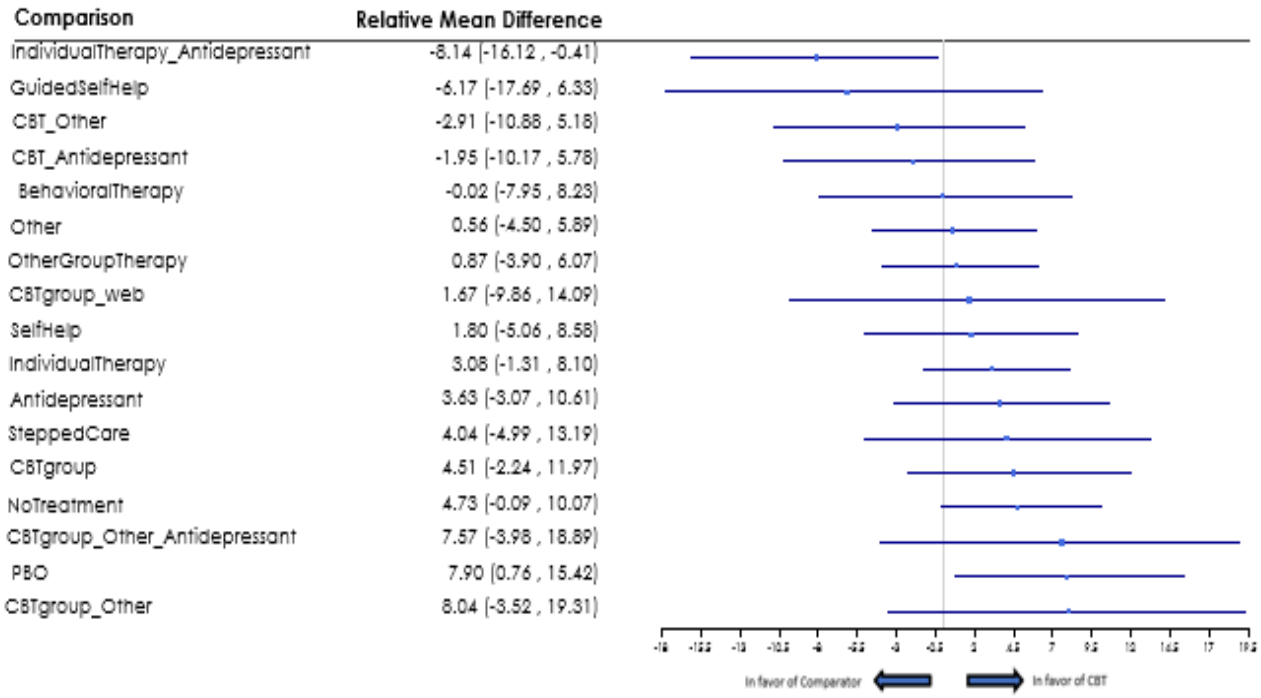
Figure C-10. Network graph of treatments for BN as compared to CBT for the outcome of change in binge-eating frequency



Note: Nodes represent a treatment. Node colors indicate broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; PBO=placebo

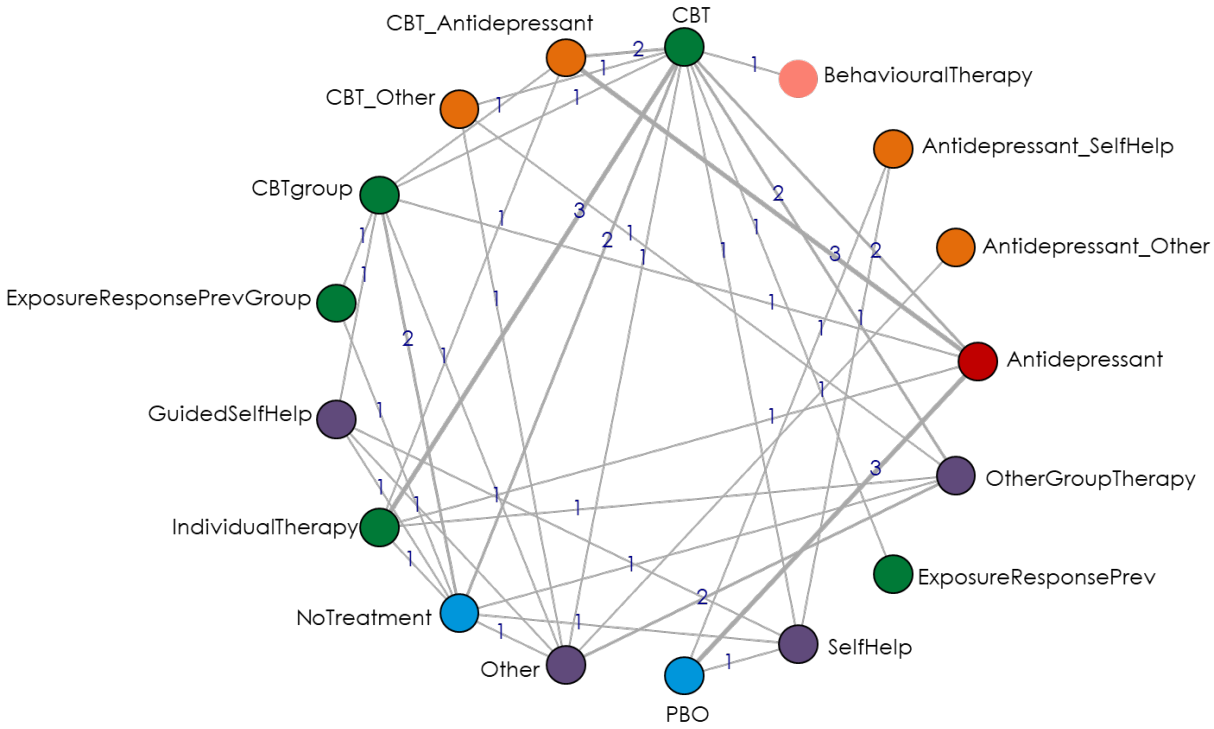
Figure C-11. Forest plot of change in binge-eating frequency at all time points as compared to CBT

In addition to other comparisons (see Table C-4), statistically significant differences are present for antidepressants compared to CBT and for antidepressants, CBT, and the combination of CBT and antidepressant as compared to placebo.



Abbreviations: CBT=cognitive-behavioral therapy; PBO=placebo

Figure C-12. Network graph of treatments for BN as compared to CBT for the outcome of vomiting frequency

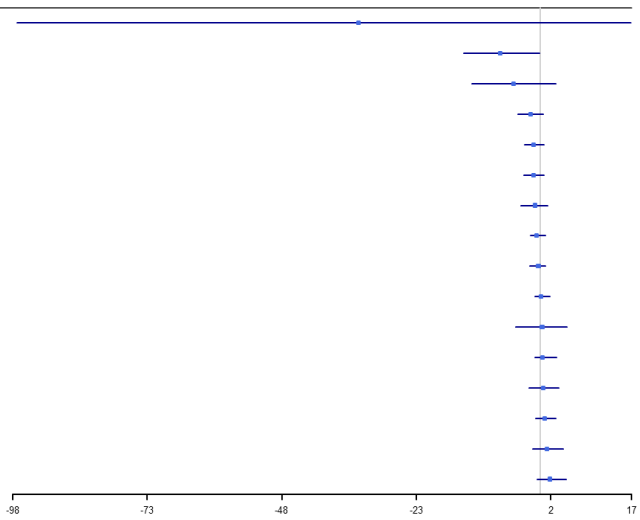


Note: Nodes represent a treatment. Node colors indicate broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; ExposureResponsePrev=Exposure and response prevention; ExposureResponsePrevGroup=group exposure and response prevention; PBO=placebo

Figure C-13. Forest plot of vomiting frequency at all time points as compared to CBT

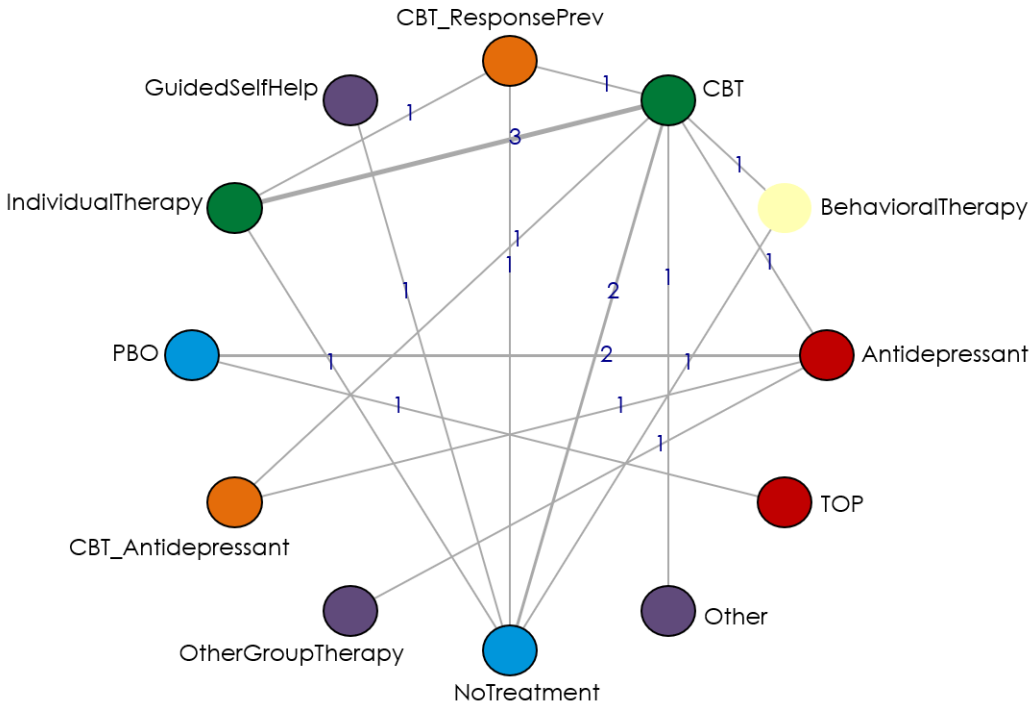
In addition to other comparisons (see Table C-4), statistically significant differences compared to placebo are present for CBT group, CBT plus other interventions, exposure and response prevention, and guided self-help.

Comparison	Relative Mean Difference
ExposureResponsePrev	-33.75 (-97.11 , 16.93)
GuidedSelfHelp	-7.35 (-14.16 , -0.12)
CBT_Antidepressant	-4.92 (-12.57 , 2.94)
ExposureResponsePrevGroup	-1.79 (-4.04 , 0.63)
CBTgroup	-1.15 (-2.81 , 0.71)
CBT_Other	-1.14 (-3.03 , 0.71)
Antidepressant_Other	-0.86 (-3.52 , 1.40)
OtherGroupTherapy	-0.57 (-1.78 , 0.97)
Other	-0.29 (-1.91 , 0.98)
NoTreatment	0.15 (-0.98 , 1.86)
BehaviouralTherapy	0.40 (-4.56 , 5.01)
IndividualTherapy	0.46 (-0.85 , 3.09)
Antidepressant_SelfHelp	0.58 (-1.97 , 3.54)
SelfHelp	0.87 (-0.81 , 2.99)
Antidepressant	1.31 (-1.26 , 4.36)
PBO	1.91 (-0.54 , 4.89)



Abbreviations: CBT=cognitive-behavioral therapy; ExposureResponsePrev=Exposure and response prevention; ExposureResponsePrevGroup=group exposure and response prevention; PBO=placebo

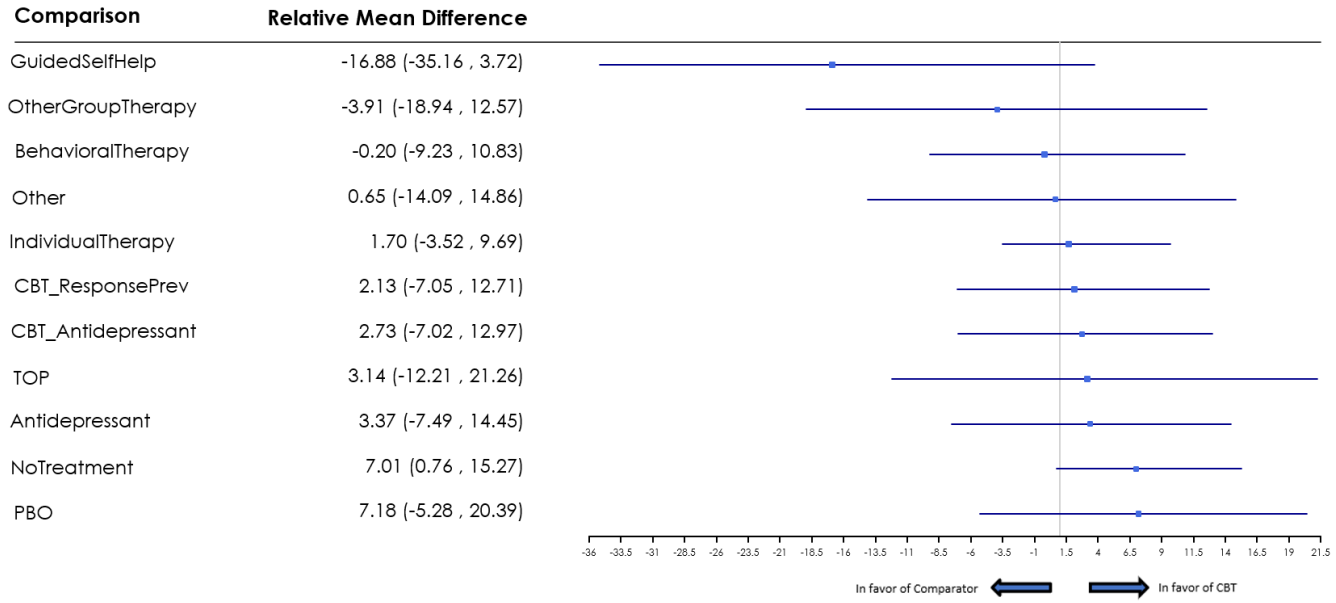
Figure C-14. Network graph of treatments for BN as compared to CBT for the outcome of purging frequency



Note: Nodes represent a treatment. Node colors indicate broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; PBO=placebo; ResponsePrev=Response prevention; TOP=topiramate

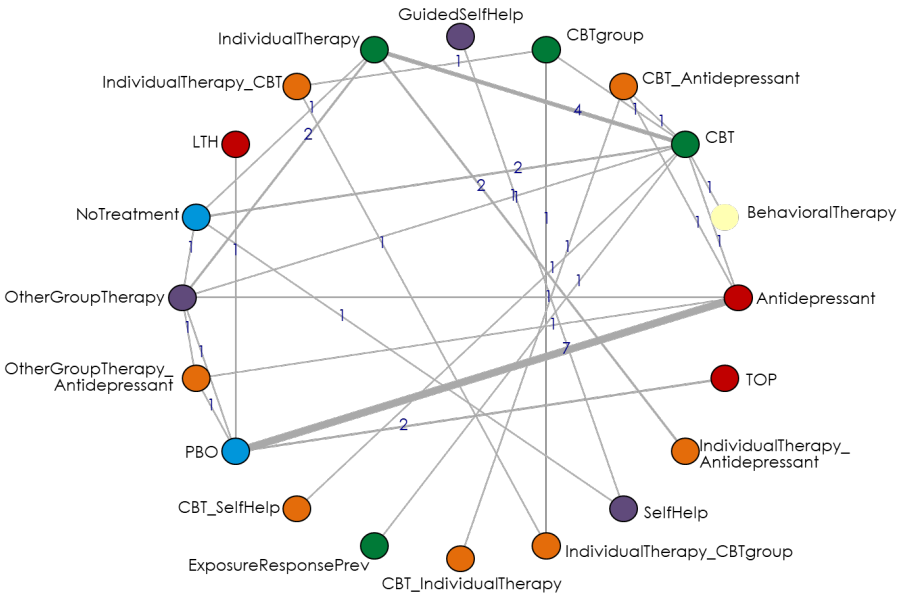
Figure C-15. Forest plot of purging frequency as compared to CBT

In addition to other comparisons (see Table C-4), statistically significant differences compared to no treatment are present for CBT and for guided self-help.



Abbreviations: CBT=cognitive-behavioral therapy; PBO=placebo; ResponsePrev=Response prevention; TOP=topiramate

Figure C-16. Network graph of treatments as compared to CBT for the outcome of study withdrawal rate

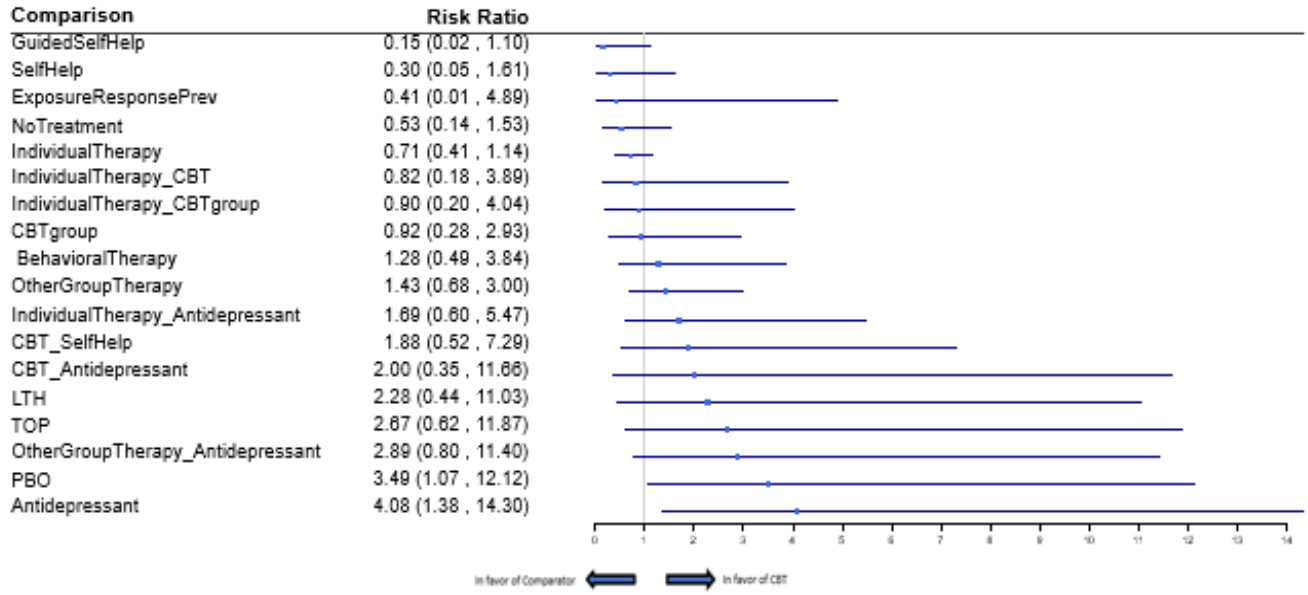


Note: Nodes represent a treatment. Node colors indicate broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.

Abbreviations: CBT=cognitive-behavioral therapy; ExposureResponsePrev=Exposure and response prevention; LTH=lithium; PBO=placebo; TOP=topiramate

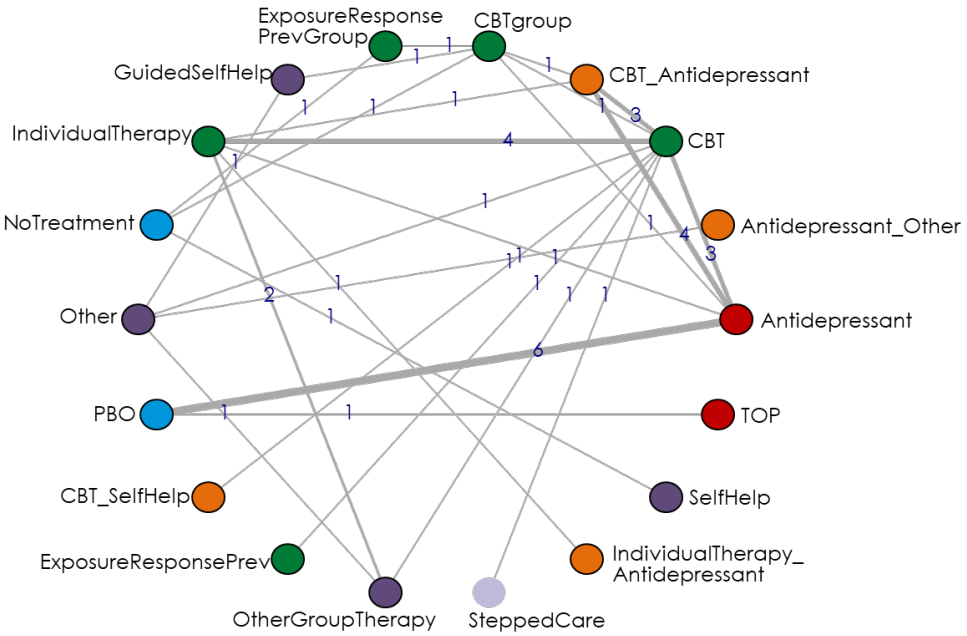
Figure C-17. Forest plot of study withdrawal as compared to CBT

In addition to other comparisons (see Table C-4), statistically significant differences are present for antidepressants compared to CBT, no treatment, guided self-help, non-CBT group therapy, or other individual therapy and for placebo compared to CBT, individual treatment, guided self-help, self-help, and no treatment.



Abbreviations: CBT=cognitive-behavioral therapy; ExposureResponsePrev=Exposure and response prevention; LTH=lithium; PBO=placebo; TOP=topiramate

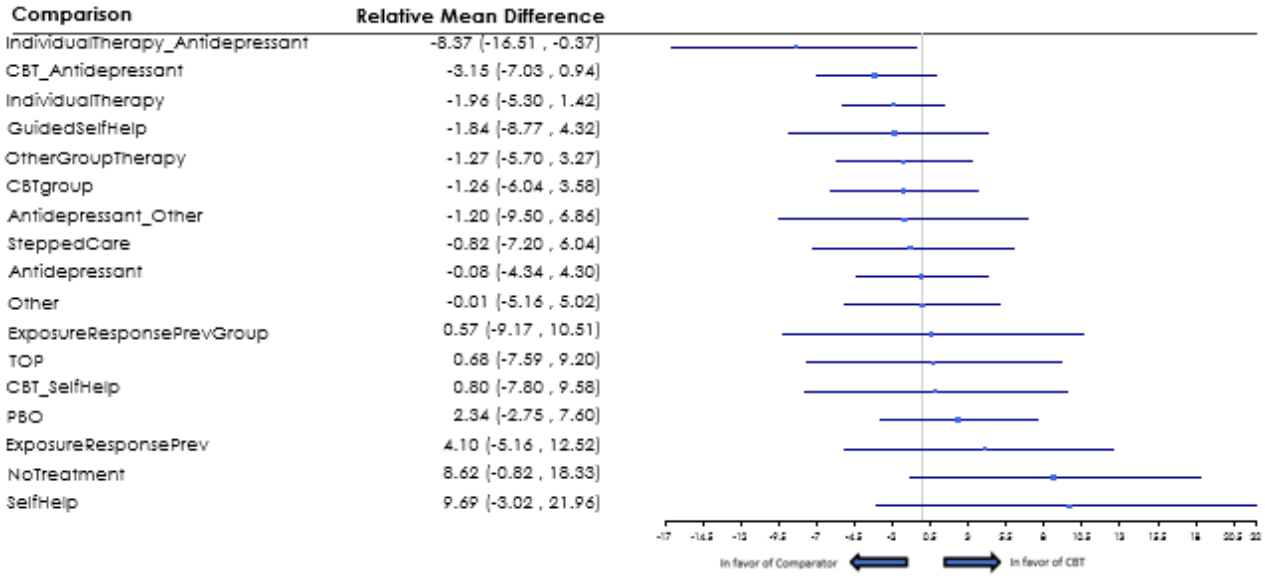
Figure C-18. Network graph of treatments for BN as compared to CBT for the outcome of depression scale scores



Note: Nodes represent a treatment. Node colors indicated broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: CBT=cognitive-behavioral therapy; ExposureResponsePrev=Exposure and response prevention; ExposureResponsePrevGroup=group Exposure and response prevention; PBO=placebo; TOP=topiramate

Figure C-19. Forest plot of change in depression scores as compared to CBT

In addition to other comparisons (see Table C-4), statistically significant differences are present for group CBT as compared to no treatment and for CBT in combination with antidepressant as compared to self-help, placebo, or no treatment.



Abbreviations: CBT=cognitive-behavioral therapy; ExposureResponsePrev=Exposure and response prevention; ExposureResponsePrevGroup=group Exposure and response prevention; PBO=placebo; TOP=topiramate

Heterogeneity Analysis of Antidepressant Effects As Compared to Placebo

Figure C-20. Heterogeneity analysis of antidepressant effects on binge-eating abstinence as compared to placebo

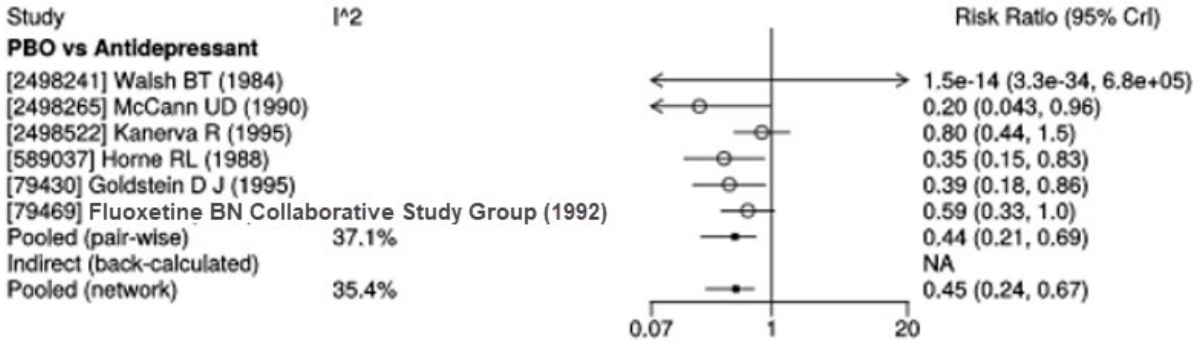


Figure C-21. Heterogeneity analysis of antidepressant effects on binge-eating frequency (binges per week and binges per month) as compared to placebo

The medication and daily dose used in each of the listed studies is Agras WS (1987) imipramine 50-300 mg; Goldstein DJ (1995) fluoxetine 60 mg; Horne RL (1988) bupropion 225-450 mg; Kanerva R (1995) fluoxetine 60 mg; Mitchell JE (2001) fluoxetine 60 mg; Romano SJ (2002) fluoxetine 60 mg; Fluoxetine Bulimia Nervosa Collaborative Study Group (1992) fluoxetine 20 mg or 60 mg; Walsh BT (1984) phenelzine 60-90 mg; Walsh BT (1988) phenelzine 60-90 mg.

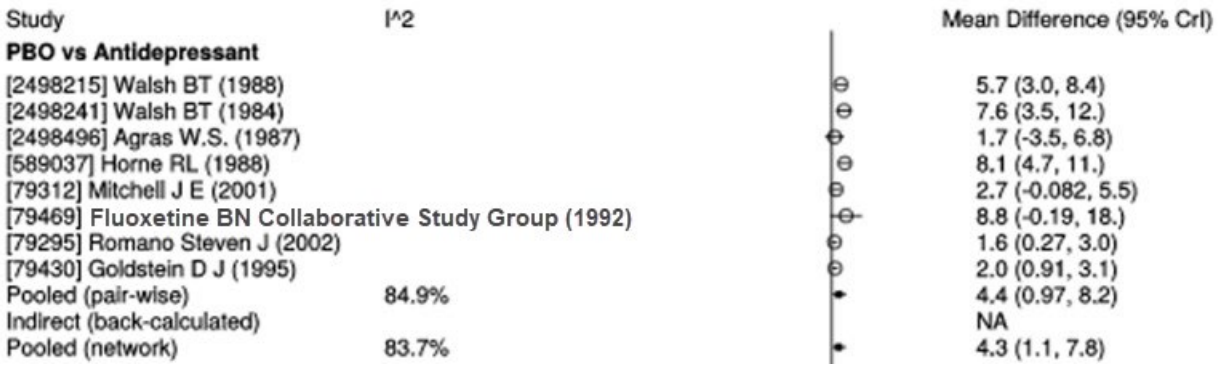


Figure C-22. Heterogeneity analysis of antidepressant effects on the change in binge-eating frequency as compared to placebo

The medication and daily dose used in each of the listed studies is Agras WS (1987) imipramine 50-300 mg; Goldstein DJ (1995) fluoxetine 60 mg; Horne RL (1988) bupropion 225-450 mg; McCann UD (1990) desmethylimipramine 25-300 mg; Mitchell JE (2001) fluoxetine 60 mg; Romano SJ (2002) fluoxetine 60 mg; Walsh BT (1984) phenelzine 60-90 mg; Walsh BT (1988) phenelzine 60-90 mg.

Detailed Review of Evidence: Cognitive-Behavioral Therapy

Table C-5. Statistically favorable effects by treatment

Intervention	Comparison	Outcomes	Statistical values
CBT	No treatment	Binge-eating abstinence	RR 4.97 (1.76, 15.29)
		Purging frequency, all units	RMD -7.01 (-15.27, -0.76)
		Purging abstinence	RR 11.15 (1.87, 132.66)
	Placebo	Binge-eating abstinence	RR 3.25 (1.37, 10.86)
		Binge-eating frequency, all units	RMD -7.90 (-15.42, -0.76)
		Binge-eating frequency, per week	RMD -7.77 (-12.30, -3.59)
		Study withdrawal	RR 0.29 (0.08, 0.94)
	Antidepressants	Binge-eating frequency, per week	RMD -4.24 (-8.13, -0.30)
		Study withdrawal	RR 0.25 (0.07, 0.73)
	Other therapy	Treatment adherence	RR 1.41 (1.01, 2.20)
	Other group therapy	Treatment adherence	RR 1.58 (1.10, 2.54)
CBT + antidepressant	No treatment	Depression scales	RMD -11.74 (-21.90, -1.84)
	Placebo	Binge-eating abstinence	RR 2.70 (1.01, 7.09)
		Depression scales	RMD -5.52 (-10.58, -0.46)
		Binge-eating frequency, all units	RMD -9.88 (-18.68, -1.75)
		Binge-eating frequency, per week	RMD -8.37 (-14.04, -2.82)
	Self-help	Depression scales	RMD -12.91 (-25.29, -0.30)
CBT group	No treatment	Eating disorder scale	RR -17.47 (-33.02, -2.14)
		Binge-eating abstinence	RR 5.36 (1.02, 26.05)
		Depression scales	RMD -9.96 (-18.74, -1.57)
	Placebo	Binge-eating abstinence	RR 3.61 (1.02, 14.84)
		Vomiting frequency, all units	RMD -3.06 (-6.04, -0.29)
		Vomiting frequency, per week	RMD -7.02 (-13.85, -0.67)
CBT + self-help	No treatment	Binge-eating abstinence	RR 4.25 (1.04, 18.93)

CBT + other	No treatment	Binge-eating frequency, per week	RMD -6.97 (-13.02, -0.96)
	Placebo	Vomiting frequency, all units	RMD -3.06 (-6.56, -0.03)
		Binge-eating frequency, per week	RMD -11.68 (-18.48, -5.63)
	Antidepressants	Binge-eating frequency, per week	RMD -8.18 (-14.55, -2.12)
	Other therapy	Treatment adherence	RR 1.46 (1.06, 2.31)
	Other group therapy	Treatment adherence	RR 1.65 (1.17, 2.63)
Individual therapy + antidepressant	CBT	Depression scales	RMD -8.37 (-16.51, -0.37)
		Binge-eating frequency, all units	RMD -8.14 (-16.12, -0.41)
		Binge-eating frequency, per week	RMD -22.68 (-30.69, -14.56)
	CBT + antidepressants	Binge-eating frequency, per week	RMD -22.00 (-31.51, -12.68)
	CBT + other	Binge-eating frequency, per week	RMD -18.80 (-27.88, -9.51)
	CBT group	Binge-eating frequency, all units	RMD -12.64 (-23.47, -3.06)
		Binge-eating frequency, per week	RMD -23.27 (-32.24, -14.54)
	CBT group + Other	Binge-eating frequency, all units	RMD -16.17 (-29.98, -2.65)
	CBT group + other + antidepressants	Binge-eating frequency, all units	RMD -15.68 (-29.89, -2.35)
Guided self-help	CBT	Vomiting frequency, all units	RMD -7.35 (-14.16, -0.12)
	CBT + antidepressants	Study withdrawal	RR 0.08 (0.01, 0.93)
	CBT + self-help	Study withdrawal	RR 0.08 (0.01, 0.84)

Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; NMA=network meta-analysis; RMD=relative mean difference; RR=relative risk

The studies of CBT generally supported its use, particularly relative to no treatment or wait list control conditions. However, there were differences in the number and duration of sessions among the studies and in the precise approach used for CBT. For example, some studies used a version of CBT-E that was focused on eating behaviors whereas other studies added modules that addressed one or more "external" maintaining mechanisms of eating disorders (e.g., clinical perfectionism, low self-esteem, marked interpersonal difficulties) when these features are severe and are disrupting treatment progress (i.e., broad CBT-E; Cooper and Fairburn 2011). In actual practice, an amalgam of CBT approaches may be used depending on patient needs.

A number of early studies compared CBT to other therapeutic approaches and to wait list or no treatment comparator groups. Freeman and colleagues (Freeman et al. 1988) randomly assigned participants to CBT (N=32), behavioral therapy (N=30), group therapy (N=30), and a wait list control condition (N=20), with all treatment groups receiving 15 weekly 1-hour sessions. Each of the active treatments was equally effective with 77% of participants achieving binge-eating abstinence at the end of active treatment; however, the groups differed in study withdrawal rates with greater attrition in the CBT and group therapy treatment arms. In comparison to no treatment, Agras and colleagues (Agras et al. 1989) assessed the effects of 14 hour-long sessions of self-monitoring, CBT, or CBT in combination with response prevention. At the end of the 4-month treatment period, CBT was superior to no treatment with 56.3% of participants achieving abstinence as compared to 5.8%. Rates of abstinence in the other groups were 23.5% and 31.2% for self-monitoring and for CBT plus response prevention, respectively. Griffiths and colleagues (Griffiths et al. 1994) randomly assigned participants to CBT (N=27), hypnobehavioral therapy (N=27), or a wait list control group (N=28). In this study, the active treatments consisted of 7 sessions of approximately 1 hour delivered over 8 weeks. Rates of binge-

eating and purging abstinence at the end of treatment were significantly greater for both active treatments than in the wait list control group (50% binge-eating abstinence with CBT, 43% with hypnobehavioral therapy, and 4.5% with wait list; for purging abstinence, 40%, 33.3%, and 4.5%, respectively). At the 9-month follow-up assessment, the active treatment groups maintained comparable outcomes, although rates of attrition were greater with hypnobehavioral therapy than with CBT (Griffiths et al. 1996). As part of a longer stepped-care intervention, Treasure and colleagues (Treasure et al. 1994) randomly assigned subjects to CBT (N=21), a self-help manual (N=41), or a wait list control condition (N=19). After 8 weeks of treatment, rates of BN remission were relatively low, but CBT and self-help were superior to the wait list condition (remission rates of 24%, 22%, and 11%, respectively). In an additional study, Davis and colleagues (Davis et al. 1999) provided all participants with 6 weekly 90-minute sessions of group psychoeducation and then randomly assigned individuals to CBT or no CBT. The findings were confounded by adjusting the number of CBT sessions based on illness severity, but remission of binge eating and purging was greater in those who received CBT (43.2%) as compared to those who did not (15.8%). Together, these studies are consistent in showing benefits of CBT for reducing binge eating and purging behaviors and superiority in achieving BN remission than wait list control conditions or no treatment. In another study (Hsu et al. 2001; N=100), CBT with or without nutritional therapy, was superior to nutritional counseling alone or a support group alone, with abstinence rates of 35%, 52%, 17%, and 21%, respectively.

Mitchell and colleagues (Mitchell et al. 1993) examined whether outcomes in BN are influenced by treatment intensity and whether a concerted focus on binge-eating and purging abstinence was beneficial. Study participants (N=143) were randomly assigned to one of four groups: high intensity CBT with high emphasis on early abstinence, high intensity CBT with low emphasis on early abstinence, low intensity CBT with high emphasis on early abstinence, and low intensity CBT with low emphasis on early abstinence. As compared to the other groups, the low intensity, low abstinence emphasis group had lower rates of overall abstinence (20.6% vs. 63.6-68.3%), binge-eating abstinence (32.4% vs. 69.7-73.2%), or vomiting abstinence (29.4% vs. 70.7-76.5%). In a secondary analysis, high intensity treatment groups had lower relapse rates once abstinence was achieved than groups that received low intensity treatment. Thus, the authors concluded that high intensity treatment may help maintain abstinence whereas an early emphasis on abstinence may help achieve abstinence.

Two RCTs have examined a focused approach to delivery of CBT as compared to a broader approach to CBT. In one trial (Ghaderi 2006), participants received 19 weekly 50-minute sessions using either the manual-based approach (N=26) of Fairburn and colleagues (Fairburn et al. 1993) or a broader approach based on an individualized assessment of the participants' needs (N=24). At the end of treatment, response or remission was seen in 92% of participants who received broad CBT and 69% who received focused CBT, but the difference was not statistically significant. Response was maintained 18 months after treatment and the two treatments remained comparable. Thompson-Brenner and colleagues (Thompson-Brenner et al. 2016) also compared an eating-focused approach to a broader approach to CBT that also addressed mood intolerance and interpersonal dysfunction. Their sample (N=50) included individuals with features of borderline personality in addition to meeting criteria for BN. Participants received a 90-minute preparatory session and then 20 sessions of 50 minutes over 20 weeks with a

tapering frequency. Although both groups showed improvements, the two treatment approaches did not differ in their efficacy with binge-purge remission occurring in 44% with focused CBT and 40% with broad CBT. Notably, however, attrition was greater in the broad CBT group (32%) as compared to the focused CBT group (16%).

Other approaches to individual therapy have also been compared to CBT. Juarascio and colleagues (Juarascio et al. 2021) conducted an RCT that compared individual CBT-E to mindfulness and acceptance-based treatment. During treatment, which consisted of twenty sessions in 20 weeks, attrition was considerable and averaged 41%. Nevertheless, both treatment groups showed significant reductions (with Cohen's d ranging from 1.25 to 2.09 from baseline to follow-up) in episodes of loss of control eating, compensatory behaviors, and reductions in global eating disorder severity as measured using the EDE. Poulsen and colleagues (Daniel et al. 2016; Folke et al. 2016; Poulsen et al. 2014) randomly assigned participants to 2 years of weekly psychoanalytic therapy or 20 CBT sessions over 5 months. At 5 months as well as at 2 years, more participants treated with CBT were abstinent from binge eating and purging than with psychoanalytic therapy (42% vs. 6% at 5 months, $p<0.01$; 44% vs. 15% at 2 years, $p=0.02$). One large two-site trial in the United States (Agras et al. 2000; Wilson et al. 2002) compared CBT (N=110) to IPT (N=110) with both treatments delivered with a tapering frequency that included a total of 19 sessions of 50 minutes each over 20 weeks. Importantly, the IPT methodology did not include any self-monitoring of behavior or any attention to weight, shape, or associated attitudes. The IPT group also differed from the CBT group in having more episodes of purging per 28 days and having more expressed eating related concerns. There were also differences in the initial sample characteristics between the two treatment sites as well as differences in the rates of study withdrawals between the sites. With these caveats, CBT was associated with greater rates of recovery (29% vs. 6%; $p<0.001$) and remission (48% vs. 28%; $p=0.008$) at the end of treatment; however, at 4-month, 8-month, and 12-month follow-up assessments, the differences between the treatment groups no longer reached statistical significance. Fairburn and colleagues (Fairburn et al. 1991, 1993) also found measured outcomes to be generally comparable between CBT and IPT, except for vomiting on which IPT had less impact. However, their study had a smaller sample with 25 participants in each group and did not use an intention-to-treat analysis making interpretation of the findings more difficult.

A third treatment arm in the study of Fairburn and colleagues (Fairburn et al. 1991, 1993) assessed effects of behavioral therapy. In the short-term, behavioral therapy had similar effects on vomiting as CBT, however, at 12-month follow-up, rates of abstinence from binge-eating and purging behaviors were much lower for behavioral therapy than for CBT or IPT, and treatment withdrawal rates were greater for those in the behavioral therapy group relative to the other treatment arms. In participants who received 8 weeks of weekly 60-minute sessions (N=47), Thackwray and colleagues (Thackwray et al. 1993) found no significant differences in abstinence rates at the end of treatment for behavioral therapy as compared to CBT or self-monitoring, although rates of abstinence at 6 months were numerically higher for CBT. Freeman and colleagues (Freeman et al. 1988) randomly assigned participants to CBT (N=32), behavioral therapy (N=30), group therapy (N=30), or a wait list control group, with active treatments consisting of 15 weeks of weekly 1-hour sessions. With the exception of the wait list control group, the outcomes in the treatment arms were comparable, although rates of study withdrawals were

lowest with the behavioral therapy treatment group. Thus, the studies that included behavioral therapy had relatively small samples, factors that might bias results, and inconsistent findings.

Exposure and response prevention is even less well studied than CBT. Leitenberg and colleagues (Leitenberg et al. 1988) randomly assigned subjects to 24 sessions of group CBT (N=12), multiple group sessions of exposure and response prevention (N=12), a single session of exposure and response prevention (N=11), or neither CBT nor exposure and response prevention (N=12). Active treatment groups had comparable outcomes at 14 weeks and at 6-month follow-up. In another study (Cooper and Steere 1995), participants received 19 treatment sessions of 50 minutes each over 18 weeks. More vomiting was noted in the exposure and response prevention group (N=16) at 12-month follow-up than in the CBT group (N=15); however, baseline vomiting rates were significantly higher in the exposure and response prevention group complicating interpretation. As discussed above, the study of Agras and colleagues (Agras et al. 1989) was also consistent with potentially detrimental effects of response prevention because a group with CBT in combination with response prevention (N=16) had a lower likelihood of achieving abstinence than a group that received CBT alone (N=17). Nevertheless, each of these studies had a small sample size and these findings cannot be viewed as definitive.

Several studies have conducted RCTs of stepped-care or other sequential approaches to treatment in an effort to optimize response. Katzman and colleagues (Katzman et al. 2010) used a two-phase study design with an initial phase of four weekly sessions and a second phase of 8 weekly sessions. One group received motivational enhancement therapy followed by individual CBT (N=79), a second group received motivational enhancement therapy followed by group CBT (N=73), and a third group received individual CBT followed by group CBT (N=73). A substantial number of participants in eachstudy arm withdrew from the study (41%, 48%, and 32%, respectively). Although improvements were noted in all groups, there did not seem to be a difference in response among the treatment strategies. Mitchell and colleagues (Mitchell et al. 2011) conducted a large RCT in which one group (N=147) received manual-based CBT (20 sessions of 50 minutes over 18 weeks) with fluoxetine (20 to 60 mg) added beginning at week 6 for participants who did not appear to be responding to CBT. The other group (N=146) was assigned to self-help in a tapering frequency of 20-minute sessions over 18 weeks with fluoxetine (20 to 60 mg) added beginning at week 6 for participants who did not appear to be responding to treatment and CBT added if response was still incomplete. At the end of treatment and at 1-year follow-up, abstinence rates were low (15% initial CBT vs. 11% initial self-help at end of treatment; 18% and 26%, respectively, at 1-year follow-up). There were also no treatment related differences in remission rates at either time point although binge-eating episodes and compensatory behaviors were significantly less in the initial self-help “stepped care” group at 1-year follow-up. These findings provide some reassurance that self-help could be used as an initial approach if other treatment is not readily available, with the caveat that additional intervention will be needed if response is not observed in a timely fashion. However, an RCT of guided self-help (N=31) as compared to CBT (N=31) suggested that individuals with high frequencies of binge-eating at baseline may do better with CBT than guided self-help although overall outcomes at the end of treatment and at long-term follow-up (mean 54.2 months) were comparable (Thiels et al. 1998, 2000, 2003).

The largest body of evidence on psychotherapy for BN is related to individual CBT, but several studies have also used group CBT or other formats for CBT delivery. Two studies with multiple treatment arms included a comparison of group CBT and a wait list control (Leitenberg et al. 1988; Sundgot-Borgen et al. 2002). Both studies were small with 12 to 15 participants per group and included 14 to 16 weeks of treatment, but both showed improvement in binge-eating and purging outcomes with group CBT but not with the wait list control condition. Another study (Bailer et al. 2004) compared group CBT (18 weekly sessions of 90 minutes; N=41) to guided self-help (self-help manual and weekly sessions of less than 20 minutes; N=40) and found sustained improvement and no difference in rates of remission or recovery at the end of treatment or at 1-year follow-up. Chen and colleagues (Chen et al. 2003) compared individual CBT (N=30) to group CBT (N=30) with 19 sessions of 50 minutes each over 4.5 months. Both treatments were comparable in terms of outcomes and study withdrawal rate at 6-month follow-up, however, more patients treated with individual CBT achieved abstinence from bulimic behaviors at the end of treatment. Nevonen and colleague (Nevonen and Broberg 2006) also compared individual CBT (N=42; weekly session of 50 to 60 minutes) to group CBT (N=44; weekly session of 2 hours). At the end of 23 weeks of treatment and at 1-year and 2.5-year follow-up assessments, rates of recovery and remission were comparable in the groups, although there did seem to be fewer binge episodes and fewer compensatory behaviors in the individual CBT group at 2.5-year follow-up. Complicating interpretation of this study was the use of IPT on an as needed basis for participants who had identifiable interpersonal issues.

A few studies have taken advantage of technological approaches to facilitate delivery of CBT. Zerwas and colleagues (Watson et al. 2017; Zerwas et al. 2017) in a large RCT of young adults (mean age 28) compared group CBT delivered in a face-to-face format (N=98) to group CBT delivered via a chat format (N=98). In addition to sixteen sessions of 90 minutes delivered over 20 weeks, participants received 2 sessions with a dietician and could receive concomitant medications or individual therapy. Study withdrawal rates were comparable in the two groups and both groups showed improvement; however, the face-to-face CBT group had greater abstinence rates at the end of treatment and a lower frequency of binge eating at follow-up. Nevertheless, a chat-based CBT format may be preferable to no treatment if face-to-face treatment is unavailable. Another large study conducted in Germany (Jacobi et al. 2017) used web-based CBT (11 sessions over 9 months) for relapse prevention after inpatient hospitalization and compared outcomes to TAU. Approximately one-third of the sample withdrew from the study and rates of vomiting were lower in the web-CBT group, although rates of abstinence from binge eating and purging were comparable between the two groups.

Grading of the Overall Supporting Body of Research Evidence for Cognitive-Behavioral Therapy in Bulimia Nervosa

- Magnitude of effect: The magnitude of effect was moderate. In the NMA, as compared to no treatment, individual CBT was associated with a reduction in binge-eating and purging frequencies by an average of 7 to 8 episodes per week. The relative likelihood of achieving abstinence from binge eating or purging was also increased by individual CBT, although CIs were wide and asymmetrical. Although there were fewer studies of group CBT, significant reductions in binge-eating and purging episodes were also observed with this approach to CBT delivery.

- Risk of bias: The risk of bias was high for 27 of the studies of CBT in BN, with a moderate risk of bias in 2 studies and a low risk of bias in 2 studies. In some instances, the method for random assignment was not well-delineated or missing data was not adequately accounted for in the analytic approach. In addition, in almost all of the studies, a high risk of bias was a result of needing to use self-reports of binge-eating and purging episodes in combination with the fact that participants were aware of the intervention that they were receiving. Even when other aspects of the study methodology were strong, this potential for confounding of results often led to a high risk of bias for the study as a whole.

- Applicability: The included studies all involve individuals with BN diagnosed using DSM criteria and treated in outpatient settings. Almost all of the studies were conducted in the US, the UK, Europe, or Australia. Although health system policies differ among these countries, the findings are expected to be generally applicable to US and Canadian patients. Study participants are primarily young, white, and female. Applicability of the evidence to older individuals and individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low rates of inclusion, limiting applicability of the findings. The studies showed heterogeneity in the number of vomiting episodes per week at baseline as well as in rates of laxative abuse. When trials reported information about the proportion of participants who had previously met criteria for a diagnosis of AN, there was substantial variability between studies. However, most studies did not report this information. Thus, the applicability of the overall findings to those with or without a prior history of AN is not clear.

- Directness: Direct. Although the majority of studies included a large number of outcome variables, almost all included outcomes related to binge eating, vomiting, response, or recovery as primary or secondary outcome measures.

- Consistency: In the studies that included TAU or wait list control as a comparator group, there was a consistent benefit of CBT. In other studies that compared CBT to other active interventions, there was also a consistent benefit for CBT, even when the two treatments being compared showed no difference in their effects.

- Precision: Imprecise. For comparisons in the NMA, CIs were wide and overlapped each other for most outcomes.

- Dose-response relationship: A single study looked at high intensity treatment as compared to low intensity treatment and found greater benefit with high intensity treatment. Nevertheless, additional confirmation is needed before reporting a definite dose-response relationship between treatment response and treatment frequency or duration.

- Confounding factors (including likely direction of effect): For all psychotherapy studies, the participant and the therapist are aware of the treatment that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention. This can present significant difficulties when self-reports of behavior are used as primary outcomes.

o Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

o Overall strength of research evidence: The overall strength of the research evidence is low. Although the studies of CBT in BN are consistent in showing a significant effect of treatment on binge-eating episodes, purging episodes, and likelihood of achieving abstinence from binge eating and purging, the high risk of bias in most of the studies contributes to a low strength of research evidence.

Detailed Review of Evidence: Serotonin Reuptake Inhibitors

Studies of antidepressant medications in the treatment of BN have primarily focused on SSRI antidepressants. A large RCT (Fluoxetine Bulimia Nervosa Collaborative Study Group 1992) compared 8 weeks of treatment with either fluoxetine 20 mg (N=129) or fluoxetine 60 mg (N=129) to placebo (N=129). Fluoxetine at a dose of 60 mg daily was associated with a greater decrease in weekly binge-eating ($p<0.001$) and vomiting episodes ($p<0.001$) than placebo, whereas fluoxetine 20 mg had intermediate effects on these outcomes. A greater reduction in weight (1.6 kg; $p<0.001$) as well as improvements in depressive symptoms ($p<0.033$) were also seen with fluoxetine 60 mg as compared to placebo. Fluoxetine treated groups had a greater number of reported adverse effects, but study withdrawal rates were comparable among the groups, suggesting that most adverse effects were tolerable. The study design did include an initial week of placebo in all groups, which may have reduced the number of individuals who would be likely to respond to placebo. A smaller study (Kanerva et al. 1995) compared 8 weeks of treatment with either 60 mg of fluoxetine daily or placebo and found reductions in binge eating in both groups, without a statistically significant difference with active treatment. Goldstein and colleagues (Goldstein et al. 1995), in a 16-week trial, also assessed the effects of 60 mg of fluoxetine (N=296) as compared to placebo (N=120). Fluoxetine at 60 mg daily was associated with significant reductions in weekly episodes of binge eating and vomiting as compared to placebo ($p=0.0002$ and $p<0.0001$, respectively). Rates of attrition were high (42.3% overall, with more study withdrawals for lack of efficacy in the placebo group than in the fluoxetine group). Rates of study withdrawal due to adverse effects were not statistically different between the groups although a larger proportion of participants treated with fluoxetine reported an adverse effect, most often insomnia, nausea, asthenia, anxiety, tremor, dizziness, or yawning. As with the Fluoxetine Bulimia Nervosa Collaborative Study Group trial, the study design incorporated an initial week of placebo treatment in all groups. Goldstein and colleagues (Goldstein et al. 1999) also conducted secondary analyses of their 16-week trial and the Fluoxetine Bulimia Nervosa Collaborative Study Group trial to determine whether improvements in BN were associated with changes in mood. Improvements in BN outcomes were found to be independent of baseline depression rating scores and unrelated to a prior or current diagnosis of a depressive disorder. Romano and colleagues (Romano et al. 2002) examined the effects of continuing 60 mg of fluoxetine daily (N=76) for up to 52 weeks as compared to a change to placebo (N=74) after an initial response of BN to fluoxetine. By the end of the study, only a small fraction of the initial sample remained although there was no difference between the groups. The time to relapse was greater in the fluoxetine continuation group and at 3 months the estimated relapse rate was 19% with fluoxetine as compared to 37% for placebo ($p<0.04$). Together, these studies suggest that fluoxetine at a dose of 60

mg daily is beneficial in the short-term treatment of BN and that it is likely to be beneficial in maintaining an initial response to treatment with fluoxetine.

Several smaller studies have examined treatment with fluoxetine in addition to other treatments for BN. Beumont and colleagues (Beumont et al. 1997) added fluoxetine 60 mg daily to weekly sessions of individual nutritional counseling (N=34) as compared to individual nutritional counseling alone (N=33). At the end of 8 weeks of treatment, both groups had shown decreases in binge episodes and vomiting episodes. These improvements were comparable between the treatment groups as were the rates of study withdrawals in the two groups. Mitchell and colleagues (Mitchell et al. 2001) randomly assigned participants to 16 weeks of treatment with fluoxetine 60 mg daily (N= 26), manual-based self-help treatment (N=22), manual-based self-help plus fluoxetine (N=21), or placebo (N=22). Fluoxetine and manual-based self-help were each associated with reductions in binge eating and vomiting but there was no synergistic effect of the two treatments or differences among the three active treatment arms in response.

CBT and group CBT were also studied in combination with fluoxetine. Goldbloom and colleagues (Goldbloom et al. 1997) randomly assigned participants to fluoxetine 60 mg (N=23), 10 sessions of CBT (N=24), or a combination of CBT and fluoxetine (N=29). Reductions in binge-eating and vomiting rates with combination treatment were greater than with fluoxetine alone but not statistically different from CBT alone; however, study withdrawal rates were also greater in the combination treatment group. A comparison of fluoxetine 60 mg (N=16), group CBT (N=19), and combination of fluoxetine and group CBT (N=18) also showed significant improvements in all groups (Jacobi et al. 2002). In this study, however, participants treated with CBT alone had greater abstinence from vomiting at 4 months of treatment and also had a higher proportion of study withdrawals.

Other SSRIs have been used clinically in patients who are unable to tolerate fluoxetine or who prefer a different medication, but studies of SSRI antidepressants other than fluoxetine have been limited. An 8-week study of citalopram 20 to 40 mg (N=10) as compared to placebo (N=10) demonstrated a significant reduction in binge-eating and purging episodes with citalopram and minimal change in these behaviors in the placebo group (Milano et al. 2005). A similar pattern was seen in a 12-week study (Milano et al. 2004) of sertraline 100 mg daily (N=10) as compared to placebo (N=10). When fluoxetine 60 mg (N=20), fluvoxamine 200 mg (N=20), and sertraline 100 mg (N=20) were compared in a 10-week trial (Milano et al. 2013), greater reductions in binge eating and vomiting were reported with fluoxetine and fluvoxamine. In an additional comparison of fluoxetine 20 to 60 mg daily (N=18) as compared to citalopram 20 to 40 mg daily (N=19), reductions in binge eating were greater with fluoxetine than with citalopram (Leombruni et al. 2006).

A number of other studies used complex study designs with multiple treatment arms and sequential addition of treatments, which made it difficult to draw specific conclusions about the benefits or adverse effects of SSRIs in the treatment of BN (Fichter et al. 1996; Mitchell et al. 2002; Schmidt et al. 2004; Walsh et al. 1997, 2000; Wilson et al. 1999).

Grading of the Overall Supporting Body of Research Evidence for Serotonin Reuptake Inhibitors in Bulimia Nervosa

o Magnitude of effect: The magnitude of effect of SSRIs in BN was low to moderate. For studies in which antidepressants (including SSRIs) were compared to placebo, binge-eating episodes were reduced by an average of 4.29 episodes per week. Participants who received an antidepressant were more than twice as likely to achieve abstinence from binge episodes as those who received placebo. In comparisons of antidepressants to CBT, the combination of CBT plus an antidepressant was typically no more effective than CBT alone, although combination treatment was superior to antidepressant alone.

o Risk of bias: The risk of bias was high for 15 of the studies of SSRIs in BN and moderate in 2 studies. In some instances, the method for random assignment was not well-delineated or missing data was not adequately accounted for in the analytic approach. In addition, in almost all of the studies, a high risk of bias was a result of needing to use self-reports of binge-eating and purging episodes. Many of these studies included psychotherapy treatment arms in which participants were aware of the intervention that they were receiving and this also affected risk of bias rating. Even when other aspects of the study methodology were strong, this potential for confounding of results led to a high risk of bias for the study as a whole.

o Applicability: The included studies all involve individuals with BN diagnosed using DSM criteria and treated in outpatient settings. Almost all of the studies were conducted in the US, the UK, Europe, or Australia. Although health system policies differ among these countries, the findings are expected to be generally applicable to US and Canadian patients. Study participants are primarily young, white, and female. Applicability of the evidence to older individuals and individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low rates of inclusion, limiting applicability of the findings. The studies showed heterogeneity in the number of vomiting episodes per week at baseline as well as in rates of laxative abuse. When trials reported information about the proportion of participants who had previously met criteria for a diagnosis of AN, there was substantial variability between studies. However, most studies did not report this information. Thus, the applicability of the overall findings to those with or without a prior history of AN is not clear.

O Directness: Direct. Although the majority of studies included a large number of outcome variables, almost all included outcomes related to binge eating, vomiting, response, or recovery as primary or secondary outcome measures.

o Consistency: In the studies of fluoxetine as compared to placebo, there was a consistent benefit of fluoxetine in participants with BN in terms of binge-eating outcomes, but the benefits in reducing vomiting episodes were not significant in the NMA. Studies of other SSRIs had smaller sample sizes and showed less consistent benefits related to either outcome.

o Precision: Imprecise. For antidepressants in the NMA, CIs were wide and overlapped each other. In a separate meta-analysis of studies of SSRIs, CIs were narrower but many included negative values.

o Dose-response relationship: There is evidence to support a dose-response effect with higher doses of fluoxetine showing greater clinical response than lower doses.

o Confounding factors (including likely direction of effect): The use of patient self-report data for frequencies of binge-eating and purging behaviors introduces a potential for confounding factors into the study. For studies that included a medication arm and a psychotherapy arm, the participant and the therapist are aware of the type of psychotherapy that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention; however, this is less likely to be a problem in placebo-controlled studies of antidepressant medications

o Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

o Overall strength of research evidence: The overall strength of the research evidence is low. Although the placebo-controlled studies of SSRIs in BN are consistent in showing a significant effect of treatment on binge-eating episodes and likelihood of achieving abstinence from binge eating, the high risk of bias in most of the studies and the lack of a significant effect on vomiting episodes or abstinence contributes to a low strength of research evidence.

Detailed Review of Evidence: Other Medications

Older studies of tricyclic antidepressants also showed reductions in symptoms but most of these studies also had small samples (Agras et al. 1987, 1992, 1994a; McCann and Agras 1990; Mitchell and Groat 1994; Mitchell et al. 1990). Of the monoamine oxidase inhibitors, phenelzine was associated with improvements in binge eating and some improvement in rates of abstinence from binge eating and purging, although side effects were more problematic (Walsh et al. 1984, 1985, 1988). In the one study of bupropion there were significant improvements in binge eating and purging; however, 4 subjects had generalized seizures. It is unclear whether this was specific to BN or to binge-eating and purging histories or due to a rapid increase in dose to a high dose of immediate release bupropion in the clinical trial. Nevertheless, there is an FDA boxed warning for bupropion as a result of this clinical trial experience and bupropion is contraindicated for use in individuals with BN.

Of medication treatments other than antidepressants, topiramate was studied in 2 trials and lithium in 1 trial. In a 10-week trial of flexibly-dosed topiramate (25 mg to 400 mg per day; mean dose 100 mg per day; N=35) in comparison to placebo (N=34), topiramate was associated with a decrease in weekly binge days, weekly purge days, binge frequency, and purge frequency (Hedges et al. 2003; Hoopes Scott et al. 2003). A comparable number of participants in the two treatment arms withdrew due to adverse effects. Another 10-week trial titrated topiramate to 250 mg daily (N=30) and, in comparison with placebo (N=30), also found topiramate to be well tolerated and associated with reductions in the frequency of binge eating and purging (Nickel et al. 2005). The sole study of lithium (Hsu et al. 1991) showed no difference between lithium (600-1,200 mg daily, mean serum level 0.62 mEq/L; N=47) and placebo (N=44) with 8 weeks of treatment.

Statement 14 – Family-Based Treatment in Adolescents and Emerging Adults With Bulimia Nervosa

APA suggests (2C) that adolescents and emerging adults with bulimia nervosa who have an involved caregiver be treated with eating disorder-focused family-based treatment.

Support for this statement comes from the expert survey (Appendix D) and from several RCTs of FBT. In the expert survey, FBT was rated as highly appropriate as an initial intervention in adolescents. The expert survey did not include questions about appropriateness of interventions in emerging adults, ages 18-26 years of age. The RCTs of FBT in BN were not included in the NMA because they did not meet the threshold of having at least 75% of the sample with DSM-defined BN. Consequently, the strength of research evidence is rated as low.

Detailed Review of Evidence: Family-Based Treatment

Schmidt and colleagues (Schmidt et al. 2007) randomly assigned participants to FBT (N=41) or individual guided self-help with CBT (CBT-GSH; N=44). Consecutively referred patients with DSM-IV defined BN (N=61) or eating disorder not otherwise specified (N=24) were invited to participate in the study if they were aged 13 to 20 and had a parent, other relative, or partner who could participate with them in the treatment. Individuals being treated with an antidepressant were able to enroll if their medication dose had been stable for at least 4 weeks. FBT lasted 6 months and included up to 13 sessions with their parent or care partner and 2 individual sessions. CBT-GSH used a manual for patients and close others (Schmidt and Treasure 1997) and incorporated 10 weekly sessions followed by 3 monthly sessions, with 2 optional sessions with their parent or care partner. The therapist's role was delineated in a clinician guide (Schmidt and Treasure 1997) and included motivating and guiding patients through the workbook and assigning and reviewing homework. Although individuals who received CBT-GSH were more likely to be abstinent from binge eating at the end of treatment than those who received FBT (41.9% vs. 25%, p=0.03), the two treatments did not differ in rates of abstinence from binge eating at the 6-month follow-up assessment (52% vs. 55%). Furthermore, rates of abstinence from vomiting did not differ for the two treatments either at the end of treatment or at the 6-month follow-up assessment (32.3% CBT-GSH vs. 28% FBT at end of treatment; 56% vs. 51.7%, respectively at 6-month follow-up). Nevertheless, both treatments were associated with significant improvements in binge eating and vomiting from the baseline assessment to the 6-month follow-up assessment. The study authors noted that some patients who would have been eligible for the study did not want family involved their care. In addition, attrition was approximately 25% during active treatment and comparable in the two groups.

Le Grange and colleagues (Le Grange et al. 2007) randomly assigned 80 individuals, aged 12 to 19, to either FBT or a manual-based form of supportive psychotherapy (SPT); an additional 25 individuals were eligible but did not wish to participate. In terms of diagnosis, 46% of participants met criteria for BN and 54% had BN symptoms. Approximately half of the sample had a concomitant mood disorder diagnosis and approximately one-third of the sample was receiving an antidepressant medication at the baseline assessment. After 20 outpatient sessions over 6 months, remission rates were significantly greater for FBT as compared to SPT (39% vs. 18%, p=0.049) and this difference was maintained at the 6-month follow-up assessment (29% vs. 10%, p=0.05). FBT also was superior to SPT on measures of eating

psychopathology as reflected by EDE and EDE-Q scores at the end of treatment but not at the 6-month follow-up assessment.

In a subsequent RCT, Le Grange and colleagues (Le Grange et al. 2015) compared FBT (N=52) to CBT that had been adapted for adolescents (N=58). A third treatment arm allocated patients to SPT (N=20) but was not included in the final statistical comparisons. Eligible participants were aged 12 to 18, lived with at least one parent, had a %EBW of at least 85%, and met criteria for DSM-IV defined BN or partial BN. Two thirds of the sample had a concomitant psychiatric disorder. The sample differed from many outpatient studies of BN in that 46% reported being from an ethnic minority group and 33% had previously been hospitalized for BN or associated medical complications. At the end of treatment, which averaged 14 sessions as well as at the 6-month follow-up assessment, individuals who received FBT were more likely to be abstinent from binge-eating and purging behaviors than those treated with CBT (39.4% vs. 19.7 % at the end of treatment; 44.0% vs. 25.4% at 6-month follow-up, respectively). By the 12-month follow-up assessment, however, the abstinence rates for the two treatments were not statistically different (48.5% vs. 32.0%, respectively). There were also no differences between the two treatments in the numbers of binge-eating episodes or numbers of purging episodes at the end of treatment or either follow-up assessment. On the other hand, more participants who received CBT were hospitalized (21%) as compared to those who received FBT (2%).

Grading of the Overall Supporting Body of Research Evidence for Family-Based Treatment of Bulimia Nervosain Adolescents and Emerging Adults

- Magnitude of effect: The magnitude of effect is moderate. As compared to SPT, FBT is associated with a greater likelihood of remission at the end of treatment and at 6-month follow-up. FBT was also associated with improvements from baseline in the studies comparing FBT to CBT-GSH and to CBT adapted for adolescents.

- Risk of bias: Of the RCTs on FBT in BN, 2 studies had a low risk of bias and 1 had a high risk of bias. A moderate or high risk of bias was most often associated with bias in measurement of outcome data.

- Applicability: The included studies all involve adolescents and some involve emerging adults with BN, diagnosed using DSM criteria. In addition, the studies also included individuals with binge eating and purging who did not meet full DSM criteria for BN. The studies were conducted in outpatient settings in the US or the UK. Study participants were typically white and female. Applicability of the evidence to individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low rates of inclusion, limiting applicability of the findings.

- Directness: Direct. Although the studies included other outcome variables, all included a measure related to binge eating or purging as a primary outcome.

- Consistency: Consistent. The studies all showed improvements with time in individuals who received FBT, although the three studies used different comparators and showed different effects relative to the comparison treatment.

- Precision: Not assessed. The studies were not included in the NMA.

- Dose-response relationship: There is insufficient information to determine whether there is a relationship between treatment response and treatment frequency or duration.

- Confounding factors (including likely direction of effect): For all psychotherapy studies, the participant and the therapist are aware of the treatment that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention.

- Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

- Overall strength of research evidence: The overall strength of research evidence is low. All three studies included a significant fraction of individuals who did not meet DSM criteria for BN. In addition, all three studies showed improvements in primary outcomes with FBT treatment, but FBT was comparable to CBT-GSH or CBT adapted for adolescents on most variables at follow-up assessments.

Detailed Review of Evidence: Other Psychotherapies

Stefini and colleagues (Stefini et al. 2017) in Germany randomly assigned 81 adolescents to a maximum of 60 sessions (mean of 36.6 sessions) of manual-based CBT or psychodynamic therapy over 12 months. Both groups showed improvement and remission rates did not differ between the treatments at the end of treatment (33% CBT vs. 30.2% psychodynamic therapy) or at 1-year follow-up. Although adherence was less with psychodynamic therapy and study withdrawal rates were greater with CBT, these differences also did not meet statistical significance.

Binge-Eating Disorder

Statement 15 – Psychotherapy in Patients With Binge-Eating Disorder

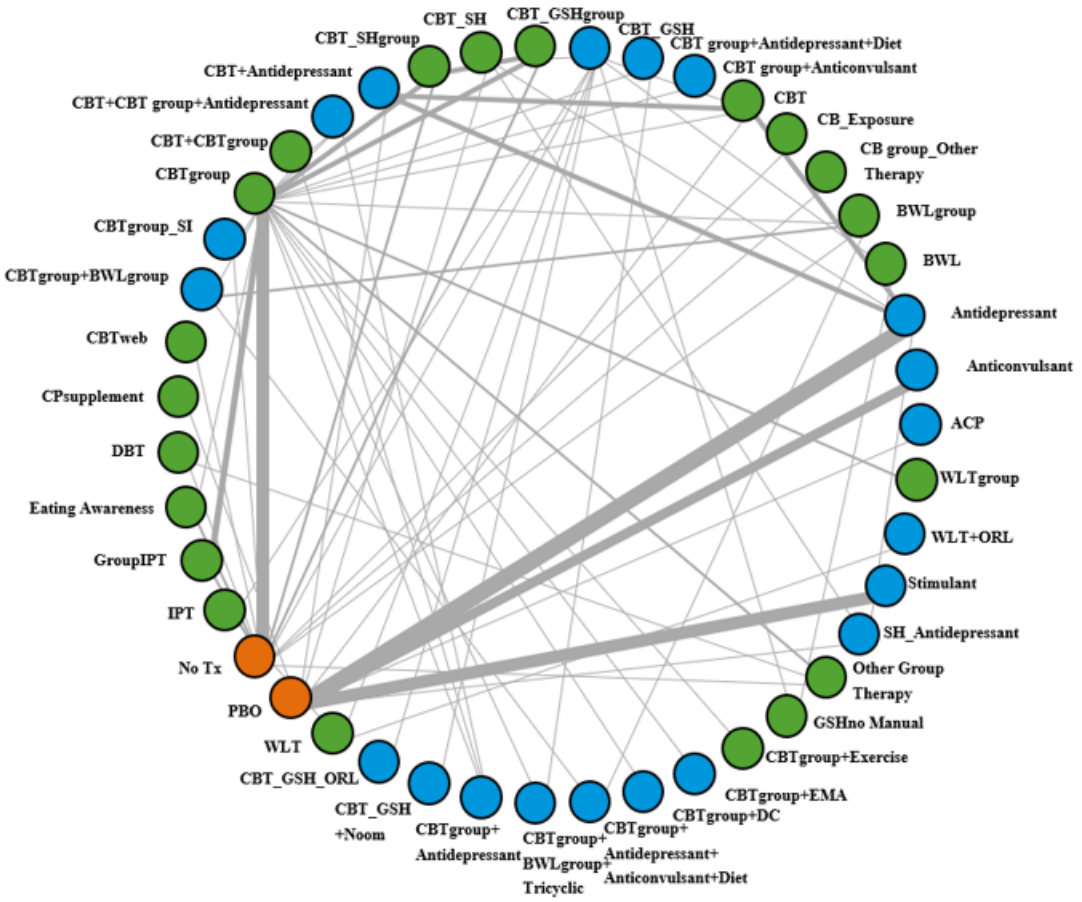
APA recommends (1C) that patients with binge-eating disorder be treated with eating disorder-focused cognitive-behavioral therapy or interpersonal therapy, in either individual or group formats.

Support for this statement comes from the expert survey (Appendix D) and from an NMA of studies of treatments of BED; however, the strength of research evidence is rated as low because of the high risk of bias of most of the studies. In the expert survey, individual CBT and psychoeducation were rated as highly appropriate for adolescents as well as adults. Nutritional rehabilitation, group therapy, individual DBT, and individual IPT were rated as moderately to highly appropriate for both adolescents and adults. For adolescents, family therapy was also rated as moderately to highly appropriate. In terms of initial interventions, psychotherapy alone was rated as highly appropriate, and a combination of medications and psychotherapy was rated as moderately to highly appropriate for both adolescents and adults.

Network Meta-Analysis of Treatments for Binge-Eating Disorder

Evidence on BED came from 76 unique RCTs (described in 81 publications), 3 non-randomized prospective studies, and 2 retrospective observational studies. Publication dates ranged from 1990 to 2017, with 14 from 1990 to 2000, 36 from 2001 to 2010, and 36 from 2011 to 2017. Most were conducted in the U.S. (44) or the United Kingdom (28) with smaller numbers of publications from other countries (Switzerland 3, Ireland 2, Austria 1, United Arab Emirates 1, unspecified country 7). For the overall NMA, the network contains 64 trials with 43 treatments and 6,887 subjects. In addition, the overall network of evidence is well connected; most treatments were connected to multiple treatments and most outcomes of interest remained connected to the network. Nevertheless, for some of the BED outcomes, networks were split into 2 or 3 distinct networks. Three studies (Hilbert and Tuschen-Caffier 2004; Leombruni et al. 2008; McIntosh et al. 2016) were endonodal and were not included in the NMA. No subgroup analyses were conducted due to the small number of studies for most comparisons. There were also an insufficient number of studies with consistent outcomes available for conducting sensitivity analyses for different durations of follow-up.

Demographic characteristics of study subjects were not reported consistently. In the 50 trials that reported age, the range was from 25 year to 52 years (mean age 42 years). in addition, there was generally a predominance of women in the trials (range 50 to 100% with 26 studies enrolling only women). Baseline mean BMI (reported in 55 trials) ranged from 26.7 to 44.6 kg/m^2 whereas baseline mean weight (reported in 37 trials) ranged from 189 to 272 lbs. Data on binge-eating frequency was variable and included independently reported rates as well as information on binge-eating frequency from the EDE Questionnaire and the EDE-I. Depression ratings also showed considerable variation; BDI scores were reported in 55 trials and ranged from 6.9 to 25.7. Thus, except for age, baseline demographic variables may contribute to heterogeneity of the findings. Heterogeneity is also likely to be increased as a result of small sample sizes in many studies as well as differences in follow-up durations and inclusion/exclusion criteria among studies.



Note: Nodes represent a treatment. Node colors indicate broader groups of the studied interventions. Labels represent included RCTs with direct comparisons for the corresponding edge. Line widths connecting the nodes are proportional to the number of studies that included a specific comparison.
Abbreviations: ACP=acamprosate; BWL=behavioral weight loss; CBT=cognitive-behavioral therapy; CPsupplement=chromium picolinate supplement; DBT=dialectical behavior therapy; DC=dietary counseling; EMA=ecological momentary assessment; GSH=guided self-help; IPT=interpersonal psychotherapy; ORL=orlistat; PBO=placebo; SH=self-help; SI=spouse involvement; TOP=topiramate; Tx=treatment; WLT=weight loss treatment

Table C-6: BED NMA feasibility and network characteristics

Outcome	Interventions : Total (NMA)	Studies: Total (NMA)	Trials per direct comparison	Total Subjects in NMA
BMI change from baseline	31	36	1-6	3,212
Weight change from baseline: Network 1	9	21	1-7	2,097
Weight change from baseline: Network 2	16	13	1-2	963
Binge-eating change from baseline (per week): Network 1	16	14	1-5	767
Binge-eating change from baseline (per week): Network 2	4	13	1-6	1,821
Binge-eating change from baseline (per month): Network 1	14	9	1	812
Binge-eating change from baseline (per month): Network 2	4	2	1	257
Binge-eating scale change from baseline (per month): Network 1	3	2	1	170
Binge-eating scale change from baseline (per month): Network 2	3	2	1	94

Binge-eating scale change from baseline (per month): Network 3	5	2	1	189
BDI score change from baseline	32	32	1-4	2,274
CGI-S change from baseline	4	11	1-4	914
CGI-I very much improved	3	12	3-5	1,608
Binge-eating abstinence	11	13	1-6	912
Remission: Network 1	8	13	1-6	680
Remission: Network 2	5	4	1	408
Remission: Network 3	7	3	1	385
Binge-eating disorder diagnosis met	7	6	1	482
Study withdrawal: Network 1	6	6	1	361
Study withdrawal: Network 2	17	19	1-4	1,395
Adherence, completed treatment: Network 1	8	13	1-6	1,909
Adherence, completed treatment: Network 2	8	5	1	511
Adherence, completed treatment: Network 3	8	4	1-2	451
Serious adverse events	3	6	2-4	1,599

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CGI-I=Clinical Global Impression-Improvement; CGI-S=Clinical Global Impression-Severity; NMA=network meta-analysis

Table C-7: Statistically favored comparisons from the BED NMA

Intervention	Comparison	Outcomes	Statistical values
Acamprosate	Anticonvulsants	Adherence, completed treatment	RR 1.84 (1.03, 3.69)
	Antidepressants	Adherence, completed treatment	RR 1.89 (1.06, 3.76)
	CBT + antidepressant	Adherence, completed treatment	RR 1.98 (1.08, 4.01)
	CBT	Adherence, completed treatment	RR 1.93 (1.03, 4.01)
Anticonvulsant	Placebo	BMI change	RMD -1.16 (-1.82, -0.47)
		Weight change	RMD -3.54 (-5.34, -1.62)
		CGI-S	RMD (-0.68 (-1.21, -0.12)
Antidepressant	Placebo	Weight change	RMD -1.58 (-3.10, -0.18)
		CGI-S	RMD -0.80 (-1.41, -0.28)
		CGI-I very much improved	RR 2.25 (1.40, 3.81)
		Remission	RR 2.03 (1.27, 3.58)
BWL-Group	No treatment	Binge eating, per month	RMD -5.73 (-10.16, -1.30)
	Antidepressants	Binge eating, per month	RMD -4.83 (-8.92, -0.79)
CBT	No Treatment	Binge eating, per month	RMD -8.88 (-12.34, -5.30)
	Placebo	Binge eating, per month	RMD -6.93 (-12.30, -1.50)
		Remission	RR 5.07 (1.45, 20.40)
	Anticonvulsants	Remission	RR 4.32 (1.04, 18.56)
	Antidepressants	Binge eating, per month	RMD -8.03 (-10.49, -5.45)
	CBT-GSH	BED diagnosis met	RR 0.61 (0.40, 0.91)
CBT + antidepressant	No Treatment	Binge eating, per month	RMD -8.87 (-12.87, -4.81)
	Placebo	Binge eating, per month	RMD -6.93 (-12.32, -1.55)
		BDI	RMD -5.02 (-9.44, -0.80)
		Remission	RR 4.51 (1.29, 18.22)
	Antidepressants	Binge eating, per month	RMD -8.00 (-10.51, -5.50)
		BDI	RMD -4.66 (-8.32, -1.12)
CBT group	No Treatment	Binge eating, per week	RMD -2.38 (-3.83, -1.00)
		Binge eating, per month	RMD -8.89 (-12.24, -5.42)
		BDI	RMD -4.00 (-7.51, -0.72)
		Binge eating Abstinence	RR 4.11 (1.44, 17.18)
	Placebo	Binge eating, per month	RMD -6.95 (-12.45, -1.38)

	Antidepressants	Binge eating, per month	RMD -8.04 (-10.82, -5.13)
	BWL group	Binge eating, per month	RMD -3.14 (-6.08, -0.18)
	CBT-GSH	BED diagnosis met	RR 0.44 (0.21, 0.93)
CBT Web	No treatment	Binge eating, per month	RMD -7.00 (-8.78, -5.19)
	Antidepressants	Binge eating, per month	RMD -6.16 (-10.88, -1.21)
CBT + CBT Group + antidepressant	CBT	BDI	RMD -7.26 (-14.32, -0.52)
	CBT Group	BDI	RMD -5.41 (-11.04, -0.31)
	GSH no manual	BDI	RMD -10.84 (-21.17, -0.34)
	Other group therapy	BDI	RMD -9.31 (-17.94, -1.36)
	No treatment	BDI	RMD -9.46 (-15.89, -3.45)
	WLT group	BDI	RMD -8.90 (-17.45, -0.77)
	Antidepressants	BDI	RMD -8.64 (-16.58, -1.14)
	CBT-GSH	BDI	RMD -7.92 (-15.65, -0.52)
	CBT + CBT group	BDI	RMD -9.05 (-17.42, -1.21)
	Placebo	BDI	RMD -9.05 (-17.42, -1.21)
CBT Group + antidepressant	No treatment	BDI	RR -5.90 (-11.54, -0.13)
CBT Group + antidepressant + anticonvulsant + diet	No treatment	Binge eating, per week	RMD -4.34 (-8.32, -0.52)
CBT group + BWL group	No Treatment	Binge eating, per month	RMD -9.92 (-14.23, -5.36)
	Placebo	Binge eating, per month	RMD -7.93 (-14.32, -1.74)
	Antidepressants	Binge eating, per month	RMD -9.05 (-13.31, -5.02)
	BWL group	Binge eating, per month	RMD -4.17 (-7.54, -1.00)
CBT group + ecological momentary assessment	CBT-GSH	BED diagnosis met	RR 0.26 (0.09, 0.71)
CBT group + exercise	No Treatment	BDI	RR -8.32 (-15.40, -1.30)
CBT group + other therapy	No Treatment	Binge eating, per week	RMD -3.35 (-6.39, -0.24)
CBT group spouse involvement	No Treatment	BDI	RR -8.12 (-14.90, -1.88)
CBT-SH	Placebo	Adherence, completed treatment	RR 1.68 (1.12, 2.67)
	Anticonvulsants	Adherence, completed treatment	RR 1.77 (1.15, 2.87)
	Antidepressants	Adherence, completed treatment	RR 1.82 (1.19, 2.95)
	CBT + antidepressant	Adherence, completed treatment	RR 1.90 (1.15, 3.18)
	CBT	Adherence, completed treatment	RR 1.87 (1.13, 3.02)
	Stimulants	Adherence, completed treatment	RR 1.69 (1.10, 2.68)
CBT-SH group	No treatment	Binge eating, per week	RMD -3.38 (-6.66, -0.32)
		Binge eating Abstinence	RR 6.35 (1.19, 57.19)
CBT-GSH	GSH No Manual	Binge eating, per month	RMD -3.31 (-6.62, -0.18)
CBT-GSH + NOOM	CBT-GSH	BDI	RR -6.36 (-12.51, -0.25)
CBT-GSH group	No treatment	Binge eating, per week	RMD -3.86 (-7.38, -0.49)
Dialectical behavior therapy	Other group therapy	Study withdrawal	RR 0.24 (0.04, 0.96)
Eating awareness	No Treatment	Binge eating, per week	RMD -5.19 (-8.48, -1.97)
		Binge eating, per month	RMD -8.87 (-11.45, -6.13)
		BDI	RR -6.91 (-12.10, -1.88)
	Placebo	Binge eating, per month	RMD -6.97 (-13.06, -0.80)
	Antidepressants	Binge eating, per month	RMD -8.06 (-12.01, -3.96)
	BWL Group	Binge eating, per week	RMD -4.46 (-9.01, -0.09)
Group IPT	No Treatment	Binge eating, per week	RMD -2.71 (-4.74, -0.76)
		Binge eating, per month	RMD -8.40 (-12.04, -4.55)
		Binge eating abstinence	RR 6.71 (1.29, 62.50)
	Antidepressants	Binge eating, per month	RMD -7.54 (-10.69, -4.19)
	Placebo	Binge eating, per month	RMD -6.44 (-12.14, -0.53)
Other group therapy	No Treatment	Binge eating, per month	RMD -8.43 (-13.73, -2.64)

	Antidepressants	Binge eating, per month	RMD -7.52 (-12.66, -2.12)
Stimulant	Placebo	BMI change	RMD -1.03 (-1.90, -0.15)
		Weight change	RMD -3.60 (-5.28, -2.11)
		Binge eating, per week	RMD -0.98 (-1.44, -0.49)
		CGI-I very much improved	RR 1.57 (1.12, 2.19)
WLT Group	No Treatment	Binge eating, per month	RMD -6.69 (-11.66, -1.87)
	Antidepressants	Binge eating, per month	RMD -5.85 (-10.22, -1.37)

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BWL=behavioral weight loss; CBT=cognitive-behavioral therapy; CBT-GSH=cognitive-behavioral therapy guided self-help; CBT-SH=cognitive-behavioral therapy self-help; CGI-I=Clinical Global Impression-Improvement; CGI-S=Clinical Global Impression; NMA=network meta-analysis; RMD=relative mean difference; RR=relative risk; WLT=weight loss treatment

Detailed Review of Evidence: Individual and Group Cognitive-Behavioral Therapy

In the NMA, CBT is associated with reductions in binge eating whether delivered in an individual format (RMD -8.88; 95% CI -12.34, -5.30 for binge eating per month vs. no treatment; RMD -6.93; 95% CI -12.30, -1.50 vs. placebo) or in a group format (RMD -2.38; 95% CI -3.83, -1.00 for binge eating per week and RMD -8.89; 95% CI -12.24, -5.42 for binge eating per month vs. no treatment; RMD -6.95; 95% CI -12.45, -1.38 for binge eating per month vs. placebo). In addition, CBT is associated with a greater likelihood of remission as compared to placebo (RR 5.07; 95% CI 1.45, 20.40), whereas group CBT is associated with a greater likelihood of binge-eating abstinence as compared to no treatment (RR 4.11; 95% CI 1.44, 17.18). Scores on the BDI were also reduced in individuals who receive group CBT as compared to no treatment (RMD -4.00; 95% CI -7.51, -0.72).

Several RCTs included individual CBT as one of the treatment arms. In the INTERBED study, de Zwaan and colleagues (de Zwaan et al. 2017) compared 20 weekly 50-minute individual sessions of CBT (N=89) to 11 internet modules of guided self-help with weekly email coaching (N=89). The number of binge-eating days was reduced by each of the treatments; however, the difference between the groups was significant only for the intention-to-treat analysis and not the per protocol analysis. In terms of binge-eating abstinence, which was a secondary outcome variable, CBT was superior to guided self-help at the end of treatment (61% to 36%) and at 6-month follow-up (58% to 38%). Ricca and colleagues (Ricca et al. 2010) used the approach of Fairburn (Fairburn 1995) and randomly assigned participants to 50-minute individual sessions or 60-minute group sessions with each treatment condition receiving 22 sessions during 24 weeks of treatment. Outcomes at the end of treatment and at 3.5-year follow-up were similar in reducing the number of binges per month at each time point. At the end of treatment, the rate of recovery was greater for those who received individual treatment than group treatment (33% vs. 16.7%) but recovery rates were comparable at the 3.5-year follow-up assessment (36.1% vs. 27.8%). The rate of study withdrawal was low and comparable for the two treatment formats (4.1% and 5.5% for individual and group CBT, respectively). McIntosh and colleagues (McIntosh et al. 2016) compared individual CBT using the Fairburn approach to two adaptations of CBT, schema therapy (which aimed to identify and change maladaptive cognitive schemas) and appetite focused CBT (which aimed to identify and respond to hunger and satiety cues through self-monitoring). All three treatment arms included weekly sessions for 6 months. Rates of treatment discontinuation were similar with the three treatments and rates of binge-eating abstinence were also comparable at the end of treatment and at

12-month and 24-month follow-up assessments. The characteristics of participants in this study differed from many studies of BED in that the sample was entirely female, that prior BN was common, and that a current or lifetime diagnosis of major depressive disorder was also frequent. One study compared CBT (N=27) delivered weekly for 12 weeks in 50-minute sessions to methylphenidate (N=22) in a dose of 18 to 72 mg (Quilty et al. 2019). Both treatments were associated with a decrease in objective and subjective binge-eating episodes but there was no significant difference between the groups.

Several small studies of group CBT included wait list or assessment only control conditions. Telch and colleagues administered group CBT (N=23) in 10 weekly sessions of 90 minutes each and found significant reductions from baseline in the number of binges per week and the number of binge days per week as compared to a wait list control condition (N=21; Telch et al. 1990). Group CBT remained superior to the wait list control at a follow-up assessment 10 weeks after the end of treatment. Three quarters of the participants attended at least 8 treatment sessions consistent with good adherence with treatment. Schlup and colleagues (Schlup et al. 2009) compared a wait list control comparison (N=18) to group CBT (N=18), which included 8 weekly sessions of 90 minutes each. Binge-eating episodes per week were significantly reduced with group CBT and 39% of the CBT treatment participants achieved abstinence from binge eating by the end of treatment as compared to 0% of the wait list control group. Treatment discontinuation was low in both groups. Gorin and colleagues (Gorin et al. 2003) compared a wait list control condition (N=31) to group CBT (N=32) and also included a group CBT condition with spousal involvement (N= 31). Groups consisted of 12 weekly sessions of 90 minutes each. Approximately one-third of participants withdrew from the study complicating interpretation of the results, but both active treatment groups showed improvement with no difference between group CBT with or without spousal involvement.

Peterson and colleagues (Peterson et al. 1998, 2001) compared a wait list control condition (N=11) to a therapist-led CBT group (N=16) as well as to self-help (N=15) and partial self-help (N=19) CBT groups. During the 8-week study, the active treatment conditions included 30 minutes of psychoeducation (either by a therapist or by videotape) and 30 minutes of group discussion in each of 14 sessions. Comparable reductions in objective and subjective binge-eating episodes occurred for all three active treatments immediately after treatment and at 1-month, 6-month, and 1-year follow-up assessments. In a subsequent study using a similar design (Peterson et al. 2009), active treatments included 15 sessions of 80 minutes each for 20 weeks with a tapering frequency of sessions. At the end of treatment, the therapist-led group (N=60) and therapist-assisted (N=63) groups had higher abstinence rates (51.7% and 33.3%, respectively) than self-help (N=67; 17%) or wait list control groups (N=69; 10.1%). Decreases in the number of binge episodes per month were also greater with therapist-led or therapist-assisted CBT groups than with self-help CBT or wait list control. Despite this, rates of abstinence were not statistically different among the groups at 6-month or 12-month follow-up assessments. Study withdrawal rates were also greater with self-help or therapist-assisted group CBT than with therapist-led group CBT or with the wait list control condition.

Two studies compared 12 weeks of group CBT to an assessment only control condition (Agras et al. 1995; Eldredge et al. 1997). In one study (Agras et al. 1995), significantly greater reductions were seen in binge days per week and abstinence rates were significantly greater with group CBT (N=39) than with

assessment only (N=11). After 12 weeks of group CBT, addition of IPT did not yield further improvements in those who had not yet responded to group CBT. In the other study (Eldredge et al. 1997), group CBT (N=36) was also associated with improvements in binge-eating behaviors from baseline as compared to a wait list control condition (N=10). In this study, participants who had not responded by 12 weeks of group CBT received additional group CBT, which was associated with treatment response in approximately half of initial non-responders.

Schag and colleagues (Schag et al. 2019) randomly assigned participants to self-monitoring (N=39) or group CBT (N=41) delivered in 8 weekly 90 min sessions, which was focused on reducing impulsivity. At the end of treatment, the number of binge-eating episodes in the prior 4 weeks was comparable for both treatment arms. At 3-month follow-up assessment, however, CBT was associated with fewer binge-eating episodes in the prior 2 months as well as lower levels of eating pathology and depression.

Lammers and colleagues (Lammers et al. 2020) compared outcomes of eating disorder-focused CBT (N=33) to DBT that had been adapted for BED (N=41). Both treatments were administered over 20 weeks in a group format (2 hours weekly for DBT, 90 minutes weekly for CBT); the CBT group members also received 6 90-minute group sessions for patients and their partners and up to 6 monthly sessions for relapse prevention. Participants were assigned to treatment groups in a quasi-random fashion. For inclusion, participants met criteria for BED and also had obesity and above average levels of emotional eating. At the end of treatment, as compared to DBT, the CBT group had fewer objective binge-eating episodes and lower levels of eating disorder psychopathology as measured by the EDE-Q global score. At the end of treatment and at 6-month follow-up assessment, a numerically greater proportion of individuals showed a clinically significant change with CBT than with DBT (69.9% vs. 52.9% at end of treatment; 65% vs. 45.8% at follow-up based on EDE-Q global score changes), but the two groups did not differ statistically.

Another study assigned participants to CBT or brief strategic therapy, but the 8 group sessions were delivered during a 4-week inpatient stay (Castelnuovo et al. 2011). Participants also received 8 sessions delivered individually by telephone during 6 months of outpatient treatment. Nutritional rehabilitation, a low-calorie diet, and moderate exercise were part of the intervention for both groups. Although weight change was comparable in the two groups, individuals who received the brief strategic therapy were more likely to show remission of BED at 6 months. At 1-year follow-up assessment, brief strategic therapy remained superior to CBT in reducing the frequency of binge eating and improving global functioning (Jackson et al. 2018).

Group CBT was also studied as an addition to a protein sparing modified fasting regimen and 12 weekly groups with a dietician (de Zwaan et al. 2005). Participants who received group CBT in addition to the very low-calorie diet (N=36) received an additional 90-minute group each week. At 18-month follow-up assessment, binge-eating abstinence rates were comparable in the two groups although study withdrawal rates were lower with adjunctive group CBT. Munsch and colleagues (Munsch et al. 2007, 2012) compared group CBT (N=36) to behavioral weight loss (BWL) treatment (N=36) with both treatment arms consisting of 16 weekly sessions of 90 minutes followed by 6 monthly sessions. Self-reported binge eating was less frequent in the CBT group at the end of treatment; however, rates of

abstinence from binge eating did not differ between the groups at 1-year or 6-year follow-up. Grilo and colleagues (Grilo et al. 2011) compared group CBT (N=45) to BWL therapy (N=45) and to sequential administration of the two treatment approaches (N=35). Each treatment consisted of 16 sessions of 60 minutes each delivered for 24 weeks. At 12-month follow-up, remission rates for binge-eating disorder were greatest in the CBT group (51%) as compared to the BWL group (36%) or the combined treatment group (40%). The frequency of binge-eating episodes per month showed a similar superiority of group CBT as compared to BWL therapy.

Small studies comparing group CBT to modifications of group CBT tended to show comparable effects. A study of group CBT with cognitive restructuring (N=14) and group CBT with exposure (N=14) showed comparable decreases in binges per week with 4 months of treatment (Hilbert and Tuschen-Caffier 2004). The addition of ecological momentary analysis to group CBT (N=19) did not show added benefits over group CBT alone (N=22), although both groups exhibited decreases in the number of binge-eating episodes per week during 12 weeks to treatment (Le Grange et al. 2002). As compared to 4 months of group CBT (N=17), addition of exercise sessions (N=20) and extension of treatment to 10 months, with and without exercise (N=24 and N=23, respectively) were associated with greater rates of binge abstinence, greater reductions in binge-eating days per week, and greater reductions in weight (Pendleton et al. 2002). When compared to behavioral treatment focused on eliminating binge eating (N=16), group CBT (N=21) showed better outcomes in terms of binge-eating episodes, binge-eating abstinence, and treatment discontinuation after completing 15 weekly sessions of 150 minutes each (Nauta et al. 2000, 2001). At a 1-year follow-up assessment, behavioral treatment was associated with fewer binge days per month, but study withdrawal rates were much higher for the behavioral treatment condition and CBT was superior to behavioral treatment in terms of shape, weight, and eating concerns.

Grading of the Overall Supporting Body of Research Evidence for Cognitive-Behavioral Therapy in Binge-Eating Disorder

o Magnitude of effect: The magnitude of effect is low to moderate, with some variation in the effects of CBT in BED in the NMA depending upon whether CBT is delivered in an individual or group format or whether outcomes were measured in terms of binge-eating episodes, binge-eating abstinence, or remission from BED. In the NMA, on average, CBT is associated with 6 to 9 fewer binge episodes per month. The likelihood of binge-eating abstinence or remission from BED is 4 to 5 times more likely in participants who received CBT, although the Cis were wide and asymmetrical.

O Risk of bias: The risk of bias was high for 29 of the studies of CBT in BED, with a moderate risk of bias in 1 study and a low risk of bias in 2 studies. In some instances, the method for random assignment was not well-delineated or missing data was not adequately accounted for in the analytic approach. In addition, in almost all of the studies, a high risk of bias was a result of needing to use self-reports of binge-eating episodes in combination with the fact that participants were aware of the intervention that they were receiving. Even when other aspects of the study methodology were strong, this potential for confounding of results led to a high risk of bias for the study as a whole.

O Applicability: The included studies all involve individuals with BED diagnosed using DSM criteria and treated in outpatient settings. Almost all of the studies were conducted in the US, the UK, Europe,

or Australia. Although health system policies differ among these countries, the findings are expected to be generally applicable to US and Canadian patients. Study participants are primarily young to middle-aged adults, white, and female, with a significant number of studies enrolling only women participants. Applicability of the evidence to adolescents, older individuals, and individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low rates of inclusion, limiting applicability of the findings. The studies showed heterogeneity in the number of binge-eating episodes per week at baseline, which may also influence applicability of the findings to some patients. There is also significant variability in the mean BMI values and weights of participants for the BED studies as a whole. This may also influence applicability of the findings, particularly to individuals with weights in the normal range or those with class 3 obesity.

O Directness: Direct. Although the majority of studies included a large number of outcome variables, almost all included outcomes related to binge-eating episodes, response, or recovery as primary or secondary outcome measures.

O Consistency: Consistent. Studies of CBT in BED typically find benefits of active treatment as compared to a wait list control group. Studies that compare different forms of CBT also are consistent in finding a benefit of treatment, even when the intervention and the active comparator do not differ in their effects.

O Precision: Imprecise. For comparisons in the NMA, Cis were wide, overlapped each other, and included negative values.

O Dose-response relationship: A single study compared the effects of two different lengths of CBT in BED and found a greater response but also a greater rate of treatment withdrawal in those who were randomly assigned to longer treatment. Further evidence is needed to determine whether there is a relationship between treatment response and treatment frequency or duration for individuals with BED who receive treatment with CBT.

O Confounding factors (including likely direction of effect): For all psychotherapy studies, the participant and the therapist are aware of the treatment that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention. This can present significant difficulties when self-reports of binge-eating episodes are used as primary outcomes.

O Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

O Overall strength of research evidence: The overall strength of the research evidence is low. Although the studies of CBT in BED are consistent in showing a significant effect of treatment on binge-eating episodes and the likelihood of achieving abstinence from binge eating, the high risk of bias in most of the studies contributes to a low strength of research evidence.

Detailed Review of Evidence: Individual and Group Interpersonal Psychotherapy

In the NMA, group IPT was associated with a greater likelihood of abstinence from binge eating as compared to no treatment (RR 6.71; 95% CI 1.29, 62.50). Group IPT also reduced the frequency of binge-eating episodes (RMD -2.71; 95% CI -4.74, -0.76 and RMD -8.40; 95% CI -12.04, -4.55 for binge-eating episodes per week and per month, respectively, compared to no treatment; RMD -6.44; 95% CI -12.14, -0.53 for binge-eating episodes per month, compared to placebo).

Individual IPT was not statistically different from other treatments or control conditions in the NMA, but was associated with better outcomes than BWL therapy in a large RCT. In this study, Wilson and colleagues (Wilson et al. 2010) randomly assigned participants to 21 hours of individual IPT in 20 sessions over 24 weeks (N=75), a comparable number and duration of BWL sessions that included recommendations for weekly exercise (N=64), or a guided self-help program using the Fairburn approach to CBT in which untrained graduate students provided approximately 5 hours of assistance in 10 sessions (N=66). At the end of treatment, the proportion of participants who had responded to treatment was comparable among the 3 groups; however, at 2-year follow-up assessments, the rates of response or remission were 43.9% for BWL therapy, 62.1% for guided self-help CBT, and 67.9% for IPT. In addition, rates of treatment discontinuation were significantly lower with IPT (7% vs. 28% with BWL therapy and 30% with guided self-help CBT).

Wilfley and colleagues (Wilfley et al. 1993) compared group IPT (N=18), group CBT (N=18), and a wait list control group (N=20). Although study withdrawal rates were low in all groups, there was significantly greater abstinence from binge eating with 16 weeks of active treatment (44% group IPT vs. 28% group CBT vs. 0% wait list control) and adherence was greater with IPT than with CBT (88% vs. 72% respectively). Tasca and colleagues (Tasca et al. 2006) used a similar study design but used a group interpersonal psychodynamic therapy approach rather than one derived from IPT for depression. With 16 weeks of treatment, binge-eating abstinence rates were 59.5% for group IPT (N=48), 62.2% for group CBT (N=47), and 9.1% for the wait list control group (N=40). Rates of binge-eating abstinence were maintained and were still comparable in the active treatment groups at the 68-week follow-up assessment. Rates of treatment discontinuation were also comparable for the two active treatments (Tasca et al. 2012). In a subsequent study Wilfley and colleagues (Hilbert et al. 2012; Wilfley et al. 2002) compared group IPT (N=81) to group CBT (N=81) with both treatments given in 20 weekly sessions of 90 minutes with 3 supplementary individual sessions. Rates of recovery from BED and reductions in the number of binge days per month were comparable for the two interventions at the end of treatment and at 1-year follow-up assessment. The rates of recovery remained comparable for the two treatments at 4 years; however, rates of recovery had dropped substantially as compared to the end of treatment (79% to 27.3% for group CBT vs. 59.7% to 22.2% for group IPT; Hilbert et al. 2012).

Grading of the Overall Supporting Body of Research Evidence for Interpersonal Psychotherapyin Binge-Eating Disorder

- Magnitude of effect: The magnitude of effect is moderate. In the NMA, group IPT was associated with 6 to 9 fewer binge-eating episodes per month as compared to placebo or a wait list control condition and a 6-to 7-fold increase in the likelihood of achieving abstinence from binge-eating episodes.

O Risk of bias: The risk of bias was high for all of the studies of IPT in BED. In some instances, the method for random assignment was not well-delineated or missing data was not adequately accounted for in the analytic approach. In addition, in almost all of the studies, a high risk of bias was a result of needing to use self-reports of binge-eating episodes in combination with the fact that participants were aware of the intervention that they were receiving. Even when other aspects of the study methodology were strong, this potential for confounding of results led to a high risk of bias for the study as a whole.

O Applicability: The included studies all involve individuals with BED diagnosed using DSM criteria and treated in outpatient settings in the US and Canada. Study participants are primarily young to middle-aged adults, white, and female. Applicability of the evidence to adolescents, older individuals, and individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low rates of inclusion, limiting applicability of the findings. The studies showed heterogeneity in the number of binge-eating episodes per week at baseline, which may also influence applicability of the findings to some patients. There is also some variability in the mean BMI values and weights of participants although most participants are overweight or obese.

O Directness: Direct. Although the majority of studies included a large number of outcome variables, almost all of the included outcomes related to binge-eating episodes, response, or recovery as primary or secondary outcome measures.

O Consistency: Consistent. In 2 studies of group IPT, consistent benefits of treatment were found relative to wait list control condition. In the other study, group IPT was associated with benefits, but no differences were seen between response to group IPT and group CBT.

o Precision: Imprecise. For comparisons in the NMA, CIs were wide and overlapped each other.

o Dose-response relationship: There is insufficient information to determine whether there is a relationship between treatment response and treatment frequency or duration.

o Confounding factors (including likely direction of effect): For all psychotherapy studies, the participant and the therapist are aware of the treatment that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention. This can present significant difficulties when self-reports of behavior are used as primary outcomes.

o Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

o Overall strength of research evidence: The overall strength of the research evidence is low. Although the studies of IPT in BED are consistent in showing significant effects of treatment on the numbers of binge-eating episodes and the likelihood of achieving abstinence from binge eating, the high risk of bias in all of the studies contributes to a low strength of research evidence.

Detailed Review of Evidence: Other Psychosocial Interventions

A number of studies have assessed other approaches to CBT including GSH and web-based CBT. Wagner and colleagues (Wagner et al. 2016) randomly assigned participants to a wait list control group (N=70) or to a web-based CBT program (N=69) that included 11 structured and personalized web-based writing assignments and therapist feedback during 16 weeks of treatment. At the end of treatment, the CBT treatment group had fewer binge-eating episodes per month and greater rates of recovery and remission than the wait list control group (47.8% vs. 4.3% and 14.6% vs. 0%, respectively) but a greater fraction of the CBT group withdrew from the study (27.5% vs. 8.6%). Carrard and colleagues compared a wait list control group (N=37) to web-based CBT (N=37), which was delivered in 11 modules over 6 months with assistance from coaches who monitored exercises and diaries (Carrard et al. 2011). Although a greater fraction of participants in the CBT group were abstinent from binge eating at the end of treatment, fewer than half of the participants completed all of the CBT modules. Grilo and colleagues (Grilo and Masheb 2005) compared self-monitoring of eating (N=15) to GSH using either CBT (N=37) or a BWL approach focused on moderate lifestyle changes, moderate caloric restriction, and increased physical activity (N=38). At the end of 12 weeks, none of the conditions was associated with significant weight loss but the number of binge episode per month was lowest with CBT and remission rates were greatest with CBT (46% vs. 18% for BWL and 13% for self-monitoring). Loeb and colleagues (Loeb et al. 2000) compared self-help CBT using the Fairburn book (N=20) to CBT with GSH (N=20), which added 6 coaching sessions of 30 minutes each during the 10-week study. GSH CBT was associated with a greater decrease in binge-eating episodes per month and a greater proportion of participants who achieved response or remission (50% vs. 30% with CBT self-help). Another study (Carter and Fairburn 1998) included a wait list control condition (N=24) as well as CBT self-help (N=24) and CBT guided-self-help (N=24) treatment arms. At the end of 12 weeks, wait list participants were randomly assigned to one of the other treatments and the active treatment results were pooled, complicating interpretation of the findings. Nevertheless, the active treatments were associated with a greater diminution in binge-eating episodes and greater rates of abstinence from binge eating than the wait list control condition. Addition of self-help CBT (N=24) to TAU (N=24) was not associated with a significant difference in binge-eating remission rates but did reduce the number of binge-eating episodes per month (Grilo et al. 2013). Similarly, addition of self-help CBT to a placebo condition in a multiple treatment arm medication study did not affect BED remission rates at the end of 4 months of treatment or at a 16-month follow-up assessment (Grilo et al. 2014). Only one study (Hildebrandt et al. 2017) has examined app-based approaches to self-monitoring in addition to guided-self-help CBT, but data are difficult to interpret because approximately one-third of the sample withdrew prior to the study endpoint and the sample included a mix of participants with BED and BN.

Grilo and colleagues (Grilo et al. 2020b) randomly assigned individuals with BED and obesity to 6 months of BWL treatment (16 50-60 minute sessions; N=39) or stepped care (N=152), which consisted of 15 sessions of BWL for all stepped care participants followed by CBT-GSH (11 individual sessions of 20-30 minutes) for those who did not respond to BWL. In addition, individuals in the stepped care group were randomly assigned to a weight-loss medication or placebo after 1 month of BWL. Both BWL and stepped care were associated with comparable rates of abstinence from binge eating at the end of treatment. In addition, rates of abstinence were lower in individuals who were randomly assigned to weight control

medications as compared to placebo; however, interpretation of these data are complicated by a change from sibutramine to orlistat mid-way through the study when sibutramine was removed from the market. At 6-month and 12-month follow-up assessments, there continued to be no significant difference between BWL and stepped care (Grilo et al. 2020a).

Although DBT was mentioned in the expert survey as a potential treatment for some individuals, the data on DBT from clinical trials is quite limited. In addition to the study comparing DBT to CBT described above (Lammers et al. 2020), one small study (Telch et al. 2001) compared a wait list control condition (N=22) to 20 weeks of DBT (N=22) that included 2 hours per week of group psychotherapy adapted for use in BED. Binge-eating abstinence at the end of treatment was more frequent among participants who completed DBT than those in the wait list control group (16 of 18 with DBT vs. 2 of 16 with the wait list control), however, 6 of the DBT treated participants had relapsed by 46-week follow-up assessment. Safer and colleagues (Safer et al. 2010) compared 21 weeks of treatment with DBT (N=50) to supportive Rogerian group therapy (N=51). Both treatments included 20 weekly sessions of 2 hours each. At the end of treatment, more participants in the DBT group were abstinent from binge eating (64% vs. 36%) and the rate of study withdrawal was also lower with DBT (4% vs. 33.3%);however, by the 34-week follow-up assessment binge-eating days per week and binge-eating abstinence rates were comparable for the two interventions. A small study (N=36; Klein et al. 2013) compared DBT to self-monitoring with DBT diary cards but the results are not possible to interpret because 64% of the participants in the DBT group withdrew by the end of the 16-week study. For these reasons, further study of DBT is needed before making any statements about its use for the treatment of BED.

Statement 16 – Medications in Adults With Binge-Eating Disorder

APA suggests (2C) that adults with binge-eating disorder who prefer medication or have not responded to psychotherapy alone be treated with either an antidepressant medication or lisdexamfetamine.

Support for this statement comes from the expert survey (Appendix D) and from an NMA of studies of treatments for BED (Appendix C, Statement 15); however, the strength of research evidence is rated as low because of the high risk of bias of most of the studies. In the expert survey, SSRI antidepressant medications were rated as moderately appropriate in adolescents and adults whereas lisdexamfetamine was rated as moderately appropriate in adults but less appropriate in adolescents. In addition, a combination of medication and psychotherapy were noted to be moderately appropriate in adolescents as well as adults.

Detailed Review of Evidence: Antidepressants

In the NMA, antidepressants as a group were associated with an increased likelihood of remission (RR 2.03 95% CI 1.27, 3.58) and increased likelihood of being very much improved (RR 2.25 95% CI 1.40, 3.81) compared to placebo whereas global symptom ratings were reduced (RMD -0.80 95% CI -1.41, -0.28).

Devlin and colleagues (Devlin et al. 2005) studied participants with BED who received 16 sessions of BWL treatment over 20 weeks and who had also been randomly assigned to CBT plus fluoxetine (N=28), CBT

plus placebo (N=25), fluoxetine (N=32), or placebo (N=31). Although all groups of participants exhibited reductions in binge frequency, those who received CBT in 20 weekly 45-minute sessions were more likely to achieve abstinence from binge eating whereas those who received fluoxetine (target dose 60 mg daily) had larger reductions in symptoms of depression. However, approximately one-third of participants withdrew from the study by the end of treatment. During a follow-up phase of the study, participants who had achieved a 75% decrease in binge-eating episode frequency were able to continue in monthly group sessions with ongoing fluoxetine for 18 additional months. Participants who received CBT continued to exhibit fewer binge-eating episode whereas those treated with fluoxetine continued to exhibit fewer symptoms of depression suggesting a maintenance of treatment benefits. Grilo and colleagues (Grilo et al. 2005a, 2012b) used a similar study design with a 16-week period of active treatment and an 80% study completion rate. Remission rates at the end of treatment were greater with CBT (61% with CBT plus placebo; N=28 and 50% with CBT plus fluoxetine; N=26) as compared to fluoxetine (22%; N=27) or placebo (26%; N=27). At the 12-month follow-up assessment, the superiority of CBT persisted although remission rates had fallen in all treatment groups. This study confirmed that fluoxetine alone is not as effective as CBT alone or CBT in combination with fluoxetine. Ricca and colleagues (Ricca et al. 2001) randomly assigned participants to fluoxetine (60 mg daily; N=21), fluvoxamine (300 mg daily; N=22), CBT (N=20), fluoxetine plus CBT (N=22), or fluvoxamine plus CBT (N=23). For the sample as a whole, 79% of the participants completed the study with some of the study withdrawals related to medication side effects (e.g., sleep disturbance, nausea, headache). In the groups that received CBT, body weight was modestly reduced at the end of treatment and at 1-year follow-up assessment. Scores on the EDE were reduced at the end of treatment and at 1-year follow-up in groups that had received CBT, with the greatest decrease in the group that received CBT plus fluvoxamine. In a 9-week multi-center RCT of flexibly dosed fluvoxamine (50-300 mg; mean 260 mg; N=42) as compared to placebo (N=43), fluvoxamine treatment led to greater rates of reductions in binge-eating frequency, Clinical Global Impression (CGI) severity scores, and BMI as compared to placebo (Hudson et al. 1998). However, the proportion of participants who achieved binge-eating remission did not differ between fluvoxamine and placebo in the intention-to-treat analysis. In addition, 21% of participants withdrew from the study, with adverse effects contributing to study withdrawals in the fluvoxamine treatment group. A smaller study compared 12 weeks of treatment with flexibly dosed fluvoxamine (N=9) or placebo (N=11) and found reductions in binge-eating episode frequency in both groups but no differences between the groups (Pearlstein et al. 2003).

Only one study used sertraline (100-200 mg daily; N=22) and compared it to treatment with fluoxetine (40-80 mg daily; N=20; Leombruni et al. 2008). Both groups showed significant reductions in binge frequency with 24 weeks of treatment but there were no differences between the groups. Another single study (Guerdjikova et al. 2008) compared 12 weeks of escitalopram (10-30 mg daily; mean 26.5 mg; N=21) to placebo (N=23). Weight and global severity of illness were reduced with escitalopram relative to placebo, but number of binge-eating days and frequency of binge-eating episodes did not differ between the groups. Guerdjikova and colleagues (Guerdjikova et al. 2012) studied 12 weeks of treatment with duloxetine (mean 78.7 mg daily; N=20) as compared to placebo (N=20) in participants with a co-occurring depressive disorder. Duloxetine treatment was associated with reductions in the weekly frequency of binge-eating days, binge-eating episodes, and weight relative to placebo but no

differences in depressive symptoms. Vortioxetine (20 mg daily) was compared to placebo in a 12-week trial (N=80; Grant et al. 2019). Both treatment arms were associated with reductions in binge-eating episodes with no differences between groups on efficacy measures or adverse effects. Bupropion (300 mg daily; N=31) was compared to placebo (N=30) in an 8-week trial but there were no significant differences between the groups on binge-eating frequency or rates of study withdrawal (White and Grilo 2013).

Two RCTs were conducted with tricyclic antidepressants. In an 8-week trial of impramine (75 mg daily; N=15) as compared to placebo (N=16), both groups had a decrease in the frequency of binge-eating episodes but there was no difference between the groups (Laederach-Hofmann et al. 1999). Weight loss was modest during the trial but was slightly greater in the imipramine treated group. Agras and colleagues (Agras et al. 1994b) used desipramine in one treatment arm of a sequential treatment trial. In this 9-month trial, weight loss treatment was compared to 3 months of CBT followed by 6 months of weight loss treatment; in the third treatment arm, the same sequence of CBT and weight loss treatment was used but desipramine (up to 300 mg nightly) was added in the final 6 months of treatment. At the end of treatment, binge-eating frequencies had decreased in all groups but there were no differences among the treatments. In addition, the desipramine group had lost more weight but this difference was modest. Although neither tricyclic antidepressant worsened weight gain in these studies, there was no benefit on binge-eating behaviors.

Grading of the Overall Supporting Body of Research Evidence for Antidepressants in Binge-Eating Disorder

o Magnitude of effect: The magnitude of effect is low. In the NMA, antidepressant treatment was associated with a greater likelihood of being very much improved or experiencing remission or symptom reduction. Nevertheless, in individual studies, antidepressant treatment did not always result in greater improvements than placebo.

o Risk of bias: The risk of bias was high for 10 of the studies of antidepressants in BED, with a moderate risk of bias in 2 studies. This was a result of needing to use self-reports of binge-eating episodes. In addition, a number of studies included a psychotherapy treatment arm and participants were aware of the intervention that they were receiving. Even when other aspects of the study methodology were strong, this potential for confounding of results led to a high risk of bias for the study as a whole.

o Applicability: The included studies all involve individuals with BED diagnosed using DSM criteria and treated in outpatient settings. Doses of antidepressant medications used were consistent with those typically used in clinical practice. Almost all of the studies were conducted in the US, the UK, Europe, or Australia. Although health system policies differ among these countries, the findings are expected to be generally applicable to US and Canadian patients. Study participants are primarily young to middle-aged adults, white, and female, with a number of studies enrolling only women participants. Applicability of the evidence to adolescents, older individuals, and individuals of other genders is unclear but likely to be diminished. Similarly, information on race, ethnicity, and other demographic characteristics of participants is often not reported but when it is noted, historically under-represented groups have low

rates of inclusion, limiting applicability of the findings. The studies showed heterogeneity in the number of binge-eating episodes per week at baseline, which may also influence applicability of the findings to some patients. There is also significant variability in the mean BMI values and weights of participants for the BED studies as a whole. This may also influence applicability of the findings, particularly to individuals with weights in the normal range or those with class 3 obesity.

o Directness: Direct. Almost all of the studies included outcomes related to binge-eating episodes, response, or recovery as primary or secondary outcome measures.

o Consistency: Inconsistent. In individual studies, antidepressant treatment did not always result in greater improvements than placebo.

o Precision: Imprecise. For comparisons in the NMA, CIs were wide and overlapped each other.

o Dose-response relationship: There is insufficient information to determine whether there is a relationship between treatment response and treatment frequency or duration.

o Confounding factors (including likely direction of effect): The use of patient self-report data for frequencies of binge-eating introduces a potential for confounding factors into the study. For studies that included a medication arm and a psychotherapy arm, the participant and the therapist are aware of the type of psychotherapy that is being received. Enthusiasm about a treatment (or conversely, lack of enthusiasm about a comparative intervention) could influence participants' response in favor of the intervention. However, this is less likely to be a problem in placebo-controlled studies of antidepressant medications

o Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

o Overall strength of research evidence: The overall strength of the research evidence is low. The studies of antidepressants in BED show inconsistent effects of treatment on binge-eating episodes and the high risk of bias in most of the studies contributes to a low strength of research evidence.

Detailed Review of Evidence: Lisdexamfetamine

In the expert survey, treatment with lisdexamfetamine was rated as mildly to moderately appropriate for adults and inappropriate to mildly appropriate in adolescents. In the NMA, treatment with a stimulant medication was associated with modest reductions in BMI (RMD -1.03; 95% CI -1.90, -0.15), weight (RMD -3.60; 95% CI -5.28, -2.11), and binge-eating episodes per week (RMD -0.98; 95% CI -1.44, -0.49) as well as an increased likelihood of being very much improved on a clinical global rating (CGI-Global Improvement RR 1.57; 95% CI 1.12, 2.19).

The majority of clinical trials of stimulants in BED have examined the effects of lisdexamfetamine on binge-eating behaviors. McElroy and colleagues (McElroy et al. 2015b, 2016b), in a multicenter trial in the US, compared placebo (N=64) to 3 dosages of lisdexamfetamine (30 mg, N=66; 50 mg, N=65; 70 mg, N=65). The dose of lisdexamfetamine was titrated during the initial 3 weeks of treatment and participants remained on the final dose for 8 weeks. The primary outcome of the study was a log

transformed measure of binge-eating days per week and this outcome was reduced in participants who received 50 mg or 70 mg of lisdexamfetamine as compared to placebo. Rates of binge-eating cessation were also greater at doses of 50 mg or 70 mg. Those treatment groups also experienced reductions in BED severity and modest reductions in weight. Although the number of adverse effects were greater in participants receiving lisdexamfetamine, study withdrawal rates were comparable. Two additional multi-center studies were conducted by McElroy and colleagues (McElroy et al. 2016a). One 12-week study compared lisdexamfetamine (N=192) to placebo (N=191) and used a dose of 50 to 70 mg, titrated based on initial response and tolerability. Rates of binge abstinence were higher with lisdexamfetamine (40% vs. 14.1% with placebo), more participants were rated as improved or very much improved (86% vs. 47% with placebo), and weight loss was also greater (mean loss of 6.25 kg with lisdexamfetamine vs. 0.1 kg with placebo). Both lisdexamfetamine and placebo treatment were associated with a decrease in binge days per week, but the magnitude was greater with lisdexamfetamine than placebo. The second study used the same design (N=195 in each group) and also found greater rates of abstinence from binge eating (36.2% vs. 13.1%), greater rates of being much improved or very much improved (86% vs. 43%), and greater amounts of mean weight loss (5.57 kg vs. 0.15 kg) with lisdexamfetamine as compared to placebo. A small (N=50) 12-week placebo-controlled trial (Guerdjikova et al. 2016) of flexibly-dosed lisdexamfetamine (20 to 70 mg daily, mean dose 59.6 mg) conducted in the US found a greater proportion of lisdexamfetamine-treated participants were much improved or very much improved (87% vs. 61% with placebo) with a greater proportion of those participants losing at least 7% of their body weight (26% vs. 0% with placebo). Side effects such as dry mouth, insomnia, and jitteriness were more prominent with lisdexamfetamine than placebo, but treatment discontinuation rates were comparable. Together, these studies suggest modest benefits of lisdexamfetamine; however, study participants were recruited primarily from primary care practices and are unlikely to be representative of individuals with BED in other contexts. A final study of lisdexamfetamine (Hudson et al. 2017) included an open label period of treatment with 50 to 70 mg of lisdexamfetamine after which participants were randomly assigned to continuation of the medication (N=137) or a change to placebo (N=138). The time to relapse after randomization was greater in those who continued lisdexamfetamine, relapse was less frequent (3.7% with lisdexamfetamine vs. 32.1% with placebo), and study discontinuation rates were lower with continuation of lisdexamfetamine (25.6% vs. 63.8% with placebo).

One study of armodafinil (McElroy et al. 2015a) compared 150 to 250 mg of armodafinil daily (mean dose 216.7 mg; N=30) to placebo (N=30). Both groups exhibited decreases in binge days per week and binge-eating episodes per week but the magnitude of these changes with treatment were comparable for armodafinil and placebo. Furthermore, almost half of the sample withdrew from the study, with no difference in armodafinil as compared to placebo.

Grading of the Overall Supporting Body of Research Evidence for Lisdexamfetamine in Binge-Eating Disorder

- Magnitude of effect: The magnitude of effect is low. In the NMA, individuals treated with lisdexamfetamine had approximately 1 less episode of binge eating per week than those treated with placebo and had a reduction in weight of 3 to 4 lbs.

- Risk of bias: The risk of bias was high for all of the studies of lisdexamfetamine in BED. This was a result of needing to use self-reports of binge-eating episodes. Even when other aspects of the study methodology were strong, this potential for confounding of results led to a high risk of bias for the study as a whole.

- Applicability: The included studies all involve obese individuals with BED diagnosed using DSM criteria and treated in outpatient primary care settings. Doses of lisdexamfetamine that were used in the studies were consistent with those typically used in clinical practice. All of the studies are conducted in the US and Europe. Although health system policies differ among these countries, the findings are expected to be generally applicable to US and Canadian patients. Study participants are primarily young to middle-aged adults, white, and female. Applicability of the evidence to adolescents, older individuals, and individuals of other genders is unclear but likely to be diminished. Information on race and ethnicity was provided and non-white participants made up about 20% of the sample; however, results were not analyzed by demographic subgroup making the applicability of the study conclusions unclear to these individuals. The studies focused on individuals who were obese with mean values of BMI consistent with class 3 obesity for most participants. This may also influence applicability of the findings, particularly to individuals who are overweight or have weights in the normal range. It is also unclear whether the findings are relevant to patients seen in specialty settings rather than in primary care.

- Directness: Direct. The studies included outcomes related to binge-eating episodes, response, or recovery as primary or secondary outcome measures.

- Consistency: Consistent. Although the studies used somewhat different outcome measures, they were consistent in showing modest benefit for lisdexamfetamine.

- Precision: Imprecise. For comparisons in the NMA, CIs were wide and overlapped each other.

- Dose-response relationship: There appears to be a dose-response relationship for lisdexamfetamine in BED, with higher doses (50-70 mg/day) being associated with a greater clinical response than lower doses (30 mg/day) or placebo.

- Confounding factors (including likely direction of effect): The use of patient self-report data for frequencies of binge-eating behaviors introduces a potential for confounding factors into the study. However, in the placebo-controlled studies of lisdexamfetamine, this is less likely to be a problem than in psychotherapy studies of BED for which participants and therapists are aware of the treatment that they are receiving.

- Publication bias: Although there is no specific evidence to suggest publication bias, it may be present given the tendency for positive findings to be published more often than negative ones.

- Overall strength of research evidence: Although the studies of lisdexamfetamine in BED are consistent in showing an effect of treatment on binge-eating episodes and on body weight, the high risk of bias in all of the studies contributes to a low strength of research evidence.

Detailed Review of Evidence: Topiramate

Topiramate was not suggested for use in BED because the potential benefits did not seem to outweigh the potential side effects for the majority of patients. Nevertheless, some clinicians have found topiramate to be beneficial and tolerable, particularly at doses of 125 mg or less. In the expert survey, topiramate was rated as inappropriate to mildly appropriate for adolescents and mildly to moderately appropriate for adults. Three studies have compared topiramate to placebo. McElroy and colleagues (McElroy et al. 2003) used a flexible dose of topiramate that was titrated over 10 weeks (25 to 600 mg daily; median dose 212 mg per day). At the end of 14 weeks of treatment, both topiramate and placebo groups had a decrease in the frequency of days with binge episodes and this was more pronounced in the topiramate group (93% decrease with topiramate vs. 46% decrease with placebo). Topiramate-treated participants also lost more weight than those treated with placebo but side effects including confusion, paresthesias, and dysgeusia were also greater with topiramate as compared to placebo. In a larger multicenter trial (McElroy et al. 2007b), topiramate (N=204) was titrated over 8 weeks using flexible dosing (25 to 400 mg daily; median dose 300 mg daily). As in the smaller study, there was a greater reduction in weight, BMI, binge days per week, and binge episodes per week with topiramate than with placebo, but side effects (e.g., confusion, memory impairment, paresthesias, dysgeusia, upper respiratory tract infections) were also greater with topiramate than with placebo. Claudino and colleagues (Claudino et al. 2007) used a slightly different clinical trial design in which all participants received group CBT and, after a 2-to-5-week run-in period, participants were assigned to placebo (N=36) or topiramate (N=37). The dose of topiramate was slowly titrated at 25 mg every 2 weeks to a target dose of 200 mg daily with additional adjustment to 300 mg daily based on response. Decreases in binge episode frequency did not differ between the two groups but more topiramate-treated participants achieved remission from BED (83.8% vs. 61.1% for placebo). In addition, weight loss was greater with topiramate (6.8 kg vs. 0.9 kg with placebo) and a greater proportion of topiramate treated participants lost more than 10% of their body weight (36% vs. 11.1% with placebo). However, as in the other studies, rates of paresthesias and dysgeusia were greater with topiramate as compared to placebo.

Night Eating Syndrome

Evidence on the treatment of night eating syndrome is limited to two small U.S. studies of an SSRI. In each study the mean age of participants was approximately 45 years with half to two-thirds of the samples being women. An 8-week, government funded RCT (total N=34) of flexibly dosed sertraline (50-200 mg/day) showed greater improvement in night eating symptoms, CGI severity, and quality of life ratings as compared to placebo as well as a greater amount of weight loss (O'Reardon et al. 2006). Rates of attrition were low with 1 subject in each group withdrawing for lack of efficacy and no study withdrawals due to adverse effects. A 12-week, industry funded RCT (total N=40) compared escitalopram (10 mg/day for 4 weeks followed by 20 mg/day for 8 weeks) to placebo (Vander Wal et al. 2012). No differences were noted in total scores on the Night Eating Questionnaire and the two groups did not differ in rates of remission or response. One subject who was treated with escitalopram stopped treatment due to adverse effects but there was no other study attrition reported. These limited findings do not allow any conclusions to be drawn on treatment of night eating syndrome.

Avoidant/Restrictive Food Intake Disorder

No studies of treatments for ARFID met the inclusion criteria for the systematic review. Since ARFID was defined in DSM-5 (American Psychiatric Association 2013), the literature includes case reports of treatment as well as case series, retrospective chart review studies, and pilot prospective trials aimed at assessing treatment feasibility in individuals with ARFID. This limited literature suggests a possible role of CBT adapted for ARFID (Dumont et al. 2019; Spettigue et al. 2018; Thomas et al. 2020, 2021), FBT adapted for ARFID (Lock et al. 2019), Supportive Parenting for Anxious Childhood Emotions adapted for ARFID (Shimshoni et al. 2020), Young Adult Temperament Based Treatment with Supports (Knatz Peck et al. 2021), or an intensive multidisciplinary feeding intervention (Sharp et al. 2016; Volkert et al. 2021). Medications with potential utility in treatment of ARFID include SSRIs (Mahr et al. 2021), hydroxyzine (Mahr et al. 2021), and olanzapine (Brewerton and D'Agostino 2017). Nevertheless, these and other pharmacotherapies and psychotherapies require more rigorous clinical trials before specific recommendations about ARFID treatment will be possible.

Appendix D. Findings from Expert Survey on Evaluation and Treatment of Patients With an Eating Disorder

Background

An expert opinion survey was fielded to 338 experts on treatment of eating disorders. These experts were identified through a blind, "snowball" nomination process. The experts were first peer-nominated by current and past chairs of academic departments of psychiatry, directors of psychiatry residency programs in the United States, leadership of other relevant medical organizations, and the members of the APA Assembly and Board of Trustee. Then, the experts nominated identified additional experts, and the process was repeated twice. The nominators were asked to identify two types of experts to participate in the survey: researchers and clinicians. "Research experts" were defined as individuals who have significant research activities, scholarly publications, or academic reputation in the treatment of eating disorders, particularly AN, BN, and BED. "Clinical experts" were defined as individuals who have substantial clinical experience in the treatment of eating disorders. The experts were contacted via email to complete the survey online.

The response rate for the survey was 56.2% (190/338); 11.8% of the responses were partial, meaning that at least one question in the main sections on assessment and treatment was completed. The experts who responded to the survey comprised approximately 51.1% clinical experts, 23.2% research experts, 23.7% experts in both categories, and 2.1% unspecified experts.

About half of the experts who responded to the survey, 54.2%, were nominated once, 20.5% were nominated twice, and the remainder were nominated up to 20 times.

Figure D- 1. Categories Experts Nominated

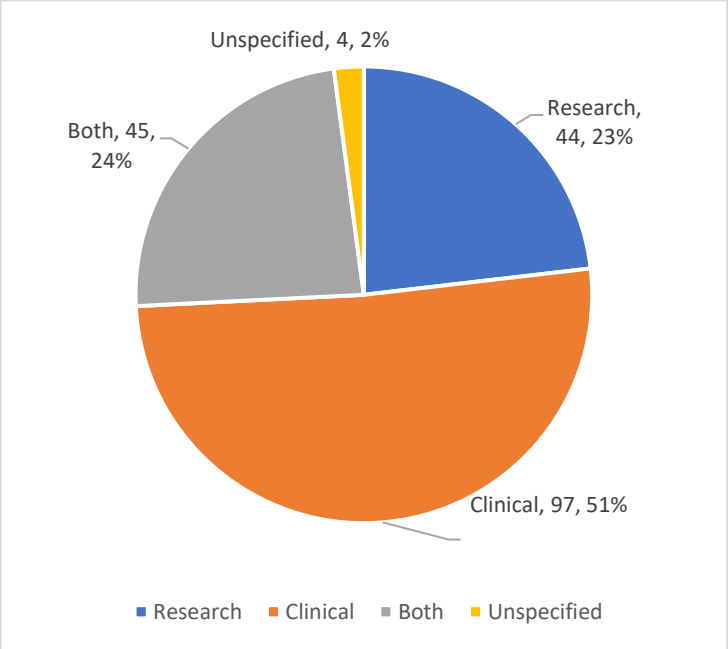
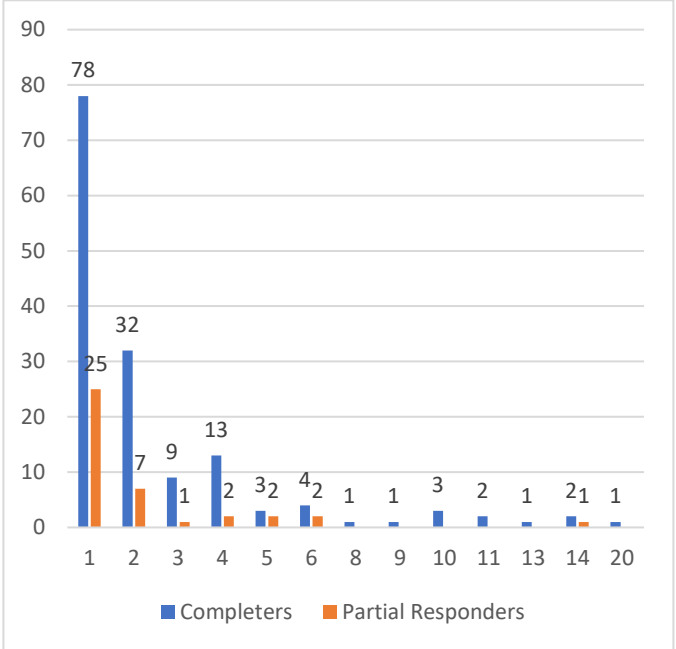
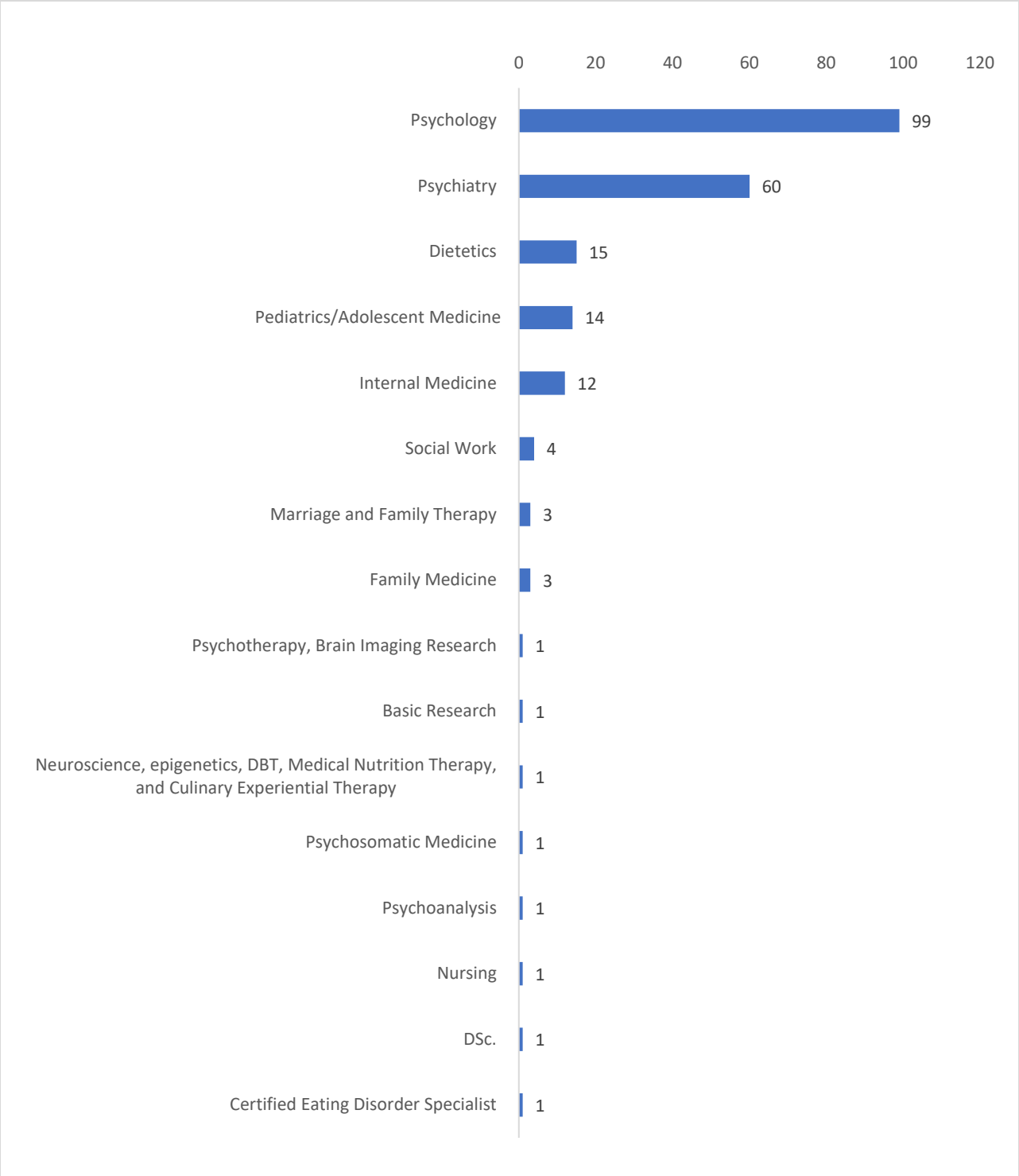


Figure D- 2. Number of Times Experts Nominated



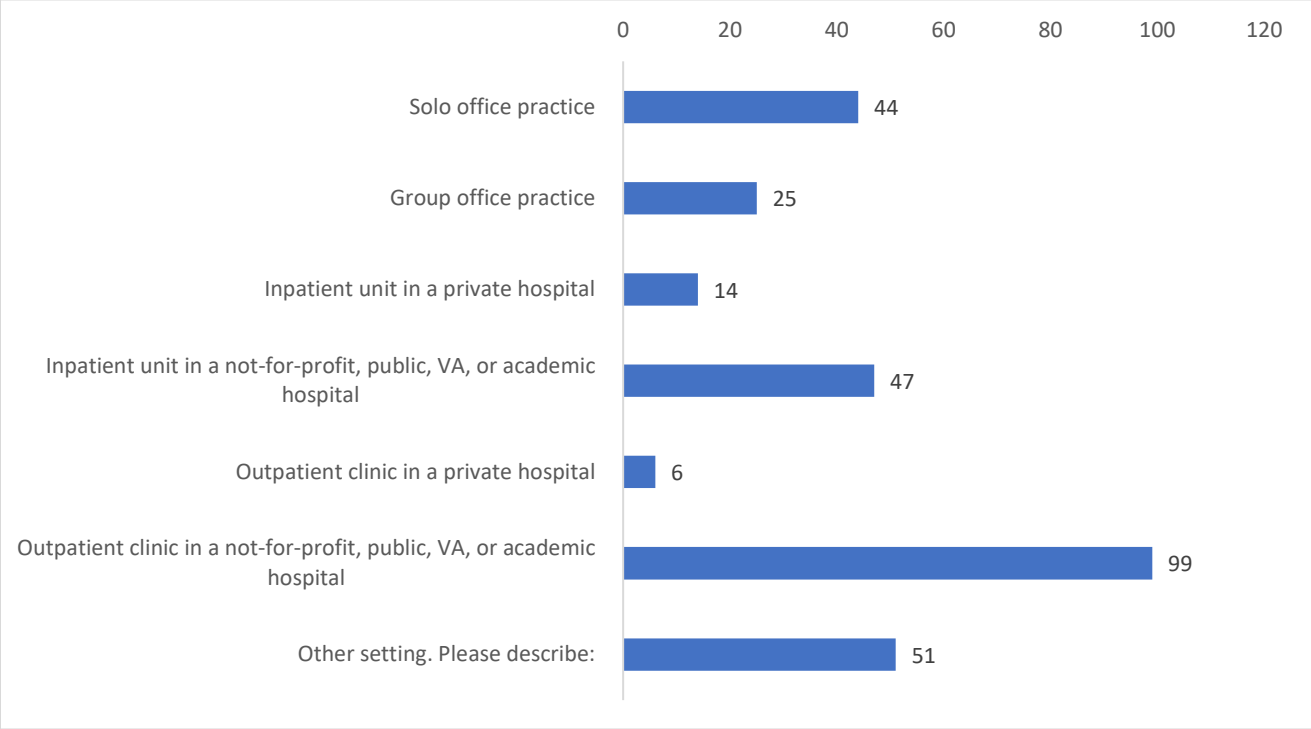
Section I. Clinical Expertise of Survey Respondents

Figure D- 3. Disciplines that describe their professional training, background, and focus of practice or research



Note: survey respondents were allowed to check multiple options

Figure D- 4. Clinical practice setting



Note: survey respondents were allowed to check multiple options

Figure D- 5. Years in practice, not including training

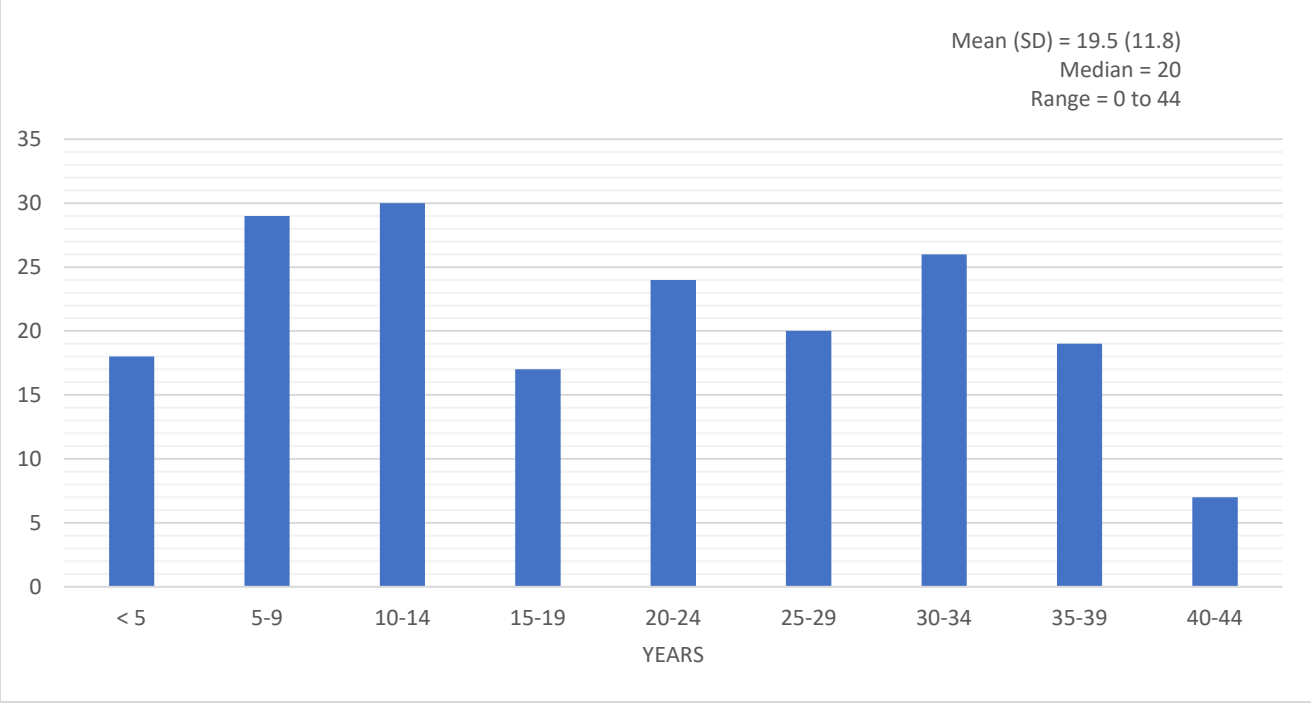
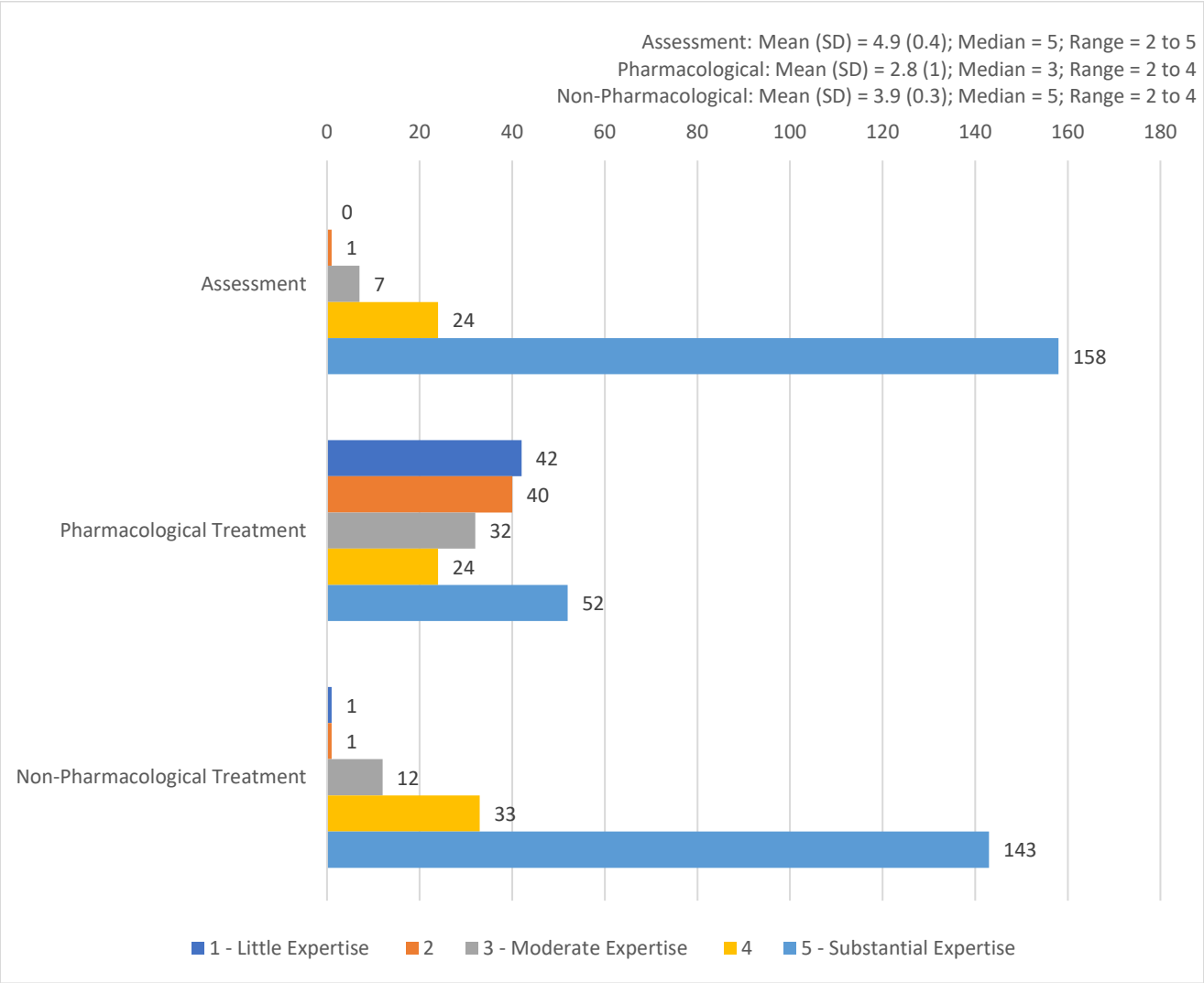


Figure D- 6. Degree of expertise in the assessment and treatment of individuals with eating disorders



Section II. Assessment and Determination of a Treatment Setting

Table D- 1. Appropriateness as part of the initial assessment of an individual with eating disorder symptoms

	Median	Mean	SD	Min	Max	N
Patterns of restrictive eating	5	5	1.3	3	5	187
Patterns of self-induced vomiting	5	5	1.8	2	5	188
History of changes in weight (and height for adolescents)	5	4.9	1.3	3	5	187
Patterns of binge eating	5	4.9	1.8	2	5	188
Patterns of laxative use	5	4.9	1.8	2	5	187
Patterns of other compensatory or purging behaviors	5	4.9	1.2	3	5	186
Core attitudes related to weight, shape, and eating	5	4.9	1.8	2	5	186
Past treatment for eating disorder and treatment response	5	4.9	1.7	2	5	186
Current suicidal ideas, plans, or intentions	5	4.9	2.3	1	5	186
Patterns of exercise	5	4.8	1.7	2	5	188
Current or past psychiatric diagnoses (including mood disorder, anxiety disorder, OCD, PTSD, ADHD, and alcohol or other substance use disorder)	5	4.8	1.7	2	5	183
Current psychiatric symptoms (including anxiety and mood symptoms)	5	4.8	2.3	1	5	186
Current height and weight, with calculation of BMI	5	4.8	2.3	1	5	184
History of suicidal behaviors or non-suicidal self-injury	5	4.7	2.2	1	5	186
Current vital signs, including orthostatic blood pressure and temperature	5	4.7	1.6	2	5	186
Laboratory testing for electrolyte abnormality (e.g., basic metabolic panel)	5	4.7	1.7	2	5	186
Psychosocial stressors including family/relationship stressors	5	4.6	1.6	2	5	186
History of abuse or neglect (including physical, emotional, or sexual)	5	4.6	2.1	1	5	184
Menstrual history, including changes in menses	5	4.5	2.1	1	5	188
Past psychopharmacologic treatment and response	5	4.5	1.5	2	5	187
Family history of eating disorders including obesity	5	4.5	2.1	1	5	186
Laboratory testing for anemia or other hematologic abnormality (e.g., complete blood count)	5	4.5	1.5	2	5	184
Family attitudes and interactions related to eating	5	4.4	1.4	2	5	185
Family history of psychiatric illness	5	4.4	2	1	5	187
Evidence of self-injury	5	4.4	2	1	5	185
Current cardiovascular function, including peripheral vascular function	5	4.2	1.9	1	5	185

Laboratory testing for thyroid abnormality (e.g., thyroid stimulating hormone)	5	4.2	1.3	2	5	184
Bone density testing if amenorrheic for at least 6 months	5	4.2	1.9	1	5	185
Electrocardiogram	5	4.2	1.8	1	5	184
Current evidence of dermatological manifestations of eating disorders	4	3.7	1.6	1	5	183
Dental examination in individuals with a history of purging	4	3.7	1.6	1	5	185
Bone density testing regardless of menstrual status	3	3.1	1.4	1	5	182

Note: sorted by median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate

Abbreviations: ADHD=attention-deficit/hyperactivity disorder; BMI=body mass index; OCD=obsessive-compulsive disorder; PTSD; posttraumatic stress disorder; SD=standard deviation

Table D- 2. Appropriateness as factors that suggest needing a higher level of care – adolescents and adults

	Adolescents with AN or BN						Adults with AN or BN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Other evidence of medical instability (e.g., significant electrolyte imbalance, arrhythmia)	5	4.9	1.2	3	5	178	5	4.8	1.7	2	5	177
Medical complications of vomiting, including uncontrolled vomiting, hematemesis	5	4.8	1.2	3	5	179	5	4.8	1.7	2	5	179
Rapid decline in weight despite treatment	5	4.7	2.2	1	5	178	5	4.5	2.1	1	5	172
Marked orthostasis	5	4.5	2	1	5	177	5	4.3	1.9	1	5	174
Co-occurring psychiatric symptoms or diagnoses including significant alcohol or substance use or personality disorders that require a different level of care in their own right (e.g., suicidal ideation, plans, or intent)	5	4.5	2.1	1	5	179	5	4.5	2.1	1	5	177
Rapid persistent decline in oral intake	5	4.4	2	1	5	175	5	4.3	1.9	1	5	173
Poor glucose control (i.e., significant hypo- or hyper-glycemia) in an insulin dependent diabetic in association with an eating disorder	5	4.4	2	1	5	176	5	4.4	2	1	5	176
Chronic medically or functionally impairing treatment-resistant symptoms (e.g., persistence despite several months of intensive treatment)	5	4.4	2	1	5	177	4	4.1	1.8	1	5	175
Complicated pregnancy	5	4.3	1.9	1	5	168	5	4.3	1.9	1	5	168
Resistance, denial, poor motivation, and/or lack of cooperation with treatment in the presence of medically or functionally impairing symptoms	5	4.2	1.8	1	5	175	4	4	1.7	1	5	174
Co-occurring psychiatric symptoms or diagnoses including significant alcohol or substance use or personality disorders that are complicating treatment of the eating disorder	4	4	1.7	1	5	176	4	3.8	1.6	1	5	174

Lack of access to an otherwise appropriate level of care (e.g., due to lack of geographic accessibility, lack of insurance coverage)	4	3.9	1.7	1	5	173	4	3.8	1.6	1	5	171
Prior lack of response at similar or lower levels of care	4	3.8	1.6	1	5	176	4	3.6	1.5	1	5	175
Prior medical instability at a weight that is similar to the current weight	4	3.7	1.6	1	5	175	4	3.5	1.5	1	5	171
Weight that is below the individual's estimated healthy weight (e.g., less than 85% of IBW)	3	3.4	1.5	1	5	179	3	3.2	1.4	1	5	175
Poor or limited community support system	3	3.4	1.5	1	5	173	3	3.1	1.4	1	5	171
Psychosocial stressors that are impacting intake/weight	3	3.1	1.4	1	5	176	3	2.9	1.4	1	5	172

Note: sorted by adolescents with AN or BN's median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; IBW=ideal body weight; SD=standard deviation

Table D- 3. Appropriateness as factors that suggest needing a higher level of care – adults with SEED and adults with BED

	Adults with SEED						Adults with BED					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Medical complications of vomiting, including uncontrolled vomiting, hematemesis	5	4.8	1.7	2	5	179	5	4.2	1.9	1	5	138
Other evidence of medical instability (e.g., significant electrolyte imbalance, arrhythmia)	5	4.8	1.7	2	5	177	5	4.3	1.9	1	5	156
Rapid decline in weight despite treatment	5	4.5	2.1	1	5	172	3.5	3.4	1.5	1	5	144
Co-occurring psychiatric symptoms or diagnoses including significant alcohol or substance use or personality disorders that require a different level of care in their own right (e.g., suicidal ideation, plans, or intent)	5	4.5	2.1	1	5	177	5	4.3	2	1	5	172
Poor glucose control (i.e., significant hypo- or hyper-glycemia) in an insulin dependent diabetic in association with an eating disorder	5	4.4	2	1	5	176	4	3.9	1.7	1	5	168
Complicated pregnancy	5	4.3	1.9	1	5	168	4	3.7	1.6	1	5	162
Marked orthostasis	4	4.3	1.9	1	5	173	4	3.6	1.5	1	5	160
Rapid persistent decline in oral intake	4	4.3	1.9	1	5	173	3	3.2	1.4	1	5	142
Chronic medically or functionally impairing treatment-resistant symptoms (e.g., persistence despite several months of intensive treatment)	4	4.1	1.8	1	5	175	4	3.6	1.5	1	5	167
Resistance, denial, poor motivation, and/or lack of cooperation with treatment in the presence of medically or functionally impairing symptoms	4	4	1.7	1	5	174	3	3.4	1.5	1	5	164
Co-occurring psychiatric symptoms or diagnoses including significant alcohol or substance use or personality disorders that are complicating treatment of the eating disorder	4	3.8	1.6	1	5	174	4	3.5	1.5	1	5	170

Lack of access to an otherwise appropriate level of care (e.g., due to lack of geographic accessibility, lack of insurance coverage)	4	3.8	1.6	1	5	171	3	3.4	1.5	1	5	164
Prior lack of response at similar or lower levels of care	3	3.6	1.5	1	5	175	3	3.2	1.4	1	5	169
Prior medical instability at a weight that is similar to the current weight	3	3.5	1.5	1	5	171	3	3.1	1.4	1	5	138
Weight that is below the individual's estimated healthy weight (e.g., less than 85% of IBW)	3	3.2	1.4	1	5	175	3	2.7	1.4	1	5	133
Poor or limited community support system	3	3.1	1.4	1	5	171	3	2.6	1.5	1	5	164
Psychosocial stressors that are impacting intake/weight	3	2.9	1.4	1	5	172	3	2.6	1.5	1	5	166

Note: sorted by adults with SEED's median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: BED=binge-eating disorder; IBW=ideal body weight; SD=standard deviation; SEED=severe and enduring eating disorders

Section III. Appropriate Treatment of Anorexia Nervosa
Refeeding Phase of Individuals with Anorexia Nervosa
Target Weight Gains in Specific Settings

Table D- 4. Appropriateness of target weight gains for individuals with AN who require refeeding in inpatient, intensive outpatient, or partial hospital settings

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
2-3 lb/week (0.9-1.36 kg/week)	4	4.1	1.8	1	5	147	4	4	1.7	1	5	146	4	3.4	1.5	1	5	140
1-2 lb/week (0.45-0.9 kg/week)	4	3.5	1.5	1	5	149	4	3.5	1.5	1	5	153	4	3.5	1.5	1	5	145
3-4 lb/week (1.36-1.81 kg/week)	3	3.3	1.4	1	5	144	3	3.2	1.4	1	5	146	2	2.7	1.4	1	5	142
0-1 lb/week (0-0.45 kg/week)	2	2.2	1.6	1	5	145	2	2.2	1.6	1	5	144	2	2.7	1.4	1	5	146
4-5 lb/week (1.81-2.27 kg/week)	2	2	1.7	1	5	144	2	2	1.7	1	5	144	1	1.8	1.9	1	5	138
Over 5 lb/week (over 2.27 kg/week)	1	1.4	2.2	1	5	140	1	1.3	2.2	1	5	137	1	1.3	2.2	1	5	136

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: AN=anorexia nervosa; SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 7. Appropriateness of target weight gains for adolescents with AN in inpatient, intensive outpatient, or partial hospital settings

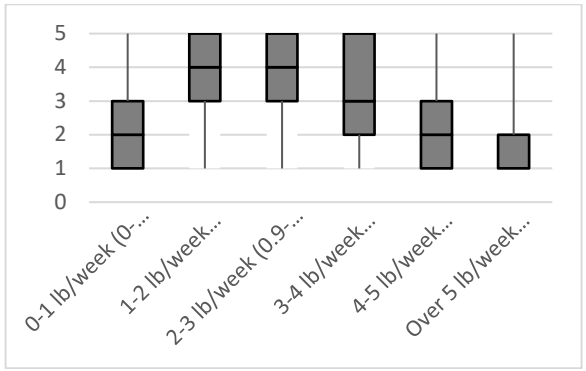


Figure D- 8. Appropriateness of target weight gains for adults with AN in inpatient, intensive outpatient, or partial hospital settings

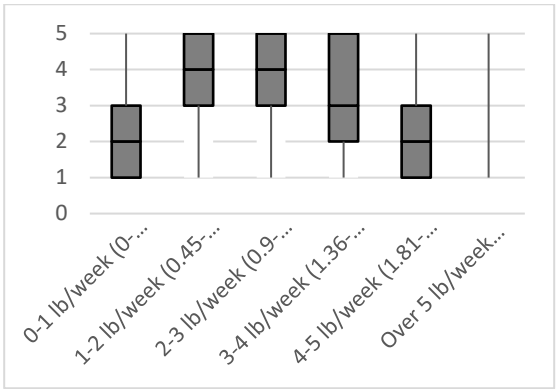


Figure D- 9. Appropriateness of target weight gains for individuals with SEAN in inpatient, intensive outpatient, or partial hospital settings

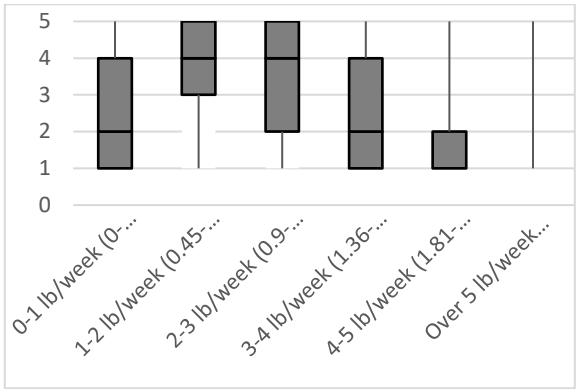


Table D- 5. Appropriateness of target weight gains for individuals with AN who require refeeding in office-based outpatient settings

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
1-2 lb/week (0.45-0.9 kg/week)	5	4.4	1.5	2	5	151	5	4.4	2	1	5	152	4.5	3.9	1.7	1	5	146
0-1 lb/week (0-0.45 kg/week)	5	3.3	1.5	1	5	149	4	3.4	1.5	1	5	149	4	3.5	1.5	1	5	147
2-3 lb/week (0.9-1.36 kg/week)	3	3.1	1.4	1	5	146	3	3	1.4	1	5	146	2	2.6	1.5	1	5	139
3-4 lb/week (1.36-1.81 kg/week)	2	2.1	1.7	1	5	145	2	2	1.7	1	5	144	1	1.7	1.9	1	5	139
4-5 lb/week (1.81-2.27 kg/week)	1	1.4	2.1	1	5	140	1	1.4	2.2	1	5	140	1	1.2	1.7	1	4	138
Over 5 lb/week (over 2.27 kg/week)	1	1.1	1.7	1	4	140	1	1.1	1.8	1	4	137	1	1.1	1.8	1	4	136

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: AN=anorexia nervosa; SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 10. Appropriateness of target weight gains for adolescents with AN in office-based outpatient settings

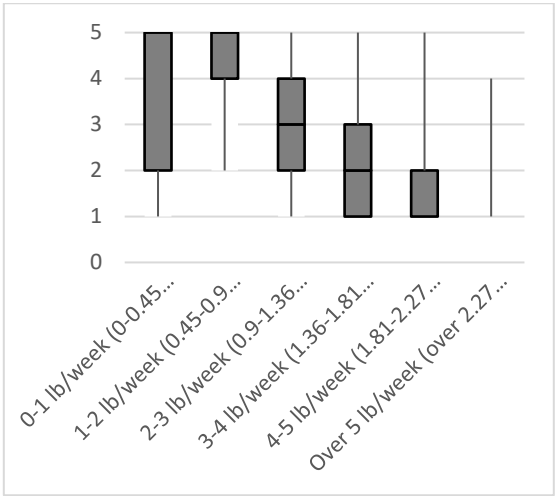


Figure D- 11. Appropriateness of target weight gains for adults with AN in office-based outpatient settings

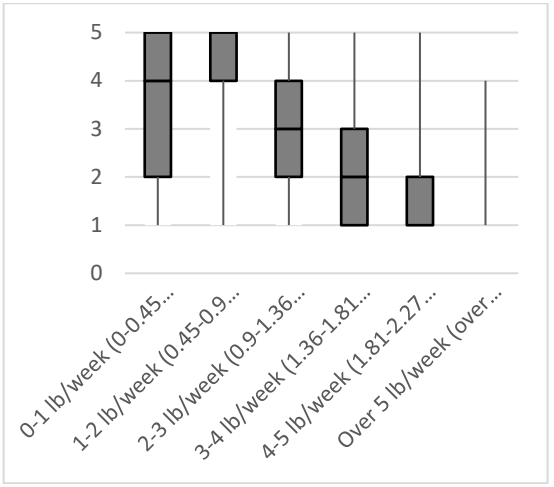
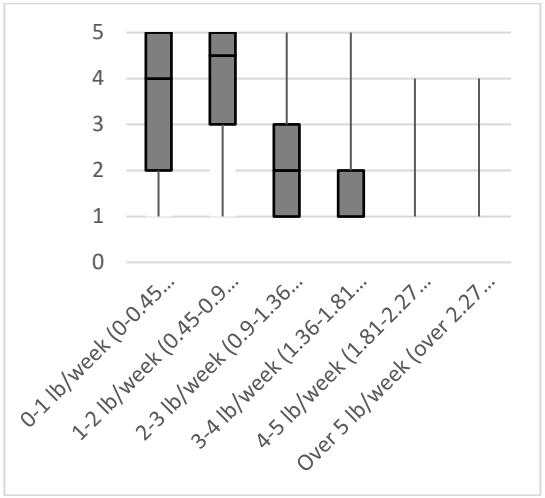


Figure D- 12. Appropriateness of target weight gains for individuals with SEAN in office-based outpatient settings



Target Kcal/Day in Specific Settings

Table D- 6. Appropriateness of target kcal/day for individuals with AN who require refeeding in inpatient, intensive outpatient, or partial hospital settings

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
40-50 kcal/kg/day	4	3.4	1.5	1	5	116	4	3.4	1.5	1	5	117	3	2.9	1.4	1	5	115
50-60 kcal/kg/day	4	3.4	1.5	1	5	115	4	3.3	1.4	1	5	114	3	2.8	1.4	1	5	113
>60 kcal/kg/day	3	3.1	1.4	1	5	116	3	3	1.4	1	5	113	2	2.6	1.5	1	5	111
30-40 kcal/kg/day	3	3.1	1.4	1	5	116	3	3	1.4	1	5	115	3	2.9	1.4	1	5	114
20-30 kcal/kg/day	2	2.2	1.6	1	5	113	2	2.1	1.7	1	5	113	2	2.5	1.5	1	5	114

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: AN=anorexia nervosa; SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 13. Appropriateness of target kcal/day for adolescents with AN who require refeeding in inpatient, intensive outpatient, or partial hospital settings

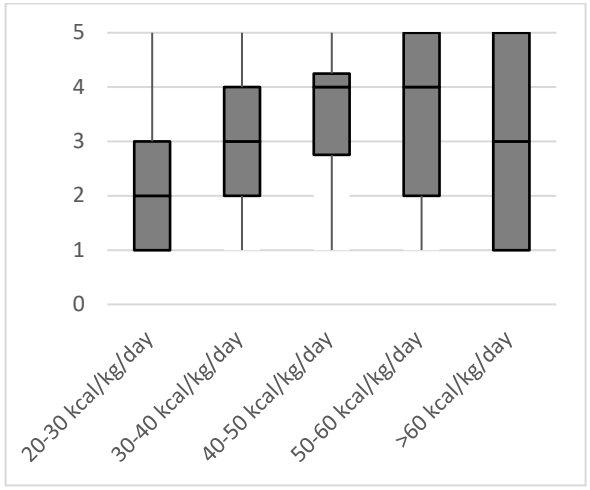


Figure D- 14. Appropriateness of target kcal/day for adults with AN who require refeeding in inpatient, intensive outpatient, or partial hospital settings

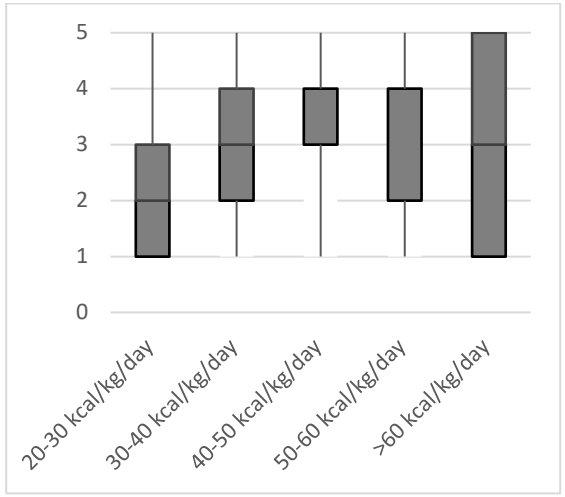


Figure D- 15. Appropriateness of target kcal/day for individuals with SEAN who require refeeding in inpatient, intensive outpatient, or partial hospital settings

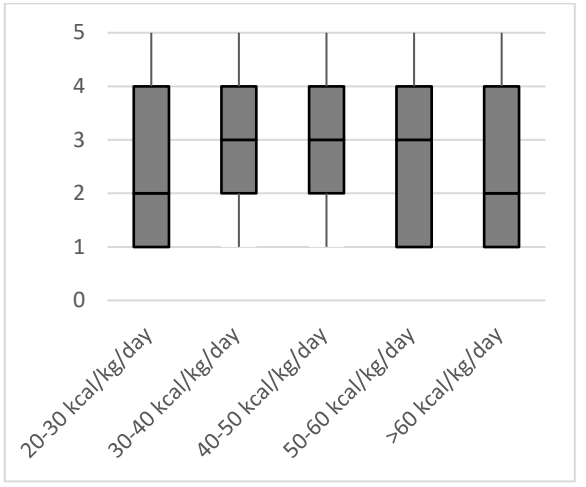


Table D- 7. Appropriateness of target kcal/day for individuals with AN who require refeeding in office-based outpatient settings

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
30-40 kcal/kg/day	3	3.2	1.4	1	5	119	3	3.3	1.4	1	5	117	3	3	1.4	1	5	114
40-50 kcal/kg/day	3	3.2	1.4	1	5	114	3	3.2	1.4	1	5	114	3	2.8	1.4	1	5	110
50-60 kcal/kg/day	3	3.1	1.4	1	5	112	3	3	1.4	1	5	114	2	2.6	1.5	1	5	112
>60 kcal/kg/day	2	2.6	1.5	1	5	113	2	2.4	1.5	1	5	110	1	2.1	1.7	1	5	109
20-30 kcal/kg/day	1	2.4	1.5	1	5	113	2	2.4	1.5	1	5	113	2	2.6	1.5	1	5	112

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: AN=anorexia nervosa; SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 16. Appropriateness of target kcal/day for adolescents with AN who require refeeding in office-based outpatient settings

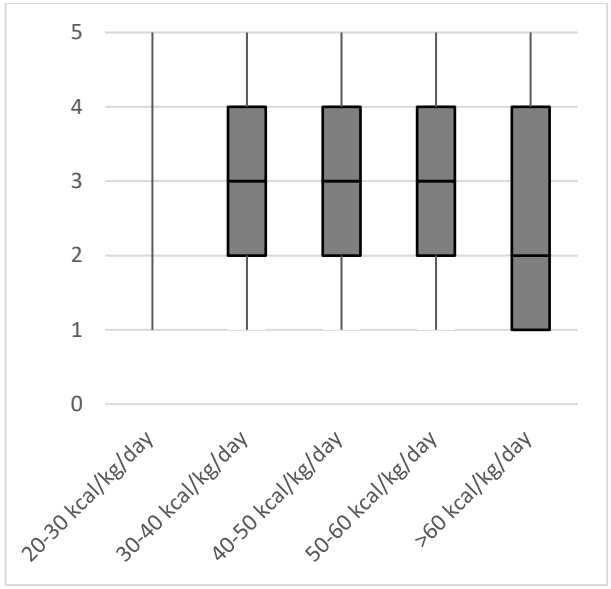


Figure D- 17. Appropriateness of target kcal/day for adults with AN who require refeeding in office-based outpatient settings

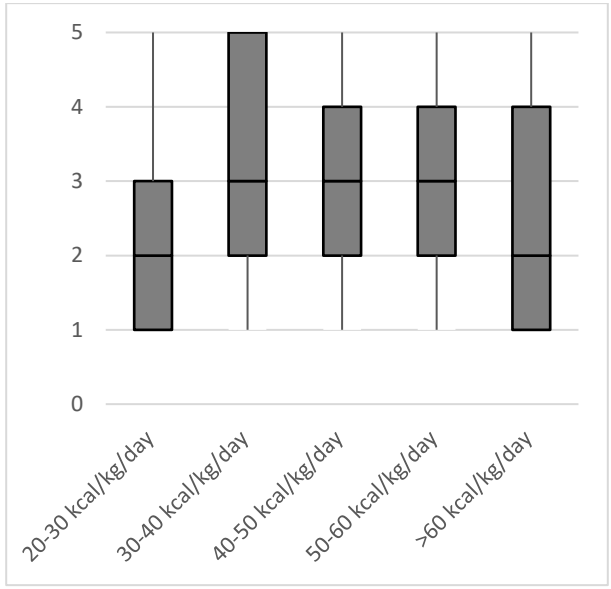
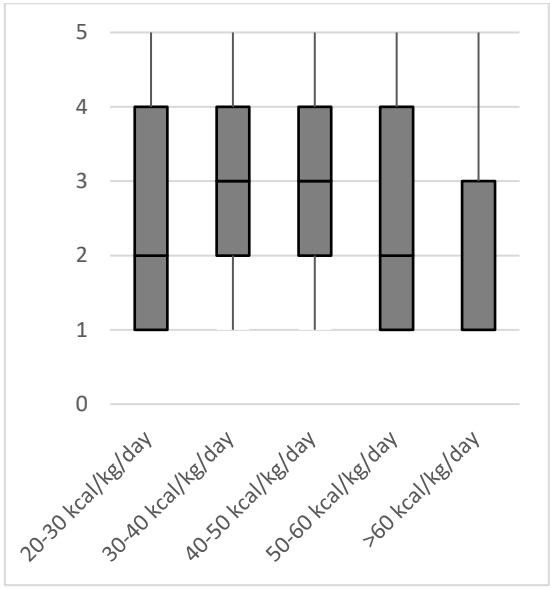


Figure D- 18. Appropriateness of target kcal/day for individuals with SEAN who require refeeding in office-based outpatient settings



Methods to Determine Daily Caloric Intake

Table D- 8. Appropriateness of methods to determine daily caloric intake for individuals with AN who require refeeding in any setting

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Calculating kcal/day to be prescribed based on initial and target weights and anticipated/recommended rate of weight gain	5	4.2	1.8	1	5	146	5	4.1	1.8	1	5	149	4	3.8	1.6	1	5	143
Using lower target goals for weight gain or kcal/kg/day for outpatients as compared to inpatients	4	3.3	1.5	1	5	147	4	3.4	1.5	1	5	146	4	3.5	1.5	1	5	146
Setting a low expectation for caloric intake initially then increasing expectations for caloric intake as treatment proceeds	2	2.7	1.4	1	5	148	3	2.9	1.4	1	5	147	3	3.1	1.4	1	5	145
Picking a number for daily caloric intake and sticking with that throughout the course of treatment	1	1.5	2.1	1	5	146	1	1.5	2.1	1	5	148	1	1.6	2	1	5	145

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: AN=anorexia nervosa; SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 19. Appropriateness of methods to determine daily caloric intake for adolescents with AN who require refeeding in any setting

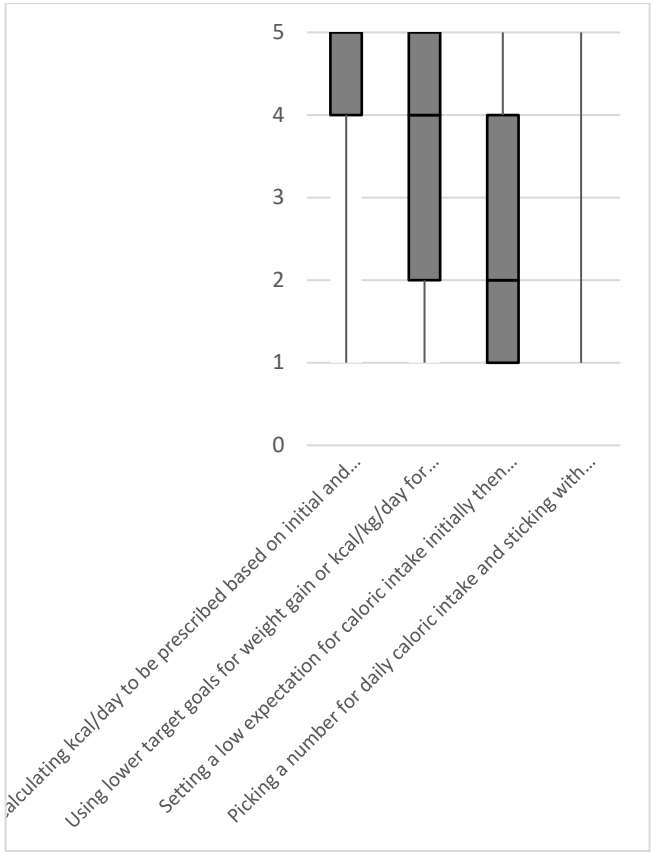


Figure D- 20. Appropriateness of methods to determine daily caloric intake for adults with AN who require refeeding in any setting

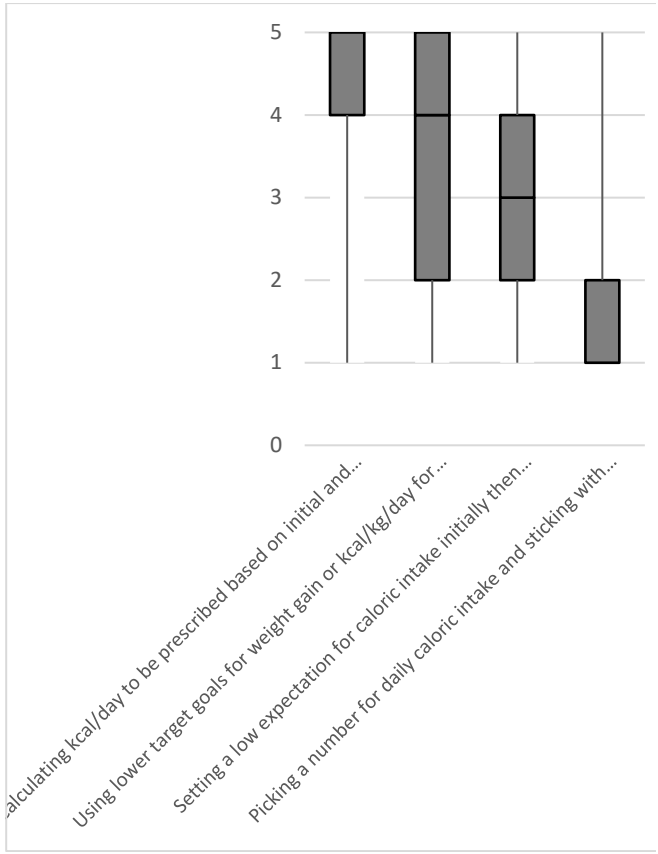
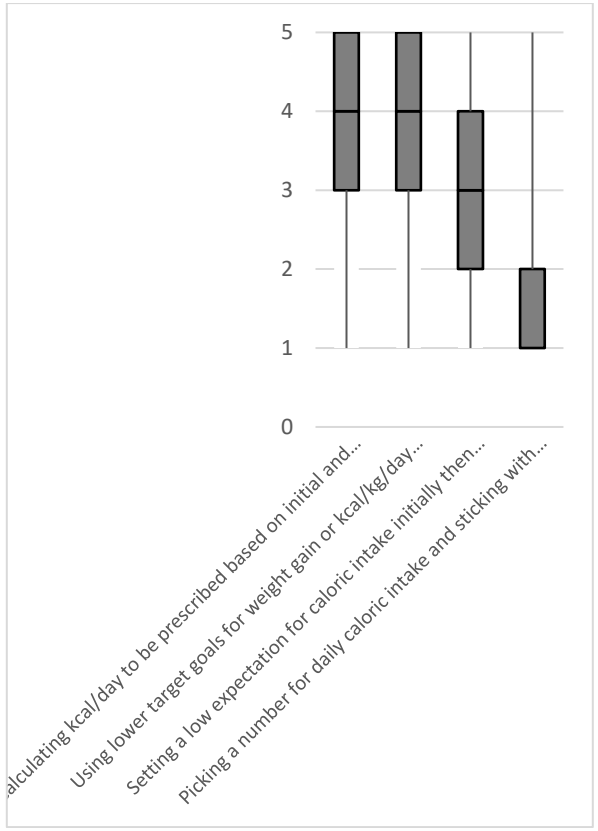


Figure D- 21. Appropriateness of methods to determine daily caloric intake for individuals with SEAN who require refeeding in any setting



Routine Assessments

Table D- 9. Appropriateness of routine assessments (e.g., at least monthly to assure physical health during refeeding) for individuals with AN who require refeeding in any setting

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Checking height, weight, calculation of BMI	5	4.9	1.8	2	5	157	5	4.8	1.7	2	5	160	5	4.7	2.2	1	5	159
Checking vital signs, including orthostatics and temperature	5	4.8	1.7	2	5	158	5	4.8	1.7	2	5	160	5	4.8	2.3	1	5	159
Assessing physical findings (e.g., peripheral edema, evidence of congestive heart failure, gastrointestinal abnormalities)	5	4.8	1.7	2	5	157	5	4.7	1.7	2	5	154	5	4.7	2.2	1	5	157
Ordering/interpreting laboratory studies (e.g., electrolytes, phosphate, magnesium, calcium)	5	4.7	2.2	1	5	157	5	4.6	2.1	1	5	158	5	4.6	2.1	1	5	157
Ordering/interpreting electrocardiogram	4	3.9	1.7	1	5	154	4	3.8	1.6	1	5	158	4	3.8	1.6	1	5	155

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: AN=anorexia nervosa; BMI=body mass index; SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Interventions

Table D- 10. Appropriateness of interventions to promote healthy weight gain in individuals with AN who require refeeding in any setting

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Nutritional rehabilitation	5	4.8	2.3	1	5	149	5	4.8	2.3	1	5	151	5	4.6	2.1	1	5	147
FBT (Maudsley approach)	5	4.7	2.2	1	5	156	3	2.6	1.5	1	5	145	2	2.3	1.6	1	5	142
Psychoeducation	5	4.6	2.1	1	5	151	5	4.7	2.2	1	5	153	5	4.5	2.1	1	5	149
Individual CBT	4	3.6	1.5	1	5	149	5	4.3	1.4	2	5	151	4	3.9	1.7	1	5	146
Group therapy	3	3.4	1.5	1	5	149	4	3.6	1.5	1	5	147	3	3.3	1.4	1	5	147
Other approaches to family or couples therapy	3	3.3	1.5	1	5	153	4	3.6	1.5	1	5	151	3	3.5	1.5	1	5	146

SSCM (i.e., including support, education, advice, praise)	3	3.1	1.4	1	5	149	4	3.5	1.5	1	5	149	4	3.8	1.6	1	5	145
NG continuous tube feeding	3	2.8	1.4	1	5	136	3	2.7	1.4	1	5	135	3	2.8	1.4	1	5	133
Supplemental overnight tube feeding	3	2.8	1.4	1	5	137	3	2.7	1.4	1	5	133	3	2.9	1.4	1	5	131
Individual IPT	3	2.8	1.4	1	5	146	3	3.3	1.5	1	5	146	3	3.2	1.4	1	5	143
NG bolus tube feeding	3	2.7	1.5	1	5	135	3	2.6	1.5	1	5	133	3	2.5	1.5	1	5	131
Individual supportive psychotherapy	3	2.7	1.5	1	5	146	3	2.9	1.4	1	5	146	3	3.2	1.4	1	5	149
2nd generation antipsychotic	3	2.6	1.5	1	5	129	3	2.7	1.4	1	5	127	3	2.8	1.4	1	5	126
SSRI	2	2.3	1.6	1	5	137	2	2.5	1.5	1	5	134	2	2.4	1.5	1	5	130
Psychodynamically informed individual therapy	2	2.1	1.7	1	5	144	2	2.4	1.5	1	5	145	2	2.3	1.6	1	5	142
Intravenous tube feeding	2	1.8	1.8	1	5	133	2	1.8	1.8	1	5	133	2	2	1.7	1	5	133
SNRI	1	1.8	1.8	1	5	126	1.5	2	1.7	1	5	124	1	2	1.8	1	5	121
Mirtazapine	1	1.8	1.9	1	5	121	2	1.9	1.8	1	5	117	2	1.9	1.8	1	5	115
Metoclopramide	1	1.8	1.9	1	5	116	1	1.9	1.8	1	5	114	1	1.9	1.8	1	5	109
Benzodiazepine	1	1.6	2	1	5	124	1	1.7	1.9	1	5	120	1	1.7	1.9	1	5	119
Self-help/12 step programs	1	1.4	2.1	1	5	141	1	1.7	2	1	5	140	1	1.7	1.9	1	5	140
Bupropion	1	1.4	2.1	1	5	123	1	1.5	2.1	1	5	116	1	1.5	2.1	1	5	121
Anticonvulsant	1	1.4	2.2	1	5	120	1	1.5	2.1	1	5	116	1	1.4	2.1	1	5	115

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: CBT=cognitive-behavioral therapy; FBT=family-based therapy; IPT=interpersonal therapy; NG=nasogastric; SD=standard deviation; SEAN=severe and enduring anorexia nervosa; SNRI=Serotonin and norepinephrine reuptake inhibitor; SSCM=specialist supportive clinical management; SSRI=selective serotonin reuptake inhibitor

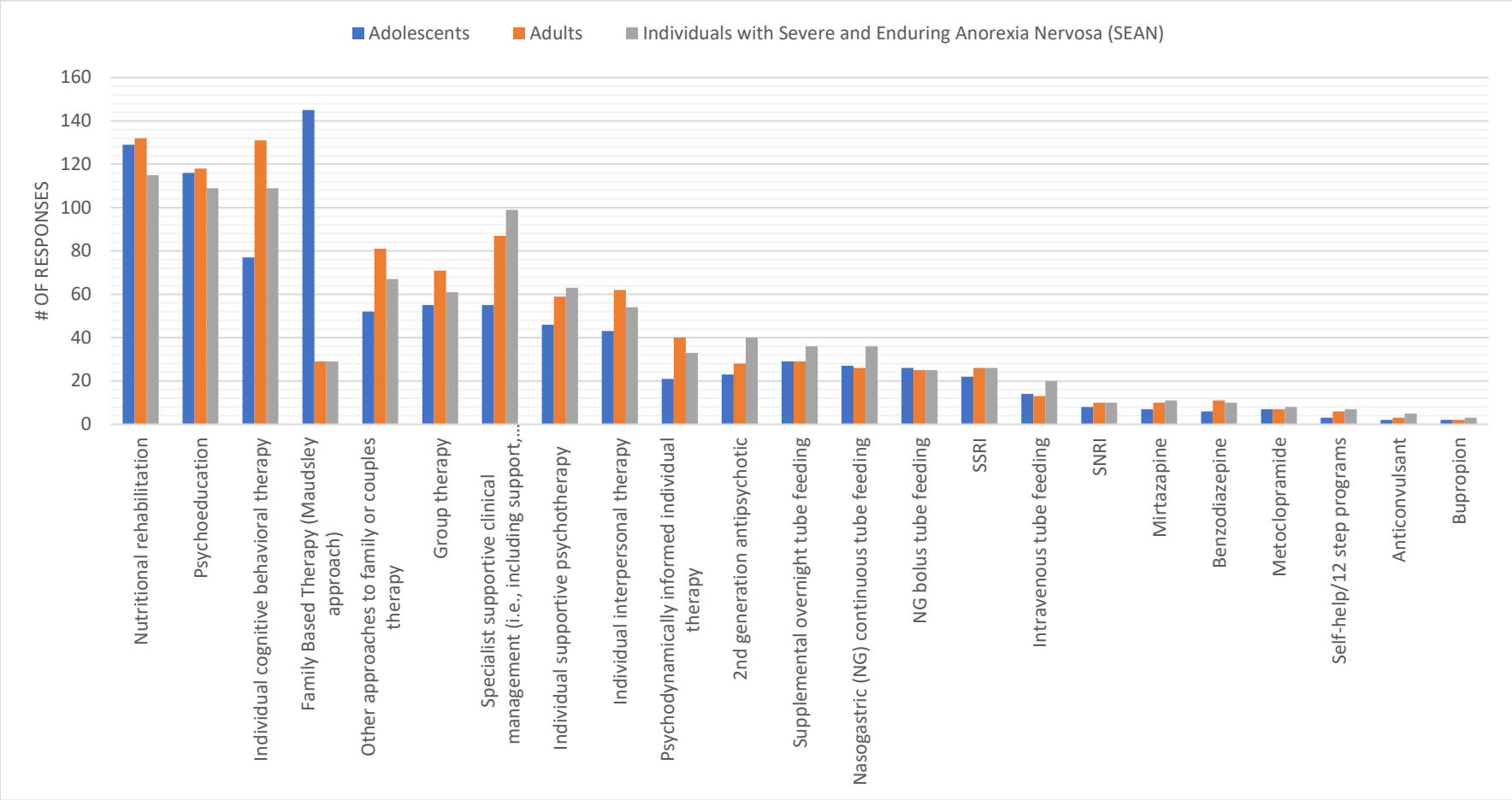
Table D- 11. Appropriateness of modalities to treat individuals with AN who require refeeding in any setting

	Adolescents						**Adults**						**Individuals with SEAN**					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Psychotherapy alone as a first line treatment	5	3.8	1.6	1	5	154	4	3.7	1.6	1	5	152	4	3.4	1.5	1	5	148

Combination medication and psychotherapy as a first line treatment	3	2.9	1.4	1	5	148	3	3.1	1.4	1	5	148	3	3.3	1.5	1	5	149
Medication alone as a first line treatment	1	1.1	2.3	1	5	149	1	1.2	2.3	1	5	147	1	1.3	2.2	1	5	145
Self-help alone as a first line treatment	1	1.1	1.2	1	3	149	1	1.2	1.7	1	4	147	1	1.3	2.2	1	5	146

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviation: SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 22. Appropriateness as initial intervention in an episode of care for individuals who require refeeding



Note: survey respondents were allowed to check multiple options

Bone Density in Anorexia Nervosa

Table D- 12. Appropriateness of interventions to improve bone density or prevent further deterioration in bone density for individuals with AN who have had at least 6 months of amenorrhea

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Calcium and vitamin D supplementation	5	4.1	1.8	1	5	127	5	4.1	1.8	1	5	129	5	4	1.7	1	5	125
Moderate exercise (if no history of compulsive exercising)	3	2.9	1.4	1	5	130	4	2.9	1.4	1	5	128	3.25	2.7	1.5	1	5	124
Hormone replacement therapy	1	1.8	1.9	1	5	121	3	1.9	1.8	1	5	121	3	1.9	1.8	1	5	117
Bisphosphonates	1	1.5	1.9	1	5	108	3	2.3	1.6	1	5	109	3	2.6	1.5	1	5	108

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviation: SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 23. Appropriateness of interventions to improve bone density or prevent further deterioration in bone density for adolescents with AN who have had at least 6 months of amenorrhea

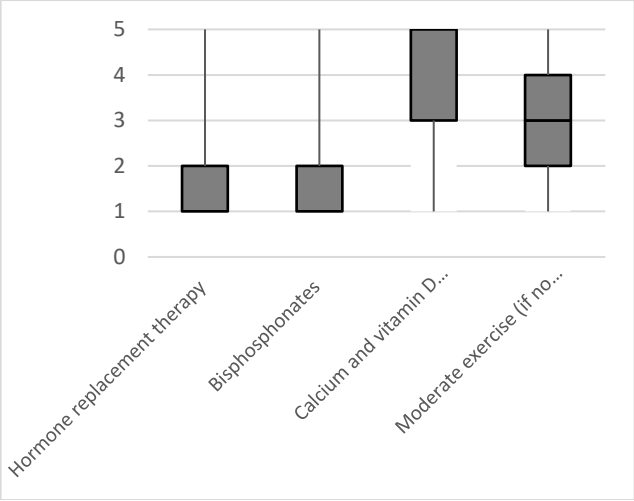


Figure D- 24. Appropriateness of interventions to improve bone density or prevent further deterioration in bone density for adults with AN who have had at least 6 months of amenorrhea

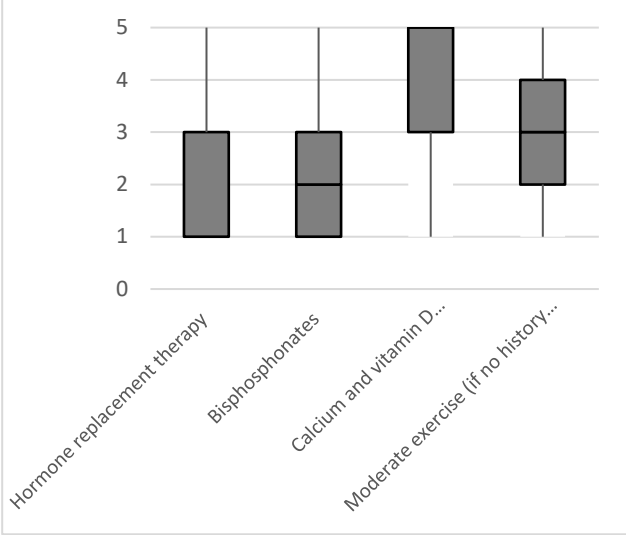
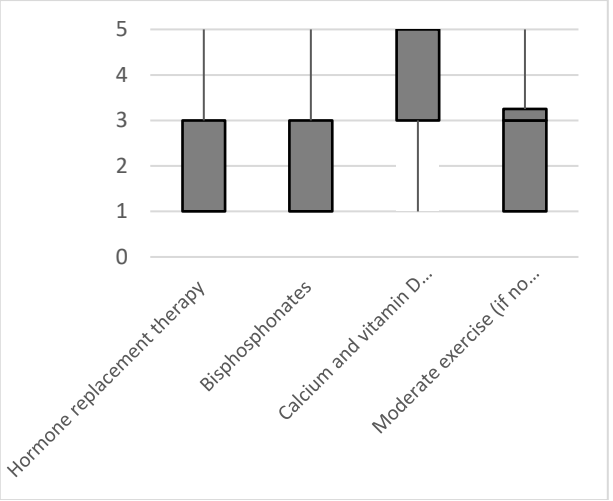


Figure D- 25. Appropriateness of interventions to improve bone density or prevent further deterioration in bone density for individuals with SEAN who have had at least 6 months of amenorrhea



Anorexia Nervosa Once Medical Stabilization and Severe Malnutrition Have Been Addressed

Table D- 13. Appropriateness of interventions to treat individuals with AN once malnutrition has been addressed

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
FBT (Maudsley approach)	5	4.5	2	1	5	148	2	2.1	1.7	1	5	128	1.5	2.1	1.7	1	5	122
Psychoeducation	5	4.4	2	1	5	141	5	4.4	2	1	5	137	5	4.3	1.9	1	5	131
Nutritional rehabilitation	5	4.3	1.9	1	5	140	5	4.3	1.9	1	5	138	5	4.1	1.8	1	5	134
Individual CBT	4	3.9	1.7	1	5	141	5	4.4	2	1	5	140	5	4.1	1.8	1	5	135
Group therapy	4	3.5	1.5	1	5	136	4	3.8	1.6	1	5	133	4	3.6	1.5	1	5	132
Other approaches to family or couples therapy	4	3.3	1.4	1	5	134	4	3.7	1.6	1	5	139	4	3.5	1.5	1	5	129
Individual IPT	3	3	1.4	1	5	135	3	3.4	1.5	1	5	137	3	3.3	1.5	1	5	128
SSCM (including support, education, advice, and praise)	3	3	1.4	1	5	130	4	3.5	1.5	1	5	132	4	3.8	1.6	1	5	131
Individual supportive psychotherapy	3	2.9	1.4	1	5	135	3	3.1	1.4	1	5	135	3	3.3	1.5	1	5	132
SSRI	3	2.7	1.4	1	5	121	3	3	1.4	1	5	119	3	3	1.4	1	5	116
Psychodynamically informed individual therapy	2	2.2	1.6	1	5	161	2	2.6	1.5	1	5	131	2	2.6	1.5	1	5	128
2nd generation antipsychotic	2	2.2	1.6	1	5	110	2	2.3	1.6	1	5	107	3	2.6	1.5	1	5	106
SNRI	2	2.1	1.7	1	5	109	2	2.3	1.6	1	5	106	2	2.4	1.5	1	5	103
Mirtazapine	1	1.7	1.9	1	5	102	1	1.8	1.9	1	5	103	1	1.9	1.8	1	5	100
Bupropion	1	1.5	2	1	5	106	1	1.7	1.9	1	5	103	1	1.7	1.9	1	5	100
Metoclopramide	1	1.5	2	1	5	100	1	1.6	2	1	5	96	1	1.7	1.9	1	5	95
Self-help/12 step programs	1	1.4	2.2	1	5	125	1	1.7	1.9	1	5	124	1	1.8	1.9	1	5	122
Benzodiazepine	1	1.4	2.1	1	5	103	1	1.6	2	1	5	103	1	1.6	2	1	5	101
Anticonvulsant	1	1.4	2.1	1	5	100	1	1.6	2	1	5	97	1	1.6	2	1	5	95

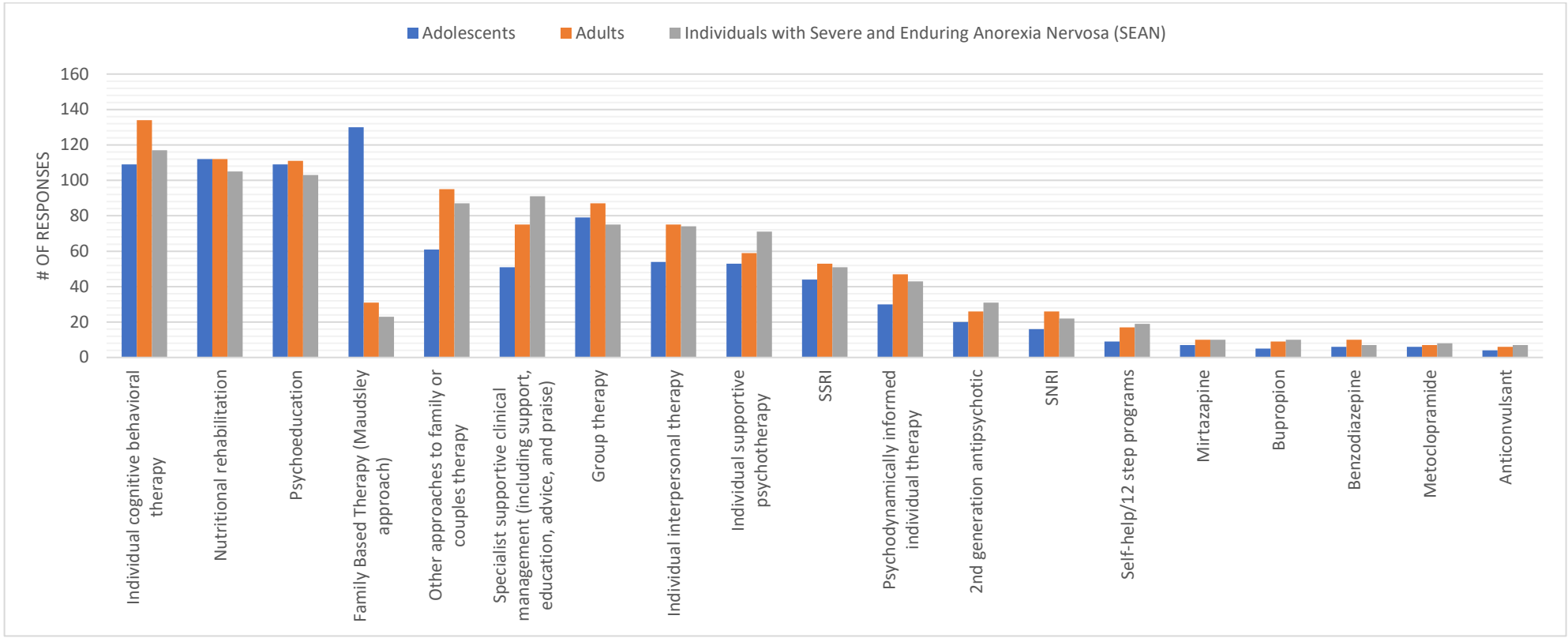
Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: CBT=cognitive-behavioral therapy; FBT=family-based therapy; IPT=interpersonal therapy; SD=standard deviation; SEAN=severe and enduring anorexia nervosa; SNRI=Serotonin and norepinephrine reuptake inhibitor; SSCM=specialist supportive clinical management; SSRI=selective serotonin reuptake inhibitor

Table D- 14. Appropriateness of modalities to treat individuals with AN once malnutrition has been addressed

	Adolescents						Adults						Individuals with SEAN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Psychotherapy alone as a first line treatment	5	4.1	1.8	1	5	142	5	4.1	1.8	1	5	141	4	3.8	1.6	1	5	139
Combination medication and psychotherapy as a first line treatment	3	3.2	1.4	1	5	140	4	3.5	1.5	1	5	141	4	3.7	1.6	1	5	139
Medication alone as a first line treatment	1	1.2	1.8	1	4	138	1	1.3	1.7	1	4	136	1	1.3	2.2	1	5	133
Self-help alone as a first line treatment	1	1.2	2.2	1	5	135	1	1.4	2.1	1	5	136	1	1.5	2.1	1	5	135

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviation: SD=standard deviation; SEAN=severe and enduring anorexia nervosa

Figure D- 26. Appropriateness as an initial intervention in an episode of care for individuals with AN once malnutrition has been addressed



Note: survey respondents were allowed to check multiple options

Section IV. Appropriate Treatment of Bulimia Nervosa

Table D- 15. Appropriateness of interventions to treat individuals with BN

	Adolescents						Adults						Individuals with multi-impulsive BN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Psychoeducation	5	4.5	2.1	1	5	140	5	4.5	2.1	1	5	138	5	4.5	2	1	5	135
Individual CBT	5	4.5	2	1	5	140	5	4.7	1.7	2	5	136	5	4.4	2	1	5	137
Nutritional rehabilitation	5	4.2	1.9	1	5	139	5	4.2	1.9	1	5	135	5	4.1	1.8	1	5	135
FBT	4	4.1	1.8	1	5	141	2	2.1	1.7	1	5	120	2	2.1	1.7	1	5	117
Group therapy	4	3.6	1.5	1	5	132	4	3.9	1.7	1	5	130	4	3.7	1.6	1	5	130
Individual DBT	4	3.6	1.5	1	5	137	4	3.9	1.7	1	5	136	5	4.5	2.1	1	5	140
SSRI	4	3.6	1.5	1	5	125	4	4.1	1.8	1	5	130	5	4.2	1.9	1	5	124
Individual IPT	3	3.1	1.4	1	5	133	4	3.8	1.6	1	5	135	3	3.4	1.5	1	5	134
Individual supportive psychotherapy	2	2.5	1.5	1	5	131	2	2.6	1.5	1	5	128	2	2.6	1.5	1	5	127
SNRI	2	2.4	1.5	1	5	110	3	2.8	1.4	1	5	112	3	2.8	1.4	1	5	108
Psychodynamically informed individual therapy	1	2	1.8	1	5	125	2	2.3	1.6	1	5	125	2	2.2	1.6	1	5	123
Couples therapy (if relevant)	1	1.8	1.9	1	5	98	4	3.4	1.5	1	5	133	3	3.4	1.5	1	5	128
Topiramate	1	1.7	2	1	5	106	2	2.2	1.6	1	5	109	3	2.4	1.5	1	5	105
Self-help/12 step programs	1	1.6	2	1	5	124	2	2.1	1.7	1	5	125	2	2.1	1.7	1	5	121
Bupropion	1	1.3	1.4	1	4	108	1	1.3	2.2	1	5	106	1	1.4	2.1	1	5	103
Lithium	1	1.3	2.1	1	5	101	1	1.5	2	1	5	99	1	1.9	1.8	1	5	99
Benzodiazepine	1	1.2	1.1	1	3	102	1	1.3	2	1	5	104	1	1.4	2.2	1	5	101

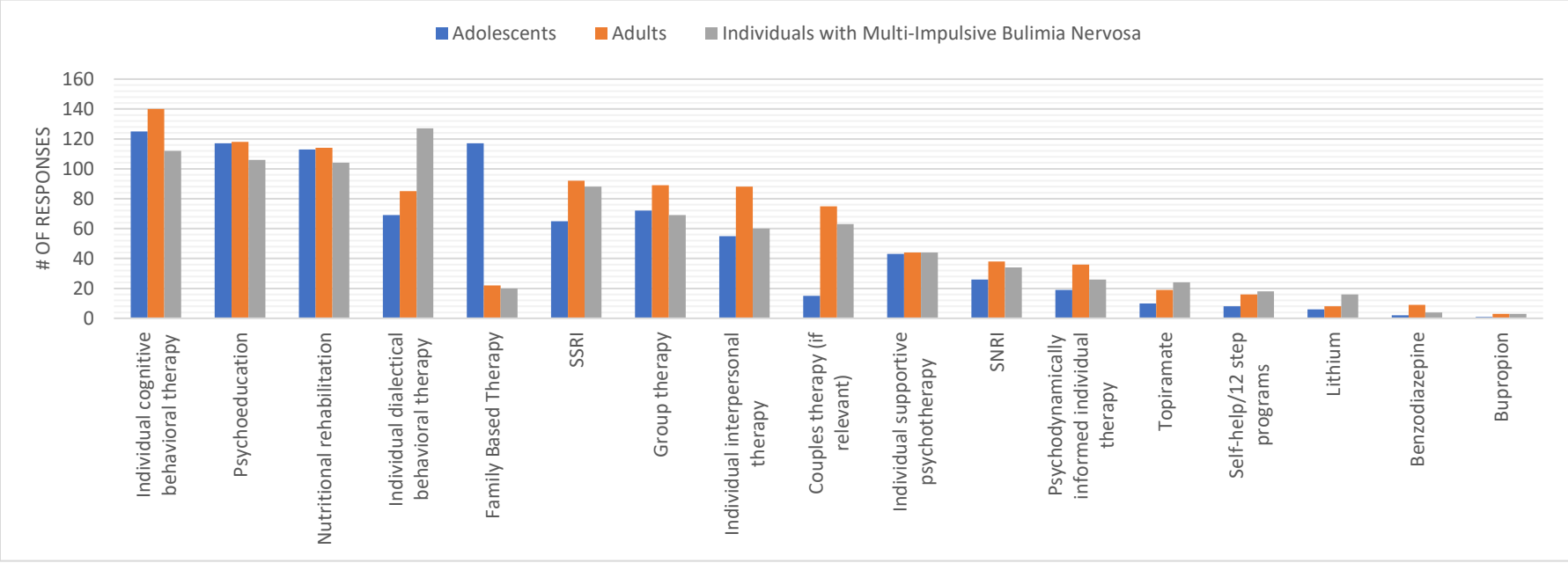
Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate; for purposes of the survey, individuals with multi-impulsive BN are generally characterized by severe dysregulation, borderline personality disorder, often concurrent bipolar disorder, PTSD, ADHD, and/or alcohol and substance abuse.
Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; DBT=dialectical behavior therapy; FBT=family-based therapy; IPT=interpersonal therapy; SD=standard deviation; SNRI=Serotonin and norepinephrine reuptake inhibitor; SSRI=selective serotonin reuptake inhibitor

Table D- 16. Appropriateness of modalities to treat individuals with BN

	Adolescents						Adults						Individuals with multi-impulsive BN					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Psychotherapy alone as a first line treatment	5	4.2	1.8	1	5	142	5	4.1	1.8	1	5	138	5	4.3	1.9	1	5	138
Combination medication and psychotherapy as a first line treatment	4	3.5	1.5	1	5	140	4	3.9	1.7	1	5	138	1	3.7	1.6	1	5	135
Self-help alone as a first line treatment	1	1.5	2.1	1	5	134	2	2	1.7	1	5	133	1	1.6	2	1	5	131
Medication alone as a first line treatment	1	1.3	2.2	1	5	137	1	1.7	1.9	1	5	134	1	1.6	2	1	5	131

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviation: BN=bulimia nervosa; SD=standard deviation

Figure D- 27. Appropriateness as an initial intervention in an episode of care of BN



Note: survey respondents were allowed to check multiple options

Section V. Appropriate Treatment of Binge-Eating Disorder

Table D- 17. Appropriateness of interventions to treat individuals with BED

	Adolescents						Adults					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Psychoeducation	5	4.6	2.1	1	5	139	5	4.6	2.1	1	5	136
Individual CBT	5	4.6	1.6	2	5	139	5	4.8	1.1	3	5	139
Nutritional rehabilitation	5	3.9	1.7	1	5	134	5	4	1.7	1	5	133
Family therapy	4	3.9	1.7	1	5	131	3	2.7	1.5	1	5	125
Group therapy	4	3.8	1.6	1	5	132	5	4.1	1.8	1	5	135
Individual DBT	4	3.6	1.5	1	5	136	4	3.8	1.6	1	5	135
Individual IPT	4	3.4	1.5	1	5	134	4	3.8	1.6	1	5	135
SSRI	3	2.9	1.4	1	5	122	3	3.3	1.4	1	5	122
Individual supportive psychotherapy	2	2.5	1.5	1	5	126	2	2.6	1.5	1	5	123
Couples therapy (if relevant)	1	2	1.7	1	5	87	3	3.4	1.5	1	5	132
Psychodynamically informed individual therapy	1	2	1.8	1	5	122	2	2.2	1.6	1	5	121
Topiramate	1	2	1.8	1	5	100	3	2.7	1.5	1	5	107
Self-help/12 step programs	1	1.9	1.8	1	5	117	2	2.3	1.6	1	5	122
Lisdexamfetamine	1	1.8	1.8	1	5	98	3	2.5	1.5	1	5	99
Bupropion/naltrexone combination therapy	1	1.7	1.9	1	5	102	2	2.2	1.6	1	5	102
Bupropion alone	1	1.6	2	1	5	98	1	1.9	1.8	1	5	96
Bariatric surgery	1	1.2	1.1	1	3	107	1	1.6	2	1	5	111
Benzodiazepine	1	1.1	1.2	1	3	95	1	1.2	2.1	1	5	96

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: BED=binge-eating disorder; CBT=cognitive-behavioral therapy; DBT=dialectical behavior therapy; IPT=interpersonal therapy; SD=standard deviation; SSRI=selective serotonin reuptake inhibitor

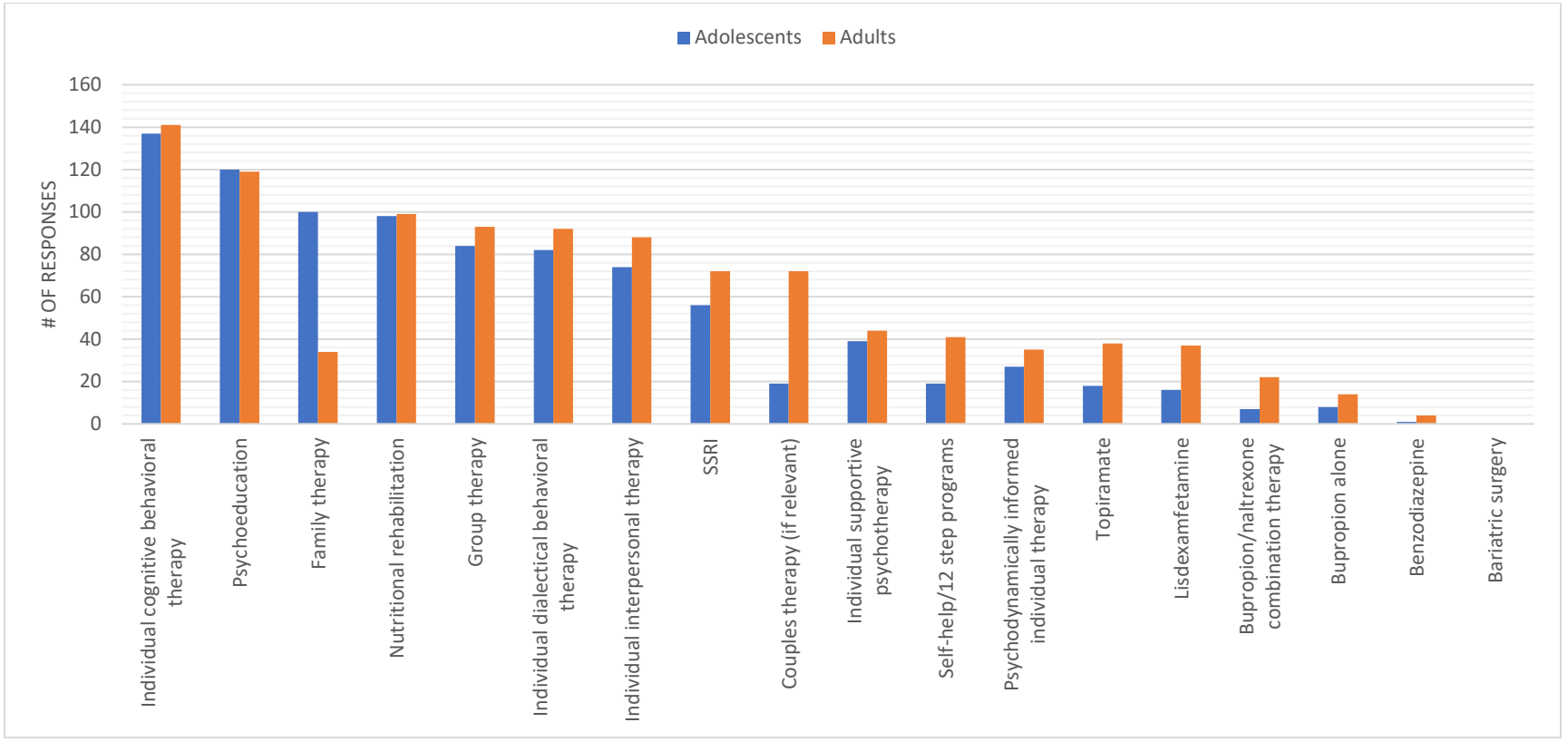
Table D- 18. Appropriateness of modalities to treat individuals with BED

	Adolescents						Adults					
	Median	Mean	SD	Min	Max	N	Median	Mean	SD	Min	Max	N
Psychotherapy alone as a first line treatment	5	4.3	1.9	1	5	138	5	4.3	1.9	1	5	136

Combination medication and psychotherapy as a first line treatment	4	3.5	1.5	1	5	135	4	3.9	1.7	1	5	136
Self-help alone as a first line treatment	1	1.8	1.9	1	5	130	2	2.4	1.5	1	5	134
Medication alone as a first line treatment	1	1.3	2.2	1	5	132	1	1.8	1.8	1	5	130

Note: sorted by adolescents' median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviation: BED=binge-eating disorder; SD=standard deviation

Figure D- 28. Appropriateness as an initial intervention in an episode of care of BED



Note: survey respondents were allowed to check multiple options

Section VI. Appropriate Treatment of Night Eating Syndrome

Table D- 19. Appropriateness of interventions to treat individuals with NES

	Median	Mean	SD	Min	Max	N
Individual CBT	5	4.5	2	1	5	123
Psychoeducation	5	4.4	2	1	5	125
Nutritional rehabilitation	4	3.7	1.6	1	5	117
SSRI	3	3.3	1.4	1	5	113
Progressive muscle relaxation	3	2.9	1.4	1	5	113
Group therapy	3	2.9	1.4	1	5	115
Individual supportive psychotherapy	3	2.7	1.4	1	5	116
Family therapy or Couples therapy	3	2.7	1.5	1	5	117
Self-help/12 step programs	1	1.8	1.8	1	5	107

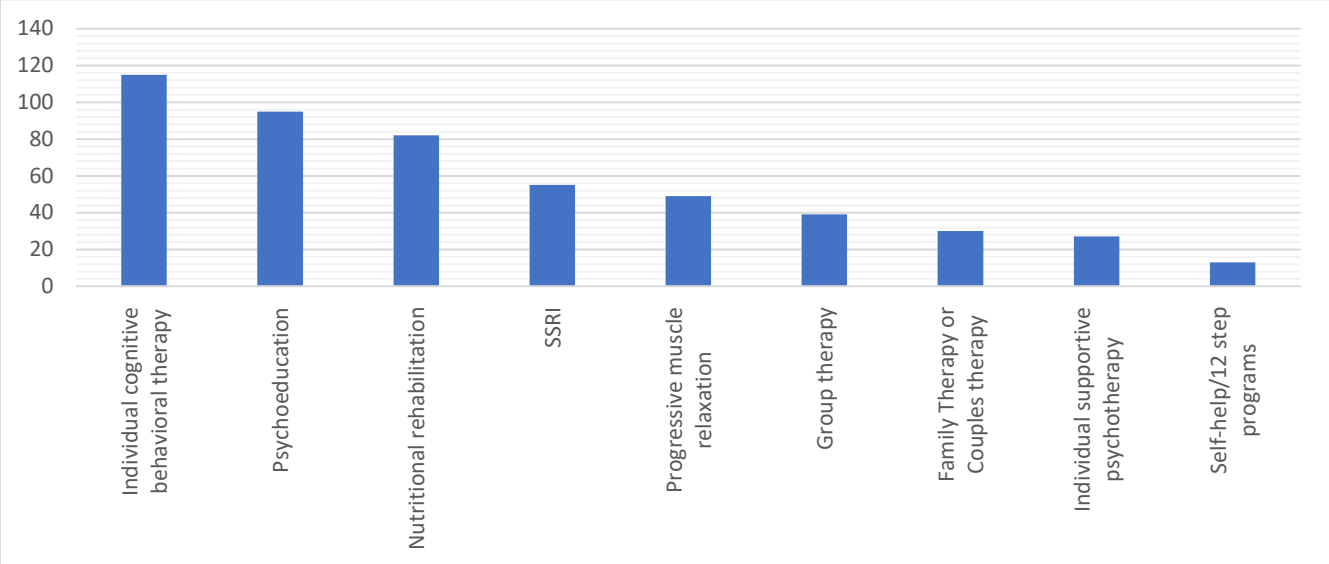
Note: sorted by median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviations: CBT=cognitive-behavioral therapy; NES=night eating syndrome; SD=standard deviation; SSRI=selective serotonin reuptake inhibitor

Table D- 20. Appropriateness of modalities to treat individuals with NES

	Median	Mean	SD	Min	Max	N
Psychotherapy alone as a first line treatment	4	3.9	1.7	1	5	119
Combination medication and psychotherapy as a first line treatment	4	3.7	1.6	1	5	121
Medication alone as a first line treatment	2	1.9	1.8	1	5	117
Self-help alone as a first line treatment	2	1.9	1.8	1	5	113

Note: sorted by median then mean; 1=inappropriate, 3=moderately appropriate, and 5=highly appropriate
Abbreviation: NES=night eating syndrome; SD=standard deviation

Figure D- 29. Appropriateness as an initial intervention to treat individuals with NES



Note: survey respondents were allowed to check multiple options

Appendix E. Evidence Tables for Individual Studies Supporting Guideline Statements

Studies marked by an asterisk were endonodal and not included in the NMA.

Anorexia Nervosa Studies Supporting Guideline Statements

Family Treatment With Parents in Charge

Compared to Family Treatment With Parents in Charge

Conjoint compared to separated

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Eisler et al. (2000*, 2007*)	Design: RCT; Follow-up/Extension Setting: Single Center: Eating Disorder Services of the Maudsley Hospital Country: United Kingdom Funding: Academic	Randomized N=40 Conjoint Family Therapy 1 yr (N=19) - Maternal EE, High (N=7) - Maternal EE, Low (N=12) Separated Family Therapy 1 yr (N=21) - Maternal EE, High (N=10) - Maternal EE, Low (N=11) Follow-up: Baseline – 6 yr	Inclusion: AN; adolescent Exclusion: NR	AN: 40 (100%) AN, Duration: 12.9 mo (SD ± 9.4) - 13.9 mo vs. 12 mo Weight – Baseline: 40 kg (SD ± 6.4) Adolescent: 40 (100%) Age: 15.5 yr (SD ± 1.6) - 15.5 yr vs. 15.5 yr Gender - Female: 39 (97.5%) - Male: 1 (2.5%) Race: NR	Conjoint and separated family therapy had comparable outcomes at the end of treatment but separated treatment appeared preferable if levels of maternal criticism were high. Weight – Baseline: 39.3 kg vs. 40.7 kg Weight, Change - Baseline – 12 mo: 6.4 kg (SD ± 6.2) vs. 9.8 kg (SD ± 6.7) (MD -3.4 kg, p=0.09) %ABW - Baseline: 72% vs. 76% - 1 yr: 82% vs. 90.5% - 6 yr: 91% (SD ± 12.2) vs. 97.7% (SD ± 9.32) (MD -6.7%, p<0.09) BMI, Change - Baseline – 12 mo: 2.4 kg/m² (SD ± 2.5) vs. 3.6	Moderate

					kg/m² (SD ± 2.4) (MD -1.2 kg/m², p=0.1) Disease Response - 12 mo - Poor: 10 (52.63%) vs. 5 (23.81%) - Good: 5 (26.32%) vs. 10 (47.62%) Disease Response - 6 yr - Poor: 4 (22.2%, N=18) vs. 2 (10%, N=20) - Good: 13 (72.2%, N=18) vs. 16 (80%, N=20) Hospitalization - Baseline – 1 yr: 3 (15.79%) vs. 1 (4.76%) Improvement in menstruation at 6 yr follow-up: 13 (72.22%, N=18) vs. 19 (95%, N=20) (p=0.02 for superiority of conjoint therapy) Attrition: 11% (2/19) vs. 10% (2/21)	

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BMI=body mass index; EE=Expressed Emotion; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Short-term compared to long-term

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Lock et al. (2005*, 2006b, 2006a*)	Designs: RCT; Follow-up/Extension Setting: NR Country: NR	Randomized N=86 FBT 6 mo (N=44) FBT 12 mo (N=42)	Inclusion: 12-18 years of age; AN Exclusion: Severe physical health problems likely to affect weight; diabetes mellitus; severe psychiatric illnesses that	AN: 86 (100%) - Restricting: 70 (81.4%) - Binge-eating and purging: 16 (19%)	No differences in overall outcomes were noted between 6 mo and 12 mo of treatment with FBT, although there was a suggestion that 12 mo of treatment was more beneficial with non-intact families or	Low

	Funding: Government	Follow-up: 3.96 yr (Mean) Follow-up analysis (N=60) - 32 vs. 28	would interfere with treatment; psychosis; lack of response to family treatment	AN, Duration: 11.3 mo (SD ± 10.4) vs. 12 mo (SD ± 9.9) Age: 15.2 yr (SD ± 1.6) vs. 15.2 yr (SD ± 1.7) Gender - Female: 39 (89%) vs. 38 (91%) - Male: 5 (11%) vs. 4 (9%) Race - Caucasian: 32 (73%) vs. 32 (76%) - Asian: 6 (14%) vs. 2 (5%) - Biracial: 3 (4%) - Native American: 1 (2%) vs. 0 (0%) Ethnicity - Hispanic/Latino: 4 (9%) vs. 6 (14%) - Other: 1 (2%) vs. 2 (5%)	participants with more severe eating-related obsessive-compulsive symptoms. Weight - Baseline: 44.6 kg (SD ± 5.5) vs. 46.7 kg (SD ± 7.2) Weight, Change - Baseline – 6 mo: 6.1 kg vs. 4.8 kg - Baseline – 12 mo: 7.5 kg vs. 6.6 kg BMI – Baseline->12 mo: 17->19.5 kg/m² vs. 17.3->19.5 kg/m² Disease Response, Remission - 12 mo: 21 (60%, N=35) vs. 21 (63.64%, N=33) Hospitalization: 10 (23%) vs. 9 (21%) Hospitalization, Duration: 20 d vs. 16 d Treatment Discontinuation: 2 (4.55%) vs. 7 (16.67%) With longer term follow-up of a mean of 4 years, no differences in outcomes were noted for 6 GSH vs 12 mo of FBT. Outcomes at follow-up: BMI: 20.57 kg/m² (SD ± 2.03) vs. 20.74 kg/m² (SD ± 2.25)	

					BMI > 20 kg/m²: 24 (64.86%) vs. 20 (58.82%) %IBW > 90%: 32 (86.49%) vs. 31 (91.18%) Amenorrhea: 3 (8.11%) vs. 1 (2.94%) Menstruation, Resumed, In the Previous 6 mo: 20 (62.5%, N=32) vs. 18 (64.29%, N=28) Hospitalization, None:31 (86.11%, N=36) vs. 26 (83.87%, N=31) Hospitalization >= 3: 2 (5.56%, N=36) vs. 0 (0%, N=31) Residential Treatment: 4 (10.81%) vs. 1 (3.03%, N=33) Attrition: 9% (4/44) vs. 24% (10/42)	

Abbreviations: AN=anorexia nervosa; FBT=family-based treatment; BMI=body mass index; IBW=ideal body weight; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation

Compared to +/- intensive parental coaching

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Lock et al. (2015b)*	Design: RCT Setting: Multi-center Country: United States Funding: Government	Randomized N=45 FBT 6 mo (N=10) FBT +/- IPC 6 mo (N=35) - IPC, None (N=23) - IPC, Yes (for those with weight gain below 2.3kg at wk 4) (N=12)	Inclusion: 12-18 years of age; AN; medically stable for outpatient treatment; stable dose of psychotropic medication for at least 8 weeks; taking a psychotropic medication for a comorbid psychiatric condition; living with family Exclusion: Physical, psychotic or other mental illness requiring hospitalization; dependent on drugs or alcohol; physical conditions known to influence eating or weight; previous FBT	AN: 45 (100%) AN, Duration: 4.3 mo (SD ± 1.6) vs. 12.6 mo (SD ± 13.7) - IPC, None: 9.8 mo (SD ± 9) - IPC, Yes: 18 mo (SD ± 19.4) Age 12 yr-18 yr: 45 (100%) Age: 14.3 yr (SD ± 1.5) vs. 14.6 yr (SD ± 1.4) Gender - Female: 9 (90%) vs. 5 (14.3%): - Male: 1 (10%) vs. 30 (85.7%) Race - Caucasian: 9 (90%) vs. 28 (80%) - Asian: 1 (10%) vs. 4 (11.4%) - Mixed: 0 (0%) vs. 3 (8.6%)	Outcomes did not differ for the initial randomly assigned groups. Poor early responders achieved comparable weight gain to early responders by the end of treatment, but the study design was unbalanced and lacked statistical power. Weight - 6 mo: 114.4 lbs (SD ± 12.9, N=8) vs. 111.6 lbs (SD ± 13.5, N=33) (MD 2.8 lbs, p=0.598) - IPC, None vs. Yes: 111.5 lbs (SD ± 16.1, N=21) vs. 111.7 lbs (SD ± 8) (MD -0.2 lbs, p=0.955) BMI - Baseline: 16.1 kg/m² (SD ± 1.1) vs. 16.2 kg/m² (SD ± 0.9) - IPC, None vs. Yes: 16.1 kg/m² (SD ± 0.8) vs. 16.4 kg/m² (SD ± 0.9) BMI - 6 mo: 18.9 kg/m² (SD ± 1.2, N=8) vs. 19 kg/m² (SD ± 1.4, N=33) (MD -0.1 kg/m², p=0.735) - IPC, None vs. Yes: 18.9 kg/m² (SD ± 1.6, N=21) vs. 19.3 kg/m² (SD ± 0.9) (MD -0.4 kg/m², p=0.487) %IBW – Baseline: 82.8% (SD ± 3.8) vs. 82.4% (SD ± 3.2): - IPC, None vs. Yes: 82% (SD ± 3.3) vs. 83.2% (SD ± 2.9) %EBW - 6 mo: 96.5% (SD ± 4.7, N=8) vs. 95.7% (SD ± 7.2, N=33) (MD 0.8%, p=0.759) - IPC, None vs. Yes: 95.1% (SD ± 7.6, N=21) vs. 96.7%	High

					(SD ± 6.5) (MD -1.6%, p=0.552) Attrition: 20% (2/10) vs. 20% (7/35) - IPC, None vs. Yes: 22% (5/23) vs. 17% (2/12)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; FBT=family-based treatment; EBW= expected body weight; IBW=ideal body weight; IPC=intensive parental coaching; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Family-based treatment with art therapy compared to family-based treatment with cognitive remediation therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Lock et al. (2018)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=30 Art Therapy + FBT 9 mo (N=15) CRT + FBT 9 mo (N=15)	Inclusion: 12-18 years of age; AN; medically stable for outpatient treatment; Yale Brown Cornell Eating Disorder Scale score > 1; children's Yale Brown Obsessive Compulsive Scale score > 8; obsessive compulsive Exclusion: Associated physical, psychotic, or other mental illness requiring hospitalization; current dependence on drugs or alcohol; physical conditions known to influence eating or weight; scores below the normal range in the Wechsler Abbreviated Scale of Intelligence; family history of child abuse or neglect; current child abuse or neglect; diabetes mellitus; pregnancy	AN: 30 (100%) AN, Duration: 10.38 mo (SD ± 12.75) - 8.47 mo (SD ± 5.46) vs. 12.43 mo (SD ± 17.59) Age 12 yr-18 yr: 30 (100%) Age: 14.49 yr (SD ± 1.64) - 14.55 yr (SD ± 1.48) vs. 14.42 yr (SD ± 1.83) Gender - Female: 14 (93.3%) vs. 13 (86.7%) - Male: 1 (6.7%) vs. 2 (13.3%) Race - Caucasian: 9 (60%) vs. 9 (60%) - Asian: 3 (20%) vs. 2 (13.3%) - Mixed: 3 (20%) vs. 4 (26.7%)	In adolescents with AN and high levels of obsessive-compulsive features, FBT in combination with either art therapy or cognitive remediation therapy was associated with improvements in weight-related outcomes and reductions in cognitive inefficiencies. BMI – Baseline: 16.32 kg/m² (SD ± 1.2) vs. 16.37 kg/m² (SD ± 1) BMI, Change - Baseline – 9 mo: 2.1 kg/m² (SD ± 1.38, N=11) vs. 1.51 kg/m² (SD ± 0.95, N=12) (MD 0.59 kg/m², p=0.24) Percent Estimated Body Weight – Baseline: 83.17% (SD ± 4.63) vs. 83.96% (SD ± 4.04) Percent Estimated Body Weight, Change - Baseline – 9 mo: 8.77% (SD ± 6.22, N=11) vs.	High

				Ethnicity, Hispanic/Latino: - 5 (33%) vs. 4 (26.7%)	6.39% (SD ± 5.1, N=12) (MD 2.38%, p=0.32) Attrition: 33% (15) vs. 13% (2/15)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CRT=cognitive remediation therapy; FBT=family-based treatment; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation

Videoconferencing compared to online guided self-help program

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Lock et al. (2021)	Designs: RCT Setting: Multi-center Country: United States Funding: Government	Randomized N=40 FBT online guided self-help 4-6 mo (N=20) FBT via videoconferencing NR (N=20) Follow-up: 3 mo	Inclusion: 12-18 years of age; AN; ≤ 88% EBW Exclusion: NR	AN: 40 (100%) Age: 14.8 yr (SD ± 1.86) vs. 14.9 yr (SD ± 1.82) Gender - Female: 18 (90%) vs. 16 (80%) - Male: 2 (10%) vs. 4 (20%) Race - Caucasian: 18 (90%) vs. 17 (85%) - Asian: 1 (5%) vs. 1 (5%) Ethnicity - Hispanic/Latino: 2 (10%) vs. 2 (10%)	Across both treatment groups, the average change in percent of estimated ABW from baseline to end of treatment was 9.28 percentage points (SD ± 6.21). Percent of Estimated ABW-Baseline->End of Treatment: 80.55 (SD ± 4.38)->90.80 (SD ± 7.16, N=18) vs. 84.47 (SD ± 4.26)->92.97 (SD ± 7.33, N=19) BMI - Baseline->End of Treatment: 16.02 (SD ± 1.20)->18.27 kg/m² (SD ± 1.70, N=18 vs. 16.84 (SD ± 0.93)->18.81 kg/m² (SD ± 1.55, N=19) Weight, Remission – End of Treatment: 5 (27.8%) vs. 9 (45%) Remission – End of Treatment: 2 (11%) vs. 6 (30%)	Moderate

					Hospitalization: 4 (22%) vs. 1 (5%) Hospitalization, Duration: 19 d vs. 22 d Attrition: 10% (2/20) vs.15% (3/20)	

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BMI=body mass index; d=day; EBW=expected body weight; FBT=family-based treatment; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation

Compared to Family Treatment Without Parents in Charge

Compared to systematic family therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (2014) (RIAN); Lock et al. (2016)	Design: RCT; Post-hoc Analysis Setting: Multi-center Country: United States Funding: Government	Randomized N=164 FBT 36 wk (N=82) Systemic Family Therapy 36 wk (N=82) Current analysis (N=158) - 78 vs. 80 Follow-up: Baseline – 88 wk	Inclusion: Adolescents; AN; %IBW <=87%; 12-18 years of age Exclusion: Current psychotic illness; intellectual disability that would prohibit the use of psychotherapy; bipolar disorder; dependence on drugs or alcohol; previous family therapy for AN; taking medications that may induce weight loss; medical instability; weight at or below 75% of the IBW	AN: 164 (100%) AN, Duration 13.5 mo (SD ± 13.9, N=158) - 11.6 mo (SD ± 9.8, N=78) vs. 15.4 mo (SD ± 16.9, N=80) %IBW <= 87%: 164 (100%) %IBW: 81.9% (N=158) Age 12 yr-18 yr: 164 (100%) Age: 15.3 yr (SD ± 1.8, N=158) - 15.1 yr (SD ± 1.7, N=78) vs. 15.6 yr (SD ± 1.8, N=80) Gender	FBT and systemic family therapy did not differ in the primary outcomes of %IBW or remission and did not differ in eating disorder symptoms or co-occurring conditions at 36-wk or at 88-wk follow-up. FBT showed significantly shorter hospital days/admission: 8.3 d/admission (N=78) vs. 21 d/admission (N=80) (MD -12.7 d/admission, p=0.02) %IBW - Baseline: 82.2% (SD ± 3.8, N=78) vs. 81.7% (SD ± 3.7, N=80) - 36 wk: 92.1% (N=78) vs. 91.1% (N=80) (MD 1%, p=0.31) - 88 wk: 94.6% (N=78) vs. 93.3% (N=80) (MD 1.3%, p=0.31)	Low

				- Female: 67 (85.9%, N=78) vs. 74 (92.5%, N=80) - Male: 11 (14.1%) vs. 6 (7.5%) Race: NR	Disease Response, Remission - 36 wk: 26 (33.1%, N=78) vs. 20 (25.3%, N=80) (p=0.22) - 88 wk: 32 (40.7%, N=78) vs. 31 (39%, N=80) (p=0.84) Hospitalization, Sum – Baseline – 1 yr: 369 d vs. 655 d Adverse Events, Serious – Baseline – 36 wk: 12 (15.4%, N=78) vs. 20 (25%, N=80) Study Withdrawal, Adverse Events, Serious – Baseline – 36 wk: 3 (3.85%, N=78) vs. 7 (8.75%, N=80) Attrition:25% (20/82) vs. 25% (20/82)	

Abbreviations: AN=anorexia nervosa; d=day; FBT=family-based treatment; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; RIAN=Research in Anorexia Nervosa; SD=standard deviation; wk=week; wks=weeks; yr=year

Compared to parent focused treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Le Grange et al. (2016)*; Allan et al. (2018)*	Design: RCT; Extension Setting: Single center: specialist pediatric eating disorders program within a tertiary public hospital Country: Australia	Randomized N=107 FBT 6 mo (N=55) Parent-Focused Treatment 6 mo (N=52) Follow-up: Baseline – 18 mo	Inclusion: AN; 12-18 years of age; living with at least 1 parent available to undertake treatment; weight was <95% median BMI Exclusion: Medical instability; current psychotic disorder; drug or alcohol dependence; acute suicidality; physical condition	AN: 107 (100%) AN, Duration: 10.5 mo (SD ± 8.8, N=106) - 11 mo (SD ± 9.4) vs. 10 mo (SD ± 8.1, N=51) BMI: 16.5 kg/m² (SD ± 1.3, N=106)	Remission rates were greater with parent-focused treatment than FBT at the end of treatment but comparable at both follow-up times. Median percent BMI did not differ between the groups at any time point. Disease Response, Remission	Low

	Funding: Non-profit and government		influencing eating; cancer; previous FBT for AN; psychotropic medication <8 weeks; physical condition influencing weight	- 16.3 kg/m² (SD ± 1.2) vs. 16.7 kg/m² (SD ± 1.4, N=51) BMI, Median Percent < 95%: 107 (100%) BMI, Median Percent: 81.9% (SD ± 6.1, N=106) Age: 15.5 yr (SD ± 1.5, N=106) - 15.4 yr (SD ± 1.3) vs. 15.7 yr (SD ± 1.6, N=51) Gender - Female: 49 (89.1%) vs. 44 (86.3%, N=51) - Male: 6 (10.9%) vs. 7 (13.7%) Race: NR	- 6 mo: 12 (21.8%, N=55) vs. 22 (43.1%, N=51) (OR 0.33, 95% CI 0.13 – 0.81, p=0.016) - 12 mo: 12 (21.8%; N=55) vs. 20 (39.2%; N=51) (OR 2.48; 95% CI 0.989-6.22; p=0.053) - 18 mo: 16 (29.1%, N=55) vs. 19 (37.3%, N=51) (OR 0.7194, 95% CI 0.31 – 1.67, p=0.444) Hospitalization - Baseline: 21 (38.2%) vs. 18 (35.3%, N=51) - Baseline – 6 mo: 13 (23.6%, N=55) vs. 6 (11.8%, N=51) Rehospitalizations - Baseline – 18 mo: 10 (18.18%, N=55) vs. 5 (9.8%, N=51) Study Withdrawal, Hospitalization >= 2 - Baseline – 6 mo: 2 (3.64%, N=55) vs. 2 (3.92%, N=51) Attrition: 16% (9/55) vs. 15% (8/52)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CI=confidence interval; FBT=family-based treatment; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to adolescent-focused therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Lock et al. (2010); Ciao et al. (2015); Le Grange et al. (2014a; 2014b)	Design: RCT; Follow-up; Extension Setting: Multi-center Country: United States Funding: Government	Randomized N=121 FBT 12 mo (N=61) AFT 12 mo (N=60) Follow-up: Baseline – 24 mo; Baseline – 4 yr 4 yr Follow-up N=79 4 yr Follow-up Mean Duration: 3.3 yr (SD ± 1.33, N=36) vs. 3.21 yr (SD ± 1.26, N= 43)	Inclusion: AN; 12-18 years of age; live with their parents or legal guardians; IBW <86% Exclusion: Current psychotic disorder; dependence on drugs or alcohol; physical condition known to influence eating or weight; diabetes mellitus; pregnancy; previous treatment with FBT or AFT	AN: 121 (100%) AN, Duration: 11.3 mo (SD ± 8.6) - 12.3 mo (SD ± 8.5) vs. 10.3 mo (SD ± 8.7) BMI: 16.1 kg/m² (SD ± 1.1) %IBW: 82% %EBW: 80.4% (SD ± 3.6) Age 12 yr-18 yr: 121 (100%) Age: 14.4 yr (SD ± 1.6) - 14.1 yr (SD ± 1.7) vs. 14.7 yr (SD ± 1.5) Gender - Female: 54 (89%) vs. 56 (93%) - Male: 7 (11%) vs. 4 (7%) Race - Caucasian: 45 (74%) vs. 47 (78%) - Asian: 7 (12%) vs. 6 (10%) - Black or African American: 0 (0%) vs. 1 (2%) Ethnicity - Hispanic/Latino: 6 (10%) vs. 3 (5%) - Minority: 16 (26%) vs. 13 (22%)	- FBT was associated with significantly greater remission rates at 18- and 24-mo follow-ups:18 mo: 40% vs. 18% (N=88, p=0.029) - 24 mo: 22 (49%, N=44) vs. 11 (23%, N=49) (p=0.024) BMI - 12 mo: 31.4 Percentile (SD ± 21.87) vs. 23.4 Percentile (SD ± 21.69) - 18 mo: 31.4 Percentile (SD ± 27.34) vs. 29.1 Percentile (SD ± 26.34) - 24 mo: 32.2 Percentile (SD ± 26.55) vs. 29 Percentile (SD ± 26.34) %EBW < 95% - 6 mo: 45% vs. 61% - 12 mo: 38% vs. 55% - 18 mo: 27% vs. 44% - 24 mo: 23% vs. 40% %EBW > 95% - Baseline – 24 mo: 41 (67.2%) vs. 33 (55%) %EBW - 1 yr: 94.23% (SD ± 9.49) vs. 93.06% (SD ± 13.72) - 4 yr: 94.43% (SD ± 12.1, N=36) vs. 93.84% (SD ± 10.34, N=43) (MD 0.59%, p=0.82) FBT was associated with significantly less hospitalization at 24 mo.: 9 (15%) vs. 22 (37%) (p=0.02).	Moderate

					Study Withdrawal, All-Cause - Baseline – 24 mo: 9 (14.75%) vs. 3 (5%) Attrition at 12-mo follow-up: 28% (17/61) vs. 18% (11/60)	

Abbreviations: AFT=adolescent-focused therapy; AN=anorexia nervosa; BMI=body mass index; d=day; EBW=expected body weight; FBT=family-based treatment; IBW= ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to +/- intensive parental coaching

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Lock et al. (2015b)*	Design: RCT Setting: Multi-center Country: United States Funding: Government	Randomized N=45 FBT 6 mo (N=10) FBT +/- IPC 6 mo (N=35) - IPC, None (N=23) - IPC, Yes (for those with weight gain below 2.3kg at wk 4) (N=12)	Inclusion: 12-18 years of age; AN; medically stable for outpatient treatment; stable dose of psychotropic medication for at least 8 weeks; taking a psychotropic medication for a comorbid psychiatric condition; living with family Exclusion: Physical,psychotic, or other mental illness requiring hospitalization; dependent on drugs or alcohol; physical conditions known to influence eating or weight; previous FBT	AN: 45 (100%) AN, Duration: 4.3 mo (SD ± 1.6) vs. 12.6 mo (SD ± 13.7) - IPC, None: 9.8 mo (SD ± 9) - IPC, Yes: 18 mo (SD ± 19.4) Age 12 yr-18 yr: 45 (100%) Age: 14.3 yr (SD ± 1.5) vs. 14.6 yr (SD ± 1.4) Gender - Female: 9 (90%) vs. 5 (14.3%): - Male: 1 (10%) vs. 30 (85.7%) Race - Caucasian: 9 (90%) vs. 28 (80%) - Asian: 1 (10%) vs. 4 (11.4%) - Mixed: 0 (0%) vs. 3 (8.6%)	Outcomes did not differ for the initial randomly assigned groups. Poor early responders achieved comparable weight gain to early responders by the end of treatment, but the study design was unbalanced and lacked statistical power. Weight - 6 mo: 114.4 lbs (SD ± 12.9, N=8) vs. 111.6 lbs (SD ± 13.5, N=33) (MD 2.8 lbs, p=0.598) - IPC, None vs. Yes: 111.5 lbs (SD ± 16.1, N=21) vs. 111.7 lbs (SD ± 8) (MD -0.2 lbs, p=0.955) BMI - Baseline: 16.1 kg/m² (SD ± 1.1) vs. 16.2 kg/m² (SD ± 0.9) - IPC, None vs. Yes: 16.1 kg/m² (SD ± 0.8) vs. 16.4 kg/m² (SD ± 0.9) BMI - 6 mo: 18.9 kg/m² (SD ± 1.2, N=8) vs. 19 kg/m² (SD ±	High

					1.4, N=33) (MD -0.1 kg/m², p=0.735) - IPC, None vs. Yes: 18.9 kg/m² (SD ± 1.6, N=21) vs. 19.3 kg/m² (SD ± 0.9) (MD -0.4 kg/m², p=0.487) %IBW – Baseline: 82.8% (SD ± 3.8) vs. 82.4% (SD ± 3.2): - IPC, None vs. Yes: 82% (SD ± 3.3) vs. 83.2% (SD ± 2.9) %EBW - 6 mo: 96.5% (SD ± 4.7, N=8) vs. 95.7% (SD ± 7.2, N=33) (MD 0.8%, p=0.759) - IPC, None vs. Yes: 95.1% (SD ± 7.6, N=21) vs. 96.7% (SD ± 6.5) (MD -1.6%, p=0.552) Attrition: 20% (2/10) vs. 20% (7/35) - IPC, None vs. Yes: 22% (5/23) vs. 17% (2/12)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; FBT=family-based treatment; EBW=expected body weight; IBW=ideal body weight; IPC=intensive parental coaching; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to Treatment As Usual

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Gabel et al. (2014)	Design: Retrospective Cohort Study Setting: Single Center: Hospital for Sick Children Country: Canada Funding: NR	Current analysis N=50 Multiple Family Therapy + TAU 1 yr (N=25) TAU 1 yr (N=25)	Inclusion: Adolescents; AN; underwent treatment in the eating disorders program at the Hospital for Sick Children between 2002 and 2010 Exclusion: NR	AN: 50 (100%) %IBW: 78.4% (SD ± 9.77) Adolescent: 50 (100%) Age: 14.1 yr (SD ± 1.87)	Multiple family therapy showed significantly greater %IBW than TAU at 1-yr follow-up: - Baseline: 77.72% vs. 79.11% - 1 yr: 99.6% (SD ± 7.27) vs. 95.4% (SD ± 6.88) (MD 4.2%, p<0.05)	Not determined due to study design

				Gender, Female: 50 (100%) Race: NR	Attrition: NR	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; IBW=ideal body weight; MD=mean difference; NR=not reported; SD=standard deviation; TAU=treatment as usual; yr=year

Compared to Other Psychotherapy

Compared to cognitive-behavioral therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Ball and Mitchell (2004)	Design: RCT Setting: Outpatient: Eating Disorders Unit at Prince Henry Hospital Country: Australia Funding: Non-profit	Randomized N=25 CBT 12 mo (N=13) Behavioral Family Therapy 12 mo (N=12) Follow-up: Baseline – 18 mo	Inclusion: Female; 13-23 years of age; AN; currently living with their family Exclusion: BMI < 13.5 kg/m^2; currently receiving other psychological or pharmacological treatments; comorbid physical disorder or psychiatric disorder; current drug abuse or alcohol abuse; self-harming behavior over the past 12 months; other indications for hospitalization; severe physical complications; suicidal ideation; recent history of untreated physical trauma; recent history of psychological trauma; recent history of sexual abuse	AN: 25 (100%) - Restricting type: 7 (53.8%) vs. 9 (75%) - Binge-eating and purging type: 6 (46.2%) vs. 3 (25%) History of hospitalization: 4 (30.8%) vs. 3 (25%) No history of hospitalization: 9 (69.2%) vs. 9 (75%) Age 13 yr-23 yr: 25 (100%) - 18.45 yr (SD ± 2.57) vs. 17.58 yr (SD ± 3.37) Gender, Female: 25 (100%) Race: NR	Disease response and change in BMI did not differ in individuals treated with CBT vs. behavioral family therapy. Disease Response, Good - 12 mo: 15 (60%) - 18 mo: 15 (60%) BMI - Baseline: 16.06 kg/m² (SD ± 1.58) vs. 16.45 kg/m² (SD ± 0.85) BMI, Change: - Baseline – 12 mo: 2.67 kg/m² (SD ± 1.28, N=9) vs. 2.54 kg/m² (SD ± 1.57, N=9) - Baseline – 18 mo: 2.49 kg/m² (SD ± 1.31, N=9) vs. 3.2 kg/m² (SD ± 1.55, N=9) Hospitalization - Baseline – 12 mo: 3 (16%, N=18) Attrition: 31% (4/13) vs. 25% (3/12)	High

Nyman-Carlsson et al. (2020)	Design: RCT Setting: Outpatient Country: Sweden Funding: Non-profit	Randomized N=78 CBT 18 mo (N=38) Family + Individual Therapy 18 mo (N=40) Follow-up: Baseline – 36 mo Current Analysis (N=74) - 37 vs. 37	Inclusion: Female; 17-24 years of age; AN; BMI < 17.5; parents' participation Exclusion: Critical medical status; current suicidal thoughts and/or suicidal behavior; current alcohol or substance abuse; on-going psychotherapeutic or psychotropic treatments	AN: 78 (100%) - Binge-purge type: 16 (43%) vs. 12 (32%) - Restrictive type: 21 (57%) vs. 25 (68%) AN Duration: 31.6 mo (SD ± 24.1) vs. 26.8 mo (SD ± 24.4) Age 17 yr-24 yr: 78 (100%) - 19.1 yr (SD ± 1.9) vs. 18.7 yr (SD ± 2.0) Gender, Female: 78 (100%) Race: NR	BMI increased significantly from baseline to post-treatment in both groups (p=0.0001): 16.49->19.61 kg/m² for CBT vs. 16.54->19.33 kg/m² for Family + Individual Therapy Remission – Post-Treatment: 28 (75.7%) vs. 28 (75.7%) Both groups decreased significantly (all p=0.001) from baseline to post-treatment in eating disorder-specific symptoms and general psychological symptoms, as measured by the EDI-3, GPMC, and BDI. Attrition: 3% (1/38) vs. 8% (3/40)	High

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BMI=body mass index; CBT=cognitive-behavioral therapy; EDI=Eating Disorder Inventory; GPMC=General Psychological Maladjustment Composite; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to family group psychoeducation

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Geist et al. (2000)	Design: RCT Setting: Inpatient: The Hospital for Sick Children Country: Canada Funding: Non-profit	Randomized N=25 Family Therapy 4 mo (N=12) Family Group Psychoeducation 4 mo (N=13)	Inclusion: Adolescents; AN; current weight <90% of IBW; requiring hospitalization; AN, severe; self-imposed food restriction; female Exclusion: Under 12 years of age; male; older than 17.4 years; immediate suicide risk; psychotic features; Individual therapy in the community; family therapy in the community; BN; previous	AN: 25 (100%) - Restricting type: 19 (76%) %IBW < 90%: 25 (100%) %IBW: 78.4% (SD ± 9.77) Weight: 41.1 kg (SD ± 7) vs. 41.1 kg (SD ± 6.3)	Both family therapy and family group psychoeducation were associated with improvements in %IBW but there was no significant difference between the treatments on %IBW or measures of eating pathology. %IBW - Baseline: 77.7% vs. 77.2% (SD ± 11.1) - 4 mo: 91.3% (SD ± 7.3) vs. 96.3% (SD ± 8.2)	Moderate

			admissions to the inpatient eating disorder program; risk for self-harm	Adolescent: 25 (100%) Age: 14.3 yr (SD ± 1.5) vs. 14.9 yr (SD ± 1.7) Gender, Female: 25 (100%) Race: NR	- Hospital discharge: 89.1% vs. 90.4% Hospitalization, Duration - Baseline – 4 mo: 46.3 d (SD ± 22.7) vs. 40.8 d (SD ± 22.2) Attrition: 0% vs. 0%	

Abbreviations: AN=anorexia nervosa; d=day; IBW=ideal body weight; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to ego-oriented individual therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Robin et al. (1994, 1995, 1999)	Design: RCT Setting: Outpatient Country: United States Funding: Government	Randomized N=24 BFST (N=12) Ego-Oriented Individual Therapy (N=12) Treatment: 15.9 mo (Mean) Follow-up: 12 mo	Inclusion: Caucasian; adolescents; AN; female; resided at home with one or both parents Exclusion: Bulimic features	AN: 24 (100%) Weight: 85.4 lbs (SD ± 12.7, N=11) vs. 91 lbs (SD ± 23.1, N=11) Adolescent: 24 (100%) Age: 14.7 yr (SD ± 2.7, N=11) vs. 13.9 yr (SD ± 2.1, N=11) Gender, Female: 24 (100%) Race, Caucasian: 24 (100%)	Significantly greater BMI change was associated with BFST than with ego-oriented individual therapy, but other outcomes did not differ. BMI Regression Analysis: Baseline to 15.9 mo (mean): 5.1 kg/m² (SD ± 1.6, N=11) vs. 2.7 kg/m² (SD ± 2.2, N=11) (MD 2.4 kg/m², p<0.01) Menstruation, Resumed - End of Treatment: 10 (89%, N=11) vs. 7 (60%, N=11) Hospitalization - 15.9 mo (Mean): 3 (27.27%) vs. 5 (45.45%) Attrition: 8% (1/12) vs. 8% (1/12)	High

Abbreviations: AN=anorexia nervosa; BFST=behavioral family systems therapy; BMI=body mass index; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation

Compared to individual supportive therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Russell et al. (1987); Dare et al. (1990); Eisler et al. (1997)	Design: RCT; Follow-up/Extension Setting: Outpatient Country: United Kingdom Funding: NR	Randomized N= 80 Family Therapy 1 yr (N=41) - Age > 18 yr (N=24) - Age <= 18 yr (N=17) - AN, Duration < 3 yr and Age, at Disease Onset <= 18 yr (N=10) - AN, Duration > 3 yr or Age, at Disease Onset <= 18 yr (N=10) - Age, at Disease Onset > 18 yr (N=7) Individual Supportive Therapy 1 yr (N=39) - Age > 18 yr (N=24) - Age <= 18 yr (N=15) - AN, Duration < 3 yr and Age, at Disease Onset <= 18 yr (N=11) - AN, Duration > 3 yr or Age, at Disease Onset <= 18 yr (N=9) - Age, at Disease Onset > 18 yr (N=7) Follow-up: 5.2 yr (Mean, SD ± 2.1; N=77)	Inclusion: AN/BN severe and requiring hospitalization Exclusion: NR	AN: 27 (75%, N=36) vs. 27 (72.97%, N=37) BN: 9 (25%, N=36) vs. 10 (27.03%, N=37) Age: 14 – 55 - <= 18 yr: 17 (41.46%) vs. 15 (38.46%) - 18 yr: 24 (58.54%) vs. 24 (61.54%) Gender, Unknown: 80 (100%) Race: NR	Outcome in terms of disease response and % change in ABW was better with family therapy in individuals with an illness duration < 3 yr who were <=18 at illness onset, but better with individual therapy in those with illness onset >18 yr. Disease Response - 1 yr - Good: 8 (19.51%) vs. 6 (15.38%) - Intermediate: 7 (17.07%) vs. 5 (12.82%) - Poor: 21 (51.22%) vs. 26 (66.67%) AN subgroup with Duration < 3 yr and Age at Disease Onset <= 18 yr: %ABW - Hospital Admission: 67% vs. 65% - Baseline: 89% vs. 88% - 1 yr: 93% vs. 80% - 5.2 yr (Mean): 103.4 (SD 13.2, N=10) vs. 94.4 (SD 16.8; N=9) %ABW, Change - Baseline – 1 yr: 25.5% vs. 15.5% (MD 10%, p<0.01) Disease Response - 1 yr - Good: 6 (60%) vs. 1 (9.09%) (p<0.02)	High

					- Intermediate: 3 (30%) vs. 1 (9.09%) - Poor: 1 (10%) vs. 9 (81.82%) (p<0.002) Disease Response - 5.2 yr (Mean, SD ± 2.1) - Good:9 (90%) vs. 4 (36%) - Intermediate:0 (0%) vs.2 (18%) - Poor:1 (10%) vs. 5 (45%) AN subgroup with Age at Disease Onset <= 18 yr and Duration > 3 yr: %ABW - Hospital Admission: 67% vs. 65% - Baseline:91% vs. 92% - 1 yr: 82% vs. 80% - 5.2 yr (Mean): 86.9% (SD ± 11.9) vs. 95.7% (SD ± 11.5) Disease Response - 1 yr - Good: 2 (20%) vs. 2 (22.22%) - Intermediate: 2 (20%) vs. 1 (11.11%) - Poor: 6 (60%) vs. 6 (66.67%) Disease Response - 5.2 yr (Mean, SD ± 2.1) - Good: 3 (30%, N=10) vs. 1 (11.11%, N=9) - Intermediate: 1 (10%, N=10) vs. 4 (44.44%, N=9) - Poor: 6 (60%, N=10) vs. 4 (44.44%, N=9)	

					AN subgroup with Age at Disease Onset > 18 yr: %ABW - Hospital Admission: 65% vs. 60% - Baseline: 85% vs. 86% - 1 yr: 71% vs. 79% - 5.2 yr (Mean): 93.7% (SD ± 18, N=7) vs. 97.5% (SD ± 9, N=7) %ABW, Change - Baseline – 1 yr: 5.4% vs. 19.9% (MD -14.5%, p<0.01) Disease Response - 1 yr - Good: 0 (0%) vs. 2 (28.57%) - Intermediate: 1 (14.29%) vs. 1 (14.29%) - Poor: 6 (85.71%) vs. 4 (57.14%) Disease Response - 5.2 yr (Mean, SD ± 2.1) - Good: 2 (28.57%, N=7) vs. 4 (57.14%, N=7) - Intermediate: 2 (28.57%, N=7) vs. 2 (28.57%, N=7) - Poor: 3 (42.86%, N=7) vs. 1 (14.29%, N=7) Attrition: 37% (15/41) vs. 33% (13/39) for original study and 13% (10/80) overall at 5-yr follow-up	

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BN=bulimia nervosa; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to family therapy with body awareness therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Wallin et al. (2000)	Design: RCT Setting: Single Center: University Hospital of Lund Country: Sweden Funding: NR	Randomized N=26 Body Awareness Therapy + Family Therapy (N=13) Family Therapy (N=13) Treatment Duration: NR Follow-up: Baseline – 2 yr	Inclusion: Teenage; AN; female Exclusion: NR	AN: 26 (100%) AN, Duration: 11.6 mo - 15.4 mo (SD ± 15.6) vs. 8.2 mo (SD ± 3.3) BMI: 15.1 kg/m² (SD ± 1.9) vs. 15.8 kg/m² (SD ± 1.6) Age 13 yr-19 yr: 26 (100%) Gender, Female: 26 (100%) Race: NR	Addition of body awareness therapy to family therapy was not associated with any difference in weight related outcomes. %EBW – Baseline: 72.5% (SD ± 8.3) vs. 75.3% (SD ± 8.3) %EBW - 2 yr (both groups): 90.9% (p<0.0001) Recovery - Baseline - 2 yr: 8 (61.5%) vs. 9 (69.2%) Hospitalization: 4 (30.77%) vs. 4 (30.77%) Hospitalization, Duration: 54.3 d (SD ± 52.6) vs. 50 d (SD ± 61.6) Attrition: NR	High

Abbreviations: AN=anorexia nervosa; BMI=body mass index; d=day; EBW=expected body weight; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to inpatient treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Wallin and Holmer (2021)	Design: Non-RCT Setting: Inpatient; Child and Adolescent Mental	N=185 Family Treatment Apartment (N=115)	Inclusion: AN Exclusion: NR	AN, Restrictive Type: 185 (100%)	Readmissions due to weight loss within 6 mo from discharge were less for family treatment apartment than inpatient	Not determined due to

	Health Service in Malmö Country: Sweden Funding: Non-profit	Inpatient Treatment (N=70) Follow-up: 15.5 yr (SD ± 5.0) vs. 12.6 yr (SD ± 4.0) Current Analysis (N=68) - 43 vs. 25		Age at Admission: 14.5 yr (SD ± 2.1, N=43) vs. 15.1 yr (SD ± 1.6, N=25) Gender: - Female: 40 (93%, N=43) vs. 23 (92%, N=25) - Male: 3 (7%, N=43) vs. 2 (8%, N=25) Race: NR	treatment (2, 4.7% vs. 8, 32.0%; p=0.017). Duration of Admission - Baseline-End of Treatment: 42.1 d (SD ± 20.4) vs. 75.7 d (SD ± 66.4) (p=0.007) %EBW - Admission: 76.8% (SD ± 9.8) (SD ± 8.3) vs. 76.4% (SD ± 10.2) %EBW - Discharge: 80.8% (SD ± 10.0) vs. 88.1% (SD ± 11.8) (p=0.013) Weight Gain - Baseline-End of Treatment: 0.29 kg/wk (SD ± 0.63) vs. 0.69 kg/wk (SD ± 0.53) (p=0.011) Attrition: 63% (72/115) vs. 64% (45/70)	study design

Abbreviations: AN=anorexia nervosa; BMI=body mass index; EBW=expected body weight; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Family Therapies Without Parents in Charge Compared to Family-Based Treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (2014) (RIAN); Lock et al. (2016)	Design: RCT; Post-hoc Analysis Setting: Multi-center	Randomized N=164 FBT 36 wk (N=82) Systemic Family Therapy 36 wk (N=82) Current analysis (N=158)	Inclusion: Adolescents; AN; %IBW <=87%; 12-18 years of age Exclusion: Current psychotic illness; intellectual disability that would prohibit the use of psychotherapy; bipolar disorder;	AN: 164 (100%) AN, Duration 13.5 mo (SD ± 13.9, N=158) - 11.6 mo (SD ± 9.8, N=78) vs. 15.4 mo (SD ± 16.9, N=80)	FBT and systemic family therapy did not differ in the primary outcomes of %IBW or remission and did not differ in eating disorder symptoms or co-occurring conditions at 36-wk or at 88-wk follow-up.	Low

	Country: United States Funding: Government	- 78 vs. 80 Follow-up: Baseline - 88 wk	dependence on drugs or alcohol; previous family therapy for AN; taking medications that may induce weight loss; medical instability; weight at or below 75% of the IBW	%IBW <= 87%: 164 (100%) %IBW: 81.9% (N=158) Age 12 yr-18 yr: 164 (100%) Age: 15.3 yr (SD ± 1.8, N=158) - 15.1 yr (SD ± 1.7, N=78) vs. 15.6 yr (SD ± 1.8, N=80) Gender - Female: 67 (85.9%, N=78) vs. 74 (92.5%, N=80) - Male: 11 (14.1%) vs. 6 (7.5%) Race: NR	FBT showed significantly shorter hospital days/admission: 8.3 d/admission (N=78) vs. 21 d/admission (N=80) (MD -12.7 d/admission, p=0.02) %IBW - Baseline: 82.2% (SD ± 3.8, N=78) vs. 81.7% (SD ± 3.7, N=80) - 36 wk: 92.1% (N=78) vs. 91.1% (N=80) (MD 1%, p=0.31) - 88 wk: 94.6% (N=78) vs. 93.3% (N=80) (MD 1.3%, p=0.31) Disease Response, Remission - 36 wk: 26 (33.1%, N=78) vs. 20 (25.3%, N=80) (p=0.22) - 88 wk: 32 (40.7%, N=78) vs. 31 (39%, N=80) (p=0.84) Hospitalization, Sum - Baseline - 1 yr: 369 d vs. 655 d Adverse Events, Serious - Baseline - 36 wk: 12 (15.4%, N=78) vs. 20 (25%, N=80) Study Withdrawal, Adverse Events, Serious - Baseline - 36 wk: 3 (3.85%, N=78) vs. 7 (8.75%, N=80) Attrition:25% (20/82) vs. 25% (20/82)	

Herscovici et al. (2017)*	Design: RCT Setting: Outpatient: Universidad del Salvador Country: Argentina Funding: NR	Randomized N=23 Family Therapy + Family Meal Intervention 6 mo (N=11) Family Therapy 6 mo (N=12) Follow-up: Baseline - 12 mo	Inclusion: Aged 12-20 years; AN Exclusion: Require hospitalization	AN: 23 (100%) AN, Duration: 21.9 mo (SD ± 11.9) vs. 21.1 mo (SD ± 12) Weight: 42.9 kg (SD ± 7.3) Amenorrhea: 9 (90%, N=10) vs. 10 (100%, N=10) Amenorrhea, Duration: 17.7 mo (SD ± 29.8) vs.10.2 mo (SD ± 16.6) Bulimic Symptoms: 5 (45%) vs. 3 (25%) %EBW: 77.8% (SD ± 8.9) Age: 17.1 yr (SD ± 2.3) - 16.9 yr (SD ± 3.1) vs. 17.3 yr (SD ± 1.3) Gender - Female: 11 (100%) vs. 11 (92%) - Male: 0 (0%) vs. 1 (8%) Race, Caucasian: 23 (100%)	The majority of individuals in both groups improved but resumption of menstruation by 6 mo was more likely in the group receiving the family meal intervention in addition to family therapy: 8 (80%, N=10) vs. 3 (27%, N=11) (p=0.03). Weight - Baseline->6 mo->12 mo: 42->45.7->49.4 kg vs. 43.7->48.1->51.6 kg %EBW - Baseline->6 mo->12 mo: 80.1->86.6->91.7% vs. 75.7->82.9->86.4% Bulimic Symptoms, Developed Binges - 12 mo: NR (N=5) vs. 1 (33.33%, N=3) Attrition: 18% (2/11) vs. 0% (0/12)	Moderate

Abbreviations: AN=anorexia nervosa; d=day; EBW=expected body weight; FBT=family-based treatment; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; RIAN=Research in Anorexia Nervosa; SD=standard deviation; wk=week; yr=year

Compared to Treatment As Usual

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Godart et al. (2012)	Design: RCT Setting: Single Center: Institut Mutualiste Montsouris Department of Psychiatry Country: France Funding: Government and academic	Randomized N=60 Family Therapy + TAU 18 mo (N=30) TAU 18 mo (N=30)	Inclusion: AN; female; adolescent; AN duration <=3 years at admission to the hospital; under 19 years old at illness onset; hospitalized in inpatient unit for AN; AN, severely ill; 13-21 years of age Exclusion: Previously received family therapy; any metabolic pathology interfering with eating; any metabolic pathology interfering with digestion; diabetes; psychotic disorder	AN: 60 (100%) - Severe: 60 (100%) - Binge-eating and purging type: 5 (16.67%) vs. 3 (10%) AN, Duration <= 3 yr: 60 (100%) AN, Duration: 16.6 mo (SD ± 6.8) - 17.1 mo (SD ± 8.3) vs. 16.1 mo (SD ± 5.2) BMI: 16.9 kg/m² (SD ± 1.1) Amenorrhea: 60 (100%) Hospitalization: 60 (100%) Age, At Onset < 19 yr: 60 (100%) Adolescent: 60 (100%) Age: 16.6 yr (SD ± 1.6 -) - 16.4 yr (SD ± 1.7) vs. 16.6 yr (SD ± 1.7) Gender, Female: 60 (100%) Race: NR	Addition of family therapy to TAU was associated with greater rates of treatment response, achieving a BMI ≥ the 10th percentile, and resumption of menstruation as compared to TAU alone. Disease Response, Good or Intermediate - 18 mo: 12 (40%) vs. 5 (17.2%, N=29) (OR 3.2, 95% CI 0.9 - 10) Disease Response, Relapse - Baseline – 18 mo: 10 (33.3%) vs. 14 (48.3%, N=29) BMI - Baseline: 17 kg/m² (SD ± 1.2) vs. 16.9 kg/m² (SD ± 1) BMI, Change - Baseline – 18 mo: 0.8 kg/m² (SD ± 1.52) vs. 0.5 kg/m² (SD ± 1.84, N=29) Amenorrhea - 18 mo: 11 (36.7%) vs. 19 (65.5%, N=29) (OR 0.3, 95% CI 0.1 – 0.9) Rehospitalizations, AN - Baseline – 18 mo: 10 (33.3%) vs. 14 (48.3%, N=29) (OR 0.53, 95% CI 0.19 – 1.25) Attrition: 13% (4/30) vs. 10% (3/30)	Low

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CI=confidence interval; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; yr=year

Compared to Other Psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Dare et al. (2001)	Design: RCT Setting: Single Center: Maudsley Hospital Country: United Kingdom Funding: Academic and non-profit	Randomized N= 84 Family Therapy 1 yr (N=22) Cognitive Analytic Therapy 7 mo (N=22) Focal Psychoanalytic Psychotherapy 1 yr (N=21) Low Contact Routine Treatment 1 yr (N=19)	Inclusion: AN, restricting or binge-purging types; adults Exclusion: Mental or physical state considered so dangerous as to require urgent admission to hospital; serious suicidal risk; extremely low weight; hypoglycemia; syncope; potassium less than 2.5 mMol/L; sodium less than 130 mMol/L	AN: 84 (100%) AN, Duration: 6.3 yr (SD ± 5.9) - 5.8 yr (SD ± 4.9) vs. 6.7 yr (SD ± 7.6) vs. 6.7 yr (SD ± 5.9) vs. 6.1 yr (SD ± 5) %ABW: 72.8% (SD ± 7.1) vs. 77.3% (SD ± 8.1) vs. 72.8% (SD ± 7.6) vs. 73.9% (SD ± 7.9) Age >= 18 yr: 84 (100%) Age: 26.3 yr (SD ± 6.7) - 26.6 yr (SD ± 7.6) vs. 27.2 yr (SD ± 7.6) vs. 26.7 yr (SD ± 6.4) vs. 24.3 yr (SD ± 4.5) Gender - Female: 20 (91%) vs. 22 (100%) vs. 21 (100%) vs. 19 (100%) - Male: Family Therapy 1 yr - 2 (9%) Race: NR	Responses with family therapy and focal psychoanalytic psychotherapy were better than with routine treatment. Cognitive analytic therapy had a shorter treatment duration than other groups and showed a non-significant trend to better outcomes than routine treatment. Disease Response - Baseline – 1 yr - Recovery: 3 (13.64%) vs. 3 (13.64%) vs. 3 (14.29%) vs. 0 (0%) - Significantly Improved: 5 (22.73%) vs. 3 (13.64%) vs. 4 (19.05%) vs. 1 (5.26%) - Improvement: 1 (4.55%) vs. 1 (4.55%) vs. 4 (19.05%) vs. 4 (21.05%) - Poor: 13 (59.09%) vs. 15 (68.18%) vs. 10 (47.62%) vs. 14 (73.68%) Mortality, All-Cause - Baseline – 1 yr: 0 (0%) vs. 0 (0%) vs. 0 (0%) vs. 1 (5.26%) Hospitalization - Baseline – 1 yr: 3 (13.64%) vs. 2 (9.09%) vs. 2 (9.52%) vs. 5 (26.32%)	Moderate

					Attrition: 27% (6/22) vs. 41% (9/22) vs. 43% (9/21) vs. 32% (6/19)	
Hall and Crisp (1987)	Design: RCT Setting: Outpatient Country: NR Funding: NR	Randomized N=30 Dietary Advice (N=15) Psychotherapy (N=15) Follow-up: Baseline – 1 yr	Inclusion: Female; AN; severe AN; 13-27 years of age; Social Classes I-III; weight <85% of matched population mean weight; amenorrhea; AN duration between 6-72 months Exclusion: Married	AN, Severe: 30 (100%) AN, Duration 6 mo-72 mo: 30 (100%) - 24.5 mo vs. 29.7 mo Amenorrhea: 30 (100%) Amenorrhea, Duration: 20.1 mo vs. 27.5 mo %AMPW < 85%: 30 (100%) Age 13 yr-27 yr: 30 (100%) - 19.57 yr vs. 19.55 yr Gender, Female: 30 (100%) Race: NR	Both groups showed improvement with treatment and changes in weight did not differ significantly between groups, whereas psychosocial and sexual adjustment scores were higher in the psychotherapy group vs. dietary advice. Weight - Baseline->1 yr: 39.54->46 kg vs. 41->45.1 kg Weight, Desired, Change - Baseline – 1 yr: 3.5 kg vs. 7 kg Amenorrhea - 1 yr: 10 (66.67%) vs. 8 (53.33%) Hospitalization - Baseline – 1 yr: 1 (6.67%) vs. 1 (6.67%) Study Withdrawal - Baseline – 1 yr: NR vs. 1 (6.67%) Attrition: 27% (4/15) vs. 7% (1/15)	Moderate

Abbreviations: ABW=average body weight; AMPW=average-matched population weight; AN=anorexia nervosa; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Cognitive-Behavioral Therapy Eating Focused Compared to Cognitive-Behavioral Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Dalle Grave et al. (2013a)	Design: RCT Setting: Inpatient: Villa Garda Hospital Country: Italy Funding: Non-profit	Randomized N=80 Focused CBT 20 wk (N=42) Complex Broad CBT 20 wk (N=38) Follow-up: Baseline – 72 wk	Inclusion: 14-65 years of age; require inpatient treatment; AN, severe Exclusion: NR	AN, Severe: 80 (100%) AN, Duration: 4 yr vs. 5 yr Requiring Hospitalization: 80 (100%) Weight: 37.4 kg (SD ± 5.4, N=72) BMI: 14.3 kg/m² (SD ± 1.8, N=72) BMI < 16 kg/m²: 63 (78.8%) Age 14 yr-65 yr: 80 (100%) Age: 23.4 yr (SD ± 6.9) - 23.1 yr (SD ± 6.8) vs. 23.7 yr (SD ± 7) Gender - Female: 40 (95.2%) vs. 38 (100%) - Male: 2 (4.8%) vs. 0 (0%) Race: NR	Both focused and complex-broad CBT were associated with improvements in weight related outcomes with no significant differences between the treatments initially or at follow-up. Weight - Baseline: 37.4 kg (SD ± 5.6, N=37) vs. 37.4 kg (SD ± 5.4, N=35) Weight, Change - Baseline – 44 wk: 8.4 kg (SD ± 4.87, N=34) vs. 10.6 kg (SD ± 5.5, N=33) - Baseline – 72 wk: 9.1 kg (SD ± 4.99, N=34) vs. 9.6 kg (SD ± 5.09, N=34) BMI - Baseline: 14.3 kg/m² (SD ± 1.8, N=37) vs. 14.3 kg/m² (SD ± 1.8, N=35) BMI, Change - Baseline – 44 wk: 3.3 kg/m² (SD ± 1.79, N=36) vs. 4 kg/m² (SD ± 1.44, N=33) - Baseline – 72 wk: 3.6 kg/m² (SD ± 1.72, N=34) vs. 3.5 kg/m² (SD ± 1.59, N=34) Attrition: 17% (7/42) vs. 13% (5/38)	Low

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to Maudsley Model of Anorexia Nervosa Treatment for Adults

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Byrne et al. (2017) (SWAN)	Design: RCT Setting: Multi-center Country: Australia Funding: NR	Randomized N=120 CBT-E 10 mo (N=39) MANTRA 10 mo (N=41) SSCM 10 mo (N=40) Follow-up: Baseline – 22 mo	Inclusion: BMI >= 14.0 and < 18.5 kg/m^2; age >=17 years; AN Exclusion: Severe physical illness; severe mental illness; severe substance dependence; current use of atypical antipsychotics; other active psychotherapy focusing on AN; acute suicide risk	AN: 120 (100%) - Restricting type: 12 (30.77%) vs. 20 (48.78%) vs. 21 (52.5%) - Binge-eating and purging type: 27 (69.2%) vs. 21 (51.2%) vs. 19 (47.5%) AN, Duration: 4 yr (SD ± 4.81) vs. 5 yr (SD ± 5.93) vs. 2 yr (SD ± 5.19) BMI: 16.7 kg/m² (SD ± 1.22) BMI >= 14 kg/m²-< 18.5 kg/m²: 120 (100%) Age >= 17 yr: 120 (100%) Age: 26.19 yr (SD ± 9.47) - 24.18 yr (SD ± 8) vs. 25.95 yr (SD ± 9) vs. 28.44 yr (SD ± 10.94) Gender - Female: 38 (97.44%) vs. 40 (97.56%) vs. 37 (92.5%) - Male: 1 (2.56%) vs. 1 (2.44%) vs. 3 (7.5%) Race: NR	CBT-E, MANTRA, and SSCM each resulted in improvements in weight-related outcomes with no significant differences among the treatments. BMI – Baseline: 16.59 kg/m² (SD ± 1.35) vs. 16.91 kg/m² (SD ± 1.11) vs. 16.58 kg/m² (SD ± 1.18) BMI, Change - Baseline – 10 mo: 2.1 kg/m² (SD ± 1.74) vs. 1.37 kg/m² vs. 1.58 kg/m² (SD ± 1.72) - Baseline – 22 mo: 2.35 kg/m² (SD ± 1.74) vs. 1.5 kg/m² vs. 1.9 kg/m² (SD ± 1.72) BMI > 18.5 kg/m² - Baseline->10 mo->22 mo: 2 (5.01%)->21 (54.11%)->23 (59%) vs. 1 (2.43%)->20 (48.1%)->18 (43.9%) vs. 2 (5.01%)->17 (42.37%)->19 (47.5%) Attrition: 33% (13/39) vs. 44% (18/41) vs. 43% (17/40)	Low

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM=Specialist Supportive Clinical Management; SWAN=Strong Without Anorexia Nervosa; yr=year

Compared to Specialist Supportive Clinical Management

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Byrne et al. (2017) (SWAN)	Design: RCT Setting: Multi-center Country: Australia Funding: NR	Randomized N=120 CBT-E 10 mo (N=39) MANTRA 10 mo (N=41) SSCM 10 mo (N=40) Follow-up: Baseline – 22 mo	Inclusion: BMI >= 14.0 and < 18.5 kg/m^2; age >=17 years; AN Exclusion: Severe physical illness; severe mental illness; severe substance dependence; current use of atypical antipsychotics; other active psychotherapy focusing on AN; acute suicide risk	AN: 120 (100%) - Restricting type: 12 (30.77%) vs. 20 (48.78%) vs. 21 (52.5%) - Binge-eating and purging type: 27 (69.2%) vs. 21 (51.2%) vs. 19 (47.5%) AN, Duration: 4 yr (SD ± 4.81) vs. 5 yr (SD ± 5.93) vs. 2 yr (SD ± 5.19) BMI: 16.7 kg/m² (SD ± 1.22) BMI >= 14 kg/m²-< 18.5 kg/m²: 120 (100%) Age >= 17 yr: 120 (100%) Age: 26.19 yr (SD ± 9.47) - 24.18 yr (SD ± 8) vs. 25.95 yr (SD ± 9) vs. 28.44 yr (SD ± 10.94) Gender - Female: 38 (97.44%) vs. 40 (97.56%) vs. 37 (92.5%) - Male: 1 (2.56%) vs. 1 (2.44%) vs. 3 (7.5%) Race: NR	CBT-E, MANTRA, and SSCM each resulted in improvements in weight-related outcomes with no significant differences among the treatments. BMI – Baseline: 16.59 kg/m² (SD ± 1.35) vs. 16.91 kg/m² (SD ± 1.11) vs. 16.58 kg/m² (SD ± 1.18) BMI, Change - Baseline – 10 mo: 2.1 kg/m² (SD ± 1.74) vs. 1.37 kg/m² vs. 1.58 kg/m² (SD ± 1.72) - Baseline – 22 mo: 2.35 kg/m² (SD ± 1.74) vs. 1.5 kg/m² vs. 1.9 kg/m² (SD ± 1.72) BMI > 18.5 kg/m² - Baseline->10 mo->22 mo: 2 (5.01%)->21 (54.11%)->23 (59%) vs. 1 (2.43%)->20 (48.1%)->18 (43.9%) vs. 2 (5.01%)->17 (42.37%)->19 (47.5%) Attrition: 33% (13/39) vs. 44% (18/41) vs. 43% (17/40)	Low

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM= Specialist Supportive Clinical Management; SWAN=Strong Without Anorexia Nervosa; yr=year

Compared to Treatment As Usual

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Zipfel et al. (2014) (ANTOP); Zeeck et al. (2018)	Design: RCT; Post-hoc Analysis Setting: Multi-center Country: Germany Funding: Government and non-profit	Randomized N=242 CBT-E 10 mo (N=80) FPT 10 mo (N=80) Optimized TAU 10 mo (N=82) BMI < 17.5 kg/m² Subgroup (N=53 vs.62) Follow-up: Baseline – 22 mo	Inclusion: Adult aged ≥18 years; female; AN or subsyndromal AN; BMI of 15-18.5 kg/m^2 Exclusion: Current substance abuse; use of neuroleptic drugs; psychotic disorder; bipolar disorder; serious unstable medical problems; ongoing psychotherapy	AN or AN, Subsyndromal: 242 (100%) - Restricting type: 42 (53%) vs. 46 (58%) vs. 43 (52%) - Binge-eating and purging type: 38 (48%) vs. 34 (43%) vs. 39 (48%) AN <= 6 yr: 49 (61%) vs. 49 (61%) vs. 50 (61%) AN > 6 yr: 31 (39%) vs. 31 (39%) vs. 32 (39%) Weight: 46.5 kg (SD ± 4.2) BMI 15 kg/m²-18.5 kg/m²: 242 (100%) BMI < 17.5 kg/m²: 53 (66%) vs. 62 (78%) vs. 56 (68%) BMI 17.5 kg/m²-18.5 kg/m²: 27 (34%) vs. 18 (23%) vs. 26 (32%) Age >= 18 yr: 242 (100%) Age: 27.4 yr (SD ± 7.9) vs. 28 yr (SD ± 8.6) vs. 27.7 yr (SD ± 8.1)	Weight related outcomes increased in all groups, without significant differences among groups; however, FPT was associated with significantly greater remission rate compared with TAU at follow-up: 28 (35%) vs. 11 (13%) (p=0.036). Among BMI <17.5 kg/m² subjects, a significantly greater increase was shown with CBT at the end of treatment compared with FPT: 17.5 kg/m² (N=53) vs. 16.9 kg/m² (N=62) (MD 0.6 kg/m², p=0.038) BMI – Baseline: 16.82 kg/m² (SD ± 1) vs. 16.57 kg/m² (SD ± 1) vs. 16.75 kg/m² (SD ± 1) BMI, Change - Baseline – 22 mo: 1.3 kg/m² (SD ± 1.16) vs. 1.64 kg/m² (SD ± 1.16) vs. 1.22 kg/m² (SD ± 1.17) Weight – Baseline: 46.33 kg (SD ± 3.9) vs. 46.37 kg (SD ± 4.3) vs. 46.71 kg (SD ± 4.4)	Low

				Gender, Female: 242 (100%) Race: NR	Weight, Change - Baseline – 22 mo: 4.67 kg (SD ± 6.68, N=65) vs. 4.93 kg (SD ± 5.19, N=58) vs. 1.89 kg (SD ± 7.33, N=46) Hospitalization, Duration - Baseline – 22 mo: 29.4 d (SD ± 55.3) vs. 19 d (SD ± 52.7) vs. 29.3 d (SD ± 54.2) Attrition: 19% (15/80) vs. 28% (22/80) vs. 44% (36/82)	

Abbreviations: AN=anorexia nervosa; ANTOP=Anorexia Nervosa Treatment of Outpatients; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; d=day; FPT=focal psychodynamic therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; yr=year

Compared to Family Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Nyman-Carlsson et al. (2020)	Design: RCT Setting: Outpatient Country: Sweden Funding: Non-profit	Randomized N=78 CBT 18 mo (N=38) Family + Individual Therapy 18 mo (N=40) Follow-up: Baseline – 36 mo Current Analysis (N=74) - 37 vs. 37	Inclusion: Female; 17-24 years of age; AN; BMI < 17.5; parents' participation Exclusion: Critical medical status; current suicidal thoughts and/or suicidal behavior; current alcohol or substance abuse; ongoing psychotherapeutic or psychotropic treatments	AN: 78 (100%) - Binge-purge subtype: 16 (43%) vs. 12 (32%) - Restrictive subtype: 21 (57%) vs. 25 (68%) AN Duration: 31.6 mo (SD ± 24.1) vs. 26.8 mo (SD ± 24.4) Age 17 yr-24 yr: 78 (100%) - 19.1 yr (SD ± 1.9) vs. 18.7 yr (SD ± 2.0) Gender, Female: 78 (100%) Race: NR	BMI increased significantly from baseline to post-treatment in both groups (p=0.0001) BMI - Baseline: 16.49 kg/m² (SD ± 0.8) vs. 16.54 kg/m² (SD ± 0.9) BM – Post-Treatment: 19.61kg/m² vs. 19.33 kg/m² Remission – Post-Treatment: 28 (75.7%) vs. 28 (75.7%) Both groups decreased significantly (all p=0.001) from baseline to post-treatment in eating disorder-specific symptoms and general	High

					psychological symptoms, as measured by the EDI-3, GPMC, and BDI. Attrition: 3% (1/38) vs. 8% (3/40)	

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BMI=body mass index; CBT=cognitive-behavioral therapy; EDI=Eating Disorder Inventory; GPMC=General Psychological Maladjustment Composite; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to Focal Psychodynamic Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Zipfel et al. (2014) (ANTOP); Zeeck et al. (2018)	Design: RCT; Post-hoc Analysis Setting: Multi-center Country: Germany Funding: Government and non-profit	Randomized N=242 CBT-E 10 mo (N=80) FPT 10 mo (N=80) Optimized TAU 10 mo (N=82) BMI < 17.5 kg/m² Subgroup (N=53 vs.62) Follow-up: Baseline – 22 mo	Inclusion: Adult aged ≥18 years; female; AN or subsyndromal AN; BMI of 15-18.5 kg/m^2 Exclusion: Current substance abuse; use of neuroleptic drugs; psychotic disorder; bipolar disorder; serious unstable medical problems; ongoing psychotherapy	AN or AN, Subsyndromal: 242 (100%) - Restricting type: 42 (53%) vs. 46 (58%) vs. 43 (52%) - Binge-eating and purging type: 38 (48%) vs. 34 (43%) vs. 39 (48%) AN <= 6 yr: 49 (61%) vs. 49 (61%) vs. 50 (61%) AN > 6 yr: 31 (39%) vs. 31 (39%) vs. 32 (39%) Weight: 46.5 kg (SD ± 4.2) BMI 15 kg/m²-18.5 kg/m²: 242 (100%) BMI < 17.5 kg/m²: 53 (66%) vs. 62 (78%) vs. 56 (68%)	Weight related outcomes increased in all groups, without significant differences among groups; however, FPT was associated with significantly greater remission rate compared with TAU at follow-up: 28 (35%) vs. 11 (13%) (p=0.036). Among BMI <17.5 kg/m² subjects, significantly greater increase was shown with CBT at the end of treatment compared with FPT: 17.5 kg/m² (N=53) vs. 16.9 kg/m² (N=62) (MD 0.6 kg/m², p=0.038) BMI – Baseline: 16.82 kg/m² (SD ± 1) vs. 16.57 kg/m² (SD ± 1) vs. 16.75 kg/m² (SD ± 1) BMI, Change - Baseline – 22 mo: 1.3 kg/m² (SD ± 1.16) vs. 1.64 kg/m² (SD ±	Low

				BMI 17.5 kg/m²-18.5 kg/m²: 27 (34%) vs. 18 (23%) vs. 26 (32%) Age >= 18 yr: 242 (100%) Age: 27.4 yr (SD ± 7.9) vs. 28 yr (SD ± 8.6) vs. 27.7 yr (SD ± 8.1) Gender, Female: 242 (100%) Race: NR	1.16) vs. 1.22 kg/m² (SD ± 1.17) Weight – Baseline: 46.33 kg (SD ± 3.9) vs. 46.37 kg (SD ± 4.3) vs. 46.71 kg (SD ± 4.4) Weight, Change - Baseline – 22 mo: 4.67 kg (SD ± 6.68, N=65) vs. 4.93 kg (SD ± 5.19, N=58) vs. 1.89 kg (SD ± 7.33, N=46) Hospitalization, Duration - Baseline – 22 mo: 29.4 d (SD ± 55.3) vs. 19 d (SD ± 52.7) vs. 29.3 d (SD ± 54.2) Attrition: 19% (15/80) vs. 28% (22/80) vs. 44% (36/82)	

Abbreviations: AN=anorexia nervosa; ANTOP=Anorexia Nervosa Treatment of Outpatients; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; d=day; FPT=focal psychodynamic therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; yr=year

Other Forms of Cognitive-Behavioral Therapy Compared to Cognitive-Behavioral Therapy Eating Focused

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Dalle Grave et al. (2013a)	Design: RCT Setting: Inpatient: Villa Garda Hospital Country: Italy	Randomized N=80 Focused CBT 20 wk (N=42)	Inclusion: 14-65 years of age; require inpatient treatment; AN, severe Exclusion: NR	AN, Severe: 80 (100%) AN, Duration: 4 yr vs. 5 yr Requiring Hospitalization: 80 (100%)	Both focused and complex-broad CBT were associated with improvements in weight related outcomes with no significant differences	Low

	Funding: Non-profit	Complex Broad CBT 20 wk (N=38) Follow-up: Baseline – 72 wk		Weight: 37.4 kg (SD ± 5.4, N=72) BMI: 14.3 kg/m² (SD ± 1.8, N=72) BMI < 16 kg/m²: 63 (78.8%) Age 14 yr-65 yr: 80 (100%) Age: 23.4 yr (SD ± 6.9) - 23.1 yr (SD ± 6.8) vs. 23.7 yr (SD ± 7) Gender - Female: 40 (95.2%) vs. 38 (100%) - Male: 2 (4.8%) vs. 0 (0%) Race: NR	between the treatments initially or at follow-up. Weight - Baseline: 37.4 kg (SD ± 5.6, N=37) vs. 37.4 kg (SD ± 5.4, N=35) Weight, Change - Baseline – 44 wk: 8.4 kg (SD ± 4.87, N=34) vs. 10.6 kg (SD ± 5.5, N=33) - Baseline – 72 wk: 9.1 kg (SD ± 4.99, N=34) vs. 9.6 kg (SD ± 5.09, N=34) BMI - Baseline: 14.3 kg/m² (SD ± 1.8, N=37) vs. 14.3 kg/m² (SD ± 1.8, N=35) BMI, Change - Baseline – 44 wk: 3.3 kg/m² (SD ± 1.79, N=36) vs. 4 kg/m² (SD ± 1.44, N=33) - Baseline – 72 wk: 3.6 kg/m² (SD ± 1.72, N=34) vs. 3.5 kg/m² (SD ± 1.59, N=34) Attrition: 17% (7/42) vs. 13% (5/38)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to Family Therapy With Parents in Charge

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Ball and Mitchell (2004)	Design: RCT Setting: Outpatient: Eating Disorders Unit at Prince Henry Hospital Country: Australia Funding: Non-profit	Randomized N=25 CBT 12 mo (N=13) Behavioral Family Therapy 12 mo (N=12) Follow-up: Baseline – 18 mo	Inclusion: Female; 13-23 years of age; AN; currently living with their family Exclusion: BMI < 13.5 kg/m^2; currently receiving other psychological or pharmacological treatments; comorbid physical disorder or psychiatric disorder; current drug abuse or alcohol abuse; self-harming behavior over the past 12 months; other indications for hospitalization; severe physical complications; suicidal ideation; recent history of untreated physical trauma; recent history of psychological trauma; recent history of sexual abuse	AN: 25 (100%) - Restricting type: 7 (53.8%) vs. 9 (75%) - Binge-eating and purging type: 6 (46.2%) vs. 3 (25%) History of hospitalization: 4 (30.8%) vs. 3 (25%) No history of hospitalization: 9 (69.2%) vs. 9 (75%) Age 13 yr-23 yr: 25 (100%) - 18.45 yr (SD ± 2.57) vs. 17.58 yr (SD ± 3.37) Gender, Female: 25 (100%) Race: NR	Disease response and change in BMI did not differ in individuals treated with CBT vs. behavioral family therapy. Disease Response, Good - 12 mo: 15 (60%) - 18 mo: 15 (60%) BMI - Baseline: 16.06 kg/m² (SD ± 1.58) vs. 16.45 kg/m² (SD ± 0.85) BMI, Change – - Baseline – 12 mo: 2.67 kg/m² (SD ± 1.28, N=9) vs. 2.54 kg/m² (SD ± 1.57, N=9) - Baseline – 18 mo: 2.49 kg/m² (SD ± 1.31, N=9) vs. 3.2 kg/m² (SD ± 1.55, N=9) Hospitalization - Baseline – 12 mo: 3 (16%, N=18) Attrition: 31% (4/13) vs. 25% (3/12)	High

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to Treatment As Usual

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Carter et al. (2009)	Design: Non-RCT	Total (N=88)	Inclusion: Female; AN; maintained a minimum BMI of 19.5 kg/m^2 for 2-3 weeks before participating in the	AN: 88 (100%) - Restricting type: 51 (58%)	The time to relapse was significantly longer with CBT vs. TAU ($p<0.05$) and fewer individuals who received	----------

	Setting: Single Center: Toronto General Hospital Country: Canada Funding: NR	CBT + Fluoxetine or Placebo 1 yr (N=46) Maintenance TAU 1 yr (N=42)	study; have control of binge eating and purging symptoms after completing a specialized hospital-based program Exclusion: NR	- Binge-eating and purging type: 37 (42%) AN, Duration: 5.05 yr (SD ± 3.99) vs. 6.08 yr (SD ± 6.24) BMI >= 19.5 kg/m²: 88 (100%) Completed Treatment, Hospitalization: 88 (100%) Age: 23.84 yr (SD ± 4.45) vs. 24.3 yr (SD ± 5.7) Gender, Female: 88 (100%) Race - Caucasian: 74 (84%) - Asian: 3 (3%) - West Indian: 1 (1%) - Middle Eastern: 1 (1%) Ethnicity - Afro-Caribbean: 2 (2%) - Hispanic/Latino: 2 (2%) - Unknown: 5 (6%)	CBT relapsed; however, attrition rates were high and part of the CBT group also received fluoxetine. % Relapse (BMI ≤ 17.5 for 3 mo): 24.4% vs. 50% Disease Response, Remission - Baseline – 1 yr: 30 (65%) vs. 14 (34%) Study Withdrawal, Symptom Worsening - Baseline – 1 yr: 8 (17.39%) vs. NR Attrition: 43% (20/46) vs. 29% (12/42)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; yr=year

Compared to Specialist Supportive Clinical Management

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose,	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

		duration, and follow-up				
McIntosh et al. (2005); Carter et al. (2011)	Design: RCT; Follow-up/Extension Setting: NR Country: New Zealand Funding: Government	Randomized N=56 CBT 20 wk (N=19) IPT 20 wk (N=21) Clinical Management + Supportive Psychotherapy 20 wk (N=16) Follow-up N=43 - 17 vs. 14 vs. 12 Follow-up: 6.7 yr (Mean, SD ± 1.2)	Inclusion: AN, current primary; 17-40 years of age; female; BMI 14.5-19.0 kg/m^2 Exclusion: BMI <14.5 kg/m^2; current severe major depression; psychoactive substance dependence; major medical illness; major neurological illness; developmental learning disorder; cognitive impairment; bipolar I disorder; schizophrenia; chronic, refractory course of AN	AN: 56 (100%) Weight: 46.4 kg (SD ± 3.9) BMI 14.5 kg/m²-19 kg/m²: 56 (100%) BMI: 17.3 kg/m² (SD ± 1.1) Age 17 yr-40 yr: 56 (100%) Gender, Female: 56 (100%) Race: NR	Supportive clinical management was superior to IPT whereas CBT did not differ from the other treatments in the primary outcomes of times to treatment discontinuation and % of individuals completing therapy. Weight-related outcomes did not differ among the 3 groups. Weight - End of treatment->Follow-up: 48.6->54.9 kg vs. 49->56.5 kg vs. 50.4->57.5 kg (SD ± 7.3) BMI - End of treatment->Follow-up: 18.1->20.2 kg/m² vs. 18.1->20.9 kg/m² vs. 18.8->21.3 kg/m² Hospitalization, Weight Loss or AN - Baseline – 20 wk minimum: 0 (0%) vs. 3 (14.29%) vs. 1 (6.25%) Study Withdrawal, All-Cause - Follow-up: 7 (36.84%) vs. 6 (28.57%) vs. 4 (25%) Attrition: 37% (7/19) vs. 43% (9/21) vs. 31% (5/16)	High
Touyz et al. (2013); Stiles-Shields et al. (2013)	Design: RCT; Post-hoc Analysis Setting: Outpatient, multi-center	Randomized N=63 CBT 8 mo (N=31) SSCM 8 mo (N=32)	Inclusion: AN; at least 18 years of age; female; AN, duration >=7 yr Exclusion: Presenting with a current manic episode or	AN: 63 (100%) - Restricting type: 47 (74.6%) AN, Duration >= 7 yr: 63 (100%)	In individuals with AN of >=7 yr duration, CBT and SSCM were both associated with improvements in weight and eating related outcomes without substantive	Low

	Country: Australia Funding: Government, academic, and non-profit	Follow-up: Baseline – 20 mo	psychosis; current alcohol or substance abuse; current alcohol or substance dependence; significant current medical illness; significant current neurological illness; seizure disorder; current engagement in psychotherapy and being unwilling to suspend such treatment for the duration of their participation in the study	AN, Duration: 16.6 yr (SD ± 8.5) Weight: 44.8 kg (SD ± 4.9) vs. 44.5 kg (SD ± 5.4) BMI: 16.2 kg/m² (SD ± 1.3) History of hospitalization: 0.3 per person (SD ± 0.5, N=9) vs. 0.6 per person (SD ± 1.6, N=19) Age >= 18 yr: 63 (100%) Age: 33.4 yr (SD ± 9.6) - 34.6 yr (SD ± 9) vs. 32.3 yr (SD ± 10) Gender, Female: 63 (100%) Race: NR	differences between the treatments. BMI - Baseline: 16.3 kg/m² (SD ± 1.3) vs. 16.1 kg/m² (SD ± 1.4) BMI, Change - Baseline – 20 mo: 0.7 kg/m² (SD ± 1.22) vs. 0.7 kg/m² (SD ± 1.29) Hospitalization - Baseline – 8 mo: 0.5 per person (SD ± 0.7, N=16) vs. 0.9 per person (SD ± 1.8, N=29) - 8 mo – 14 mo: 0.5 per person (SD ± 0.6, N=16) vs. 0.9 per person (SD ± 1.8, N=29) - 14 mo – 20 mo: 0.1 per person (SD ± 0.3, N=3) vs. 0.3 per person (SD ± 0.6, N=10) Mortality, All-Cause - Baseline – 20 mo: 2 (6.45%) vs. NR Attrition: 16% (5/31) vs. 9% (3/32)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; IPT=interpersonal psychotherapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM=specialist supportive clinical management; wk=week; yr=year

Compared to Other Psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose,	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and	Outcome measures, main results, and overall percent attrition	Risk of bias

		duration, and follow-up		race, and baseline clinical features (e.g., BMI)		
Gowers et al. (2007, 2010) (TOuCAN); Byford et al. (2007)	Design: RCT; Follow-up/Extension Setting: Mixed: inpatient psychiatric units; specialized outpatient; community Child and Adolescent Mental Health Service (CAMHS) Country: United Kingdom Funding: Government	Randomized N=167 Specialist Inpatient Therapy (N=57) Specialist Outpatient Therapy (N=55) General Outpatient Therapy (N=55) Follow-up: Baseline – 5 yr	Inclusion: 12-18 years of age; diagnosis of AN; food restriction with or without compensatory behaviors; weight below 85% of that expected within 1 mo of assessment; intense fear of gaining weight or under influence of weight or shape on self-evaluation; primary or secondary amenorrhea of at least 3 months, or menstruation only while on oral contraceptives Exclusion: Severe intellectual disability; severe comorbid physical conditions affecting digestion or metabolism; chronic comorbid physical conditions affecting digestion or metabolism; EDNOS	AN: 167 (100%) - Restricting type: 127 (76%) - Binge-eating and purging type: 40 (24%) AN, Duration: 13 mo - < 15 mo: 41 (72%) vs. 34 (62%) vs. 36 (65%) - > 15 mo: 13 (23%) vs. 16 (29%) vs. 18 (33%) - Unknown: 3 (5%) vs. 5 (9%) vs. 1 (2%) %EBW < 85%, In the Previous 1 mo: 167 (100%) Amenorrhea, Duration >= 3 mo or Menstruation, With OCP: 167 (100%)) Age 12 yr-18 yr: 167 (100%) Age: 14.9 yr (N=161) - 14.88 yr (SD ± 1.46) vs. 15.09 yr (SD ± 1.22) vs. 14.97 yr (SD ± 1.4) Gender - Female: 153 (92%) - Male: 14 (8%) Race: NR	Findings did not show superiority of inpatient treatment as compared to general or specialist outpatient treatment. Adherence was less in the inpatient group and protocols were not consistently followed, potentially confounding results. BMI - Baseline: 15.3 kg/m² (SD ± 1.6) vs. 15.3 kg/m² (SD ± 1.6) vs. 15.5 kg/m² (SD ± 1.6) BMI, Change - Baseline – 1 yr: NR (N=52) vs. 2.6 kg/m² (SD ± 1.57, N=52) vs. 2.8 kg/m² (SD ± 1.95, N=50) - Baseline – 2 yr: NR (N=52) vs. 3.4 kg/m² (SD ± 19.91, N=50) vs. 3.9 kg/m² (SD ± 1.95, N=48) Disease Response, Good – - 1 yr: 12 (21%, N=56) vs. 8 (15%, N=54) vs. 10 (19%, N=54) - 2 yr: 19 (36%, N=53) vs. 13 (25%, N=53) vs. 20 (37%, N=54) - 5 yr: 22 (67%, N=33) vs. 20 (57%, N=35) vs. 17 (61%, N=28)	High

					Study Withdrawal - Baseline – 2 yr: 4 (7.02%) vs. 2 (3.64%) vs. 1 (1.82%) Attrition: 51% (29/57) vs. 26% (14/55) vs. 31% (17/55)	
Lock et al. (2013)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=46 CBT 24 wk (N=23) CRT 8 wk > CBT 8 wk – 24 wk (N=23) Follow-up: Baseline – 1 yr	Inclusion: >16 years of age; AN; currently at or below 90% of mean percentile BMI for gender and height at the time of recruitment; on a stable dose of psychotropic medications for a minimum of 2 months Exclusion: Current psychotic disorder; current dependence on drugs or alcohol; previous CBT or cognitive remediation therapy for AN	AN: 46 (100%) - Binge-eating and purging type: 17 (73.91%) vs. 16 (69.57%) AN, Duration: 6.4 yr (SD ± 5.8) - 5.9 yr (SD ± 6.2) vs. 6.8 yr (SD ± 5.4) BMI, Mean Percentile <= 90 percentile: 46 (100%) BMI: 17.5 kg/m² (SD ± 1.2) Age > 16 yr: 46 (100%) Age: 22.7 yr (SD ± 5.9) - 23 yr (SD ± 6.8) vs. 22.5 yr (SD ± 4.9) Gender - Female: 20 (87%) vs. 21 (91%) - Male: 3 (13%) vs. 2 (9%) Race - Caucasian: 19 (83%) vs. 14 (61%) - Asian: 2 (9%) vs. 3 (13%) - Other: 1 (4%) vs. 3 (13%)	The group receiving initial CRT followed by CBT had comparable weight outcomes as the group that received CBT throughout, although initial attrition was greater in the CBT group. BMI – Baseline: 17.8 kg/m² (SD ± 1.1) vs. 17.1 kg/m² (SD ± 1.2) BMI, Change - Baseline – 8 wk: 0.216 kg/m² (SD ± 1.04) vs. 0.574 kg/m² (SD ± 0.91) (MD -0.358 kg/m², 95% CI -0.977 – 0.261) - Baseline – 24 wk: 0.686 kg/m² (SD ± 1.34) vs. 0.512 kg/m² (SD ± 1.39) (MD 0.174 kg/m², 95% CI -0.649 – 0.997) Study Withdrawal - Baseline – 8 wk: 8 (35%) vs. 4 (17%) Attrition: 33% (7/23) vs. 35% (8/23)	Low

				Ethnicity, Hispanic/Latino: 1 (4%) vs. 3 (13%)		
McIntosh et al. (2005); Carter et al. (2011)	Design: RCT; Follow-up/Extension Setting: NR Country: New Zealand Funding: Government	Randomized N=56 CBT 20 wk (N=19) IPT 20 wk (N=21) Clinical Management + Supportive Psychotherapy 20 wk (N=16) Follow-up N=43 - 17 vs. 14 vs. 12 Follow-up: 6.7 yr (Mean, SD ± 1.2)	Inclusion: AN, current primary; 17-40 years of age; female; BMI 14.5-19.0 kg/m^2 Exclusion: BMI <14.5 kg/m^2; current severe major depression; substance dependence; major medical illness; major neurological illness; developmental learning disorder; cognitive impairment; bipolar I disorder; schizophrenia; chronic, refractory course of AN	AN: 56 (100%) Weight: 46.4 kg (SD ± 3.9) BMI: 14.5 kg/m²-19 kg/m²: 56 (100%) BMI: 17.3 kg/m² (SD ± 1.1) Age: 17 yr-40 yr: 56 (100%) Gender, Female: 56 (100%) Race: NR	Supportive clinical management was superior to IPT whereas CBT did not differ from the other treatments in the primary outcomes of times to treatment discontinuation and % of individuals completing therapy. Weight-related outcomes did not differ among the 3 groups. Weight - End of treatment->Follow-up: 48.6->54.9 kg vs. 49->56.5 kg vs. 50.4->57.5 kg (SD ± 7.3) BMI - End of treatment->Follow-up: 18.1->20.2 kg/m² vs. 18.1->20.9 kg/m² vs. 18.8->21.3 kg/m² Hospitalization, Weight Loss or AN - Baseline – 20 wk minimum: 0 (0%) vs. 3 (14.29%) vs. 1 (6.25%) Study Withdrawal, All-Cause - Follow-up: 7 (36.84%) vs. 6 (28.57%) vs. 4 (25%) Attrition: 37% (7/19) vs. 43% (9/21) vs. 31% (5/16)	High
Pike et al. (2003)	Design: RCT Setting: Outpatient: New York State Psychiatric Institute	Randomized N=33 CBT 1 yr (N=18)	Inclusion: 18-45 years of age; AN; successfully completed inpatient hospitalization at New York State Psychiatric Institute; achievement of at least 90%	AN: 33 (100%) - Restricting type: 10 (56%) vs. 6 (40%) AN, Duration: 7.6 yr (SD ± 5.9) vs. 7.3 yr (SD ± 5.8)	The CBT group had a longer time to relapse and a lower rate of relapse than	Moderate

	Country: United States Funding: Government	Nutritional Counseling 1 yr (N=15)	of IBW for a minimum of 2 weeks; normalization of eating Exclusion: NR	%IBW >= 90%, Minimum >= 2 wk: 33 (100%) Completed Treatment, Hospitalization: 33 (100%) Age 18 yr-45 yr: 33 (100%) - 26.1 yr (SD ± 6.2) vs. 24.3 yr (SD ± 6.9) Gender, Female: 33 (100%) Race: NR	the nutritional counseling group. Disease Response - Baseline – 1 yr - Good: 8 (44%) vs. 1 (7%) - Complete Response: 3 (17%) vs. 0 (0%) Attrition: 0% (0/18) vs. 20% (3/15)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; CI=confidence interval; CRT=cognitive remediation therapy; EBW=expected body weight; EDNOS=eating disorder not otherwise specified; IBW=ideal body weight; IPT=interpersonal psychotherapy; MD=mean difference; mo=month; NR=not reported; OCP=oral contraceptive pill; RCT=randomized controlled trial; SD=standard deviation; TOuCAN=Treatment Outcome for Child and adolescent Anorexia Nervosa; wk=week; yr=year

Maudsley Model of Anorexia Nervosa Treatment for Adults Compared to Enhanced Cognitive-Behavioral Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Byrne et al. (2017) (SWAN)	Design: RCT Setting: Multi-center Country: Australia Funding: NR	Randomized N=120 CBT-E 10 mo (N=39) MANTRA 10 mo (N=41) SSCM 10 mo (N=40) Follow-up: Baseline – 22 mo	Inclusion: BMI >= 14.0 and < 18.5 kg/m^2; age >=17 years; AN Exclusion: Severe physical illness; severe mental illness; severe substance dependence; current use of atypical antipsychotics; other active psychotherapy focusing on AN; acute suicide risk	AN: 120 (100%) - Restricting type: 12 (30.77%) vs. 20 (48.78%) vs. 21 (52.5%) - Binge-eating and purging type: 27 (69.2%) vs. 21 (51.2%) vs. 19 (47.5%) AN, Duration: 4 yr (SD ± 4.81) vs. 5 yr (SD ± 5.93) vs. 2 yr (SD ± 5.19)	CBT-E, MANTRA, and SSCM each resulted in improvements in weight-related outcomes with no significant differences among the treatments. BMI – Baseline: 16.59 kg/m² (SD ± 1.35) vs. 16.91 kg/m² (SD ± 1.11) vs. 16.58 kg/m² (SD ± 1.18) BMI, Change	Low

				BMI: 16.7 kg/m² (SD ± 1.22) BMI >= 14 kg/m²-< 18.5 kg/m²: 120 (100%) Age >= 17 yr: 120 (100%) Age: 26.19 yr (SD ± 9.47) - 24.18 yr (SD ± 8) vs. 25.95 yr (SD ± 9) vs. 28.44 yr (SD ± 10.94) Gender - Female: 38 (97.44%) vs. 40 (97.56%) vs. 37 (92.5%) - Male: 1 (2.56%) vs. 1 (2.44%) vs. 3 (7.5%) Race: NR	- Baseline – 10 mo: 2.1 kg/m² (SD ± 1.74) vs. 1.37 kg/m² vs. 1.58 kg/m² (SD ± 1.72) - Baseline – 22 mo: 2.35 kg/m² (SD ± 1.74) vs. 1.5 kg/m² vs. 1.9 kg/m² (SD ± 1.72) BMI > 18.5 kg/m² - Baseline->10 mo->22 mo: 2 (5.01%)->21 (54.11%)->23 (59%) vs. 1 (2.43%)->20 (48.1%)->18 (43.9%) vs. 2 (5.01%)->17 (42.37%)->19 (47.5%) Attrition: 33% (13/39) vs. 44% (18/41) vs. 43% (17/40)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM=Specialist Supportive Clinical Management; SWAN=Strong Without Anorexia Nervosa; yr=year

Compared to Specialist Supportive Clinical Management

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Byrne et al. (2017) (SWAN)	Design: RCT Setting: Multi-center Country: Australia Funding: NR	Randomized N=120 CBT-E 10 mo (N=39) MANTRA 10 mo (N=41) SSCM 10 mo (N=40)	Inclusion: BMI >= 14.0 and < 18.5 kg/m^2; age >=17 years; AN Exclusion: Severe physical illness; severe mental illness; severe substance dependence; current use of atypical antipsychotics; other active psychotherapy	AN: 120 (100%) - Restricting type: 12 (30.77%) vs. 20 (48.78%) vs. 21 (52.5%) - Binge-eating and purging type: 27 (69.2%) vs. 21 (51.2%) vs. 19 (47.5%)	CBT-E, MANTRA, and SSCM each resulted in improvements in weight-related outcomes with no significant differences among the treatments. BMI – Baseline: 16.59 kg/m² (SD ± 1.35) vs. 16.91 kg/m²	Low

		Follow-up: Baseline – 22 mo	focusing on AN; acute suicide risk	AN, Duration: 4 yr (SD ± 4.81) vs. 5 yr (SD ± 5.93) vs. 2 yr (SD ± 5.19) BMI: 16.7 kg/m² (SD ± 1.22) BMI >= 14 kg/m²-< 18.5 kg/m²: 120 (100%) Age >= 17 yr: 120 (100%) Age: 26.19 yr (SD ± 9.47) - 24.18 yr (SD ± 8) vs. 25.95 yr (SD ± 9) vs. 28.44 yr (SD ± 10.94) Gender - Female: 38 (97.44%) vs. 40 (97.56%) vs. 37 (92.5%) - Male: 1 (2.56%) vs. 1 (2.44%) vs. 3 (7.5%) Race: NR	(SD ± 1.11) vs. 16.58 kg/m² (SD ± 1.18) BMI, Change - Baseline – 10 mo: 2.1 kg/m² (SD ± 1.74) vs. 1.37 kg/m² vs. 1.58 kg/m² (SD ± 1.72) - Baseline – 22 mo: 2.35 kg/m² (SD ± 1.74) vs. 1.5 kg/m² vs. 1.9 kg/m² (SD ± 1.72) BMI > 18.5 kg/m² - Baseline->10 mo->22 mo: 2 (5.01%)->21 (54.11%)->23 (59%) vs. 1 (2.43%)->20 (48.1%)->18 (43.9%) vs. 2 (5.01%)->17 (42.37%)->19 (47.5%) Attrition: 33% (13/39) vs. 44% (18/41) vs. 43% (17/40)	
Schmidt et al. (2012)	Design: RCT Setting: Outpatient: Eating Disorders Outpatient Service of the South London and Maudsley National Health Service Foundation Trust Country: United Kingdom Funding: Government	Randomized N=71 MANTRA 6 mo (N=34) SSCM 6 mo (N=37) Follow-up: Baseline – 12 mo	Inclusion: Aged 18 years or over; AN or EDNOS; BMI of <18.5 kg/m^2 Exclusion: Life-threatening AN requiring immediate in-patient treatment; intellectual disability; severe mental illness; severe physical illness needing treatment in its own right; psychosis; diabetes mellitus; substance dependence; pregnancy	AN or EDNOS: 71 (100%) AN - Restricting type: 14 (41.2%) vs. 11 (29.7%) - Binge-eating and purging type: 11 (32.4%) vs. 13 (35.1%) EDNOS - Restricting: 9 (26.5%) vs. 11 (29.7%) - Binge-eating and purging: 0 (0%) vs. 2 (5.4%)	MANTRA and SSCM were both associated with improvements in weight-related outcomes but there were no differences in outcomes between the 2 groups. Weight - Baseline: 44.9 kg (SD ± 5.7) vs. 43.7 kg (SD ± 4.5) Weight, Change - Baseline – 12 mo: 3.23 kg (SD ± 4.62) vs. 3.81 kg (SD ± 4.74)	Low

				AN or EDNOS, Duration: 80.6 mo (± 71.8) Age: 26.6 yr (SD ± 7.9, N=70) - 25.6 yr (SD ± 6.9) vs. 27.5 yr (SD ± 8.7, N=36) Gender - Female: 31 (91.18%) vs. 35 (94.59%) - Male: 3 (8.82%) vs. 2 (5.41%) Race and Nationality - Caucasian: 29 (85.3%) vs. 28 (75.7%) - Black or African American: 0 (0%) vs. 3 (8.1%) - Asian and British: 3 (8.7%) vs. 5 (13.5%) - Other: 2 (5.89%) vs. 1 (2.7%)	BMI - Baseline: 16.3 kg/m² (SD ± 1.3) vs. 16.4 kg/m² (SD ± 1.3) BMI, Change - Baseline – 12 mo: 1.47 kg/m² (SD ± 1.7) vs. 1.22 kg/m² (SD ± 1.83) Subjects in the MANTRA group were significantly more likely to require hospitalization but, at baseline, they were also less like to be in a partnered relationship, which may affect need for hospitalization: 7 (20.59%) vs. 0 (0%) (p=0.004) Attrition: 0% (0/34) vs. 11% (4/37) at 6 mo.; 12% (4/34) vs. 27% (10/37) at 12 mo	
Schmidt et al. (2015, 2016) (MOSAIC)	Design: RCT; Follow-up/Extension Setting: Multi-center; outpatient Country: United Kingdom Funding: Government	Randomized N=142 MANTRA +/- Dietitian Sessions +/- Carer Sessions 23 wk (Median) (N=72) SSCM +/- Dietitian Sessions +/- Carer Sessions 20 wk (Median) (N=70) BMI < 17.5 kg/m² subgroup (N=56 vs. 49)	Inclusion: AN; 18-60 years of age; BMI of 18.5 kg/m^2 or below Exclusion: Fat phobia; life-threatening AN requiring immediate inpatient treatment; insufficient knowledge of English to understand the treatment; learning disability; severe mental or physical illness which needs treatment in its own right; psychosis; diabetes mellitus; substance dependence; unstable dose of antidepressants for less than 4 weeks; received	AN: 142 (100%) AN - Restricting type: 35 (48.6%) vs. 28 (40%) - Binge-eating and purging type: 22 (30.6%) vs. 22 (31.4%) EDNOS: 15 (20.8%) vs. 20 (28.6%) AN, Duration: 8.3 yr (SD ± 7.3, N=134) - 9.3 yr (SD ± 7.9, N=67) vs. 7.2 yr (SD ± 6.5, N=67)	MANTRA and SSCM were both associated with improvements in weight-related outcomes but there were no differences in outcomes between the 2 groups, either at the end of treatment or at follow-up. Weight - Baseline: 44.8 kg (SD ± 4.5) vs. 45.4 kg (SD ± 5.4) Weight, Change - Baseline – 12 mo: 4.96 kg (SD ± 6.23) vs. 3.54 kg (SD ± 5.74)	Low

		Follow-up: Baseline – 24 mo Follow-up (N=57 vs. 47)	MANTRA in past year; receiving treatment elsewhere	Weight: 45.1 kg (SD ± 4.9) BMI <= 18.5 kg/m²: 142 (100%) BMI: 16.6 kg/m² (SD ± 1.2) Age 18 yr-60 yr: 142 (100%) Age: 26.7 yr (SD ± 7.7) - 27.5 yr (SD ± 8.1) vs. 25.9 yr (SD ± 7.1) Gender - Female: 72 (100%) vs. 67 (95.71%) - Male: 0 (0%) vs. 3 (4.29%) Race: NR	- Baseline – 24 mo: 6.02 kg (SD ± 9.86) vs. 5.93 kg (SD ± 9.6) BMI - Baseline: 16.6 kg/m² (SD ± 1.2) vs. 16.6 kg/m² (SD ± 1.3) BMI, Change - Baseline – 12 mo: 1.83 kg/m² (SD ± 3.01) vs. 1.44 kg/m² (SD ± 2.97) - BMI < 17.5 kg/m² subgroup: 1.98 kg/m² vs. 1.28 kg/m² (MD 0.7 kg/m², 95% CI -0.19 – 1.58) BMI, Change - Baseline – 24 mo: 2.25 kg/m² (SD ± 3.54) vs. 2.16 kg/m² (SD ± 3.43) - BMI < 17.5 kg/m² subgroup: 2.48 kg/m² vs. 2.04 kg/m² (MD 0.45 kg/m², 95% CI -0.154 – 0.65) Disease Response, Recovery - Baseline – 24 mo: 18 (32.15%, N=56) vs. 13 (28.3%, N=46) Mortality - Baseline – 12 mo: NR vs. 1 (1.43%) Study Withdrawal - Baseline – 24 mo: 11 (7.75%) Attrition: 25% (18/72) vs. 41% (29/70)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; CI=confidence interval; EDNOS=eating disorder not otherwise specified; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; MD=mean difference; MOSAIC=Maudsley Outpatient Study of

Treatments for Anorexia Nervosa and Related Conditions; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM=Specialist Supportive Clinical Management; SWAN=Strong Without Anorexia Nervosa; wk=week; yr=year

Specialist Supportive Clinical Management Compared to Enhanced Cognitive-Behavioral Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Byrne et al. (2017) (SWAN	Design: RCT Setting: Multi-center Country: Australia Funding: NR	Randomized N=120 CBT-E 10 mo (N=39) MANTRA 10 mo (N=41) SSCM 10 mo (N=40) Follow-up: Baseline – 22 mo	Inclusion: BMI >= 14.0 and < 18.5 kg/m^2; age >=17 years; AN Exclusion: Severe physical illness; severe mental illness; severe substance dependence; current use of atypical antipsychotics; other active psychotherapy focusing on AN; acute suicide risk	AN: 120 (100%) - Restricting type: 12 (30.77%) vs. 20 (48.78%) vs. 21 (52.5%) - Binge-eating and purging type: 27 (69.2%) vs. 21 (51.2%) vs. 19 (47.5%) AN, Duration: 4 yr (SD ± 4.81) vs. 5 yr (SD ± 5.93) vs. 2 yr (SD ± 5.19) BMI: 16.7 kg/m² (SD ± 1.22) BMI >= 14 kg/m²-< 18.5 kg/m²: 120 (100%) Age >= 17 yr: 120 (100%) Age: 26.19 yr (SD ± 9.47) - 24.18 yr (SD ± 8) vs. 25.95 yr (SD ± 9) vs. 28.44 yr (SD ± 10.94) Gender - Female: 38 (97.44%) vs. 40 (97.56%) vs. 37 (92.5%)	CBT-E, MANTRA, and SSCM each resulted in improvements in weight-related outcomes with no significant differences among the treatments. BMI – Baseline: 16.59 kg/m² (SD ± 1.35) vs. 16.91 kg/m² (SD ± 1.11) vs. 16.58 kg/m² (SD ± 1.18) BMI, Change - Baseline – 10 mo: 2.1 kg/m² (SD ± 1.74) vs. 1.37 kg/m² vs. 1.58 kg/m² (SD ± 1.72) - Baseline – 22 mo: 2.35 kg/m² (SD ± 1.74) vs. 1.5 kg/m² vs. 1.9 kg/m² (SD ± 1.72) BMI > 18.5 kg/m² - Baseline->10 mo->22 mo: 2 (5.01%)->21 (54.11%)->23 (59%) vs. 1 (2.43%)->20 (48.1%)->18 (43.9%) vs. 2 (5.01%)->17 (42.37%)->19 (47.5%)	Low

				- Male: 1 (2.56%) vs. 1 (2.44%) vs. 3 (7.5%) Race: NR	Attrition: 33% (13/39) vs. 44% (18/41) vs. 43% (17/40)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM=Specialist Supportive Clinical Management; SWAN=Strong Without Anorexia Nervosa; yr=year

Compared to Cognitive-Behavioral Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
McIntosh et al. (2005); Carter et al. (2011)	Design: RCT; Follow-up/Extension Setting: NR Country: New Zealand Funding: Government	Randomized N=56 CBT 20 wk (N=19) IPT 20 wk (N=21) Clinical Management + Supportive Psychotherapy 20 wk (N=16) Follow-up N=43 - 17 vs. 14 vs. 12 Follow-up: 6.7 yr (Mean, SD ± 1.2)	Inclusion: AN, current primary; 17-40 years of age; female; BMI 14.5-19.0 kg/m^2 Exclusion: BMI <14.5 kg/m^2; current severe major depression; substance dependence; major medical illness; major neurological illness; developmental learning disorder; cognitive impairment; bipolar I disorder; schizophrenia; chronic, refractory course of AN	AN: 56 (100%) Weight: 46.4 kg (SD ± 3.9) BMI 14.5 kg/m²-19 kg/m²: 56 (100%) BMI: 17.3 kg/m² (SD ± 1.1) Age 17 yr-40 yr: 56 (100%) Gender, Female: 56 (100%) Race: NR	Supportive clinical management was superior to IPT whereas CBT did not differ from the other treatments in the primary outcomes of times to treatment discontinuation and % of individuals completing therapy. Weight-related outcomes did not differ among the 3 groups. Weight - End of treatment->Follow-up: 48.6->54.9 kg vs. 49->56.5 kg vs. 50.4->57.5 kg (SD ± 7.3) BMI - End of treatment->Follow-up: 18.1->20.2 kg/m² vs. 18.1->20.9 kg/m² vs. 18.8->21.3 kg/m² Hospitalization, Weight Loss or AN - Baseline – 20 wk minimum: 0 (0%) vs. 3 (14.29%) vs. 1 (6.25%)	High

					Study Withdrawal, All-Cause - Follow-up: 7 (36.84%) vs. 6 (28.57%) vs. 4 (25%) Attrition: 37% (7/19) vs. 43% (9/21) vs. 31% (5/16)	
Touyz et al. (2013); Stiles-Shields et al. (2013)	Design: RCT; Post-hoc Analysis Setting: Outpatient, multi-center Country: Australia Funding: Government, academic, and non-profit	Randomized N=63 CBT 8 mo (N=31) SSCM 8 mo (N=32) Follow-up: Baseline – 20 mo	Inclusion: AN; at least 18 years of age; female; AN, duration >=7 yr Exclusion: Presenting with a current manic episode or psychosis; current alcohol or substance abuse; current alcohol or substance dependence; significant current medical illness; significant current neurological illness; seizure disorder; current engagement in psychotherapy and being unwilling to suspend such treatment for the duration of their participation in the study	AN: 63 (100%) - Restricting type: 47 (74.6%) AN, Duration >= 7 yr: 63 (100%) AN, Duration: 16.6 yr (SD ± 8.5) Weight: 44.8 kg (SD ± 4.9) vs. 44.5 kg (SD ± 5.4) BMI: 16.2 kg/m² (SD ± 1.3) History of hospitalization: 0.3 per person (SD ± 0.5, N=9) vs. 0.6 per person (SD ± 1.6, N=19) Age >= 18 yr: 63 (100%) Age: 33.4 yr (SD ± 9.6) - 34.6 yr (SD ± 9) vs. 32.3 yr (SD ± 10) Gender, Female: 63 (100%) Race: NR	In individuals with AN of >=7 yr duration, CBT and SSCM were both associated with improvements in weight and eating related outcomes without substantive differences between the treatments. BMI - Baseline: 16.3 kg/m² (SD ± 1.3) vs. 16.1 kg/m² (SD ± 1.4) BMI, Change - Baseline – 20 mo: 0.7 kg/m² (SD ± 1.22) vs. 0.7 kg/m² (SD ± 1.29) Hospitalization - Baseline – 8 mo: 0.5 per person (SD ± 0.7, N=16) vs. 0.9 per person (SD ± 1.8, N=29) - 8 mo – 14 mo: 0.5 per person (SD ± 0.6, N=16) vs. 0.9 per person (SD ± 1.8, N=29) - 14 mo – 20 mo: 0.1 per person (SD ± 0.3, N=3) vs. 0.3 per person (SD ± 0.6, N=10)	Low

					Mortality, All-Cause - Baseline – 20 mo: 2 (6.45%) vs. NR Attrition: 16% (5/31) vs. 9% (3/32)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; IPT=interpersonal psychotherapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM=specialist supportive clinical management; wk=week; yr=year

Compared to Maudsley Model of Anorexia Nervosa Treatment for Adults

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Byrne et al. (2017) (SWAN)	Design: RCT Setting: Multi-center Country: Australia Funding: NR	Randomized N=120 CBT-E 10 mo (N=39) MANTRA 10 mo (N=41) SSCM 10 mo (N=40) Follow-up: Baseline – 22 mo	Inclusion: BMI >= 14.0 and < 18.5 kg/m^2; age >=17 years; AN Exclusion: Severe physical illness; severe mental illness; severe substance dependence; current use of atypical antipsychotics; other active psychotherapy focusing on AN; acute suicide risk	AN: 120 (100%) - Restricting type: 12 (30.77%) vs. 20 (48.78%) vs. 21 (52.5%) - Binge-eating and purging type: 27 (69.2%) vs. 21 (51.2%) vs. 19 (47.5%) AN, Duration: 4 yr (SD ± 4.81) vs. 5 yr (SD ± 5.93) vs. 2 yr (SD ± 5.19) BMI: 16.7 kg/m² (SD ± 1.22) BMI >= 14 kg/m²-< 18.5 kg/m²: 120 (100%) Age >= 17 yr: 120 (100%) Age: 26.19 yr (SD ± 9.47)	CBT-E, MANTRA, and SSCM each resulted in improvements in weight-related outcomes with no significant differences among the treatments. BMI – Baseline: 16.59 kg/m² (SD ± 1.35) vs. 16.91 kg/m² (SD ± 1.11) vs. 16.58 kg/m² (SD ± 1.18) BMI, Change - Baseline – 10 mo: 2.1 kg/m² (SD ± 1.74) vs. 1.37 kg/m² vs. 1.58 kg/m² (SD ± 1.72) - Baseline – 22 mo: 2.35 kg/m² (SD ± 1.74) vs. 1.5 kg/m² vs. 1.9 kg/m² (SD ± 1.72) BMI > 18.5 kg/m² - Baseline->10 mo->22 mo: 2 (5.01%)->21 (54.11%)->23	Low

				- 24.18 yr (SD ± 8) vs. 25.95 yr (SD ± 9) vs. 28.44 yr (SD ± 10.94) Gender - Female: 38 (97.44%) vs. 40 (97.56%) vs. 37 (92.5%) - Male: 1 (2.56%) vs. 1 (2.44%) vs. 3 (7.5%) Race: NR	(59%) vs. 1 (2.43%)->20 (48.1%)->18 (43.9%) vs. 2 (5.01%)->17 (42.37%)->19 (47.5%) Attrition: 33% (13/39) vs. 44% (18/41) vs. 43% (17/40)	
Schmidt et al. (2012)	Design: RCT Setting: Outpatient: Eating Disorders Outpatient Service of the South London and Maudsley National Health Service Foundation Trust Country: United Kingdom Funding: Government	Randomized N=71 MANTRA 6 mo (N=34) SSCM 6 mo (N=37) Follow-up: Baseline – 12 mo	Inclusion: Aged 18 years or over; AN or EDNOS; BMI of <18.5 kg/m^2 Exclusion: Life-threatening AN requiring immediate in-patient treatment; intellectual disability; severe mental illness; severe physical illness needing treatment in its own right; psychosis; diabetes mellitus; substance dependence; pregnancy	AN or EDNOS: 71 (100%) AN - Restricting type: 14 (41.2%) vs. 11 (29.7%) - Binge-eating and purging type: 11 (32.4%) vs. 13 (35.1%) EDNOS - Restricting: 9 (26.5%) vs. 11 (29.7%) - Binge-eating and purging: 0 (0%) vs. 2 (5.4%) AN or EDNOS, Duration: 80.6 mo (± 71.8) Age: 26.6 yr (SD ± 7.9, N=70) - 25.6 yr (SD ± 6.9) vs. 27.5 yr (SD ± 8.7, N=36) Gender - Female: 31 (91.18%) vs. 35 (94.59%) - Male: 3 (8.82%) vs. 2 (5.41%)	MANTRA and SSCM were both associated with improvements in weight-related outcomes but there were no differences in outcomes between the 2 groups. Weight - Baseline: 44.9 kg (SD ± 5.7) vs. 43.7 kg (SD ± 4.5) Weight, Change - Baseline – 12 mo: 3.23 kg (SD ± 4.62) vs. 3.81 kg (SD ± 4.74) BMI - Baseline: 16.3 kg/m² (SD ± 1.3) vs. 16.4 kg/m² (SD ± 1.3) BMI, Change - Baseline – 12 mo: 1.47 kg/m² (SD ± 1.7) vs. 1.22 kg/m² (SD ± 1.83) Subjects in the MANTRA group were significantly more likely to require hospitalization but, at baseline, they were also	Low

				Race and Nationality - Caucasian: 29 (85.3%) vs. 28 (75.7%) - Black or African American: 0 (0%) vs. 3 (8.1%) - Asian and British: 3 (8.7%) vs. 5 (13.5%) Other: 2 (5.89%) vs. 1 (2.7%)	less like to be in a partnered relationship, which may affect need for hospitalization: 7 (20.59%) vs. 0 (0%) (p=0.004) Attrition: 0% (0/34) vs. 11% (4/37) at 6 mo.; 12% (4/34) vs. 27% (10/37) at 12 mo	
Schmidt et al. (2015, 2016) (MOSAIC)	Design: RCT; Follow-up/Extension Setting: Multi-center' outpatient Country: United Kingdom Funding: Government	Randomized N=142 MANTRA +/- Dietitian Sessions +/- Carer Sessions 23 wk (Median) (N=72) SSCM +/- Dietitian Sessions +/- Carer Sessions 20 wk (Median) (N=70) BMI < 17.5 kg/m² subgroup (N=56 vs. 49) Follow-up: Baseline – 24 mo Follow-up (N=57 vs. 47)	Inclusion: AN; 18-60 years of age; BMI of 18.5 kg/m^2 or below Exclusion: Fat phobia; life-threatening AN requiring immediate inpatient treatment; insufficient knowledge of English to understand the treatment; learning disability; severe mental or physical illness which needs treatment in its own right; psychosis; diabetes mellitus; substance dependence; unstable dose of antidepressants for less than 4 weeks; received MANTRA in past year; receiving treatment elsewhere	AN: 142 (100%) AN - Restricting type: 35 (48.6%) vs. 28 (40%) - Binge-eating and purging type: 22 (30.6%) vs. 22 (31.4%) EDNOS: 15 (20.8%) vs. 20 (28.6%) AN, Duration: 8.3 yr (SD ± 7.3, N=134) - 9.3 yr (SD ± 7.9, N=67) vs. 7.2 yr (SD ± 6.5, N=67) Weight: 45.1 kg (SD ± 4.9) BMI <= 18.5 kg/m²: 142 (100%) BMI: 16.6 kg/m² (SD ± 1.2) Age 18 yr-60 yr: 142 (100%) Age: 26.7 yr (SD ± 7.7)	MANTRA and SSCM were both associated with improvements in weight-related outcomes but there were no differences in outcomes between the 2 groups, either at the end of treatment or at follow-up. Weight - Baseline: 44.8 kg (SD ± 4.5) vs. 45.4 kg (SD ± 5.4) Weight, Change - Baseline – 12 mo: 4.96 kg (SD ± 6.23) vs. 3.54 kg (SD ± 5.74) - Baseline – 24 mo: 6.02 kg (SD ± 9.86) vs. 5.93 kg (SD ± 9.6) BMI - Baseline: 16.6 kg/m² (SD ± 1.2) vs. 16.6 kg/m² (SD ± 1.3) BMI, Change - Baseline – 12 mo: 1.83 kg/m² (SD ± 3.01) vs. 1.44 kg/m² (SD ± 2.97) - BMI < 17.5 kg/m² subgroup: 1.98 kg/m² vs. 1.28 kg/m² (MD 0.7	Low

				- 27.5 yr (SD ± 8.1) vs. 25.9 yr (SD ± 7.1) Gender - Female: 72 (100%) vs. 67 (95.71%) - Male: 0 (0%) vs. 3 (4.29%) Race: NR	kg/m², 95% CI -0.19 – 1.58) BMI, Change - Baseline – 24 mo: 2.25 kg/m² (SD ± 3.54) vs. 2.16 kg/m² (SD ± 3.43) - BMI < 17.5 kg/m² subgroup: 2.48 kg/m² vs. 2.04 kg/m² (MD 0.45 kg/m², 95% CI -0.154 – 0.65) Disease Response, Recovery - Baseline – 24 mo: 18 (32.15%, N=56) vs. 13 (28.3%, N=46) Mortality - Baseline – 12 mo: NR vs. 1 (1.43%) Study Withdrawal - Baseline – 24 mo: 11 (7.75%) Attrition: 25% (18/72) vs. 41% (29/70)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; CI=confidence interval; EDNOS=eating disorder not otherwise specified; MANTRA=Maudsley Model of Anorexia Nervosa Treatment for Adults; MD=mean difference; MOSAIC=Maudsley Outpatient Study of Treatments for Anorexia Nervosa and Related Conditions; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSCM=Specialist Supportive Clinical Management; SWAN=Strong Without Anorexia Nervosa; wk=week; yr=year

Compared to Interpersonal Psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
McIntosh et al. (2005); Carter et al. (2011)	Design: RCT; Follow-up/Extension Setting: NR	Randomized N=56 CBT 20 wk (N=19)	Inclusion: AN, current primary; 17-40 years of age; female; BMI 14.5-19.0 kg/m^2	AN: 56 (100%) Weight: 46.4 kg (SD ± 3.9)	Supportive clinical management was superior to IPT whereas CBT did not differ from the other	High

	Country: New Zealand Funding: Government	IPT 20 wk (N=21) Clinical Management + Supportive Psychotherapy 20 wk (N=16) Follow-up N=43 - 17 vs. 14 vs. 12 Follow-up: 6.7 yr (Mean, SD ± 1.2)	Exclusion: BMI <14.5 kg/m^2; current severe major depression; psychoactive substance dependence; major medical illness; major neurological illness; developmental learning disorder; cognitive impairment; bipolar I disorder; schizophrenia; chronic, refractory course of AN	BMI 14.5 kg/m²-19 kg/m²: 56 (100%) BMI: 17.3 kg/m² (SD ± 1.1) Age 17 yr-40 yr: 56 (100%) Gender, Female: 56 (100%) Race: NR	treatments in the primary outcomes of times to treatment discontinuation and % of individuals completing therapy. Weight-related outcomes did not differ among the 3 groups. Weight - End of treatment->Follow-up: 48.6->54.9 kg vs. 49->56.5 kg vs. 50.4->57.5 kg (SD ± 7.3) BMI - End of treatment->Follow-up: 18.1->20.2 kg/m² vs. 18.1->20.9 kg/m² vs. 18.8->21.3 kg/m² Hospitalization, Weight Loss or AN - Baseline – 20 wk minimum: 0 (0%) vs. 3 (14.29%) vs. 1 (6.25%) Study Withdrawal, All-Cause - Follow-up: 7 (36.84%) vs. 6 (28.57%) vs. 4 (25%) Attrition: 37% (7/19) vs. 43% (9/21) vs. 31% (5/16)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; IPT=interpersonal psychotherapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Experienced Caregivers Helping Others Compared to Treatment As Usual

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Hodsoll et al. (2017); Salerno et al. (2016)	Design: RCT Setting: Multi-center Country: United Kingdom Funding: Government	Randomized N=149 ECHO + TAU (pooled) (N=99) - Guided ECHO + TAU 12 mo (N=50) - Unguided ECHO + TAU 12 mo (N=49) TAU 12 mo (N=50)	Inclusion: 13-20 years of age; AN or atypical AN	AN or AN, Atypical: 149 (100%) - AN: 38 (76%) vs. 33 (67%) vs. 41 (82%) - AN, Atypical: 12 (24%) vs. 16 (33%) vs. 9 (18%) AN, Duration: 12 mo (SD ± 60) vs. 13 mo (SD ± 80) vs. 15 mo (SD ± 78.52) BMI: 17 kg/m² (SD ± 2.2) %ABW: 82.9% (SD ± 11.2) Age 13 yr-20 yr: 149 (100%) Age: 16.7 yr (SD ± 2.4) vs. 17.2 yr (SD ± 2) vs. 16.9 yr (SD ± 2.1) Gender - Female: 44 (88%) vs. 45 (92%) vs. 48 (96%) - Male: 6 (12%) vs. 4 (8%) vs. 2 (4%) Race - Caucasian: 46 (92%) vs. 47 (96%) vs. 47 (96%) - Asian, Biracial, or Other: 4 (8%) vs. 1 (2%) vs. 3 (6%) Ethnicity, Missing: 0 (0%) vs. 1 (2%) vs. 0 (0%)	Addition of ECHO to TAU was associated with small benefits in patient weights, increases in caregiver skill, and reductions in caregiver time; however, use of the ECHO materials was low even in those who remained in the study. BMI - Baseline->12 mo: 17->18.9 kg/m² vs. 16.9->19.7 kg/m² vs. 17.3->18.9 kg/m² %ABW - Baseline->12 mo: 82.6% (N=44)->87.1% (N=39) vs. 81% (N=46)->91.2% (N=42) vs. 83.9% (N=46) >88.6% (N=37) Hospitalization, Higher Intensity Care - Baseline – 6 mo: ECHO + TAU (pooled) vs. TAU: 12 (12%) vs. 8 (16%) Hospitalization, Higher Intensity Care - 6 mo – 12 mo: ECHO + TAU (pooled) vs. TAU: 9 (9%) vs. 4 (8%) Attrition at 12 mo follow-up: 18% (9/50) vs. 12% (6/49) vs. 26% (13/50)	Moderate

Hibbs et al. (2015); Magill et al. (2016)	Design: RCT/Follow-up/Extension of RCT Setting: Multi-center Country: United Kingdom Funding: NR	Randomized N= caregivers 268; patients 178 ECHO 24 mo (N=86) TAU 24 mo (N=92)	Inclusion: hospitalized for AN; aged 12 years or older Exclusion: Severe comorbidity at time of admission; severe learning disability at time of admission; severe physical illness at time of admission; severe psychosis at time of admission	AN: 178 (100%) AN, Duration > 3 yr – Discharge: 123 (69%, N=178) AN > 6 yr – Discharge: 83 (47%, N=178) Age >= 12 yr: 178 (100%) Age: 27 yr (N=178) Adolescent: 11 Gender, Unknown: 178 (100%) Race: NR	The ECHO and TAU groups did not show statistically significant differences in patient or caregiver outcomes although improvements were seen in both groups. BMI - Admission: 14.4 kg/m² (N=86) vs. 14.4 kg/m² (N=92) - Discharge: 17.5 kg/m² (N=82) vs. 17.5 kg/m² (N=80) - 12 mo: 17.33 kg/m² (2.90; N=65) vs. 17.05 kg/m² (SD 2.51; N=64) - 24 mo: 19.3 kg/m² (SD ± 7.46, N=61) vs. 18.6 kg/m² (SD ± 5.22, N=58) (MD 0.8 kg/m², p=0.13) Mortality, All-Cause - Baseline – 24 mo: 1 (0.75%) vs. 1 (0.75%) Attrition: 29% (25/86) vs. 37% (34/92)	Low

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BMI=body mass index; ECHO=experienced caregivers helping others; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; yr=year

Relapse Prevention

Compared to Treatment As Usual

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Fichter et al. (2012, 2013)	Design: RCT; Follow-up Setting: Multi-center Country: Germany Funding: Government	Randomized N=258 Internet-Based Relapse Prevention Program 9 mo (N=128) TAU 9 mo (N=130) Follow-up: Baseline – 18 mo Follow-up (N=210) - 92 vs. 118	Inclusion: Aged 16 years or older; female; AN or EDNOS type 1; positive course of inpatient treatment Exclusion: History of forced feeding; other serious mental impairments; other physical impairments; acute or chronic organic psychosis; acute or chronic schizophrenic psychosis; marked suicidal ideation or suicidal behavior; premature or irregular discharge from inpatient treatment; history of long inpatient stays without a clinically significant weight gain	AN or EDNOS, Type 1: 258 (100%) - Binge-eating and purging type: 56 (44.1%) vs. 63 (48.1%) Completed Treatment, Inpatient: 258 (100%) Age >= 16 yr: 258 (100%) Age: 23.8 yr (SD ± 6.5) vs. 24.1 yr (SD ± 5.6) Gender, Female: 258 (100%) Race: NR	For the groups as a whole, weight change was minimal. Those who received relapse prevention and completed the internet-based program gained more weight than those receiving TAU, but relapse prevention program adherence was low, overall. BMI - Baseline: 17.8 kg/m² (SD ± 1.4) vs. 17.7 kg/m² (SD ± 1.2) BMI, Change - Baseline – 9 mo: 0.47 kg/m² (=88) vs. 0.02 kg/m² (N=117) - Baseline – 18 mo: 0.86 kg/m² (N=88) vs. 0.61 kg/m² (N=117) Study Withdrawal, All-Cause - Baseline – 18 mo: 31 (24.2%) vs. 8 (6.2%) Attrition: 24% (31/128) vs. 6% (8/130)	Low

Abbreviations: AN=anorexia nervosa; BMI=body mass index; EDNOS=eating disorder not otherwise specified; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; yr=year

Individual Dynamic Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Dare et al. (2001)	Design: RCT Setting: Single Center: Maudsley Hospital	Randomized N= 84 Family Therapy 1 yr (N=22)	Inclusion: AN, restricting or binge-purging types; adults Exclusion: Mental or physical state considered so	AN: 84 (100%) AN, Duration: 6.3 yr (SD ± 5.9)	Responses with family therapy and focal psychoanalytic psychotherapy were better than with routine treatment.	Moderate

	Country: United Kingdom Funding: Academic and non-profit	Cognitive Analytic Therapy 7 mo (N=22) Focal Psychoanalytic Psychotherapy 1 yr (N=21) Low Contact Routine Treatment 1 yr (N=19)	dangerous as to require urgent admission to hospital; serious suicidal risk; extremely low weight; hypoglycemia; syncope; potassium less than 2.5 mMol/L; sodium less than 130 mMol/L	- 5.8 yr (SD ± 4.9) vs. 6.7 yr (SD ± 7.6) vs. 6.7 yr (SD ± 5.9) vs. 6.1 yr (SD ± 5) %ABW: 72.8% (SD ± 7.1) vs. 77.3% (SD ± 8.1) vs. 72.8% (SD ± 7.6) vs. 73.9% (SD ± 7.9) Age >= 18 yr: 84 (100%) Age: 26.3 yr (SD ± 6.7) - 26.6 yr (SD ± 7.6) vs. 27.2 yr (SD ± 7.6) vs. 26.7 yr (SD ± 6.4) vs. 24.3 yr (SD ± 4.5) Gender - Female: 20 (91%) vs. 22 (100%) vs. 21 (100%) vs. 19 (100%) - Male: Family Therapy 1 yr – 2 (9%) Race: NR	Cognitive analytic therapy had a shorter treatment duration than other groups and showed a non-significant trend to better outcomes than routine treatment. Disease Response – Baseline – 1 yr - Recovery: 3 (13.64%) vs. 3 (13.64%) vs. 3 (14.29%) vs. 0 (0%) - Significantly Improved: 5 (22.73%) vs. 3 (13.64%) vs. 4 (19.05%) vs. 1 (5.26%) - Improvement: 1 (4.55%) vs. 1 (4.55%) vs. 4 (19.05%) vs. 4 (21.05%) - Poor: 13 (59.09%) vs. 15 (68.18%) vs. 10 (47.62%) vs. 14 (73.68%) Mortality, All-Cause – Baseline – 1 yr: 0 (0%) vs. 0 (0%) vs. 0 (0%) vs. 1 (5.26%) Hospitalization – Baseline – 1 yr: 3 (13.64%) vs. 2 (9.09%) vs. 2 (9.52%) vs. 5 (26.32%) Attrition: 27% (6/22) vs. 41% (9/22) vs. 43% (9/21) vs. 32% (6/19)	
Robin et al. (1994, 1995, 1999)	Design: RCT Setting: Outpatient	Randomized N=24	Inclusion: Caucasian; adolescents; AN; female;	AN: 24 (100%)	Significantly greater BMI change was associated with BFST than with ego-oriented individual therapy,	High

	Country: United States Funding: Government	BFST (N=12) Ego-Oriented Individual Therapy (N=12) Treatment: 15.9 mo (Mean) Follow-up: 12 mo	resided at home with one or both parents Exclusion: Bulimic features	Weight: 85.4 lbs (SD ± 12.7, N=11) vs. 91 lbs (SD ± 23.1, N=11) Adolescent: 24 (100%) Age: 14.7 yr (SD ± 2.7, N=11) vs. 13.9 yr (SD ± 2.1, N=11) Gender, Female: 24 (100%) Race, Caucasian: 24 (100%)	but other outcomes did not differ. BMI Regression Analysis: Baseline to 15.9 mo (mean): 5.1 kg/m² (SD ± 1.6, N=11) vs. 2.7 kg/m² (SD ± 2.2, N=11) (MD 2.4 kg/m², p<0.01) Menstruation, Resumed – End of Treatment: 10 (89%, N=11) vs. 7 (60%, N=11) Hospitalization – 15.9 mo (Mean): 3 (27.27%) vs. 5 (45.45%) Attrition: 8% (1/12) vs. 8% (1/12)	
Treasure et al. (1995)	Design: RCT Setting: Outpatient: Eating Disorder Clinic at the Maudsley Clinic Country: United Kingdom Funding: Non-profit	Randomized N=30 Cognitive Analytical Therapy 20 wk (N=14) Educational Behavioral Therapy 20 wk (N=16) Follow-up: Baseline – 1 yr	Inclusion: AN; aged 18 years or older Exclusion: Inpatient treatment because of extreme, rapid weight loss with additional symptoms and signs of severe emaciation; proximal myopathy; marrow suppression; hypoglycemia	AN: 30 (100%) Amenorrhea, Duration: 63.1 mo (SD ± 77) vs. 50.1 mo (SD ± 60) History of Hospitalization: 3 (21.43%) vs. 6 (37.5%) Age > 18 yr: 30 (100%) Age: 24.7 yr (SD ± 5) vs. 25.3 yr (SD ± 7) Gender - Female: 29 (96.67%) - Male: 1 (3.33%) Race: NR	Both treatment groups showed similar improvement on weight-related outcomes but subjectively reported improvement was greater in the group that received cognitive analytical therapy as compared to educational behavioral therapy. Weight – Baseline: 42.9 kg (SD ± 5) vs. 42.2 kg (SD ± 4) Weight, Change – Baseline – 1 yr: 6.9 kg (SD ± 4.3) vs. 6.7 kg (SD ± 5.2	Moderate

					BMI – Baseline: 15.6 kg/m² (SD ± 2.1) vs. 15 kg/m² (SD ± 1) BMI, Change – Baseline – 1 yr: 2.9 kg/m² (SD ± 1.63) vs. 2.4 kg/m² (SD ± 2.41) Disease Response – 1 yr - Good: 6 (42%) vs. 5 (31%) - Intermediate: 5 (36%) vs. 3 (19%) - Poor: 3 (22%) vs. 8 (50%) Attrition: 29% (4/14) vs. 38% (6/16)	
Zipfel et al. (2014) (ANTOP); Zeeck et al. (2018)	Design: RCT; Post-hoc Analysis Setting: Multi-center Country: Germany Funding: Government and non-profit	Randomized N=242 CBT-E 10 mo (N=80) FPT 10 mo (N=80) Optimized TAU 10 mo (N=82) BMI < 17.5 kg/m² Subgroup (N=53 vs.62) Follow-up: Baseline – 22 mo	Inclusion: Adult aged ≥18 years; female; AN or subsyndromal AN; BMI of 15-18.5 kg/m^2 Exclusion: Current substance abuse; use of neuroleptic drugs; psychotic disorder; bipolar disorder; serious unstable medical problems; ongoing psychotherapy	AN or AN, Subsyndromal: 242 (100%) - Restricting type: 42 (53%) vs. 46 (58%) vs. 43 (52%) - Binge-eating and purging type: 38 (48%) vs. 34 (43%) vs. 39 (48%) AN <= 6 yr: 49 (61%) vs. 49 (61%) vs. 50 (61%) AN > 6 yr: 31 (39%) vs. 31 (39%) vs. 32 (39%) Weight: 46.5 kg (SD ± 4.2) BMI 15 kg/m²-18.5 kg/m²: 242 (100%)	Weight related outcomes increased in all groups, without significant differences among groups; however, FPT was associated with significantly greater remission rate compared with TAU at follow-up: 28 (35%) vs. 11 (13%) (p=0.036). Among BMI <17.5 kg/m² subjects, significantly greater increase was shown with CBT at the end of treatment compared with FPT: 17.5 kg/m² (N=53) vs. 16.9 kg/m² (N=62) (MD 0.6 kg/m², p=0.038) BMI – Baseline: 16.82 kg/m² (SD ± 1) vs. 16.57	Low

				BMI < 17.5 kg/m²: 53 (66%) vs. 62 (78%) vs. 56 (68%) BMI 17.5 kg/m²-18.5 kg/m²: 27 (34%) vs. 18 (23%) vs. 26 (32%) Age >= 18 yr: 242 (100%) Age: 27.4 yr (SD ± 7.9) vs. 28 yr (SD ± 8.6) vs. 27.7 yr (SD ± 8.1) Gender, Female: 242 (100%) Race: NR	kg/m² (SD ± 1) vs. 16.75 kg/m² (SD ± 1) BMI, Change - Baseline – 22 mo: 1.3 kg/m² (SD ± 1.16) vs. 1.64 kg/m² (SD ± 1.16) vs. 1.22 kg/m² (SD ± 1.17) Weight – Baseline: 46.33 kg (SD ± 3.9) vs. 46.37 kg (SD ± 4.3) vs. 46.71 kg (SD ± 4.4) Weight, Change - Baseline – 22 mo: 4.67 kg (SD ± 6.68, N=65) vs. 4.93 kg (SD ± 5.19, N=58) vs. 1.89 kg (SD ± 7.33, N=46) Hospitalization, Duration - Baseline – 22 mo: 29.4 d (SD ± 55.3) vs. 19 d (SD ± 52.7) vs. 29.3 d (SD ± 54.2) Attrition: 19% (15/80) vs. 28% (22/80) vs. 44% (36/82)	

Abbreviations: ABW=average body weight; AN=anorexia nervosa; ANTOP=Anorexia Nervosa Treatment of Outpatients; BFST=behavioral family systems therapy; BMI=body mass index; CBT-E=enhanced cognitive-behavioral therapy; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; wk=week; yr=year

Other Therapies

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Crisp et al. (1991);	Design: RCT; Follow-up/Extension	Randomized N=90	Inclusion: AN, severe; female; duration of AN of less than ten years; living	AN, Severe: 90 (100%)	When individual outpatient therapy was compared with no treatment at 1-yr and 2-yr	High

Gowers et al. (1994)	Setting: Inpatient and outpatient Country: NR Funding: Industry and non-profit	Group Outpatient Therapy + Dietary Counselling (N=20) Individual Outpatient Therapy + Family Therapy + Dietary Counselling (N=20) Multidisciplinary Inpatient Therapy (N=30) No Treatment (N=20) Follow-up: Baseline – 2 yr	within outpatient reach of the service Exclusion: NR	AN, Duration < 10 yr: 90 (100%) AN, Duration: 39 mo - 27.5 mo (SD ± 25.8) vs. 33.4 mo (SD ± 25.9) vs. 41 mo (SD ± 30.17) vs. 53.5 mo (SD ± 52.9) Weight: 40.2 kg (SD ± 6) vs. 40.3 kg (SD ± 3.8) vs. 40.8 kg (SD ± 6.1) vs. 41 kg (SD ± 6.1) BMI: 15.52 kg/m² (SD ± 1.44) vs. 15.84 kg/m² (SD ± 1.67) Age: 22 yr (N=90) - 19.7 yr (SD ± 2.6) vs. 21.2 yr (SD ± 5.1) vs. 23.2 yr (SD ± 4.9) vs. 21.9 yr (SD ± 4.5) Gender, Female: 90 (100%) Race: NR	follow-ups, significant Improvements were shown in weight, BMI, and %MMPW: Weight - 1 yr: 48.76 kg (SD ± 6.2) vs. 43.92 kg (SD ± 8) (MD 4.84 kg, p<0.05) - 2 yr: 52.51 kg (SD ± 8.5) vs. 46.24 kg (SD ± 8.6) (MD 6.27 kg, p<0.05) BMI - 1 yr: 18.97 kg/m² (SD ± 2) vs. 16.93 kg/m² (SD ± 2.8) (MD 2.04 kg/m², p<0.05) - 2 yr: 20.09 kg/m² (SD ± 2.8) vs. 17.83 kg/m² (SD ± 3.2) (MD 2.26 kg/m², p<0.05) %MMPW - Baseline: 74.5% (SD ± 6.9) vs. 75% (SD ± 8.5) - 1 yr: 88.9% (SD ± 11.7) vs. 79.5% (SD ± 14.1) (MD 9.4%, p<0.05) - 2 yr: 94.5% (SD ± 14) vs. 83% (SD ± 15.4) (MD 11.5%, p<0.05) Hospitalization - Baseline – 1 yr: NR vs. NR vs. 18 (60%) vs. NR Attrition: 15% (3/20) vs. 10% (2/20) vs. 40% (12/30) vs. 5% (1/20)	
Geist et al. (2000)	Design: RCT Setting: Inpatient: The Hospital for Sick Children	Randomized N=25 Family Therapy 4 mo (N=12)	Inclusion: Adolescents; AN; current weight <90% of IBW; requiring hospitalization; AN, severe;	AN: 25 (100%) - Restricting type: 19 (76%)	Both family therapy and family group psychoeducation were associated with improvements in %IBW but there was no significant difference between	Moderate

	Country: Canada Funding: Non-profit	Family Group Psychoeducation 4 mo (N=13)	self-imposed food restriction; female Exclusion: Under 12 years of age; male; older than 17.4 years; immediate suicide risk; psychotic features; Individual therapy in the community; family therapy in the community; BN; previous admissions to the inpatient eating disorder program; risk for self-harm	%IBW < 90%: 25 (100%) %IBW: 78.4% (SD ± 9.77) Weight: 41.1 kg (SD ± 7) vs. 41.1 kg (SD ± 6.3) Adolescent: 25 (100%) Age: 14.3 yr (SD ± 1.5) vs. 14.9 yr (SD ± 1.7) Gender, Female: 25 (100%) Race: NR	the treatments on %IBW or measures of eating pathology. %IBW - Baseline: 77.7% vs. 77.2% (SD ± 11.1) - 4 mo: 91.3% (SD ± 7.3) vs. 96.3% (SD ± 8.2) - Hospital discharge: 89.1% vs. 90.4% Hospitalization, Duration - Baseline – 4 mo: 46.3 d (SD ± 22.7) vs. 40.8 d (SD ± 22.2) Attrition: 0% vs. 0%	
Hall and Crisp (1987)	Design: RCT Setting: Outpatient Country: NR Funding: NR	Randomized N=30 Dietary Advice (N=15) Psychotherapy (N=15) Follow-up: Baseline – 1 yr	Inclusion: Female; AN; severe AN; 13-27 years of age; Social Classes I-III; weight <85% of matched population mean weight; amenorrhea; AN duration between 6-72 months Exclusion: Married	AN, Severe: 30 (100%) AN, Duration 6 mo-72 mo: 30 (100%) - 24.5 mo vs. 29.7 mo Amenorrhea: 30 (100%) Amenorrhea, Duration: 20.1 mo vs. 27.5 mo %MMPW < 85%: 30 (100%) Age 13 yr-27 yr: 30 (100%) - 19.57 yr vs. 19.55 yr Gender, Female: 30 (100%) Race: NR	Both groups showed improvement with treatment and changes in weight did not differ significantly between the groups, whereas psychosocial and sexual adjustment scores were higher in the psychotherapy group vs. dietary advice. Weight - Baseline->1 yr: 39.54->46 kg vs. 41->45.1 kg Weight, Desired, Change - Baseline – 1 yr: 3.5 kg vs. 7 kg Amenorrhea - 1 yr: 10 (66.67%) vs. 8 (53.33%) Hospitalization - Baseline – 1 yr: 1 (6.67%) vs. 1 (6.67%)	Moderate

					Study Withdrawal - Baseline – 1 yr: NR vs. 1 (6.67%) Attrition: 27% (4/15) vs. 7% (1/15)	
Lock et al. (2013)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=46 CBT 24 wk (N=23) CRT 8 wk > CBT 8 wk – 24 wk (N=23) Follow-up: Baseline – 1 yr	Inclusion: >16 years of age; AN; currently at or below 90% of mean percentile BMI for gender and height at the time of recruitment; on a stable dose of psychotropic medications for a minimum of 2 months Exclusion: Current psychotic disorder; current dependence on drugs or alcohol; previous CBT or cognitive remediation therapy for AN	AN: 46 (100%) - Binge-eating and purging type: 17 (73.91%) vs. 16 (69.57%) AN, Duration: 6.4 yr (SD ± 5.8) - 5.9 yr (SD ± 6.2) vs. 6.8 yr (SD ± 5.4) BMI, Mean Percentile <= 90 percentile: 46 (100%) BMI: 17.5 kg/m² (SD ± 1.2) Age > 16 yr: 46 (100%) Age: 22.7 yr (SD ± 5.9) - 23 yr (SD ± 6.8) vs. 22.5 yr (SD ± 4.9) Gender - Female: 20 (87%) vs. 21 (91%) - Male: 3 (13%) vs. 2 (9%) Race - Caucasian: 19 (83%) vs. 14 (61%) - Asian: 2 (9%) vs. 3 (13%) - Other: 1 (4%) vs. 3 (13%)	The group receiving initial CRT followed by CBT had comparable weight outcomes as the group that received CBT throughout, although initial attrition was greater in the CBT group. BMI – Baseline: 17.8 kg/m² (SD ± 1.1) vs. 17.1 kg/m² (SD ± 1.2) BMI, Change - Baseline – 8 wk: 0.216 kg/m² (SD ± 1.04) vs. 0.574 kg/m² (SD ± 0.91) (MD -0.358 kg/m², 95% CI -0.977 – 0.261) - Baseline – 24 wk: 0.686 kg/m² (SD ± 1.34) vs. 0.512 kg/m² (SD ± 1.39) (MD 0.174 kg/m², 95% CI -0.649 – 0.997) Study Withdrawal - Baseline – 8 wk: 8 (35%) vs. 4 (17%) Attrition: 33% (7/23) vs. 35% (8/23)	Low

				Ethnicity, Hispanic/Latino: 1 (4%) vs. 3 (13%)		
Lock et al. (2018)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=30 Art Therapy + FBT 9 mo (N=15) CRT + FBT 9 mo (N=15)	Inclusion: 12-18 years of age; AN; medically stable for outpatient treatment; Yale Brown Cornell Eating Disorder Scale score > 1; children's Yale Brown Obsessive Compulsive Scale score > 8; obsessive compulsive Exclusion: Associated physical illness that necessitated hospitalization; psychotic illness; other mental illness; mental illness requiring hospitalization; current dependence on drugs or alcohol; physical conditions known to influence eating or weight; scores below the normal range in the Wechsler Abbreviated Scale of Intelligence; family history of child abuse or neglect; current child abuse or neglect; diabetes mellitus; pregnancy	AN: 30 (100%) AN, Duration: 10.38 mo (SD ± 12.75) - 8.47 mo (SD ± 5.46) vs. 12.43 mo (SD ± 17.59) Age 12 yr-18 yr: 30 (100%) Age: 14.49 yr (SD ± 1.64) - 14.55 yr (SD ± 1.48) vs. 14.42 yr (SD ± 1.83) Gender - Female: 14 (93.3%) vs. 13 (86.7%) - Male: 1 (6.7%) vs. 2 (13.3%) Race - Caucasian: 9 (60%) vs. 9 (60%) - Asian: 3 (20%) vs. 2 (13.3%) - Mixed: 3 (20%) vs. 4 (26.7%) Ethnicity, Hispanic/Latino: 5 (33%) vs. 4 (26.7%)	In adolescents with AN and high levels of obsessive-compulsive features, FBT in combination with either art therapy or cognitive remediation therapy was associated with improvements in weight-related outcomes and reductions in cognitive inefficiencies. BMI – Baseline: 16.32 kg/m² (SD ± 1.2) vs. 16.37 kg/m² (SD ± 1) BMI, Change - Baseline – 9 mo: 2.1 kg/m² (SD ± 1.38, N=11) vs. 1.51 kg/m² (SD ± 0.95, N=12) (MD 0.59 kg/m², p=0.24) Percent Estimated Body Weight – Baseline: 83.17% (SD ± 4.63) vs. 83.96% (SD ± 4.04) Percent Estimated Body Weight, Change - Baseline – 9 mo: 8.77% (SD ± 6.22, N=11) vs. 6.39% (SD ± 5.1, N=12) (MD 2.38%, p=0.32) Attrition: 33% (15) vs. 13% (2/15)	High
Madden et al. (2015)*	Design: RCT Setting: Multi-center	Randomized N=82 Inpatient Medical Stabilization 21.73 d	Inclusion: Aged between 12 and 18 years; AN of less than 3 years' duration; AN	AN: 82 (100%) - Restricting type: 29 (70.73%) vs. 28 (68.29%) - Binge-eating and purging type: 12	The %EBW was greater for the weight restoration group at the end of hospitalization but not at other time points. Groups did not differ significantly in initial days of hospitalization but	Low

	Country: Australia Funding: Government	(Mean) > Outpatient FBT12 mo (N=41) Inpatient Weight Restoration 36.89 d (Mean) > Outpatient FBT 12 mo (N=41) Follow-up: Baseline – 12 mo	Exclusion: AN illness duration of more than 3 years; evidence of psychosis; mania; substance abuse; significant intercurrent medical illnesses	(29.27%) vs. 13 (31.71%) AN, Duration < 3 yr: 82 (100%) AN, Duration: 7.39 mo (SD ± 5.42) vs. 7.85 mo (SD ± 6.89) %EBW: 78.26% (SD ± 6.35) History of Hospitalization: 3 (7.32%) vs. 2 (4.88%) Age 12 yr-18 yr: 82 (100%) Age: 14.89 yr (SD ± 1.46) - 14.89 yr (SD ± 1.36) vs. 14.88 yr (SD ± 1.56) Gender - Female: 39 (95.1%) vs. 39 (95.1%) - Male: 2 (4.9%) vs. 2 (4.9%) Race - Caucasian: 31 (75.6%) vs. 37 (90%) - Asian: 7 (17.1%) vs. 3 (7.3%) Ethnicity - Other: 3 (7.3%) vs. 1 (2.4%)	significantly fewer hospitalization/rehospitalization days occurred by the end of follow-up in the medical stabilization group as compared to the weight restoration group. Hospitalization, Duration - Baseline – 12 mo - 21.73 d (SD ± 5.925, N=40) vs. 36.89 d (SD ± 17.06, N=38) (MD -15.16 d, p<0.05) Rehospitalizations - Baseline – 12 mo: 22.78 d (SD ± 41.59, N=40) vs. 27.51 d (SD ± 51.7, N=38) (MD -4.73 d, p>0.05) Hospitalization and Rehospitalizations, Duration - Baseline – 12 mo: 45.2 d vs. 65.5 d (MD -20.2 d, 95% CI -40.1 – -0.3, p=0.046) %EBW - Baseline: 77.28% (SD ± 6.67) vs. 79.25% (SD ± 5.95) - End of Treatment: 84.4% vs. 92% (MD -7.6%, 95% CI -9 – -6.1, p=0.001) Study Withdrawal, All-Cause - Hospital Discharge – 12 mo: 4 (9.76%) vs. 5 (12.2%) Attrition: 12% (5/41) vs. 20% (8/41)	
Pike et al. (2003)	Design: RCT	Randomized N=33	Inclusion: 18-45 years of age; AN; successfully completed inpatient	AN: 33 (100%) - Restricting type: 10 (56%) vs. 6 (40%)	The CBT group had a longer time to relapse and a lower	Moderate

	Setting: Outpatient: New York State Psychiatric Institute Country: United States Funding: Government	CBT 1 yr (N=18) Nutritional Counseling 1 yr (N=15)	hospitalization at New York State Psychiatric Institute; achievement of at least 90% of IBW for a minimum of 2 weeks; normalization of eating Exclusion: NR	AN, Duration: 7.6 yr (SD ± 5.9) vs. 7.3 yr (SD ± 5.8) %IBW >= 90%, Minimum >= 2 wk: 33 (100%) Completed Treatment, Hospitalization: 33 (100%) Age 18 yr-45 yr: 33 (100%) - 26.1 yr (SD ± 6.2) vs. 24.3 yr (SD ± 6.9) Gender, Female: 33 (100%) Race: NR	rate of relapse than the nutritional counseling group. Disease Response - Baseline – 1 yr - Good: 8 (44%) vs. 1 (7%) - Complete Response: 3 (17%) vs. 0 (0%) Attrition: 0% (0/18) vs. 20% (3/15)	
Treasure et al. (1995)	Design: RCT Setting: Outpatient: Eating Disorder Clinic at the Maudsley Clinic Country: United Kingdom Funding: Non-profit	Randomized N=30 Cognitive Analytical Therapy 20 wk (N=14) Educational Behavioral Therapy 20 wk (N=16) Follow-up: Baseline – 1 yr	Inclusion: AN; aged 18 years or older Exclusion: Inpatient treatment because of extreme, rapid weight loss with additional symptoms and signs of severe emaciation; proximal myopathy; marrow suppression; hypoglycemia	AN: 30 (100%) Amenorrhea, Duration: 63.1 mo (SD ± 77vs. 50.1 mo (SD ± 60) History of Hospitalization: 3 (21.43%) vs. 6 (37.5%) Age > 18 yr: 30 (100%) Age: 24.7 yr (SD ± 5) vs. 25.3 yr (SD ± 7) Gender - Female: 29 (96.67%) - Male: 1 (3.33%) Race: NR	Both treatment groups showed similar improvement on weight-related outcomes but subjectively reported improvement was greater in the group that received cognitive analytical therapy as compared to educational behavioral therapy. Weight - Baseline: 42.9 kg (SD ± 5) vs. 42.2 kg (SD ± 4) Weight, Change - Baseline – 1 yr: 6.9 kg (SD ± 4.3) vs. 6.7 kg (SD ± 5.2 BMI - Baseline: 15.6 kg/m² (SD ± 2.1) vs. 15 kg/m² (SD ± 1)	Moderate

					BMI, Change - Baseline – 1 yr: 2.9 kg/m² (SD ± 1.63) vs. 2.4 kg/m² (SD ± 2.41) Disease Response - 1 yr - Good: 6 (42%) vs. 5 (31%) - Intermediate: 5 (36%) vs. 3 (19%) - Poor: 3 (22%) vs. 8 (50%) Attrition: 29% (4/14) vs. 38% (6/16)	
Wallin et al. (2000)	Design: RCT Setting: Single Center: University Hospital of Lund Country: Sweden Funding: NR	Randomized N=26 Body Awareness Therapy + Family Therapy (N=13) Family Therapy (N=13) Treatment Duration: NR Follow-up: Baseline – 2 yr	Inclusion: Teenage; AN; female Exclusion: NR	AN: 26 (100%) AN, Duration: 11.6 mo - 15.4 mo (SD ± 15.6) vs. 8.2 mo (SD ± 3.3) BMI: 15.1 kg/m² (SD ± 1.9) vs. 15.8 kg/m² (SD ± 1.6) Age 13 yr-19 yr: 26 (100%) Gender, Female: 26 (100%) Race: NR	Addition of body awareness therapy to family therapy was not associated with any difference in weight related outcomes. %EBW – Baseline: 72.5% (SD ± 8.3) vs. 75.3% (SD ± 8.3) %EBW - 2 yr (both groups): 90.9% (p<0.0001) Recovery - Baseline – 2 yr: 8 (61.5%) vs. 9 (69.2%) Hospitalization: 4 (30.77%) vs. 4 (30.77%) Hospitalization, Duration: 54.3 d (SD ± 52.6) vs. 50 d (SD ± 61.6) Attrition: NR	High

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; CI=confidence interval; CRT=cognitive remediation therapy; d=day; EBW=expected body weight; EDNOS=eating disorder not otherwise specified; FBT=family-based treatment; IBW=ideal body weight; MD=mean difference; mo=month; %MMPW=percent mean matched-population weight; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; wk=week; yr=year

Bulimia Nervosa Studies Supporting Guideline Statements

Cognitive-Behavioral Therapy

Compared to Wait-List Control/Treatment As Usual/No Treatment

Compared to wait-list control

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1989)	Design: RCT Setting: NR Location: NR Funding: Government	Randomized N=77 CBT + Response Prevention Therapy 4 mo (N=17) CBT 4 mo (N=22) Self-Monitoring Therapy 4 mo (N=19) WLC 4 mo (N=19) Follow-up: Baseline – 10 mo Current Analysis (N=67) - 16 vs. 17 vs. 16 vs. 18	Inclusion: Female; BN Exclusion: Age below 18 years; above 65 years; concurrent pharmacological or psychological treatment for bulimia; concurrent AN, schizophrenia, bipolar affective disorder, or unipolar affective disorder; concurrent drug abuse; concurrent alcoholism; medical disorders such as significant hepatic disease; medical disorders such as renal disease or major cardiac disease; pregnancy; abnormal values of serum potassium	BN: 77 (100%) BN, Duration: 8.8 yr (± 6.6) Purging: 12.2/wk (SD ± 8.3, N=16) vs. 11.1/wk (SD ± 6, N=17) vs. 12.3/wk (SD ± 8.3, N=16) vs. 13.8/wk (SD ± 8.4, N=18) History of AN: 13 (17%) Age: 29.2 yr (SD ± 8.6) Gender, Female: 77 (100%) Race: NR	In contrast to the WLC group, all treatment groups improved. Purging, % Change, Baseline – 4 mo: -52.8% (N=16) vs. -78.2% (N=17) vs. -63.6% (N=16) vs. -8.9% (N=18) CBT was statistically superior to no treatment at 4 mo in terms of purging abstinence but differences from other groups were not significant: 31.2% (N=16) vs. 56.3% (N=17) vs. 23.5% (N=16) vs. 5.8% (N=18). Attrition: 6% (1/16) vs. 23% (5/22) vs. 16% (3/19) vs.5% (1/19)	High
Freeman et al. (1988)	Design: RCT Setting: NR Country: United Kingdom Funding: Non-profit	Randomized N=112 CBT 15 wk (N=32) Behavior Therapy 15 wk (N=30) Group Therapy 15 wk (N=30)	Inclusion: BN; women; aged 18 and over; binged at least 3 times in the previous mo; established bulimia Exclusion: History of psychotic illness	BN: 112 (100%) BN, Duration: 6 yr (SD ± 4.9) Binge Eating >= 3 episodes, In the Previous 1 mo: 112 (100%)	Active treatments were equally effective with 77% achieving binge-eating abstinence at the end of treatment. Scores on a number of eating related rating scales were also improved with some statistical differences between treatments on individual scale items.	High

		WLC 15 wk (N=20) Follow-up: Baseline – 1 yr		BN, Age at Onset: 18.2 yr (SD ± 4.6) Age >= 18 yr: 112 (100%) Age: 24.2 yr (SD ± 5.6) Gender, Female: 112 (100%) Race: NR	Binge Eating - Baseline: 6.2/wk vs. 4.6/wk vs. 6.3/wk vs. 5.7/wk - Change - Baseline – 15 wk: -4.9/wk vs. -4/wk vs. -5.5/wk vs. -2/wk Vomiting, Self-Induced - Baseline: 7.4/wk vs. 3.6/wk vs. 8.9/wk vs. 8/wk - Change - Baseline – 15 wk: -6.4/wk vs. -3.3/wk vs. -8.3/wk vs.-1.7/wk Laxative Abuse - Baseline: 6.2 tablets/wk vs. 5.1 tablets/wk vs. 14.6 tablets/wk vs. 10.4 tablets/wk - 15 wk: 1.3 tablets/wk vs. 0 tablets/wk vs. 4.3 tablets/wk vs. 13.5 tablets/wk Attrition: 34% (11/32) vs. 17% (5/30) vs. 37% (11/30) vs. 20% (4/20)	
Griffiths et al. (1994, 1996)	Design: RCT; Follow-up Setting: Outpatient Country: Australia Funding: NR	Randomized N=78 CBT 8 wk (N=23) Hypnobehavioral Treatment 8 wk (N=27) WLC 8 wk (N=28) Follow-up: Baseline – 12 mo Follow-up (N=72) 19 vs. 21 vs. 22	Inclusion: BN; female; 17-50 years of age; BMI 18-26kg/m^2; agreeable not to seek additional treatment for their eating disorder during the research Exclusion: More than 2 previous inpatient admissions for treatment of an eating disorder; concurrent pharmacological or psychological treatment; coexisting major psychiatric disorder other than a depressive state; coexisting major psychiatric disorder other than an anxiety state; coexisting major psychiatric disorder other than a personality disorder;	BN: 78 (100%) BN, Symptomatic, Duration: 6.19 yr (SD ± 5.08) - 5.4 yr (SD ± 2.31, N=19) vs. 3.31 yr (SD ± 2.99, N=21) vs. NR (N=22) BN, Objective, Symptomatic, Duration: 4.54 yr (SD ± 5.15) Bulimic Episodes, Objective: 14.18 d/mo (SD ± 7.78) Binge Eating: 3.18 d (SD ± 1.49, N=20) vs. 3.95 d (SD ±	Abstinence rates were: 10 (50%, N=20) vs. 9 (43%, N=21) vs. 1 (4.5%, N=22) for binge eating; and 8 (40%, N=20) vs. 7 (33.3%, N=21) vs. 1 (4.5%, N=22) for purging. There were no statistical differences in outcomes among the groups. 9-mo follow-up continued to show no differences in outcomes between active treatment groups. Binge Eating Episodes - Baseline: 4.73/2 wks (SD ± 2.79,	High

			physically dependent on drugs; physically dependent on alcohol; indications for hospitalization because of risk of suicide; indications for hospitalization because of risk of poor physical health	1.67, N=21) vs. 4.77 d (SD ± 1.83, N=22) Purging: 3.38 d (SD ± 2.29, N=20) vs. 3.86 d (SD ± 2.46, N=21) vs. 5.27 d (SD ± 2, N=22) Vomiting, Self-Induced: 68 (87.2%) Vomiting, Self-Induced: 15.76 d/mo (SD ± 10.4) Laxative Abuse: 32 (41%) Laxative Abuse: 4.69 d/mo (SD ± 8.67) Diuretics: 8 (11%) Exercise, Excessive: 53 (67.9%) History of AN: 20 (25.6%) BMI 18 kg/m²-26 kg/m²: 78 (100%) Age 17 yr-50 yr: 78 (100%) Age: 25.91 yr (SD ± 5.73) Gender, Female: 78 (100%) Race: NR	N=20) vs. 6.38/2 wks (SD ± 6.12, N=21) vs. 9.82/2 wks (SD ± 9.49, N=22) Binge Eating Episodes, Change - Baseline – 11 mo: -3.64/mo (SD ± 4.91, N=25) vs. -3.37/mo (SD ± 3.36, N=23) vs. NR Purging Episodes - Baseline: 6.48/2 wks (SD ± 7.43, N=20) vs. 8.55/2 wks (SD ± 9.94, N=21) vs. 11.77/2 wks (SD ± 9.87, N=22) Purging Episodes, Change - Baseline – 11 mo: -2.05/mo (SD ± 6.04, N=25) vs. -2.24/mo (SD ± 5.33, N=23) vs. NR Vomiting, Self-Induced - Baseline - 6.02 d/mo (SD ± 9.33, N=25) vs. 5.63 d/mo (SD ± 8.22, N=23) vs. NR - 9.5/30 days (SD ± 12.88, N=25) vs. 7.62/30 days (SD ± 10.43, N=23) vs. NR Vomiting, Self-Induced, Change - Baseline – 11 mo - -5.76 d/mo (SD ± 7.8, N=25) vs. -5.48 d/mo (SD ± 6.17, N=23) vs. NR - -9.15/30 days (SD ± 9.89, N=25) vs. -7.46/30 days (SD ± 7.57, N=23) vs. NR Treatment Discontinuation - Baseline – 8 wk: 5 (26.32%,	

					N=19) vs. 5 (23.81%, N=21) vs. NR Attrition: 16% (6/38) vs. 23% (9/40) vs. NR	
Leitenberg et al. (1988)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=59 Current Analysis (N=47) CBT 10 wk > 14 wk (N=12) Exposure Plus Response-Prevention Single Setting 10 wk > 14 wk (N=11) Exposure Plus Response-Prevention Multiple Setting 10 wk > 14 wk (N=12) WLC 14 wk (N=12) Follow-up: Baseline – 41 wk	Inclusion: Women; 18–45-years of age; within 80-120% of their normal weight; BN; vomited an average of 3 times a wk Exclusion: Abuse of laxatives; signs of alcoholism; signs of psychosis; signs of serious suicide risk; involved in concurrent treatment	BN: 59 (100%) BN, Duration: 6.94 yr (N=47) - 5.6 yr (SD ± 4.2) vs. 10 yr (SD ± 9.6) vs. 7.7 yr (SD ± 4.8) vs. 4.7 yr (SD ± 4.2) Vomiting 3 episodes/wk: 59 (100%) Vomiting: 12.13/wk (N=47) %EBW 80%-120%: 59 (100%) Age 18 yr-45 yr: 59 (100%) Age: 26 yr (N=47) - 25 yr (SD ± 3.4) vs. 28 yr (SD ± 10.1) vs. 27 yr (SD ± 5.7) vs. 24 yr (SD ± 5.3) Gender, Female: 59 (100%) Race: NR	At the end of treatment (which all patients completed) and at 6-mo follow-up, treatment groups improved significantly on most outcomes with minimal change in the WLC group. However, there were no statistical differences in outcomes among the groups, likely due to the sample size. Vomiting - Baseline: 8.57/wk (SD ± 4.5) vs. 13.81/wk (SD ± 8.1) vs. 10.21/wk (SD ± 8.4) vs. 16.04/wk (SD ± 8.7) Vomiting, % Change - Baseline – 17 wk: -40% vs. -73% vs. -67% vs. NR - Baseline – 41 wk: -39% vs. -62% vs. -85% (N=10) vs. NR Vomiting, Abstinence - 17 wk: 1 (8.33%) vs. 4 (36.36%) vs. 4 (33.33%) vs. 0 (0%) - 41 wk: 4 (33.33%) vs. 2 (18.18%) vs. 5 (50%, N=10) vs. 0 (0%) Overall Attrition: 20% (12/59)	High
Sundgot-Borgen et al. (2002)	Design: RCT Setting: NR	Randomized N=64 Group CBT 16 wk (N=16)	Inclusion: Normal weight; female; BN; 18-29 years of age Exclusion: History of AN; history of other psychiatric disorders; history of somatic disorders;	BN: 64 (100%)	Group CBT was superior to nutritional counseling on vomiting episodes/wk at 42 wk (3.5/wk vs. 7.06/wk, MD -3.56/wk, p<0.001) and 22 mo	Moderate

	Country: Norway Funding: NR	Nutrition Counseling Therapy 16 wk (N=17) Exercise 16 wk (N=15) WLC 16 wk (N=16) Follow-up: Baseline – 94 wk Current Analysis (N=58) 14 vs. 17 vs. 12 vs. 15	treatment for eating disorders 6 months before entering present study; use of medication	BN, Duration: 5 yr (SD ± 1.6) vs. 5 yr (SD ± 2.3) vs. 7 yr (SD ± 3.7) vs. 6 yr (SD ± 3.8) Vomiting: 8.6/wk (SD ± 4.68) vs. 8.2/wk (SD ± 4.34) vs. 7.8/wk (SD ± 3.39) vs. 5.6/wk (SD ± 3.15) Weight, Normal: 64 (100%) BMI: 20 kg/m² (SD ± 1.9) vs. 21 kg/m² (SD ± 2.1) vs. 21 kg/m² (SD ± 2) vs. 22 kg/m² (SD ± 2.5) Age 18 yr-29 yr: 64 (100%) Age: 22 yr (SD ± 2.7) vs. 22 yr (SD ± 2.9) vs. 23 yr (SD ± 2.3) vs. 23.2 yr (SD ± 3.2) Gender, Female: 64 (100%) Race: NR	(2.71/wk vs. 7.18/wk, MD -4.47/wk, p<0.001). Exercise was superior to other treatment conditions in affecting scores on specific rating scale items (e.g., body dissatisfaction, drive for thinness). Laxative Abuse Baseline: 2.3/wk (SD ± 1.8) vs. NR vs. NR vs. NR 16 wk: 2.1/wk (SD ± 1.7) vs. NR vs. 0.85/wk (SD ± 0.99) vs. NR - CBT 16 wk vs. Exercise 16 wk: MD 1.25/wk (p<0.02) 42 wk: 2.57/wk (SD ± 2.1) vs. NR vs. 0/wk (SD ± 0) vs. NR - CBT 16 wk vs. Exercise 16 wk: MD 2.57/wk (p<0.0001) 22 mo: 3.1/wk (SD ± 2.4) vs. NR vs. 0.08/wk (SD ± 0.28) vs. NR - CBT 16 wk vs. Exercise 16 wk: MD 3.02/wk (p<0.0001) Attrition: 13% (2/16) vs. 0% (1/17) vs. 20% (3/15) vs. 6% (1/16)	
Treasure et al. (1994)	Design: RCT Setting: Outpatient: Maudsley Hospital Country: United Kingdom Funding: NR	Randomized N=81 CBT 8 wk (N=21) SH Manual Therapy 8 wk (N=41) WLC 8 wk (N=19)	Inclusion: BN or atypical BN Exclusion: Severe comorbidity; diabetes mellitus; high risk of suicide; dependence on alcohol	BN or BN, Atypical: 81 (100%) BMI: 26.8 kg/m² (SD ± 7) vs. 24 kg/m² (SD ± 5.9) vs. 23.3 kg/m² (SD ± 6.7) History of AN: 1 (5%) vs. 9 (21%) vs. 6 (30%)	Rates of full remission were less in the WLC group (11%) vs. CBT (24%) or SH (22%) at the end of treatment. CBT was associated with a reduced frequency of binge eating, vomiting, and other compensatory behaviors; SH reduced the frequency of binge eating and compensatory	High

				Age: 26 yr (SD ± 6.6) vs. 25.7 yr (SD ± 5.8) vs. 26 yr (SD ± 6.7) Gender, Female: 81 (100%) Race: NR	behaviors but not vomiting. No changes were seen in the WLC group. Binge Eating, Abstinence - 8 wk: 7 (35%, N=20) vs. 11 (31%, N=35) vs. 3 (17%) Vomiting, Abstinence - 8 wk: 4 (29%, N=14) vs. 7 (24%, N=29) vs. 2 (15%, N=13) Binge Eating, Physician Assessment – Baseline->8 wk: 4->1 units vs. 3->1 units vs. 3->3 units Vomiting, Physician Assessment - Baseline->8 wk: 3->0 units vs. 3->1 units vs. 1->1 units Dietary Restraint, Physician Assessment – Baseline->8 wk: 3->1 units vs. 3->2 units vs. 3->2 units Overall Attrition: 0% (0/81)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EBW=expected body weight; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; WLC=wait-list control; yr=year

Compared to treatment as usual

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Jacobi et al. (2017)	Design: RCT	Randomized N=253	Inclusion: Women; BN; aged 17 years or older; successfully completed inpatient treatment; reduction of binge eating by at least 50%; reduction of	BN: 253 (100%) BN, Duration - Hospital Admission: 7.2 yr	CBT group had fewer vomiting episodes (46% lower; 4.3	Low

	Setting: Inpatient: Psychosomatic Hospitals Country: Germany Funding: Government	Web-Based CBT 9 mo (N=126) TAU 9 mo (N=127) Follow-up: Baseline – 18 mo Per Protocol (N=150) - 68 vs. 82	compensatory behaviors by at least 50% Exclusion: BMI below 17.5 kg/m^2 during inpatient treatment; unfit to participate in a web-based program; psychotic symptoms; acute suicidality; severe personality disorder	- 6.62 yr (SD ± 5.59) vs. 7.65 yr (SD ± 6.28) Binge Eating, Decrease >= 50%: 253 (100%) Compensatory Behavior, Decrease >= 50%: 253 (100%) Binge Eating or Compensatory Behaviors - -2 wk – Baseline: 114 (45.1%) - 58 (46%) vs. 56 (44.1%) Vomiting - -3 mo – Hospital Admission: 18.1/wk (SD ± 19.67) vs. 18.73/wk (SD ± 20.44) - -2 wk – Baseline: 0.63/wk (SD ± 1.47) vs. 0.8/wk (SD ± 2.16) History of AN: 41 (32.5%) vs. 58 (45.6%) BMI < 17.5 kg/m²: 0 (0%, N=253) Age >= 17 yr: 253 (100%) Age: 25.67 yr (SD ± 7.18) vs. 26.26 yr (SD ± 6.92) Gender, Female: 253 (100%) Race: NR	episodes/wk vs. 7.9/wk, MD -3.6/wk, p=0.003). At the end of treatment, abstinence rates did not differ (21.4% vs. 18.9%, p=0.44). Compensatory Behaviors - -3 mo – Hospital Admission: 22.57/wk (SD ± 20.31) vs. 23.39/wk (SD ± 20.13) - -2 wk – Baseline: 1.49/wk (SD ± 2.48) vs. 1.71/wk (SD ± 2.96) - 9 mo: 6.8/wk vs. 9.8/wk - 18 mo: 7.2/wk vs. 11.1/wk Rehospitalizations - Baseline – 9 mo: 6 (7.1%, N=85) vs. 2 (2.4%, N=82) - 9 mo – 18 mo: 9 (11.6%, N=77) vs. 7 (8.4%, N=83) Attrition: 37% (47/126) vs. 33% (42/127)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; wk=week; yr=year

Compared to no further treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Davis et al. (1999)	Design: RCT Setting: Single center, outpatient: Eating Disorder Outpatient Clinic of The Toronto Hospital Country: Canada Funding: Government	Randomized N=56 Group Psychoeducation > Individual CBT 16 wk (N=37) Group Psychoeducation > No Treatment 16 wk (N=19) Follow-up: Baseline – 32 wk	Inclusion: BN; female; 18-41 years of age; 85-125% of matched population mean weight; minimum 6-month duration of BN Exclusion: Ongoing psychopharmacological or psychological treatment; immediate suicidal risk; psychosis; acute medical instability; previous exposure to one of the manual-based treatments under study	BN: 56 (100%) Percent ABW, Matched-Population 85%-125%: 56 (100%) Age 18 yr-41 yr: 56 (100%) Gender, Female: 56 (100%) Race: NR	Rates of remission of binge eating, purging, and both binge eating and purging were significantly greater in the group that received CBT as compared to group psychoeducation alone (51.4%, 54.1%, and 43.2% respectively with CBT vs. 26.3%, 21.1%, and 15.8%; $p<0.05$). Improvement in the CBT group was maintained 16 weeks after stopping treatment. Study Withdrawal – Varies: 2 (5.41%) vs. 0 (0%) Overall Attrition: 21% (15/71)	High

Abbreviations: ABW=average body weight; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; wk=week; yr=year

Compared to Cognitive-Behavioral Therapy

Group compared to individual cognitive-behavioral therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Chen et al. (2003)	Design: RCT Setting: Outpatient Country: Australia Funding: Government and non-profit	Randomized N=60 Group CBT 4.5 mo (N=30) Individual CBT 4.5 mo (N=30)	Inclusion: Female; 18 years or older; BN; BMI between 19 and 27 kg/m^2 Exclusion: Currently receiving treatment for BN; suicide risk; medically compromised; unable to be present for the study; lived	BN: 60 (100%) - Non-purging type: 5 (8%) - Purging type: 55 (92%) BN, Duration: 9.6 yr (SD ± 7.26)	More individual CBT patients were abstinent from bulimic behaviors at posttreatment (0/30 group CBT vs. 6/30 with individual CBT, $p<0.01$), though, the treatments were equivalent at later follow-up times (1 vs. 5,	High

		Follow-up: Baseline – 11 mo	more than 1.5 hours away from the University of Sydney	Binge Eating, Objective: 30.12/28 days (SD ± 24.54) Vomiting, Self-Induced: 55 (92%) Vomiting: 36.54/28 days (SD ± 42.06, N=55) BMI 19 kg/m²-27 kg/m²: 60 (100%) BMI: 22.19 kg/m² (SD ± 2.81) Age >= 18 yr: 60 (100%) Age: 25.8 yr (SD ± 7.24) Gender, Female: 60 (100%) Race: NR	p=0.09 at 8 mo; 3 vs. 4, p=0.69 at 11 mo). Improvements were shown in both treatment arms on multiple outcomes including measures of binge eating, purging, and overexercising. Treatments were comparable on most measures. Binge Eating, Change - Baseline – 5 mo: -18.33/28 days (SD ± 24.02) vs. -33.97/28 days (SD ± 35.19) - Baseline – 11 mo: -18.7/28 days (SD ± 23.02) vs. -32.53/28 days (SD ± 36.23) Vomiting - Baseline: 31.2/28 days (SD ± 34.08) vs. 41.7/28 days (SD ± 48.79) Vomiting, Change - Baseline – 5 mo: -12.37/28 days (SD ± 38.27) vs. -32.97/28 days (SD ± 39.11) - Baseline – 11 mo: -20/28 days (SD ± 29.3) vs. -28.9/28 days (SD ± 38.46) Exercise, Excessive - 5 mo: 5.1/28 days (SD ± 8.97) vs. 2.53/28 days (SD ± 6.31) - 11 mo: 3.2/28 days (SD ± 7.17) vs. 2.47/28 days (SD ± 9.52) Attrition: 27% (8/30) vs. 27% (8/30)	
Katzman et al. (2010)	Design: RCT Setting: Outpatient	Randomized N=225	Inclusion: BN or EDNOS	BN or EDNOS: 225 (100%) EDNOS: 60 (26.67%)	Significant improvements were noted across outcomes for each treatment with no apparent	High

	Country: NR Funding: NR	Individual CBT 4 wk > Group CBT 12 wk (N=73) Motivational Enhancement Therapy 4 wk > Group CBT 12 wk (N=73) Motivational Enhancement Therapy 4 wk > Individual CBT 12 wk (N=79) Follow-up: Baseline – 2.5 yr	Exclusion: Pregnancy; diabetes mellitus; severe mental illness; schizophrenia; bipolar illness; severe learning disability; inability to commit to treatment from the outset; referral for assessment only	Binge Eating: 3.6 units (SD ± 1.4) - 3.6 units (SD ± 1.4) vs. 3.7 units (SD ± 1.4) vs. 3.5 units (SD ± 1.5) Vomiting: 3.4 units (SD ± 1.7) - 3.3 units (SD ± 1.6) vs. 3.7 units (SD ± 1.6) vs. 3.3 units (SD ± 1.7) Laxative Abuse: 1.8 units (SD ± 1.6) - 1.7 units (SD ± 1.3) vs. 1.8 units (SD ± 1.4) vs. 1.9 units (SD ± 1.4) Age: 29.3 yr (SD ± 7.5) 27.8 yr (SD ± 6.3) vs. 28.9 yr (SD ± 8.1) vs. 31 yr (SD ± 7.7) Gender, Female: 225 (100%) Race: NR	differences in response among them. Binge Eating, Abstinence - Baseline: 2 (5%, N=40) vs. 1 (2.7%, N=37) vs. 0 (0%, N=39) - 12 wk: 8 (40%, N=20) vs. 8 (24.2%, N=33) vs. 5 (25%, N=20) - 2.5 yr: 12 (57.2%, N=21) vs. 5 (38.5%, N=13) vs. 8 (40%, N=20) Vomiting, Abstinence - Baseline: 12 (26.7%, N=45) vs. 6 (16.7%, N=36) vs. 8 (17.8%, N=45) - 12 wk: 8 (40%, N=20) vs. 8 (24.2%, N=33) vs. 5 (25%, N=20) - 2.5 yr: 12 (57.1%, N=21) vs. 8 (38.5%, N=21) vs. 8 (40%, N=20) Laxative Abuse, Abstinence - Baseline: 22 (53.7%, N=41) vs. 17 (54.8%, N=31) vs. 26 (66.7%, N=39) - 12 wk: 14 (82.4%, N=17) vs. 20 (71.4%, N=28) vs. 13 (72.2%, N=18) - 2.5 yr: 16 (84.2%, N=19) vs. 12 (92.3%, N=13) vs. 18 (81.8%, N=22) Attrition: 32% (19/60) vs. 48% (29/61) vs. 43% (31/72)	
Nevonen and Broberg (2006)	Design: RCT	Randomized N=86 Group CBT + IPT 23 wk (N=44)	Inclusion: BN; female; 18-24 years of age; BMI > 18 kg/m² Exclusion: Current psychotic disorder; current receipt of psychopharmacologic	BN: 86 (100%) - Purging type, vomiting, self-induced: 63 (73%) - Non-purging type, exercise or non-purging type, fasting: 23 (27%)	Outcomes did not differ at the end of treatment or at 1-yr follow-up. There was a benefit for individual therapy at 2.5 years in	High

	Setting: Outpatient: Queen Silvia Children's Hospital Country: Sweden Funding: Government	Individual CBT + IPT 23 wk (N=42) Follow-up: Baseline – 2.5 yr Per Protocol (N=63) - 32 vs. 31	medication; current receipt of psychotherapy; current alcohol abuse; current drug abuse	BN, Duration: 5.1 yr (SD ± 2.9) vs. 4.5 yr (SD ± 2.8) BMI > 18 kg/m²: 86 (100%) BMI: 21.7 kg/m² (SD ± 2.1) - 21.5 kg/m² (SD ± 2.1) vs. 21.9 kg/m² (SD ± 2.1) Age 18 yr-24 yr: 86 (100%) Age: 20.7 yr - 21.1 yr (SD ± 2) vs. 20.3 yr (SD ± 2) Gender, Female: 86 (100%) Nationality, Swedish: 29 (67%) vs. 27 (65%)	terms of episodes of binge eating and compensatory behaviors but not in change from baseline rates of binge eating or compensatory behaviors. Binge Eating - Baseline-1 yr: 3.9->1.6 vs. 3.7->0.9 d/wk - 2.5 yr: 1.8 vs. 0.8 d/wk (MD 1 d/wk, p<0.05) Compensatory Behaviors - Baseline-1 yr: 2.9->1.5 vs. 3.9->0.8 d/wk - 2.5 yr: 1 d/wk vs. 0.4 d/wk (MD 0.6 d/wk, p<0.05) Rates of recovery and remission were not statistically different but ITT remission with 55% with group CBT at 2.5 years vs. 79% with individual CBT. Study Withdrawal - Baseline – 2.5 yr: 13 (29.55%) vs. 4 (9.52%) Attrition: 27% (12/44) vs. 26% (11/42)	

Abbreviations: BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; EDNOS=eating disorder not otherwise specified; IPT=interpersonal psychotherapy; ITT=intention-to-treat; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Individualized (broad) compared to manual-based (focused) cognitive-behavioral therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Ghaderi (2006)*	Design: Randomized Non-Controlled Trial Setting: Outpatient: Therapy Office Country: Sweden Funding: Non-profit	Randomized N=50 CBT Individualized (broad) 12 wk (N=24) CBT Manual-Based (focused) 12 wk (N=26) Follow-up: Baseline – 9 mo	Inclusion: Binge and purge severity criteria for BN Exclusion: Psychotropic medication; current psychosocial treatments for eating disorders; current AN; BMI < 18; younger than 18 years of age; pregnancy; substance abuse; obstacles for committing to the study; very severe and recurring depression; risk of suicide; psychotic disorders; bipolar disorders; BDI>45	BN, Purging Type: 50 (100%) BN, Duration: 9.2 yr (SD ± 9.2) - 10.9 yr (SD ± 6.5) vs. 7.6 yr (SD ± 5.8) AN: 0 (0%, N=50) BMI < 18 kg/m²: 0 (0%, N=50) BMI: 25 kg/m² (SD ± 5.1) - 25.9 kg/m² (SD ± 4.2) vs. 24.2 kg/m² (SD ± 5.7) Age < 18 yr: 0 (0%, N=50) Age: 27.2 yr (SD ± 7.8) Gender, Unknown: 50 (100%) Race: NR	Both groups showed comparable improvement. Those in the broad CBT group had somewhat greater decreases in binge eating at the end of treatment (18->0.6/28 days vs. 12->1.5 days) but not at the follow-up assessment (1.3/28 days vs. 1.5 days). Response or remission was seen in 92% (N=22) with broad CBT and 69% (N=18) with focused CBT at the end of treatment. Vomiting, Self-Induced - Baseline: 15/28 days (SD ± 20.4) vs. 12.8/28 days (SD ± 17.6) - 12 wk: 2.5/28 days (SD ± 7) vs. 2.9/28 days (SD ± 4.7) - 9 mo: 6.8/28 days (SD ± 20.2) vs. 3.1/28 days (SD ± 5.6) Exercise, Excessive - Baseline: 6.1/28 days (SD ± 8.7) vs. 11/28 days (SD ± 10.3) - 12 wk: 0.6/28 days (SD ± 2) vs. 3.1/28 days (SD ± 4.7) - 9 mo: 0/28 days (SD ± 0) vs. 2.4/28 days (SD ± 4.5) (MD -2.4/28 days, p=0.005) Overall Attrition: 4% (2/50)	High
Thompson-Brenner et al. (2016)*	Design: RCT Setting: Single Center: Center for Anxiety and	Randomized N=50 Broad CBT-E 20 wk (N=25)	Inclusion: BN; >=8 binge/purge episodes in the 28 days prior to intake; score of >=5 on the Diagnostic Interview for Borderlines-Revised; current clinical levels of borderline	BN: 50 (100%) - Purging type: 23 (92%) vs. 24 (96%) - Non-purging type: 2 (8%) vs. 1 (4%)	Both treatments were associated with improvement but there were no significant differences between the 2 interventions.	Low

	Related Disorders (CARD) Country: United States Funding: Government	Focused CBT-E 20 wk (N=25) Follow-up: Baseline – 11 mo	personality disorder; recent history of clinical affective problems; diagnosis of at least one mood or anxiety disorder episode in the past two years; female; 18-65 years of age Exclusion: Present serious suicide risk; current substance dependence; schizophrenia precluding CBT; bipolar I disorder precluding CBT; cognitive dysfunction precluding CBT	Binge Eating and Purging >= 8 episodes, In the Previous 28 d: 50 (100%) History of AN: 3 (12%) vs. 6 (24%) BMI: 23.65 kg/m² (SD ± 3.52) Age 18 yr-65 yr: 50 (100%) Age: 25.63 yr (SD ± 8.13) - 25.75 yr (SD ± 8.15) vs. 25.52 yr (SD ± 8.28) Gender, Female: 50 (100%) Race - Caucasian: 41 (82%) - Asian: 4 (8%) - Black or African American: 1 (2%) - Native American: 1 (2%) Ethnicity, Hispanic/Latino: 3 (6%)	Binge Eating, Objective - Baseline->20 wk->11 mo - 27.84->9.55->7.4/mo vs. 28.04->8->8.58/mo Purging - Baseline->20 wk->11 mo - 31.8->8.8->7.8/mo vs. 37.88->8.24->14.5/mo Binge Eating, Objective and Purging, Remission - 20 wk: 10 (40%) vs. 11 (44%) - 11 mo: 7 (46.7%, N=15) vs. 7 (36.8%, N=19) Attrition: 32% (8/25) vs. 16% (4/25)	

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BMI=body mass index; BN=bulimia nervosa; CBT-E=enhanced cognitive-behavioral therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Group cognitive-behavioral psychotherapy high/low intensity compared to high/low abstinence

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Mitchell et al. (1993)*; Crosby et al. (1993)*	Design: RCT; Post-hoc Analysis Setting: Single Center: Eating Disorders Clinic at the University of Minnesota	Randomized N=143 Cognitive-Behavioral Group Psychotherapy HIHE 12 wk (N=33)	Inclusion: BN; binge eating; minimum age 18; female; minimum of 85% IBW; self-induced vomiting and/or laxative abuse Exclusion: Currently receiving pharmacotherapy or	BN: 143 (100%) Binge Eating: 143 (100%) BN, Duration: 8.8 yr (SD ± 5.7) vs. 8.6 yr (SD ± 6.1) vs.	At 12 wk, treatment with LILE was associated with lower rates of abstinence overall (20.6% vs. 63.6-68.3%), binge-eating abstinence (32.4% vs. 69.7-73.2%), or vomiting abstinence (29.4% vs. 70.7-76.5%).	High

	Country: United States Funding: NR	Cognitive-Behavioral Group Psychotherapy HILE 12 wk (N=35) Cognitive-Behavioral Group Psychotherapy LIHE 12 wk (N=41) Cognitive-Behavioral Group Psychotherapy LILE 12 wk (N=34)	psychotherapy for BN; currently receiving pharmacotherapy or psychotherapy for any other psychiatric condition; concurrent medical or psychiatric condition that would preclude safe outpatient therapy; diagnosed as having bipolar affective disorder or schizophrenia; actively abusing drugs; actively abusing alcohol	7.8 yr (SD ± 5) vs. 9.1 yr (SD ± 7.6) Binge Eating: 9.02/wk (SD ± 5.43) vs. 10.3/wk (SD ± 6.97) vs. 8.24/wk (SD ± 5.84) vs. 8.66/wk (SD ± 4.76) Vomiting, Self-Induced or Laxative Abuse: 143 (100%) Vomiting: 9.41/wk (SD ± 7.06) vs. 10.8/wk (SD ± 9.19) vs. 10.6/wk (SD ± 8.34) vs. 9.63/wk (SD ± 7.15) %IBW >= 85%: 143 (100%) Age >= 18 yr: 143 (100%) Age: 25.9 yr - 25.8 yr (SD ± 6.8) vs. 26.4 yr (SD ± 5.7) vs. 25.6 yr (SD ± 6) vs. 25.7 yr (SD ± 6.8) Gender, Female: 143 (100%) Race: NR	HIHE had equal or better outcomes than the other treatments. High intensity treatment groups had lower relapse rates after achieving abstinence than low intensity groups. Eating Disorder, Abstinence - 12 wk: 21 (63.6%) vs. 24 (67.6%) vs. 28 (68.3%) vs. 7 (20.6%) Binge Eating, Abstinence - 12 wk: 23 (69.7%) vs. 25 (70.6%) vs. 30 (73.2%) vs. 11 (32.4%) Vomiting, Abstinence - 12 wk: 24 (72.7%) vs. 27 (76.5%) vs. 29 (70.7%) vs. 10 (29.4%) Attrition: 12% (4/33) vs. 17% (5/35) vs. 12% (5/41) vs. 18% (6/34)	

Abbreviations: BN=bulimia nervosa; IBW=ideal body weight; HIHE=High Intensity+High Emphasis on Early Abstinence; HILE=High Intensity+Low Emphasis on Early Abstinence; LIHE=Low Intensity+High Emphasis on Early Abstinence; LILE=Low Intensity+Low Emphasis on Early Abstinence; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Manual-based compared to stepped care

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Mitchell et al. (2011)	Design: RCT Setting: Multi-center Country: NR Funding: Government, industry, and non-profit	Randomized N=293 Manual-based Individual CBT > Fluoxetine (for non-responders) 18 wk (N=147) Stepped Care (Supervised SH > Fluoxetine for non-responders > CBT for non-responders) 18 wk (N=146) Follow-up: Baseline – 1 yr Follow-up (N=197) - 103 vs. 94	Inclusion: Purging or non-purging BN; 18 years or older Exclusion: Current active psychotherapy for their eating disorder; alcohol or drug misuse in the previous 6 months; alcohol or drug dependence in the previous 6 months; acute suicidal risk; medical illness that would preclude safe study participation; history of psychotic disorder	BN: 293 (100%) - Purging type: 280 (96%) - Non-purging type: 13 (4%) Binge Eating, Objective: 25/28 days (SD ± 16.296, N=228) Compensatory Behaviors: 43/28 days (SD ± 26.667, N=228) BMI: 23.3 kg/m² (SD ± 4.9, N=228) - 23.3 kg/m² (SD ± 4.5) vs. 23.5 kg/m² (SD ± 5.3) Age >= 18 yr: 293 (100%) Age: 29.7 yr - 29.5 yr (SD ± 8) vs. 29.8 yr (SD ± 9.8) Gender, Unknown: 293 (100%) Race, Black, African American, or Native American or Ethnicity, Hispanic/Latino: 32 (14%, N=228) - 24 (16%) vs. 18 (12%)	The treatments had similar responses at 18 wk but stepped care was more effective at the 1-yr post-treatment follow-up in reducing binge eating (MD 7/28 days, p<0.05) and compensatory behaviors (MD 10/28 days, p<0.05) in terms of episodes/28 days. Binge Eating, Objective - Baseline->18 wk->1 yr - 27->4->10/28 days vs. 27->8->3/28 days Compensatory Behaviors - Baseline->18 wk->1 yr - 44->12->15/28 days vs. 43->19->5/28 days Binge Eating and Purging, Abstinence - 18 wk: 22 (15%) vs. 16 (11%) - 1 yr: 26 (18%) vs. 38 (26%) Attrition: 23% (24/147) vs. 29% (42/146)	High

Abbreviations: BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; yr=year

Compared to cognitive-behavioral therapy with guided self-help

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Thiels et al. (1998, 2000, 2003)	Design: RCT; Follow-up/Extension Setting: Outpatient Country: Germany Funding: Government and academic	Randomized N=62 CBT 16 wk (N=31) GSH + CBT 16 wk (N=31) Follow-up: Baseline – 4 yr Follow-up (N=26) - 13 vs. 13	Inclusion: BN; aged 15 years or older Exclusion: NR	BN: 62 (100%) BN, Duration: 8.5 yr (SD ± 9.2, N=30) vs. 6.1 yr (SD ± 5.6) History of AN: 15 (48.39%) vs. 13 (41.94%) AN, Concurrent: 1 (3.23%) vs. 0 (0%) BMI: 21.95 kg/m² (SD ± 3.56) - 21.31 kg/m² (SD ± 3.11, N=30) vs. 22.57 kg/m² (SD ± 3.89) Age >= 15 yr: 62 (100%) Age: 28.7 yr (SD ± 9.1) vs. 27.5 yr (SD ± 6.9) Gender, Unknown: 62 (100%) Race: NR	Groups were comparable at baseline except GSH had more subjects with prior treatment for psychiatric diagnoses. Both treatments led to significant improvements in outcomes that continued to the follow-up assessment. There was no difference in abstinence from binge eating or vomiting at follow-up. For longer term follow-up, 45% of the original sample were located at an average of 54.2 months of follow-up. Both groups showed comparable rates of abstinence from binge eating, vomiting, or using laxatives. Attrition: 13% (4/31) vs. 29% (9/31)	High

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; GSH=guided self-help; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

In-person compared to web group

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Zerwas et al. (2017); Watson et al. (2017)	Design: RCT; Secondary Analysis Setting: Multi-center	Randomized N=196 Group CBT 12 wk > 20 wk (face-to-face) (N=98)	Inclusion: BN; age 18 years or older; BMI ≥ 18.5 kg/m^2; English speaking; private access to the Internet Exclusion: Major medical condition that would interfere	BN: 196 (100%) BN, Duration: 9.5 yr (SD ± 8.8, N=90) vs. 9.5 yr (SD ± 8.9, N=89)	The percent with abstinence increased from baseline to the end of treatment and from the end of treatment to follow-up in both groups. Face-to-face was superior to online chat at the end of treatment but not at	Low

	Country: United States Funding: Government and non-profit	Group CBT 12 wk > 20 wk (online-chat) (N=98) Follow-up: Baseline – 12 mo Current Analysis (N=179) - 90 vs. 89	with treatment; type 1 diabetes mellitus; alcohol or drug dependence in the last 3 months; psychosis; schizophrenia; bipolar I disorder; current significant suicidal ideation reported during the clinical assessment	BN, Age of Onset: 18.3 yr (SD ± 5.4, N=90) vs. 18.6 yr (SD ± 5.6, N=89) BMI >= 18.5 kg/m²: 196 (100%) BMI: 24.2 kg/m² (SD ± 4.7, N=90) vs. 24.1 kg/m² (SD ± 5.7, N=89) Age >= 18 yr: 196 (100%) Age: 27.5 yr (SD ± 9.1, N=90) vs. 28.5 yr (SD ± 9.3, N=89) Gender - Female: 88 (98%, N=90) vs.87 (98%, N=89) - Male: 2 (2%, N=90) vs. 2 (2%, N=89) Race - Caucasian: 77 (86%, N=90) vs. 75 (84%, N=89) - Black or African American: 5 (6%, N=90) vs. 6 (7%, N=89) - Asian: 1 (1%, N=90) vs. 4 (4%, N=89) - Native Hawaiian/Pacific Islander: 1 (1%, N=90) vs. 0 (0%, N=89) - Other: 6 (7%, N=90) vs. 4 (4%, N=89) Ethnicity, Hispanic/Latino: 4 (4.44%, N=90) vs. 4 (4.49%, N=89)	follow-up: 21% (N=90) vs.14% (N=89) at 20 wk; 26% (N=90) vs. 30% (N=89) at 12 mo. Binge Eating - Baseline: 24.3/28 days (SD ± 17.1, N=90) vs. 27.8/28 days (SD ± 22.5, N=89) Binge Eating, % Change - Baseline – 20 wk: -54% (SD ± 95.8, N=90) vs. -56.6% (SD ± 55.5, N=89) - Baseline – 12 mo: -50.1% (SD ± 134, N=90) vs. -59.4% (SD ± 60.1, N=89) Purging - Baseline: 26.8/28 days (SD ± 20.7, N=90) vs. 31.7/28 days (SD ± 34.2, N=89) Purging, % Change - Baseline – 20 wk: -54% (SD ± 95.8, N=90) vs. -56.6% (SD ± 55.5, N=89) - Baseline – 12 mo: -50.1% (SD ± 133.9, N=90) vs. -59.4% (SD ± 60.1, N=89) Adverse Events - Baseline – 12 mo: 3 (3.33%, N=90) vs. NR Study Withdrawal, Lost to Follow-Up - Baseline – 20 wk: 26 (26.53%) vs. 38 (38.78%) - Baseline – 12 mo: 48 (48.98%) vs. 40 (40.82%) Attrition: 57% (51/90) vs. 61% (54/89)	

Abbreviations: BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to Guided Self-Help/Self-Help

Compared to guided self-help

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Bailer et al. (2004)	Design: RCT Setting: Outpatient: Department of General Psychiatry at the University Hospital of Psychiatry in Vienna Country: Austria Funding: Government	Randomized N=81 Group CBT 18 wk (N=41) GSH Therapy 18 wk (N=40) Follow-up: Baseline – 70 wk Follow-up (N=55) - 30 vs. 25	Inclusion: BN; aged 17 years and older Exclusion: Medically unstable; at severe suicide risk	BN: 81 (100%) BN, Age at Onset: 17.7 yr (SD ± 3.2) vs. 17.3 yr (SD ± 2.3) History of AN: 17 (41.4%) vs. 9 (22.5%) BMI: 20.69 kg/m² (SD ± 2.44) vs. 21.68 kg/m² (SD ± 3.15) Age >= 17 yr: 81 (100%) Age: 24.2 yr (SD ± 4.9) vs. 23.3 yr (SD ± 4.1) Gender, Unknown: 81 (100%) Race: NR	Both treatments reduced binge-eating and vomiting frequencies. Binge Eating - Baseline: 27.95/mo (SD ± 29.66) vs. 26.15/mo (SD ± 21.51) Binge Eating, Change - Baseline – 18 wk: -11.64/mo (SD ± 21.38, N=26) vs. -18.48/mo (SD ± 16.49, N=30) - Baseline – 70 wk: -14.84/mo (SD ± 21.2, N=26) vs. -18.61/mo (SD ± 15.48, N=30) Vomiting - Baseline: 30.38/mo (SD ± 32.85) vs. 21.18/mo (SD ± 22.79) Vomiting, Change - Baseline – 18 wk: -14.88/mo (SD ± 23.48, N=26) vs. -15.18/mo (SD ± 18.54, N=30) - Baseline – 70 wk: -18.49/mo (SD ± 23.47, N=30) vs. -16.56/mo (SD ± 16.51, N=25) Improvement was sustained at follow-up (36.6% remission with CBT vs. 50% for GSH), though, study completers (per protocol	High

					N=25 vs. N=23) showed higher remission rates with GSH vs. CBT (74% vs. 44%, p=0.035). Attrition: 37% (15/41) vs. 25% (10/40)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; GSH=guided self-help; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to cognitive-behavioral therapy with guided self-help

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Thiels et al. (1998, 2000, 2003)	Design: RCT; Follow-up/Extension Setting: Outpatient Country: Germany Funding: Government and academic	Randomized N=62 CBT 16 wk (N=31) GSH + CBT 16 wk (N=31) Follow-up: Baseline – 4 yr Follow-up (N=26) - 13 vs. 13	Inclusion: BN; aged 15 years or older Exclusion: NR	BN: 62 (100%) BN, Duration: 8.5 yr (SD ± 9.2, N=30) vs. 6.1 yr (SD ± 5.6) History of AN: 15 (48.39%) vs. 13 (41.94%) AN, Concurrent: 1 (3.23%) vs. 0 (0%) BMI: 21.95 kg/m² (SD ± 3.56) - 21.31 kg/m² (SD ± 3.11, N=30) vs. 22.57 kg/m² (SD ± 3.89) Age >= 15 yr: 62 (100%) Age: 28.7 yr (SD ± 9.1) vs. 27.5 yr (SD ± 6.9) Gender, Unknown: 62 (100%) Race: NR	Groups were comparable at baseline except GSH had more subjects with prior treatment for psychiatric diagnoses. Both treatments led to significant improvements in outcomes that continued to the follow-up assessment. There was no difference in abstinence from binge eating or vomiting at follow-up. For longer term follow-up, 45% of the original sample were located at an average of 54.2 months of follow-up. Both groups showed comparable rates of abstinence from binge eating, vomiting, or using laxatives. Attrition: 13% (4/31) vs. 29% (9/31)	High

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; GSH=guided self-help; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to group self-help and nutritional counselling

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Hsu et al. (2001)	Design: RCT Setting: Outpatient Country: United States Funding: Government and academic	Randomized N=100 CT 14 wk (N=26) CT + Nutritional Counselling 14 wk (N=27) Nutritional Counselling 14 wk (N=23) SH Support Group Therapy 14 wk (N=24) Follow-up: Baseline – 3.5 yr	Inclusion: BN; female; body weight within 85 to 125% IBW; 17-45 years of age; binge eating on average at least 3 times a wk in previous 6 months; vomiting on average at least 3 times a wk in previous 6 months; bulimia, severe; bulimia, persistent Exclusion: Alcohol or substance abuse in previous 12 months; psychotic features; suicide attempt within last 6 months; currently receiving psychotropic medication	BN, Severe: 100 (100%) BN, Persistent: 100 (100%) BN, Duration: 5.7 yr (SD ± 4.5) - 5.5 yr (SD ± 3.2) vs. 5.9 yr (SD ± 3.7) vs. 5 yr (SD ± 4.4) vs. 6.4 yr (SD ± 6.3) Binge Eating >= 3 episodes/wk, In the Previous 6 mo: 100 (100%) Vomiting >= 3 episodes/wk, In the Previous 6 mo: 100 (100%) History of AN: 10 (38%) vs. 11 (41%) vs. 9 (39%) vs. 11 (46%) %IBW 85%-125%: 100 (100%) %ABW: 112.2% (SD ± 9.5) Age 17 yr-45 yr: 100 (100%) Age: 24.5 yr (SD ± 6.4)	All treatment conditions led to decreases in binge/vomit episodes at the end of treatment: Binge Eating - Baseline: 7.2/wk vs. 12.1/wk vs. 12.3/wk vs. 12.2/wk - Change: -4.92/wk vs. -9.41/wk vs. -8.39/wk vs. -5.79/wk Vomiting - Baseline: 7.7/wk vs. 13.4/wk vs. 13.3/wk vs.14.5/wk - Change: -5.73/wk vs. -10.56/wk vs. -9.43/wk vs. -4.58/wk Combined treatment had higher rates of bulimic abstinence than the SH support group: 9 (35%) vs. 14 (52%) vs. 4 (17%) vs. 5 (20.83%) - CT + Nutritional Counselling vs. Nutritional Counselling: p=0.011 - CT + Nutritional Counselling vs. SH Support Group Therapy: p=0.022 CT (alone or with nutritional counseling) had better rates of	High

				- 23.3 yr (SD ± 5) vs. 24.1 yr (SD ± 5.3) vs. 24.2 yr (SD ± 5.6) vs. 26.5 yr (SD ± 9.1) Gender, Female: 100 (100%) Race: NR	study retention (85-89% vs. 54-61%) and was associated with greater benefits on dysfunctional attitudes and self-control. - CT vs. SH Support Group Therapy: p=0.019 - CT + Nutritional Counselling vs. Nutritional Counselling: p=0.021 - CT + Nutritional Counselling vs. SH Support Group Therapy: p=0.006 Attrition: 15% (4/26) vs. 11% (3/27) vs. 39% (9/23) vs. 46% (11/24)	

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CT=cognitive therapy; IBW=ideal body weight; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; yr=year

Compared to stepped care (supervised self-help)

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Mitchell et al. (2011)	Design: RCT Setting: Multi-center Country: NR Funding: Government, industry, and non-profit	Randomized N=293 Manual-based Individual CBT > Fluoxetine (for non-responders) 18 wk (N=147) Stepped Care (Supervised SH > Fluoxetine for non-responders > CBT for non-responders) 18 wk (N=146)	Inclusion: Purging or non-purging BN; 18 years or older Exclusion: Current active psychotherapy for their eating disorder; alcohol or drug misuse in the previous 6 months; alcohol or drug dependence in the previous 6 months; acute suicidal risk; medical illness that would preclude safe study participation; history of psychotic disorder	BN: 293 (100%) - Purging type: 280 (96%) - Non-purging type: 13 (4%) Binge Eating, Objective: 25/28 days (SD ± 16.296, N=228) Compensatory Behaviors: 43/28 days (SD ± 26.667, N=228) BMI: 23.3 kg/m² (SD ± 4.9, N=228) - 23.3 kg/m² (SD ± 4.5) vs. 23.5 kg/m² (SD ± 5.3) Age >= 18 yr: 293 (100%)	The treatments had similar responses at 18 wk but stepped care was more effective at the 1-yr post-treatment follow-up in reducing binge eating (MD 7/28 days, p<0.05) and compensatory behaviors (MD 10/28 days, p<0.05) in terms of episodes/28 days. Binge Eating, Objective - Baseline->18 wk->1 yr - 27->4->10/28 days vs. 27->8->3/28 days Compensatory Behaviors - Baseline->18 wk->1 yr	High

		Follow-up: Baseline – 1 yr Follow-up (N=197) - 103 vs. 94		Age: 29.7 yr - 29.5 yr (SD ± 8) vs. 29.8 yr (SD ± 9.8) Gender, Unknown: 293 (100%) Race, Black, African American, or Native American or Ethnicity, Hispanic/Latino: 32 (14%, N=228) - 24 (16%) vs. 18 (12%)	- 44->12->15/28 days vs. 43->19->5/28 days Binge Eating and Purging, Abstinence - 18 wk: 22 (15%) vs. 16 (11%) - 1 yr: 26 (18%) vs. 38 (26%) Attrition: 23% (24/147) vs. 29% (42/146)	

Abbreviations: BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; yr=year

Compared to self-help manual therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Treasure et al. (1994)	Design: RCT Setting: Outpatient: Maudsley Hospital Country: United Kingdom Funding: NR	Randomized N=81 CBT 8 wk (N=21) SH Manual Therapy 8 wk (N=41) WLC 8 wk (N=19)	Inclusion: BN or atypical BN Exclusion: Severe comorbidity; diabetes mellitus; high risk of suicide; dependence on alcohol	BN or BN, Atypical: 81 (100%) BMI: 26.8 kg/m² (SD ± 7) vs. 24 kg/m² (SD ± 5.9) vs. 23.3 kg/m² (SD ± 6.7) History of AN: 1 (5%) vs. 9 (21%) vs. 6 (30%) Age: 26 yr (SD ± 6.6) vs. 25.7 yr (SD ± 5.8) vs. 26 yr (SD ± 6.7) Gender, Female: 81 (100%) Race: NR	Rates of full remission were less in the WLC group (11%) vs. CBT (24%) or SH (22%) at the end of treatment. CBT was associated with a reduced frequency of binge eating, vomiting, and other compensatory behaviors; SH reduced the frequency of binge eating and compensatory behaviors but not vomiting. No changes were seen in the WLC group. Binge Eating, Abstinence - 8 wk: 7 (35%, N=20) vs. 11 (31%, N=35) vs. 3 (17%)	High

					Vomiting, Abstinence - 8 wk: 4 (29%, N=14) vs. 7 (24%, N=29) vs. 2 (15%, N=13) Binge Eating, Physician Assessment – Baseline->8 wk: 4->1 units vs. 3->1 units vs. 3->3 units Vomiting, Physician Assessment - Baseline->8 wk: 3->0 units vs. 3->1 units vs. 1->1 units Dietary Restraint, Physician Assessment – Baseline->8 wk: 3->1 units vs. 3->2 units vs. 3->2 units Overall Attrition: 0% (0/81)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; WLC=wait-list control; yr=year

Compared to Response Prevention

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1989)	Design: RCT Setting: NR Location: NR Funding: Government	Randomized N=77 CBT + Response Prevention Therapy 4 mo (N=17) CBT 4 mo (N=22) Self-Monitoring Therapy 4 mo (N=19) WLC 4 mo (N=19)	Inclusion: Female; BN Exclusion: Age below 18 years; above 65 years; concurrent pharmacological or psychological treatment for bulimia; concurrent AN, schizophrenia, bipolar affective disorder, or unipolar affective disorder; concurrent drug abuse; concurrent alcoholism; medical disorders such as significant hepatic disease; medical disorders such as renal	BN: 77 (100%) BN, Duration: 8.8 yr (SD ± 6.6) Purging: 12.2/wk (SD ± 8.3, N=16) vs. 11.1/wk (SD ± 6, N=17) vs. 12.3/wk (SD ± 8.3, N=16) vs. 13.8/wk (SD ± 8.4, N=18) History of AN: 13 (17%)	In contrast to the WLC group, all treatment groups improved. Purging, % Change, Baseline – 4 mo: -52.8% (N=16) vs. -78.2% (N=17) vs. -63.6% (N=16) vs. -8.9% (N=18) CBT was statistically superior to no treatment at 4 mo in terms of purging abstinence but differences from other groups were not significant: 31.2%	High

		Follow-up: Baseline – 10 mo Current Analysis (N=67) - 16 vs. 17 vs. 16 vs. 18	disease or major cardiac disease; pregnancy; abnormal values of serum potassium	Age: 29.2 yr (SD ± 8.6) Gender, Female: 77 (100%) Race: NR	(N=16) vs. 56.3% (N=17) vs. 23.5% (N=16) vs. 5.8% (N=18). Attrition: 6% (1/16) vs. 23% (5/22) vs. 16% (3/19) vs.5% (1/19)	
Cooper and Steere (1995)	Design: RCT Setting: Single Center: a local BN clinic Country: United Kingdom Funding: Government	Randomized N=31 Cognitive Behavioral Condition 18 wk (N=15) - Binge Eating, Abstinence N=6 - Purging, Abstinence N=7 Exposure and Response Prevention Condition 18 wk (N=16) - Binge Eating, Abstinence N=7 - Purging, Abstinence N=6 Follow-up: Baseline – 12 mo	Inclusion: BN; purged immediately after binge eating Exclusion: NR	BN, Purging Type: 31 (100%) Bulimic Episodes: 21.9/mo (SD ± 12.3, N=13) vs. 30.4/mo (SD ± 19.4, N=14) Vomiting, Abstinence: 2 Percent ABW, Matched-Population: 98.5% (SD ± 11.5, N=13) vs. 99.3% (SD ± 11, N=14) Age: 23.8 yr Gender, Female: 27 (100%) Race: NR	There were significantly more vomiting episodes at 12 mo follow-up in the exposure and response prevention group (23.4/mo (N=12) vs. 4.3/mo (N=12) with cognitive behavioral condition, MD -19.1/mo, p<0.007) but baseline mean rates also differed (79.9 vs. 36.1 episodes, respectively). Binge Eating, Abstinence - 18 wk: 6 (46.15%, N=13) vs. 7 (50%, N=14) Binge Eating, Relapse - 18 wk – 12 mo - Binge Eating, Abstinence subgroup: 0 (0%, N=6) vs. 5 (71.43%, N=7) (p<0.04) Purging, Abstinence - 14 wk – 18 wk: 7 (54%, N=13) vs. 6 (43%, N=14) Purging, Relapse - 18 wk – 12 mo - Purging, Abstinence subgroup: 1 (14.29%, N=7) vs. 5 (83.33%, N=6) (p<0.1) Disease Response, Remission - 18 wk: 6 (46%, N=13) vs. 7 (50%, N=14) Attrition: 13% (2/15) vs. 12.5% (2/16)	High

Leitenberg et al. (1988)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=59 Current Analysis (N=47) CBT 10 wk > 14 wk (N=12) Exposure Plus Response-Prevention Single Setting 10 wk > 14 wk (N=11) Exposure Plus Response-Prevention Multiple Setting 10 wk > 14 wk (N=12) WLC 14 wk (N=12) Follow-up: Baseline – 41 wk	Inclusion: Women; 18-45-years of age; within 80-120% of their normal weight; BN; vomited an average of 3 times a wk Exclusion: Abuse of laxatives; signs of alcoholism; signs of psychosis; signs of serious suicide risk; involved in concurrent treatment	BN: 59 (100%) BN, Duration: 6.94 yr (N=47) - 5.6 yr (SD ± 4.2) vs. 10 yr (SD ± 9.6) vs. 7.7 yr (SD ± 4.8) vs. 4.7 yr (SD ± 4.2) Vomiting 3 episodes/wk: 59 (100%) Vomiting: 12.13/wk (N=47) %EBW 80%-120%: 59 (100%) Age 18 yr-45 yr: 59 (100%) Age: 26 yr (N=47) - 25 yr (SD ± 3.4) vs. 28 yr (SD ± 10.1) vs. 27 yr (SD ± 5.7) vs. 24 yr (SD ± 5.3) Gender, Female: 59 (100%) Race: NR	At the end of treatment (which all patients completed) and at 6-mo follow-up, treatment groups improved significantly on most outcomes with minimal change in the WLC group. However, there were no statistical differences in outcomes among the groups, likely due to the sample size. Vomiting - Baseline: 8.57/wk (SD ± 4.5) vs. 13.81/wk (SD ± 8.1) vs. 10.21/wk (SD ± 8.4) vs. 16.04/wk (SD ± 8.7) Vomiting, % Change - Baseline – 17 wk: -40% vs. -73% vs. -67% vs. NR - Baseline – 41 wk: -39% vs. -62% vs. -85% (N=10) vs. NR Vomiting, Abstinence - 17 wk: 1 (8.33%) vs. 4 (36.36%) vs. 4 (33.33%) vs. 0 (0%) - 41 wk: 4 (33.33%) vs. 2 (18.18%) vs. 5 (50%, N=10) vs. 0 (0%) Overall Attrition: 20% (12/59)	High

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EBW=expected body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Compared to Interpersonal Psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Agras et al. (2000); Wilson et al. (2002)	Design: RCT; Post-hoc Analysis Setting: Multi-center: Columbia University; Stanford University Country: United States Funding: Government and non-profit	Randomized N=220 CBT 20 wk (N=110) - Columbia University: 54 - Stanford University: 56 IPT 20 wk (N=110) - Columbia University: 56 - Stanford University: 54 Follow-up: Baseline – 17 mo Follow-up (N=129) - 65 vs. 64	Inclusion: BN Exclusion: Receiving anti-depressants; severe physical conditions that would interfere with treatment; severe psychiatric conditions that would interfere with treatment; psychosis; current AN; current psychotherapeutic treatment of any type; all psychotropic medication; pregnancy; received an adequate trial of CBT; IPT for BN	BN: 220 (100%) Binge Eating, Duration: 11.5 yr (SD ± 7.5) vs. 11.4 yr (SD ± 7.6) - Columbia University: 12.1 yr (SD ± 8.1) vs. 9.6 yr (SD ± 6.5) - Stanford University: 10.8 yr (SD ± 6.9) vs. 13.2 yr (SD ± 8.1) Purging, Duration: 10 yr (SD ± 7.2) vs. 9.7 yr (SD ± 6.4) - Columbia University: 9.9 yr (SD ± 7.3) vs. 8 yr (SD ± 5) - Stanford University: 10.1 yr (SD ± 7.1) vs. 11.6 yr (SD ± 7.1) Binge Eating, Objective: 20/28 days (SD ± 23.704, N=65) vs. 23.5/28 days (SD ± 20, N=64) Purging: 30/28 days (SD ± 23.704, IQR Difference ± 32, N=65) vs. 42/28 days (SD ± 40, N=64) Vomiting, Self-Induced: 220 (100%) History of AN: 26 (24%) vs. 26 (24%) - Columbia University: 15 (28%) vs. 18 (32%) - Stanford University: 11 (20%) vs. 9 (17%) Age: 28.3 yr (SD ± 7) vs. 27.9 yr (SD ± 7.5)	At baseline, the IPT group had higher eating concern scores and greater purging rates. At 1-yr follow-up, outcomes did not differ among completers: 40% (N=65) recovered with CBT vs. 27% (N=64) with IPT. CBT was superior at the end of treatment in: recovery (29% vs. 6%, p<0.001); remission (48% vs. 28%, p=0.003); and meeting community norms for eating attitudes/behaviors (41% vs. 27%, p=0.04). Superiority of CBT was even greater among treatment completers at 20 wk: (45% (N=65) vs. 8% (N=64), p=0.001). Binge Eating, Objective, Change - Baseline – 5 mo: -20/28 days (SD ± 21.28, N=65) vs. -18.5/28 days (SD ± 14.68, N=64) (MD -35%, p=0.01) - Baseline – 17 mo: -20/28 days (SD ± 19.26, N=65) vs. -21.5/28 days (SD ± 14.32, N=64) Binge Eating, Objective, % Change - Baseline – 5 mo: -86% (N=65) vs. -51% (N=64) - Baseline – 72 wk: -72% (N=65) vs. -70% (N=64) (MD -2%, p=0.8) Purging, Change - Baseline – 5 mo: -29/28 days (SD ± 20, N=65) vs. -	High

				Gender, Female: 110 (100%) Race - Caucasian: 87 (79%) vs. 81 (74%) - Black or African American: 7 (6%) vs. 7 (6%) - Asian: 4 (4%) vs. 7 (6%) - Native American: 1 (1%) vs. 0 (0%) Ethnicity, Hispanic/Latino: 11 (10%) vs. 14 (13%)	28.5/28 days (SD ± 28.86, N=64) - Baseline – 17 mo: -27/28 days (SD ± 17.91, N=65) vs. -35/28 days (SD ± 29.57, N=64) Purging, % Change - Baseline – 5 mo: -84% (N=65) vs. -50% (N=64) (MD -34%, p=0.001) - Baseline – 72 wk: -61% (N=65) vs. -62% (N=64) (MD 1%, p=0.99) Attrition: 33% (37/110) vs. 26% (29/110)	
Fairburn et al. (1991, 1993)	Design: RCT; Follow-up Setting: Outpatient Country: United Kingdom Funding: Non-profit	Randomized N=75 CBT 18 wk (N=25) IPT 18 wk (N=25) Behavior Therapy 18 wk (N=25) Follow-up: Baseline – 16 mo	Inclusion: Female; aged 17 years or older; complained of having lost control over eating; used either self-induced vomiting, laxatives, or extreme dieting to control their shape or weight; BN previous 6 months; BMI greater than 17 kg/m^2 Exclusion: Significantly underweight; major coexisting psychiatric problems that required inpatient treatment; severe depressive illness; amphetamine psychosis; alcohol dependence	BN, In the Previous 6 mo: 75 (100%) BN, Duration: 4.4 yr (SD ± 4.11) Bulimic Episodes, Objective - 23.7 d/mo (SD ± 17.1) - 16.5/28 days (N=60) Vomiting, Self-Induced or Laxative Abuse or Diet, Extreme: 75 (100%) Vomiting, Self-Induced: 56 (75%) - 28.9 d/mo (SD ± 21.07, N=56) Laxative Abuse: 26 (35%) - 14.7 d/mo (SD ± 14.23, N=26) History of AN: 27 (34%)	At baseline, vomiting episodes were significantly more frequent in those who received CBT vs. IPT (mean of 28.5 episodes/28 days vs. 16.4 episodes/28 days). Effects of the treatments were comparable for binge episodes and laxative abuse, but IPT had less impact on vomiting than CBT or behavior therapy (p=0.03): - Vomiting - Baseline: 28.5/28 days (N=21) vs. 16.4/28 days (N=21) vs. 18.5/28 days (N=18) - Vomiting, Change: -27/28 days vs. -10.9/28 days vs. -17.6/28 days At 12-mo follow-up, rates of subjects with no binge eating, vomiting, or laxative abuse were 36% CBT vs. 44% IPT vs. 20% behavior therapy:	High

				BMI > 17 kg/m²: 75 (100%) BMI: 22.2 kg/m² (SD ± 3.25, 95% CI 21.5 – 23) Age >= 17 yr: 75 (100%) Age: 24.2 yr (SD ± 6.71, 95% CI 22.5 – 25.6) Gender, Female: 75 (100%) Race: NR	- CBT vs. Behavioral Therapy: OR 2.49 (95% CI 1.34 – 4.62, p<0.05) Study discontinuation at 12-mo follow-up differed by treatment (20% CBT vs. 32% IPT vs. 48% behavior therapy; p=0.04 for CBT vs. behavior therapy comparison). Attrition: 16% (4/25) vs. 12% (3/25) vs. 24% (6/25)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; CI=confidence interval; d=day; EBW=expected body weight; IPT=interpersonal psychotherapy; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to Behavior Therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Fairburn et al. (1991, 1993)	Design: RCT; Follow-up Setting: Outpatient Country: United Kingdom Funding: Non-profit	Randomized N=75 CBT 18 wk (N=25) IPT 18 wk (N=25) Behavior Therapy 18 wk (N=25) Follow-up: Baseline – 16 mo	Inclusion: Female; aged 17 years or older; complained of having lost control over eating; used either self-induced vomiting, laxatives, or extreme dieting to control their shape or weight; BN previous 6 months; BMI greater than 17 kg/m^2 Exclusion: Significantly underweight; major coexisting psychiatric problems that required inpatient treatment; severe depressive illness; amphetamine psychosis; alcohol dependence	BN, In the Previous 6 mo: 75 (100%) BN, Duration: 4.4 yr (SD ± 4.11) Bulimic Episodes, Objective - 23.7 d/mo (SD ± 17.1) - 16.5/28 days (N=60) Vomiting, Self-Induced or Laxative Abuse or Diet, Extreme: 75 (100%) Vomiting, Self-Induced: 56 (75%) - 28.9 d/mo (SD ± 21.07, N=56)	At baseline, vomiting episodes were significantly more frequent in those who received CBT vs. IPT (mean of 28.5 episodes/28 days vs. 16.4 episodes/28 days). Effects of the treatments were comparable for binge episodes and laxative abuse, but IPT had less impact on vomiting than CBT or behavior therapy (p=0.03): - Vomiting - Baseline: 28.5/28 days (N=21) vs. 16.4/28 days (N=21) vs. 18.5/28 days (N=18)	High

				Laxative Abuse: 26 (35%) - 14.7 d/mo (SD ± 14.23, N=26) History of AN: 27 (34%) BMI > 17 kg/m²: 75 (100%) BMI: 22.2 kg/m² (SD ± 3.25, 95% CI 21.5 – 23) Age >= 17 yr: 75 (100%) Age: 24.2 yr (SD ± 6.71, 95% CI 22.5 – 25.6) Gender, Female: 75 (100%) Race: NR	- Vomiting, Change: -27/28 days vs. -10.9/28 days vs. -17.6/28 days At 12-mo follow-up, rates of subjects with no binge eating, vomiting, or laxative abuse were 36% CBT vs. 44% IPT vs. 20% behavior therapy: - CBT vs. Behavioral Therapy: OR 2.49 (95% CI 1.34 – 4.62, p<0.05) Study discontinuation at 12-mo follow-up differed by treatment (20% CBT vs. 32% IPT vs. 48% behavior therapy; p=0.04 for CBT vs. behavior therapy comparison). Attrition: 16% (4/25) vs. 12% (3/25) vs. 24% (6/25)	
Freeman et al. (1988)	Design: RCT Setting: NR Country: United Kingdom Funding: Non-profit	Randomized N=112 CBT 15 wk (N=32) Behavior Therapy 15 wk (N=30) Group Therapy 15 wk (N=30) WLC 15 wk (N=20) Follow-up: Baseline – 1 yr	Inclusion: BN; women; aged 18 and over; binged at least 3 times in the previous mo; established bulimia Exclusion: History of psychotic illness	BN: 112 (100%) BN, Duration: 6 yr (SD ± 4.9) Binge Eating >= 3 episodes, In the Previous 1 mo: 112 (100%) BN, Age at Onset: 18.2 yr (SD ± 4.6) Age >= 18 yr: 112 (100%) Age: 24.2 yr (SD ± 5.6) Gender, Female: 112 (100%) Race: NR	Active treatments were equally effective with 77% achieving binge-eating abstinence at the end of treatment. Scores on a number of eating related rating scales were also improved with some statistical differences between treatments on individual scale items. Binge Eating - Baseline: 6.2/wk vs. 4.6/wk vs. 6.3/wk vs. 5.7/wk - Change - Baseline – 15 wk: -4.9/wk vs. -4/wk vs. -5.5/wk vs. -2/wk Vomiting, Self-Induced - Baseline: 7.4/wk vs. 3.6/wk vs. 8.9/wk vs. 8/wk	High

					- Change - Baseline – 15 wk: -6.4/wk vs. -3.3/wk vs. -8.3/wk vs.-1.7/wk Laxative Abuse - Baseline: 6.2 tablets/wk vs. 5.1 tablets/wk vs. 14.6 tablets/wk vs. 10.4 tablets/wk - 15 wk: 1.3 tablets/wk vs. 0 tablets/wk vs. 4.3 tablets/wk vs. 13.5 tablets/wk Attrition: 34% (11/32) vs. 17% (5/30) vs. 37% (11/30) vs. 20% (4/20)	
Thackwray et al. (1993)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=47 CBT 8 wk (N=NR) Behavioral Treatment 8 wk (N=NR) Nonspecific Self-monitoring Treatment 8 wk (N=NR) Follow-up: Baseline – 8 mo	Inclusion: BN; female Exclusion: Current involvement in treatment for BN; pregnancy; severe renal problems; cardiac problems	BN: 47 (100%) BN, Duration: 6.7 yr (SD ± 7.28) Binge Eating and/or Purging: 5.53/wk (SD ± 3.37) Age: 31.3 yr (SD ± 10.41) Gender, Female: 47 (100%) Race: NR	Although differences were not significant, abstinence rates at the end of treatment were 92% with CBT, 100% with behavioral treatment, and 69% with self-monitoring. At the 6-mo follow-up, rates were 69%, 38%, and 15%, respectively. Binge Eating and/or Purging – Baseline->8 wk->8 mo - 5.4->0.6->0.4/wk vs. 5.6->0->0.6/wk vs. 5.6->1->2.7/wk Binge Eating 1/wk and/or Purging 1/wk - 8 mo: 23% vs. 62% vs. 15% Binge Eating > 1/wk and/or Purging > 1/wk - 8 mo: 8% vs. 0% vs. 69% Overall Attrition: 17% (8/47)	High

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; CI=confidence interval; d=day; IPT=interpersonal psychotherapy; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Compared to Nutritional Counseling

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Hsu et al. (2001)	Design: RCT Setting: Outpatient Country: United States Funding: Government and academic	Randomized N=100 CT 14 wk (N=26) CT + Nutritional Counselling 14 wk (N=27) Nutritional Counselling 14 wk (N=23) SH Support Group Therapy 14 wk (N=24) Follow-up: Baseline – 3.5 yr	Inclusion: BN; female; body weight within 85 to 125% IBW; 17-45 years of age; binge eating on average at least 3 times a wk in previous 6 months; vomiting on average at least 3 times a wk in previous 6 months; bulimia, severe; bulimia, persistent Exclusion: Alcohol or substance abuse in previous 12 months; psychotic features; suicide attempt within last 6 months; currently receiving psychotropic medication	BN, Severe: 100 (100%) BN, Persistent: 100 (100%) BN, Duration: 5.7 yr (SD ± 4.5) - 5.5 yr (SD ± 3.2) vs. 5.9 yr (SD ± 3.7) vs. 5 yr (SD ± 4.4) vs. 6.4 yr (SD ± 6.3) Binge Eating >= 3 episodes/wk, In the Previous 6 mo: 100 (100%) Vomiting >= 3 episodes/wk, In the Previous 6 mo: 100 (100%) History of AN: 10 (38%) vs. 11 (41%) vs. 9 (39%) vs. 11 (46%) %IBW 85%-125%: 100 (100%) %ABW: 112.2% (SD ± 9.5) Age 17 yr-45 yr: 100 (100%) Age: 24.5 yr (SD ± 6.4) - 23.3 yr (SD ± 5) vs. 24.1 yr (SD ± 5.3) vs. 24.2 yr (SD ± 5.6) vs. 26.5 yr (SD ± 9.1)	All treatment conditions led to decreases in binge/vomit episodes at the end of treatment: Binge Eating - Baseline: 7.2/wk vs. 12.1/wk vs. 12.3/wk vs. 12.2/wk - Change: -4.92/wk vs. -9.41/wk vs. -8.39/wk vs. -5.79/wk Vomiting - Baseline: 7.7/wk vs. 13.4/wk vs. 13.3/wk vs.14.5/wk - Change: -5.73/wk vs. -10.56/wk vs. -9.43/wk vs. -4.58/wk Combined treatment had higher rates of bulimic abstinence than the SH support group: 9 (35%) vs. 14 (52%) vs. 4 (17%) vs. 5 (20.83%) - CT + Nutritional Counselling vs. Nutritional Counselling: p=0.011 - CT + Nutritional Counselling vs. SH Support Group Therapy: p=0.022 CT (alone or with nutritional counseling) had better rates of study retention (85-89% vs. 54-61%) and was associated with	High

				Gender, Female: 100 (100%) Race: NR	greater benefits on dysfunctional attitudes and self-control. - CT vs. SH Support Group Therapy: p=0.019 - CT + Nutritional Counselling vs. Nutritional Counselling: p=0.021 - CT + Nutritional Counselling vs. SH Support Group Therapy: p=0.006 Attrition: 15% (4/26) vs. 11% (3/27) vs. 39% (9/23) vs. 46% (11/24)	
Sundgot-Borgen et al. (2002)	Design: RCT Setting: NR Country: Norway Funding: NR	Randomized N=64 Group CBT 16 wk (N=16) Nutrition Counseling Therapy 16 wk (N=17) Exercise 16 wk (N=15) WLC 16 wk (N=16) Follow-up: Baseline – 94 wk Current Analysis (N=58) 14 vs. 17 vs. 12 vs. 15	Inclusion: Normal weight; female; BN; 18-29 years of age Exclusion: History of AN; history of other psychiatric disorders; history of somatic disorders; treatment for eating disorders 6 months before entering present study; use of medication	BN: 64 (100%) BN, Duration: 5 yr (SD ± 1.6) vs. 5 yr (SD ± 2.3) vs. 7 yr (SD ± 3.7) vs. 6 yr (SD ± 3.8) Vomiting: 8.6/wk (SD ± 4.68) vs. 8.2/wk (SD ± 4.34) vs. 7.8/wk (SD ± 3.39) vs. 5.6/wk (SD ± 3.15) Weight, Normal: 64 (100%) BMI: 20 kg/m² (SD ± 1.9) vs. 21 kg/m² (SD ± 2.1) vs. 21 kg/m² (SD ± 2) vs. 22 kg/m² (SD ± 2.5) Age 18 yr-29 yr: 64 (100%) Age: 22 yr (SD ± 2.7) vs. 22 yr (SD ± 2.9) vs. 23 yr (SD ± 2.3) vs. 23.2 yr (SD ± 3.2) Gender, Female: 64 (100%) Race: NR	Group CBT was superior to nutritional counseling on vomiting episodes/wk at 42 wk (3.5/wk vs. 7.06/wk, MD -3.56/wk, p<0.001) and 22 mo (2.71/wk vs. 7.18/wk, MD -4.47/wk, p<0.001). Exercise was superior to other treatment conditions in affecting scores on specific rating scale items (e.g., body dissatisfaction, drive for thinness). Laxative Abuse Baseline: 2.3/wk (SD ± 1.8) vs. NR vs. NR vs. NR 16 wk: 2.1/wk (SD ± 1.7) vs. NR vs. 0.85/wk (SD ± 0.99) vs. NR - CBT 16 wk vs. Exercise 16 wk: MD 1.25/wk (p<0.02) 42 wk: 2.57/wk (SD ± 2.1) vs. NR vs. 0/wk (SD ± 0) vs. NR - CBT 16 wk vs. Exercise 16 wk: MD 2.57/wk (p<0.0001)	Moderate

					22 mo: 3.1/wk (SD ± 2.4) vs. NR vs. 0.08/wk (SD ± 0.28) vs. NR - CBT 16 wk vs. Exercise 16 wk: MD 3.02/wk (p<0.0001) Attrition: 13% (2/16) vs. 0% (1/17) vs. 20% (3/15) vs. 6% (1/16)	

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive- behavioral therapy; CT=cognitive therapy; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; WLC=wait-list control; yr=year

Compared to Other Psychotherapy

Compared to self-monitoring therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1989)	Design: RCT Setting: NR Location: NR Funding: Government	Randomized N=77 CBT + Response Prevention Therapy 4 mo (N=17) CBT 4 mo (N=22) Self-Monitoring Therapy 4 mo (N=19) WLC 4 mo (N=19) Follow-up: Baseline – 10 mo Current Analysis (N=67) - 16 vs. 17 vs. 16 vs. 18	Inclusion: Female; BN Exclusion: Age below 18 years; above 65 years; concurrent pharmacological or psychological treatment for bulimia; concurrent AN, schizophrenia, bipolar affective disorder, or unipolar affective disorder; concurrent drug abuse; concurrent alcoholism; medical disorders such as significant hepatic disease; medical disorders such as renal disease or major cardiac disease; pregnancy; abnormal values of serum potassium	BN: 77 (100%) BN, Duration: 8.8 yr (SD ± 6.6) Purging: 12.2/wk (SD ± 8.3, N=16) vs. 11.1/wk (SD ± 6, N=17) vs. 12.3/wk (SD ± 8.3, N=16) vs. 13.8/wk (SD ± 8.4, N=18) History of AN: 13 (17%) Age: 29.2 yr (SD ± 8.6 Gender, Female: 77 (100%) Race: NR	In contrast to the WLC group, all treatment groups improved. Purging, % Change, Baseline – 4 mo: -52.8% (N=16) vs. -78.2% (N=17) vs. -63.6% (N=16) vs. -8.9% (N=18) CBT was statistically superior to no treatment at 4 mo in terms of purging abstinence but differences from other groups were not significant: 31.2% (N=16) vs. 56.3% (N=17) vs. 23.5% (N=16) vs. 5.8% (N=18). Attrition: 6% (1/16) vs. 23% (5/22) vs. 16% (3/19) vs.5% (1/19)	High

Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Compared to nonspecific self-monitoring treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Thackwray et al. (1993)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=47 CBT 8 wk (N=NR) Behavioral Treatment 8 wk (N=NR) Nonspecific Self-monitoring Treatment 8 wk (N=NR) Follow-up: Baseline – 8 mo	Inclusion: BN; female Exclusion: Current involvement in treatment for BN; pregnancy; severe renal problems; cardiac problems	BN: 47 (100%) BN, Duration: 6.7 yr (SD ± 7.28) Binge Eating and/or Purging: 5.53/wk (SD ± 3.37) Age: 31.3 yr (SD ± 10.41) Gender, Female: 47 (100%) Race: NR	Although differences were not significant, abstinence rates at the end of treatment were 92% with CBT, 100% with behavioral treatment, and 69% with self-monitoring. At the 6-mo follow-up, rates were 69%, 38%, and 15%, respectively. Binge Eating and/or Purging – Baseline->8 wk->8 mo - 5.4->0.6->0.4/wk vs. 5.6->0->0.6/wk vs. 5.6->1->2.7/wk Binge Eating 1/wk and/or Purging 1/wk - 8 mo: 23% vs. 62% vs. 15% Binge Eating > 1/wk and/or Purging > 1/wk - 8 mo: 8% vs. 0% vs. 69% Overall Attrition: 17% (8/47)	High

Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to group therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Freeman et al. (1988)	Design: RCT	Randomized N=112	Inclusion: BN; women; aged 18 and over; binged at least 3	BN: 112 (100%)	Active treatments were equally effective with 77% achieving	High

	Setting: NR Country: United Kingdom Funding: Non-profit	CBT 15 wk (N=32) Behavior Therapy 15 wk (N=30) Group Therapy 15 wk (N=30) WLC 15 wk (N=20) Follow-up: Baseline – 1 yr	times in the previous mo; established bulimia Exclusion: History of psychotic illness	BN, Duration: 6 yr (SD ± 4.9) Binge Eating >= 3 episodes, In the Previous 1 mo: 112 (100%) BN, Age at Onset: 18.2 yr (SD ± 4.6) Age >= 18 yr: 112 (100%) Age: 24.2 yr (SD ± 5.6) Gender, Female: 112 (100%) Race: NR	binge-eating abstinence at the end of treatment. Scores on a number of eating related rating scales were also improved with some statistical differences between treatments on individual scale items. Binge Eating - Baseline: 6.2/wk vs. 4.6/wk vs. 6.3/wk vs. 5.7/wk - Change - Baseline – 15 wk: -4.9/wk vs. -4/wk vs. -5.5/wk vs. -2/wk Vomiting, Self-Induced - Baseline: 7.4/wk vs. 3.6/wk vs. 8.9/wk vs. 8/wk - Change - Baseline – 15 wk: -6.4/wk vs. -3.3/wk vs. -8.3/wk vs.-1.7/wk Laxative Abuse - Baseline: 6.2 tablets/wk vs. 5.1 tablets/wk vs. 14.6 tablets/wk vs. 10.4 tablets/wk - 15 wk: 1.3 tablets/wk vs. 0 tablets/wk vs. 4.3 tablets/wk vs. 13.5 tablets/wk Attrition: 34% (11/32) vs. 17% (5/30) vs. 37% (11/30) vs. 20% (4/20)	

Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Compared to hypnobehavioral treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and	Outcome measures, main results, and overall percent attrition	Risk of bias

		size (N), dose, duration, and follow-up		baseline clinical features (e.g., BMI)		
Griffiths et al. (1994, 1996)	Design: RCT; Follow-up Setting: Outpatient Country: Australia Funding: NR	Randomized N=78 CBT 8 wk (N=23) Hypnobehavioral Treatment 8 wk (N=27) WLC 8 wk (N=28) Follow-up: Baseline – 12 mo Follow-up (N=72) - 19 vs. 21 vs. 22	Inclusion: BN; female; 17-50 years of age; BMI 18-26kg/m^2; agreeable not to seek additional treatment for their eating disorder during the research Exclusion: More than 2 previous inpatient admissions for treatment of an eating disorder; concurrent pharmacological or psychological treatment; coexisting major psychiatric disorder other than a depressive state; coexisting major psychiatric disorder other than a anxiety state; coexisting major psychiatric disorder other than a personality disorder; physically dependent on drugs; physically dependent on alcohol; indications for hospitalization because of risk of suicide; indications for hospitalization because of risk of poor physical health	BN: 78 (100%) BN, Symptomatic, Duration: 6.19 yr (SD ± 5.08) - 5.4 yr (SD ± 2.31, N=19) vs. 3.31 yr (SD ± 2.99, N=21) vs. NR (N=22) BN, Objective, Symptomatic, Duration: 4.54 yr (SD ± 5.15) Bulimic Episodes, Objective: 14.18 d/mo (SD ± 7.78) Binge Eating: 3.18 d (SD ± 1.49, N=20) vs. 3.95 d (SD ± 1.67, N=21) vs. 4.77 d (SD ± 1.83, N=22) Purging: 3.38 d (SD ± 2.29, N=20) vs. 3.86 d (SD ± 2.46, N=21) vs. 5.27 d (SD ± 2, N=22) Vomiting, Self-Induced: 68 (87.2%) Vomiting, Self-Induced: 15.76 d/mo (SD ± 10.4) Laxative Abuse: 32 (41%) Laxative Abuse: 4.69 d/mo (SD ± 8.67) Diuretics: 8 (11%)	Abstinence rates were: 10 (50%, N=20) vs. 9 (43%, N=21) vs. 1 (4.5%, N=22) for binge eating; and 8 (40%, N=20) vs. 7 (33.3%, N=21) vs. 1 (4.5%, N=22) for purging. There were no statistical differences in outcomes among the groups. 9-mo follow-up continued to show no differences in outcomes between active treatment groups. Binge Eating Episodes - Baseline: 4.73/2 wks (SD ± 2.79, N=20) vs. 6.38/2 wks (SD ± 6.12, N=21) vs. 9.82/2 wks (SD ± 9.49, N=22) Binge Eating Episodes, Change - Baseline – 11 mo: -3.64/mo (SD ± 4.91, N=25) vs. -3.37/mo (SD ± 3.36, N=23) vs. NR Purging Episodes - Baseline: 6.48/2 wks (SD ± 7.43, N=20) vs. 8.55/2 wks (SD ± 9.94, N=21) vs. 11.77/2 wks (SD ± 9.87, N=22) Purging Episodes, Change - Baseline – 11 mo: -2.05/mo (SD ± 6.04, N=25) vs. -2.24/mo (SD ± 5.33, N=23) vs. NR Vomiting, Self-Induced - Baseline	High

				Exercise, Excessive: 53 (67.9%) History of AN: 20 (25.6%) BMI 18 kg/m²-26 kg/m²: 78 (100%) Age 17 yr-50 yr: 78 (100%) Age: 25.91 yr (SD ± 5.73) Gender, Female: 78 (100%) Race: NR	- 6.02 d/mo (SD ± 9.33, N=25) vs. 5.63 d/mo (SD ± 8.22, N=23) vs. NR - 9.5/30 days (SD ± 12.88, N=25) vs. 7.62/30 days (SD ± 10.43, N=23) vs. NR Vomiting, Self-Induced, Change - Baseline – 11 mo - -5.76 d/mo (SD ± 7.8, N=25) vs. -5.48 d/mo (SD ± 6.17, N=23) vs. NR - -9.15/30 days (SD ± 9.89, N=25) vs. -7.46/30 days (SD ± 7.57, N=23) vs. NR Treatment Discontinuation - Baseline – 8 wk: 5 (26.32%, N=19) vs. 5 (23.81%, N=21) vs. NR Attrition: 16% (6/38) vs. 23% (9/40) vs. NR	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; wks=weeks; WLC=wait-list control; yr=year

Compared to mindfulness and acceptance-based behavioral treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Juarascio et al. (2021)	Design: RCT Setting: Outpatient Country: United States Funding: Government	Randomized N=44 CBT 20 wk (N=18) MABT 20 wk (N=26) Follow-up: Baseline – 11 mo	Inclusion: BN; BN with subjectively large binge episodes; age 18 years or older Exclusion: Medical complications; severe comorbid psychiatric or intellectual/developmental disorder; pregnancy; unstable psychiatric medication; history of bariatric surgery; other	BN: 37 (84.1%) Other Specified Feeding or Eating Disorder, BN: 7 (15.9%) Age: 35.22 yr vs. 29.77 yr Gender	Both CBT and MABT showed large reductions in symptoms that were sustained through the 6-mo follow-up. EDE, loss of control episodes: 27.44 (SD ± 19.03)->3.94 (SD ± 4.32) vs. 27.27 (SD ± 17.65)->3.84 (SD ± 6.81)	Moderate

			current eating disorder treatment.	- Female: 16 (88.9%) vs. 23 (88.5%) - Male: 2 (11.1%) vs. 3 (11.5%) Race - Caucasian: 16 (88.9%) vs. 21 (80.8%) - Black or African American: 1 (5.6%) vs. 3 (11.5%) - Asian: 1 (5.6%) vs. 2 (7.7%) Ethnicity, Hispanic/Latino: 5 (27.8%) vs. 1 (3.8%)	EDE, Compensatory behaviors: 35.83 (SD ± 29.08)->7.27 (SD ± 9.43) vs. 31.12 (SD ± 20.79)-> 5.05 (SD ± 5.76) Attrition: 44.4% (8/18) vs. 38.5% (10/26)	

Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EDE=Eating Disorder Examination; MABT=mindfulness and acceptance-based behavioral treatment; mo=month; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to motivational enhancement therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Katzman et al. (2010)	Design: RCT Setting: Outpatient Country: NR Funding: NR	Randomized N=225 Individual CBT 4 wk > Group CBT 12 wk (N=73) Motivational Enhancement Therapy 4 wk > Group CBT 12 wk (N=73) Motivational Enhancement Therapy 4 wk > Individual CBT 12 wk (N=79) Follow-up: Baseline – 2.5 yr	Inclusion: BN or EDNOS Exclusion: Pregnancy; diabetes mellitus; severe mental illness; schizophrenia; bipolar illness; severe learning disability; inability to commit to treatment from the outset; referral for assessment only	BN or EDNOS: 225 (100%) EDNOS: 60 (26.67%) Binge Eating: 3.6 units (SD ± 1.4) - 3.6 units (SD ± 1.4) vs. 3.7 units (SD ± 1.4) vs. 3.5 units (SD ± 1.5) Vomiting: 3.4 units (SD ± 1.7) - 3.3 units (SD ± 1.6) vs. 3.7 units (SD ± 1.6) vs. 3.3 units (SD ± 1.7) Laxative Abuse: 1.8 units (SD ± 1.6)	Significant improvements were noted across outcomes for each treatment with no apparent differences in response among them. Binge Eating, Abstinence - Baseline: 2 (5%, N=40) vs. 1 (2.7%, N=37) vs. 0 (0%, N=39) - 12 wk: 8 (40%, N=20) vs. 8 (24.2%, N=33) vs. 5 (25%, N=20) - 2.5 yr: 12 (57.2%, N=21) vs. 5 (38.5%, N=13) vs. 8 (40%, N=20)	High

				- 1.7 units (SD ± 1.3) vs. 1.8 units (SD ± 1.4) vs. 1.9 units (SD ± 1.4) Age: 29.3 yr (SD ± 7.5) 27.8 yr (SD ± 6.3) vs. 28.9 yr (SD ± 8.1) vs. 31 yr (SD ± 7.7) Gender, Female: 225 (100%) Race: NR	Vomiting, Abstinence - Baseline: 12 (26.7%, N=45) vs. 6 (16.7%, N=36) vs. 8 (17.8%, N=45) - 12 wk: 8 (40%, N=20) vs. 8 (24.2%, N=33) vs. 5 (25%, N=20) - 2.5 yr: 12 (57.1%, N=21) vs. 8 (38.5%, N=21) vs. 8 (40%, N=20) Laxative Abuse, Abstinence - Baseline: 22 (53.7%, N=41) vs. 17 (54.8%, N=31) vs. 26 (66.7%, N=39) - 12 wk: 14 (82.4%, N=17) vs. 20 (71.4%, N=28) vs. 13 (72.2%, N=18) - 2.5 yr: 16 (84.2%, N=19) vs. 12 (92.3%, N=13) vs. 18 (81.8%, N=22) Attrition: 32% (19/60) vs. 48% (29/61) vs. 43% (31/72)	

Abbreviations: BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EDNOS=eating disorder not otherwise specified; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to family-based treatment

Le Grange et al. (2015)	Design: RCT Setting: Multi-center Country: United States Funding: Government	Randomized N=130 FBT 6 mo (N=52) CBT 6 mo (N=58) SPT 6 mo (N=20) Follow-up: Baseline – 18 mo Current Analysis (N=109)	Inclusion: Adolescent; 12-18 years of age; BN or partial BN; EBW >85% Exclusion: Current psychotic illness or other mental illness requiring hospitalization; bipolar I disorder; depression with active suicidal thoughts and behavior; associated physical illness that necessitates hospitalization; current dependence on drugs or alcohol; current diagnosis of AN or weight less than 85% EBW; physical conditions known to	BN: 85 (100%) - BN: 18 (43.9%) vs. 19 (48.7%) - Partial BN: 23 (56.1%) vs. 20 (51.3%) BN, Duration: 19.6 mo (SD ± 19.9) vs. 18.4 mo (SD ± 14.7) %EBW: 110.6% (SD ± 27.6) vs. 108.3% (SD ±15.0) Age 12 yr-18 yr: 109 (100%)	Compared with CBT, the FBT group had higher abstinence rates: 39.4% FBT vs. 19.7% CBT (p=0.040) at the end of treatment; 44% FBT vs. 25.4% CBT (p=0.030) at 6-mo follow-up. Binge Eating, Episodes - Baseline: 17.0/mo (SD ± 22.0) vs. 17.0/mo (SD ± 29.5) - End of Treatment: 4.1/mo (SD ± 7.4) vs. 7.8/mo (SD ± 21.5)	Low

		- FBT N=51 vs. CBT N=58	influence eating or weight; previous FBT, CBT, or SPT for BN; married or emancipated minors	Age: 15.9 yr (SD ± 1.5) vs. 15.7 yr (SD ± 1.5) Gender - Female: 47 (92%) vs. 55 (95%) - Male: 4 (8%) vs. 3 (5%) Ethnicity - Minority: 50 (46%)	- 6-mo follow-up: 7.5/mo (SD ± 16.8) 6.7/mo (SD ± 16.7) - 12-mo follow-up: 6.7/mo (SD ± 19.0) 5.8/mo (SD ± 9.5) Purging, Episodes - Baseline: 28.0/mo (SD ± 28.0) vs. 33.0/mo (SD ± 37.5) - End of Treatment: 7.6/mo (SD ± 10.9) vs. 13.2/mo (SD ± 21.5) - 6-mo follow-up: 10.0/mo (SD ± 16.5) vs. 11.5/mo (SD ± 17.7) - 12-mo follow-up: 7.0/mo (SD ± 11.4) 7.0/mo (SD ± 10.8) More subjects were hospitalized in CBT (N=12, 21%) than in FBT (N=1, 2%) (p=0.015). Attrition at Posttreatment: 17% (9/52) vs. 28% (15/58)	

Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EBW=expected body weight; FBT=family-based treatment; mo=month; RCT=randomized controlled trial; SD=standard deviation; SPT=supportive psychotherapy; yr=year

Compared to psychoanalytic psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Poulsen et al. (2014); Daniel et al. (2016); Folke et al. (2016)	Design: RCT; Secondary Analysis; Post-hoc Analysis Setting: Outpatient: University of Copenhagen Country: Denmark	Randomized N=70 CBT-E 5 mo (N=36) - Binge Eating and Purging (N=6) - Binge Eating and Purging, Abstinence (N=15)	Inclusion: Age at least 18 years; being available for the duration of the longer of the 2 treatments; BN Exclusion: Severe physical or psychiatric conditions that interfere with treatment; psychosis; pregnancy; current psychotherapeutic treatment;	BN: 70 (100%) Binge Eating, Objective: 25/28 days (SD ± 21.111) Purging: 35/28 days (SD ± 29.259) Eating Disorder, Duration: 12.3 yr (SD ± 6.2, N=69)	Binge eating and purging had stopped in more patients treated with CBT than psychoanalytic therapy: 42% vs. 6% at 5 mo (OR 13.4, 95% CI 2.45 – 73.42); 44% vs. 15% at 24 mo (OR 4.34, 95% CI 1.33 – 14.21). CBT reduced significantly more binge eating and purging	High

	Funding: Government, industry, and non-profit	Psychoanalytic Psychotherapy 24 mo (N=34) Follow-up: Baseline – 24 mo Follow-up (N=51) - 28 vs. 23	difficulty speaking Danish; difficulty understanding Danish	- 11.6 yr (SD ± 6.2) vs. 13 yr (SD ± 2) History of AN: 25 (37.3%, N=67) - 13 (38.2%, N=34) vs. 12 (36.4%, N=33) BMI: 22.6 kg/m² (SD ± 2.33) - 22.94 kg/m² (SD ± 2.49) vs. 22.24 kg/m² (SD ± 2.11) Age >= 18 yr: 70 (100%) Age: 25.8 yr (SD ± 4.9) - 25.7 yr (SD ± 5.4) vs. 25.8 yr (SD ± 4.3) Gender - Female: 35 (97.2%) vs. 34 (100%) - Male: 1 (2.8%) vs. 0 (0%) Race: NR	episodes than psychoanalytic therapy both at the end of treatment and follow-up. Binge Eating, Objective: - Baseline->5 mo: 23->2.29 /28 days vs. 28->18.32/28 days (MD -16.03/28 days, p<0.001) - 24 mo: 1.14/28 days vs. 5.07/28 days (MD -3.93/28 days, p=0.038) Purging: - Baseline->5 mo: 31->4.4/28 days vs. 35->20.61/28 days (MD -16.21/28 days, p<0.001) - 24 mo: 2.02/28 days vs. 9.44/28 days (MD -7.42/28 days, p=0.009) Attrition: 22% (8/36) vs. 29% (10/34)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT-E=enhanced cognitive-behavioral therapy; CI=confidence interval; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to psychodynamic therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Stefini et al. (2017)	Design: RCT Setting: Outpatient Country: Germany Funding: NR	Randomized N=81 CBT 12 mo (N=39) PDT 12 mo (N=42) Follow-up: Baseline – 24 mo	Inclusion: Female; age 14-20 years; BN or partial BN Exclusion: AN, concurrent; severe physical or mental conditions; current psychosis; alcohol abuse; drug abuse; drug addiction; suicidality; ADHD; intelligence quotient of <80;	BN or BN, Partial: 81 (100%) BN: 63 (77.78%) - 29 (74.36%) vs. 34 (80.95%) BN, Partial: 18 (22%) - 10 (25.64%) vs. 8 (19.05%)	Adherence to core features of the specific psychotherapy was less with PDT (1.16 units (SD ± 0.38) vs. 0.92 units (SD ± 0.43)) whereas dropouts were non-significantly greater with CBT (39% vs. 21%).	High

			current psychotropic or psychotherapeutic treatment	BN, Duration: 3.8 yr (SD ± 2.8) - 3.5 yr (SD ± 1.9) vs. 4.1 yr (SD ± 3.3) Binge Eating, Objective: 16.27/28 days (SD ± 15.1) Purging, Objective: 21.99/28 days (SD ± 23.1) History of AN: 15 (18.52%) - 9 (23.08%) vs. 6 (14.29%) History of Binge Eating: (2.56%) vs. 0 (0%) Age 14 yr-20 yr: 81 (100%) Age: 18.7 yr (SD ± 1.9) - 18.8 yr (SD ± 2.3) vs. 18.6 yr (SD ± 1.4) Gender, Female: 81 (100%) Race: NR	Primary outcome of remission at 12 mo did not differ: 13 (33.3%) CBT vs. 12 (30.2%) PDT (OR 1.12, 95% CI 0.44 – 2.84, p=0.81). Rates of remission were stable or improved at 1-yr follow-up: 15 (38.5%) vs. 13 (31%) (p=0.48). For both groups, numbers meeting full BN criteria also decreased (29->15 CBT vs. 34->16 PDT). Overall effect size was 1.20. Secondary measures (binge and purge behavior, EDE, and general pathology scores) improved with both treatments. Attrition: 39% (15/39) vs. 21% (9/42)	

Abbreviations: ADHD=attention-deficit/hyperactivity disorder; AN=anorexia nervosa; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EDE=Eating Disorder Examination; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; PDT=psychodynamic therapy; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to supportive psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Le Grange et al. (2015)	Design: RCT Setting: Multi-center	Randomized N=130 FBT 6 mo (N=52)	Inclusion: Adolescent; 12-18 years of age; BN or partial BN; EBW >85% Exclusion: Current psychotic illness or other mental illness	BN: 85 (100%) - BN: 18 (43.9%) vs. 19 (48.7%) - Partial BN: 23 (56.1%) vs. 20 (51.3%)	Compared with CBT, the FBT group had higher abstinence rates: 39.4% FBT vs. 19.7% CBT (p=0.040) at the end of treatment; 44% FBT vs. 25.4%	Low

	Country: United States Funding: Government	CBT 6 mo (N=58) SPT 6 mo (N=20) Follow-up: Baseline – 18 mo Current Analysis (N=109) FBT N=51 vs. CBT N=58	requiring hospitalization; bipolar I disorder; depression with active suicidal thoughts and behavior; associated physical illness that necessitates hospitalization; current dependence on drugs or alcohol; current diagnosis of AN or weight less than 85% EBW; physical conditions known to influence eating or weight; previous FBT, CBT, or SPT for BN; married or emancipated minors	BN, Duration: 19.6 mo (SD ± 19.9) vs. 18.4 mo (SD ± 14.7) %EBW: 110.6% (SD ± 27.6) vs. 108.3% (SD ±15.0) Age 12 yr-18 yr: 109 (100%) Age: 15.9 yr (SD ± 1.5) vs. 15.7 yr (SD ± 1.5) Gender - Female: 47 (92%) vs. 55 (95%) - Male: 4 (8%) vs. 3 (5%) Ethnicity Minority: 50 (46%)	CBT (p=0.030) at 6-mo follow-up. Binge Eating, Episodes - Baseline: 17.0/mo (SD ± 22.0) vs. 17.0/mo (SD ± 29.5) - End of Treatment: 4.1/mo (SD ± 7.4) vs. 7.8/mo (SD ± 21.5) - 6-mo follow-up: 7.5/mo (SD ± 16.8) 6.7/mo (SD ± 16.7) - 12-mo follow-up: 6.7/mo (SD ± 19.0) 5.8/mo (SD ± 9.5) Purging, Episodes - Baseline: 28.0/mo (SD ± 28.0) vs. 33.0/mo (SD ± 37.5) - End of Treatment: 7.6/mo (SD ± 10.9) vs. 13.2/mo (SD ± 21.5) - 6-mo follow-up: 10.0/mo (SD ± 16.5) vs. 11.5/mo (SD ± 17.7) - 12-mo follow-up: 7.0/mo (SD ± 11.4) 7.0/mo (SD ± 10.8) More subjects were hospitalized in CBT (N=12, 21%) than in FBT (N=1, 2%) (p=0.015). Attrition at Posttreatment: 17% (9/52) vs. 28% (15/58)	
Walsh et al. (1997); Wilson et al. (1999)	Design: RCT Setting: Outpatient Country: NR	Randomized N=120 CBT + Placebo 16 wk (N=25)	Inclusion: BN; women; 18-45 years of age; self-induced vomiting as a primary method of compensating for binge eating;	BN: 120 (100%) BN, Duration: 7.91 yr (SD ± 4.7) - 8 yr vs. 7.26 yr vs. 7.55 yr vs. 9.55 yr vs. 7.36 yr	CBT reduced binge eating and vomiting more than SPT. CBT plus meds was superior to medication alone, but SPT plus meds was not.	High

	Funding: Government	CBT + Desipramine NR-300 mg 10 wk > Desipramine 200-300 mg / Fluoxetine 60 mg 16 wk (N=23) SPT + Placebo 16 wk (N=22) SPT + Desipramine NR-300 mg 10 wk > Desipramine / Fluoxetine 60 mg 16 wk (N=22) Desipramine NR-300 mg 10 wk > Desipramine 200-300 mg/ Fluoxetine 60 mg 16 wk (N=28) CBT 4 mo (pooled) (N=32) Desipramine 200-300 mg 10 wk > Fluoxetine 60 mg 4 mo (pooled) (N=32) SPT 4 mo (pooled) (N=35)	weight was between 80% and 120% of IBW Exclusion: Medically ill; evidence of cardiac conduction disease; pregnant; abused drugs or alcohol within the past yr; acutely suicidal; previous adverse reaction to desipramine or fluoxetine	Vomiting, Self-Induced: 120 (100%) %IBW 80%-120%: 120 (100%) Weight: 130 lbs (SD ± 15) BMI: 21.9 kg/m² (SD ± 2.2 History of AN: 9 (36%) vs. 6 (27%) vs. 6 (27%) vs. 7 (32%) vs. 9 (32%) Age 18 yr-45 yr: 120 (100%) Age: 26.1 yr (SD ± 4.9) - 25.8 yr vs. 26.1 yr vs. 26.9 yr vs. 28 yr vs. 24.3 yr Gender, Female: 120 (100%) Race - Caucasian: 100 (83%) - Black or African American: 7 (6%) - Asian: 6 (5%) Ethnicity, Hispanic/Latino: 7 (6%)	Binge Eating – Baseline-> 16 wk: 7.22->2.56/wk vs. 7.29->0.95 vs. 6.18->3.32 vs. 7.92->3.57 vs. 8.32->2.59 - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine at 16 wk: MD -1.64/wk (p=0.04) Vomiting, Diary – Baseline-> 16 wk: 10.8->5.6/wk vs. 10.8->1.1/wk vs. 11.9->7.5/wk vs. 10.6->5.5/wk vs. 10.5->3.7 - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine16 wk: MD -2.6/wk (p=0.01) Binge eating and depression were improved more with medication than placebo plus psychological treatment. Treatment Adherence, Treatment Sessions, Fulfilled - Baseline – 16 wk: 16.5 vs. 16.8 vs. 17.7 vs. 17.8 vs. 11.5 - CBT / Supportive Psychotherapy +/- Desipramine / Fluoxetine (pooled) vs. Desipramine / Fluoxetine: MD 5.7 (p=0.0001) Attrition: 36% (9/25) vs. 35% (8/23) vs. 27% (6/22) vs. 27% (6/22) vs. 43% (12/28)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EBW=expected body weight; FBT=family-based treatment; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SPT=supportive psychotherapy; wk=week; yr=year

Compared to integrative cognitive affective therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Wonderlich et al. (2014)	Design: RCT Setting: Outpatient Country: United States Funding: Government	Randomized N=80 CBT-E 19 wk (N=40) Integrative Cognitive-Affective Therapy 19 wk (N=40) Follow-up (N=68) - 34 vs. 34 Follow-up: Baseline – 4 mo	Inclusion: BN; compensatory behaviors; subjective bulimic episodes at least weekly for 3 months prior to enrollment; adults Exclusion: Pregnancy; lactation; BMI<18 kg/m^2; lifetime diagnosis of bipolar disorder; current diagnosis of substance use disorder; medical instability; psychiatric instability; lifetime diagnosis of psychotic disorder; acute suicide risk; current psychotherapy	BN: 80 (100%) BN, Subthreshold: 22 (27.5%) - 11 (27.5%) vs. 11 (27.5%) BN >= 1 episodes/wk, In the Previous 3 mo: 80 (100%) Compensatory Behavior: 80 (100%) Binge Eating and Purging, Abstinence: 0 (0%) vs. 1 (2.5%) BMI: 23.9 kg/m² (SD ± 5.5) Age >= 18 yr: 80 (100%) Age: 27.3 yr (SD ± 9.6) - 28.8 yr (SD ± 10.8) vs. 25.8 yr (SD ± 8.2) Gender - Female: 36 (90%) vs. 36 (90%) - Male: 4 (10%) vs. 4 (10%) Race - Caucasian: 70 (87.5%) - Asian: 5 (6.3%) - Black or African American: 1 (1.3%)	Both treatments led to significant improvements in symptoms with no significant differences between the treatments. At the end of treatment, binge eating and purging abstinence rates were 22.5% with CBT-E vs. 37.5% with integrative cognitive-affective therapy. At 4-mo follow-up, rates were 22.5% and 32.5% respectively. Binge Eating, Objective - Baseline: 22.4/wk (SD ± 21) vs. 23.2/wk (SD ± 19.6) Binge Eating, Objective, % Change - Baseline – 19 wk: -76.3% vs. -73.7% - Baseline – 4 mo: -62.1% vs. -75.9% Purging - Baseline: 30.5/wk (SD ± 32.6) vs. 30.6/wk (SD ± 27) Purging, % Change - Baseline – 19 wk: -75.7% vs. -72.9% - Baseline – 4 mo: -66.9% vs. -71.9% Attrition: 25% (10/40) vs. 15% (6/40)	High

				- Native American: 1 (1.3%) Ethnicity - Hispanic/Latino: 2 (2.5%) - Ethnicity, Other: 1 (1.3%)		

Abbreviations: BMI=body mass index; BN=bulimia nervosa; CBT-E=enhanced cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to Pharmacotherapy

Compared to fluoxetine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Goldbloom et al. (1997)	Design: RCT Setting: Single Center: Toronto Hospital Country: Canada Funding: Industry	Randomized N=76 CBT 16 wk (N=24) Fluoxetine 60 mg + CBT 16 wk (N=29) Fluoxetine 60 mg 16 wk (N=23) Follow-up: Baseline – 20 wk Current Analysis (N=38) - 14 vs. 12 vs. 12	Inclusion: Females; 18-45 years of age; 85-125% matched population mean weight; BN; binge and vomit frequency of at least twice/wk; minimum 6-mo duration of bulimia Exclusion: Psychosis; ongoing pharmacotherapy or psychotherapy; use of MAOIs within 2 weeks prior to the onset of the study treatment; immediate suicide risk	BN: 76 (100%) BN, Duration >= 6 mo: 76 (100%) Binge Eating and Purging >= 2 episodes/wk: 76 (100%) Percent Average Body Weight, Matched-Population 85%-125%: 76 (100%) Age 18 yr-45 yr: 76 (100%) Age: 25.8 yr (SD ± 5.5, N=38) Gender, Female: 76 (100%) Race: NR	All treatments led to clinically significant improvement with some benefit of CBT + fluoxetine over fluoxetine alone but not over CBT alone. Binge Eating, Objective - Baseline->20 wk: 33.6->7.4/mo vs. 29.6->1.8 vs. 21->10 - Fluoxetine + CBT vs. Fluoxetine at 20 wk: MD -8.2/mo (p<0.03) Vomiting - Baseline->20 wk: 41.8->9/mo vs. 30.9->3.3 vs. 24.6->17.3 - CBT vs. Fluoxetine at 20 wk: MD -8.3/mo (p<0.07) - Fluoxetine + CBT vs. Fluoxetine at 20 wk: MD -14/mo (p<0.03) Binge Eating or Vomiting, Abstinence - 20 wk: 6 (43%,	High

					N=14) vs. 3 (25%, N=12) vs. 2 (17%, N=12) Attrition: 33% (8/24) vs. 55% (16/29) vs. 39% (9/23)	
Jacobi et al. (2002)	Design: RCT Setting: Outpatient: Department of Psychology at the University of Hamburg Country: Germany Funding: Industry	Randomized N=89 Current Analysis N=53 Group CBT 4 mo (N=19) Group CBT + Fluoxetine 60 mg (induction 20-40 mg) 4 mo (N=18) Fluoxetine 60 mg 4 mo (induction 20-40 mg) (N=16) Follow-up: Baseline – 16 mo	Inclusion: Women; 18-65 years of age; BN; minimum of 2 episodes of binge eating and vomiting for at least 6 months prior to the beginning of the study; actual BMI 17.5-25 kg/m² Exclusion: Concurrent severe psychiatric disturbance; concurrent psychosis or depression with suicidal risk; concurrent alcohol or drug abuse; concurrent involvement in other treatment; concurrent use of other medication	BN: 89 (100%) Binge Eating and Purging >= 2 episodes, In the Previous >= 6 mo: 89 (100%) BMI 17.5 kg/m²-25 kg/m²: 89 (100%) BMI: 20.6 kg/m² (SD ± 2, N=53) Age 18 yr-65 yr: 89 (100%) Age: 26 yr (SD ± 5.8, N=53) Gender, Female: 53 (100%) Race: NR	Baseline mean binges/28 days were 54.2 in the fluoxetine group vs. 36.5 and 33.5 in the CBT and fluoxetine + CBT groups respectively. All treatments led to significant improvements in eating disorder symptoms and in other psychological disturbances. Binge eating abstinence rates for completers were highest for CBT at both post-treatment and follow-up: - 4 mo: 5 (26%) vs. 3 (17%) vs. 2 (13%) - 16 mo: 4 (40%, N=10) vs. 1 (11%, N=9) vs. 1 (13%, N=8) At the end of treatment, vomiting abstinence was greater for CBT (37%) than for fluoxetine (6%) (p=0.046) or fluoxetine + CBT (6%) (p=0.041). Drug Discontinuation, Adverse Events - Baseline – 4 mo: 5 (27.78%) vs. 4 (25%) Attrition: 42% (8/19) vs. 33% (6/18) vs. 25% (4/16)	High
Mitchell et al. (2001)	Design: RCT Setting: Single Center: University of Minnesota	Randomized N=91	Inclusion: BN; female; at least 18 years of age; at least 85% of IBW; binge eating three times a wk for the last 6 months; self-	BN: 91 (100%)	Active treatments reduced binge eating and vomiting as compared to placebo.	High

	Hospital Eating Disorder Program Country: United States Funding: Government, industry, and non-profit	Fluoxetine 60mg 16 wk (N=26) SH Manual (manual-based CBT) + Placebo 16 wk (N=22) Fluoxetine 60mg + SH Manual (manual-based CBT) 16 wk (N=21) Placebo 16 wk (N=22)	induced vomiting 3 times a wk for the last 6 months Exclusion: Receiving pharmacotherapy or psychotherapy; current medical condition that would preclude safe outpatient treatment; history of hypersensitivity to fluoxetine; prior exposure to fluoxetine in a total amount greater than 140 mg; prior exposure to fluoxetine within the preceding 5 weeks before entering the study	Binge Eating 3 episodes/wk, In the Previous 6 mo: 91 (100%) Vomiting, Self-Induced 3 episodes/wk, In the Previous 6 mo: 91 (100%) %IBW >= 85%: 91 (100%) Age >= 18 yr: 91 (100%) Age: 26.6 yr (SD ± 7.1) - 26.6 yr (SD ± 7.1) vs. 26.8 yr (SD ± 6.9) vs. 29.3 yr (SD ± 7.8) vs. 23.8 yr (SD ± 6.1) Gender, Female: 91 (100%) Race, Caucasian: 25 (100%, N=25) vs. 22 (100%) vs. 20 (95.2%) vs. 21 (95.5%)	Binge Eating – Baseline: 11.58/wk (SD ± 6.74) vs. 11.91/wk (SD ± 10.7) vs. 11.29/wk (SD ± 5.87) vs. 9.45/wk (SD ± 5.34) Binge Eating, % Change - Baseline – 16 wk: -50.3% (SD ± 52.6) vs. -59.7% (SD ± 39.6) vs. -66.8% (SD ± 29.9, N=20) vs. -32.4% (SD ± 66.7, N=21) Vomiting – Baseline: 16.81/wk (SD ± 27.72) vs. 13.86/wk (SD ± 10.81) vs. 12.43/wk (SD ± 6.92) vs. 11.77/wk (SD ± 6.67) Vomiting, % Change - Baseline – 16 wk: -52.8% (SD ± 50.7) vs. -50.2% (SD ± 55) vs. -66.7% (SD ± 31.2, N=20) vs. -22.8% (SD ± 56.1, N=21) Attrition: 4% (1/26) vs. 5% (1/22) vs. 10% (2/21) vs. 18% (4/22)	
Walsh et al. (1997); Wilson et al. (1999)	Design: RCT Setting: Outpatient Country: NR Funding: Government	Randomized N=120 CBT + Placebo 16 wk (N=25) CBT + Desipramine NR-300 mg 10 wk > Desipramine 200-300 mg / Fluoxetine 60 mg 16 wk (N=23) Supportive Psychotherapy + Placebo 16 wk (N=22) Supportive Psychotherapy +	Inclusion: BN; women; 18-45 years of age; self-induced vomiting as a primary method of compensating for binge eating; weight was between 80% and 120% of IBW Exclusion: Medically ill; evidence of cardiac conduction disease; pregnant; abused drugs or alcohol within the past yr; acutely suicidal; previous adverse reaction to desipramine or fluoxetine	BN: 120 (100%) BN, Duration: 7.91 yr (SD ± 4.7) - 8 yr vs. 7.26 yr vs. 7.55 yr vs. 9.55 yr vs. 7.36 yr Vomiting, Self-Induced: 120 (100%) %IBW 80%-120%: 120 (100%) Weight: 130 lbs (SD ± 15)	CBT reduced binge eating and vomiting more than supportive psychotherapy. CBT plus meds was superior to medication alone, but supportive psychotherapy plus meds was not. Binge Eating – Baseline-> 16 wk: 7.22->2.56/wk vs. 7.29->0.95 vs. 6.18->3.32 vs. 7.92->3.57 vs. 8.32->2.59 - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine at 16 wk: MD -1.64/wk (p=0.04)	High

		Desipramine NR-300 mg 10 wk > Desipramine / Fluoxetine 60 mg 16 wk (N=22) Desipramine NR-300 mg 10 wk > Desipramine 200-300 mg/ Fluoxetine 60 mg 16 wk (N=28) CBT 4 mo (pooled) (N=32) Desipramine 200-300 mg 10 wk > Fluoxetine 60 mg 4 mo (pooled) (N=32) Supportive Psychotherapy 4 mo (pooled) (N=35)		BMI: 21.9 kg/m² (SD ± 2.2) History of AN: 9 (36%) vs. 6 (27%) vs. 6 (27%) vs. 7 (32%) vs. 9 (32%) Age 18 yr-45 yr: 120 (100%) Age: 26.1 yr (SD ± 4.9) - 25.8 yr vs. 26.1 yr vs. 26.9 yr vs. 28 yr vs. 24.3 yr Gender, Female: 120 (100%) Race - Caucasian: 100 (83%) - Black or African American: 7 (6%) - Asian: 6 (5%) Ethnicity, Hispanic/Latino: 7 (6%)	Vomiting, Diary – Baseline-> 16 wk: 10.8->5.6/wk vs. 10.8->1.1/wk vs. 11.9->7.5/wk vs. 10.6->5.5/wk vs. 10.5->3.7 - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine16 wk: MD -2.6/wk (p=0.01) Binge eating and depression were improved more with medication than placebo plus psychological treatment. Treatment Adherence, Treatment Sessions, Fulfilled - Baseline – 16 wk: 16.5 vs. 16.8 vs. 17.7 vs. 17.8 vs. 11.5 - CBT / Supportive Psychotherapy +/- Desipramine / Fluoxetine (pooled) vs. Desipramine / Fluoxetine: MD 5.7 (p=0.0001) Attrition: 36% (9/25) vs. 35% (8/23) vs. 27% (6/22) vs. 27% (6/22) vs. 43% (12/28)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; IBW=ideal body weight; MAOI=monoamine oxidase inhibitor; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; yr=year

Compared to desipramine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1992, 1994a)	Design: RCT; Follow-up Setting: NR	Randomized N=71 Desipramine HCl 25-350 mg (titrate) 16 wk (N=12)	Inclusion: Women; aged 18-65 years; BN Exclusion: Concurrent medical condition that would preclude the use of antidepressants; evidence of conduction	BN: 71 (100%) Binge Eating: 5.5/wk (SD ± 4.6) vs. 5.9/wk (SD ± 5.1) vs. 7.5/wk (SD ± 3.4) vs. 9.3/wk	At 16 wk, both CBT and combined treatment were superior to medication given for 16 weeks in reducing binge eating and purging.	High

	Location: NR Funding: Government	Desipramine HCl 25-350 mg (titrate) 24 wk (N=12) Desipramine HCl 25-350 mg (titrate) + CBT 16 wk (N=12) Desipramine HCl 25-350 mg (titrate) + CBT 16 wk > (-) CBT 24 wk (N=12) CBT 24 wk (N=23) Follow-up: Baseline – 72 wk	disturbance on electrocardiography; current AN; drug or alcohol abuse; psychosis; depression with suicidal risk of sufficient severity to preclude the sue of antidepressants on an outpatient basis	(SD ± 5.8) vs. 8.7/wk (SD ± 7.2) Purging: 9.7/wk (SD ± 9.4) vs. 6.3/wk (SD ± 4.9) vs. 8.3/wk (SD ± 4.3) vs. 11.7/wk (SD ± 5.9) vs. 10.1/wk (SD ± 7.7) Age 18 yr-65 yr: 71 (100%) Gender, Female: 71 (100%) Race: NR	Binge Eating, % Change - Baseline – 16 wk: -34% vs. -40% vs. -67% vs. -79% vs. -81.7% - CBT vs. Desipramine 16 wk/24 wk (pooled) (MD -42.9%, p<0.005) - Desipramine + CBT 16 wk > (+/-) Desipramine 24 wk (pooled) vs. Desipramine 16 wk/24 wk (pooled) (MD -43.8%, p<0.004) Purging, % Change - Baseline – 16 wk: -52% vs. -38% vs. -69% vs. -89% vs. -82.6% - CBT vs. Desipramine 16 wk/24 wk (pooled) (MD -39.9%, p<0.004) - Desipramine 16 wk/24 wk (pooled) vs. Desipramine + CBT 16 wk > (+/-) Desipramine 24 wk (pooled) (MD 38.2%, p<0.003) At 32 wk, only combined 24-wk treatment was superior to medication given for 16 weeks (-35% vs. -45% vs. -60% vs. -90% vs. -78%). Continuing CBT appeared to prevent relapse in patients withdrawn from medication at 16 wk. At 1-yr follow-up, combined 24-wk treatment and CBT alone were significantly superior in reducing binge eating to desipramine given for 16 weeks: -22% (N=11) vs. -67% (N=9) vs. -55% (N=10) vs. -95% (N=9) vs. -72% (N=22).	

					Only 18% (2 of 11) of those receiving 16 weeks of desipramine were free of binge eating and purging at follow-up compared with 78% (7 of 9) of those receiving the combined 24-wk treatment: 2 (18%, N=11) vs. 6 (67%, N=9) vs. 4 (40%, N=10) vs. 7 (78%, N=9) vs. 12 (54%, N=22) Attrition: 8% (1/12) vs. 25% (3/12) vs. 17% (2/12) vs. 25% (3/12) vs. 4% (1/23)	
Walsh et al. (1997); Wilson et al. (1999)	Design: RCT Setting: Outpatient Country: NR Funding: Government	Randomized N=120 CBT + Placebo 16 wk (N=25) CBT + Desipramine NR-300 mg 10 wk > Desipramine 200-300 mg / Fluoxetine 60 mg 16 wk (N=23) Supportive Psychotherapy + Placebo 16 wk (N=22) Supportive Psychotherapy + Desipramine NR-300 mg 10 wk > Desipramine / Fluoxetine 60 mg 16 wk (N=22) Desipramine NR-300 mg 10 wk > Desipramine 200-300	Inclusion: BN; women; 18-45 years of age; self-induced vomiting as a primary method of compensating for binge eating; weight was between 80% and 120% of IBW Exclusion: Medically ill; evidence of cardiac conduction disease; pregnant; abused drugs or alcohol within the past yr; acutely suicidal; previous adverse reaction to desipramine or fluoxetine	BN: 120 (100%) BN, Duration: 7.91 yr (SD ± 4.7) - 8 yr vs. 7.26 yr vs. 7.55 yr vs. 9.55 yr vs. 7.36 yr Vomiting, Self-Induced: 120 (100%) %IBW 80%-120%: 120 (100%) Weight: 130 lbs (SD ± 15) BMI: 21.9 kg/m² (SD ± 2.2) History of AN: 9 (36%) vs. 6 (27%) vs. 6 (27%) vs. 7 (32%) vs. 9 (32%) Age 18 yr-45 yr: 120 (100%) Age: 26.1 yr (SD ± 4.9) - 25.8 yr vs. 26.1 yr vs. 26.9 yr vs. 28 yr vs. 24.3 yr	CBT reduced binge eating and vomiting more than supportive psychotherapy. CBT plus meds was superior to medication alone, but supportive psychotherapy plus meds was not. Binge Eating – Baseline-> 16 wk: 7.22->2.56/wk vs. 7.29->0.95 vs. 6.18->3.32 vs. 7.92->3.57 vs. 8.32->2.59 - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine at 16 wk: MD -1.64/wk (p=0.04) Vomiting, Diary – Baseline-> 16 wk: 10.8->5.6/wk vs. 10.8->1.1/wk vs. 11.9->7.5/wk vs. 10.6->5.5/wk vs. 10.5->3.7 - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine16 wk: MD -2.6/wk (p=0.01) Binge eating and depression were improved more with	High

		mg/ Fluoxetine 60 mg 16 wk (N=28) CBT 4 mo (pooled) (N=32) Desipramine 200-300 mg 10 wk > Fluoxetine 60 mg 4 mo (pooled) (N=32) Supportive Psychotherapy 4 mo (pooled) (N=35)		Gender, Female: 120 (100%) Race - Caucasian: 100 (83%) - Black or African American: 7 (6%) - Asian: 6 (5%) Ethnicity, Hispanic/Latino: 7 (6%)	medication than placebo plus psychological treatment. Treatment Adherence, Treatment Sessions, Fulfilled – Baseline – 16 wk: 16.5 vs. 16.8 vs. 17.7 vs. 17.8 vs. 11.5 - CBT / Supportive Psychotherapy +/- Desipramine / Fluoxetine (pooled) vs. Desipramine / Fluoxetine: MD 5.7 (p=0.0001) Attrition: 36% (9/25) vs. 35% (8/23) vs. 27% (6/22) vs. 27% (6/22) vs. 43% (12/28)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; wks=weeks; yr=year

Compared to imipramine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Mitchell et al.1990; Keel et al. (2002)	Design: RCT; Follow-up/Extension Setting: Outpatient: Eating Disorders Clinic; University of Minnesota Country: United States Funding: Government and non-profit	Randomized N=171 Intensive Group Therapy + Placebo 10 wk (N=34) Intensive Group Therapy + Imipramine HCl 200-300 mg 10 wk (50 mg induction) (up-titrate) (N=52) Imipramine HCl 200-300 mg 10 wk (50 mg	Inclusion: 18-40 years of age; female; IBW 80% to 120%; BN, binge eating and purging Exclusion: Current involvement in psychotherapy or pharmacotherapy for BN; concurrent medical condition that would preclude safe outpatient therapy with an antidepressant; active abuse of alcohol or drugs in the past 6 months	BN, Purging Type: 171 (100%) BN, Duration: 6.2 yr (SD ± 4) vs. 7 yr (SD ± 4.9) vs. 6.5 yr (SD ± 2.9) vs. 6.4 yr (SD ± 3.3) History of Laxative Abuse or Laxative Abuse: 62 (40%) (N=155) - 8 (24%, N=33) vs. 22 (46%, N=48) vs. 20 (44%, N=45) vs. 12 (41%, N=29)	All three active treatments led to significant reductions in binge eating and purging and improvement in mood relative to placebo. Intensive group psychotherapy had more improvement than Imipramine alone, with no benefit of combination treatment on eating behaviors (though Imipramine did help depression and anxiety.) Binge Eating - Baseline	High

		induction) (up-titrate) (N=54) Placebo 10 wk (N=31) Imipramine HCl 200-300 mg / (Intensive Group Therapy + Imipramine HCl 200-300 mg) 10 wk (pooled) (N=106) Intensive Group Therapy / (Intensive Group Therapy + Imipramine HCl 200-300 mg) 10 wk (pooled) (N=86) Placebo / Imipramine HCl 200-300 mg 10 wk (pooled) (N=85) Placebo / Intensive Group Therapy 10 wk (pooled) (N=65) Current Analysis (N=155) - 33 vs. 48 vs. 45 vs. 29 Follow-up: Baseline – 10 yr Follow-up (N=101)		%IBW 80%-120%: 171 (100%) %IBW: 97.7% (SD ± 10.2) vs. 108.2% (SD ± 12.4) vs. 106.5% (SD ± 12.8) vs. 107.6% (SD ± 11.3) History of AN: 25 (16.13%, N=155) - 10 (30%, N=33) vs. 5 (10%, N=48) vs. 8 (18%, N=45) vs. 2 (7%, N=29) Age 18 yr-40 yr: 171 (100%) Age: 22.8 yr (SD ± 4.3) vs. 24.3 yr (SD ± 5.7) vs. 24.1 yr (SD ± 4.4) vs. 24.4 yr (SD ± 5.2) Gender, Female: 171 (100%) Race: NR	- 9.2/wk (N=33) vs. 8.4/wk (N=48) vs. 7.3/wk (N=45) vs. 8/wk (N=29) - 11.9 hr/wk (N=33) vs. 10.8 hr/wk (N=48) vs. 10.3 hr/wk (N=45) vs. 10.1 hr/wk (N=29) Binge Eating, Change - Baseline – 10 wk: -8.2/wk vs. -7.7/wk vs. -3.6/wk vs. -0.2/wk - Intensive Group Therapy vs. Imipramine: MD -4.6/wk, p=0.0001 -10.6 hr/wk vs. -9.7 hr/wk vs. -5.3 hr/wk vs. -1.7 hr/wk - Intensive Group Therapy vs. Imipramine: MD -5.3 hr/wk (p=0.0001) Purging – Baseline: 13.2/wk (N=33) vs. 9.6/wk (N=48) vs. 8.6/wk (N=45) vs. 10/wk (N=29) Purging, Change - Baseline – 10 wk: -11.2/wk vs. -8.6/wk vs. -3.9/wk vs. -1.2/wk - Intensive Group Therapy vs. Imipramine: MD -7.3/wk (p=0.0001) Binge Eating – Baseline->10 yr: 6.3->2.4/d vs. 5.9->2.5/d vs. 5.9->2.5/d vs. 5.6->3.4/d Vomiting– Baseline->10 yr: 6.4->2.3/d vs. 5.4->2.6/d vs. 5.7->2.4/d vs. 5.9->3.4/d	

					Laxative Abuse – Baseline->10 yr: 1.3->1/d vs. 2->1.2/d vs. 2.1->1.4/d vs. 1.9->1.3/d Attrition: 15% (5/34) vs. 25% (13/52) vs. 43% (23/54) vs. 16% (5/31)	
Pyle et al. (1990)	Design: Follow-up of RCT (Mitchell et al. 1990) Setting: NR Country: NR Funding: Government	Randomized N=68 Imipramine 200-300 mg 12 wk (N=3) Imipramine 200-300 mg + Intensive Support Group (Group CBT + Nutritional Counseling) 12 wk (N=19) Intensive Support Group (Group CBT + Nutritional Counseling) 12 wk (N=25) Placebo + Intensive Support Group (Group CBT + Nutritional Counseling) 12 wk (N=15) Placebo 12 wk (N=6) Follow-up: Baseline – 6 mo Follow-up (N=61) - 3 vs. 18 vs. 21 vs. 13 vs. 6	Inclusion: BN; history of binge eating at least 3 times a wk for 6 months; women; 18-40 years of age; responded to intensive group psychotherapy plus imipramine or placebo or to imipramine alone; history of self-induced vomiting or laxative abuse at least 3 times a wk for 6 months Exclusion: NR	BN: 68 (100%) Binge Eating >= 3 episodes/wk, In the Previous 6 mo: 68 (100%) Vomiting, Self-Induced >= 3 episodes/wk, Duration 6 mo or Laxative Abuse >= 3 episodes/wk, Duration 6 mo: 68 (100%) Age 18 yr-40 yr: 68 (100%) Gender, Female: 68 (100%) Race: NR	Although overall 30% relapsed by 6 mo, initial treatment with intensive group psychotherapy plus placebo or imipramine was associated with a lower relapse rate than imipramine alone: 2 (67%) vs. 4 (22%, N=18) vs. 3 (14%, N=21) vs. 4 (31%, N=13) vs. 5 (83%) Binge Eating, % Change - -10 wk – 6 mo: -100% vs. -88% (N=18) vs. -92% (N=21) vs. -94% (N=13) vs. -95% Bulimic Episodes, Abstinence - 6 mo: 1 (33%) vs. 11 (61%, N=18) vs. 13 (62%, N=21) vs. 5 (38%, N=13) vs. 1 (17%) Disease Response, Remission - 6 mo: 1 (33%) vs. 13 (72%, N=18) vs. 17 (81%, N=21) vs. 7 (54%, N=13) vs. 1 (17%) Attrition: 0% (0/3) vs. 6% (1/19) vs. 19% (4/25) vs. 15% (2/15) vs. 0% (0/6)	High

Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EDE=Eating Disorder Examination; HCl=hydrochloride; hr=hour; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; wks=weeks; yr=year

Family-Based Treatment Compared to Supportive Psychotherapy

Le Grange et al. (2007)	Design: RCT Setting: Outpatient Country: United States Funding: Government	Randomized N=80 FBT 6 mo (N=41) SPT 6 mo (N=39) Follow-up: Baseline – 12 mo	Inclusion: Adolescent; 12-19 years of age; BN or partial BN; Exclusion: Physical or psychiatric disorder necessitating hospitalization; insufficient knowledge of English; current physical dependence on drugs or alcohol; current low body weight (BMI =< 17.5); current treatment for the eating disorder or current use of medication known to affect eating or weight; and physical conditions (e.g., diabetes mellitus or pregnancy) or treatments known to influence eating or weight; 50 mg or more of fluoxetine	BN: 80 (100%) - BN: 18 (43.9%) vs. 19 (48.7%) - Partial BN: 23 (56.1%) vs. 20 (51.3%) BN, Duration: 22.3 mo (SD ± 20.4) vs. 20.1 mo (SD ± 24.4) BMI: 21.8 kg/m² (SD ± 2.5) vs. 22.4 kg/m² (SD ± 3.4) Age 12 yr-19 yr: 80 (100%) Age: 16 yr (SD ± 1.7) vs. 16.1 yr (SD ± 1.6) Gender - Female: 40 (97.6%) vs. 38 (97.4%) - Male: 1 (2.4%) vs. 1 (2.6%) Race - Caucasian: 31 (75.6%) vs. 20 (51.2%) - African American: 4 (9.8%) vs. 5 (12.8%) - Other: 0 (0%) vs. 4 (10.3%) Ethnicity - Hispanic: 6 (14.6%) vs. 10 (25.6%)	Compared with SPT, remission rates were significantly higher for FBT: 16 (39%) vs. 7 (18%) (p=0.049) at post-treatment; 12 (29%) vs. 4 (10%) (p=0.05) at 6-mo follow-up. Binge Eating, Objective - Baseline: 18.4/mo (SD ± 28.1) vs. 18.9/mo (SD ± 22.3) - Post-treatment: 4.1/mo (SD ± 14.8) vs. 3.2/mo (SD ± 5.1) - 12 mo: 2.5/mo (SD ± 6.8) 5.4/mo (SD ± 13.7) Binge Eating, Subjective - Baseline: 9.9/mo (SD ± 16.6) vs. 7.6/mo (SD ± 10.1) - Post-treatment: 4.5/mo (SD ± 13.3) vs. 4.6/mo (SD ± 8.6) - 12 mo: 2.8/mo (SD ± 6.9) vs. 2.4/mo (SD ± 5.2) Vomiting - Baseline: 34.5/mo (SD ± 31.0) vs. 33.2/mo (SD ± 33.5) - Post-treatment: 4.8/mo (SD ± 9.4) vs. 17.4/mo (SD ± 26.0) - 12 mo: 10.1/mo (SD ± 21.8) vs. 14.5/mo (SD ± 27.7) Attrition: 12% (5/41) vs. 10% (4/39)	High

Abbreviations: BMI=body mass index; BN=bulimia nervosa; FBT=family-based treatment; mo=month; RCT=randomized controlled trial; SD=standard deviation; SPT=supportive psychotherapy; yr=year

Compared to Cognitive-Behavioral Therapy

Le Grange et al. (2015)	Design: RCT Setting: Multi-center Country: United States Funding: Government	Randomized N=130 FBT 6 mo (N=52) CBT 6 mo (N=58) SPT 6 mo (N=20) Follow-up: Baseline – 18 mo Current Analysis (N=109) - FBT N=51 vs. CBT N=58	Inclusion: Adolescent; 12-18 years of age; BN or partial BN; EBW >85% Exclusion: Current psychotic illness or other mental illness requiring hospitalization; bipolar I disorder; depression with active suicidal thoughts and behavior; associated physical illness that necessitates hospitalization; current dependence on drugs or alcohol; current diagnosis of AN or weight less than 85% EBW; physical conditions known to influence eating or weight; previous FBT, CBT, or SPT for BN; married or emancipated minors	BN: 85 (100%) - BN: 18 (43.9%) vs. 19 (48.7%) - Partial BN: 23 (56.1%) vs. 20 (51.3%) BN, Duration: 19.6 mo (SD ± 19.9) vs. 18.4 mo (SD ± 14.7) %EBW: 110.6% (SD ± 27.6) vs. 108.3% (SD ±15.0) Age 12 yr-18 yr: 109 (100%) Age: 15.9 yr (SD ± 1.5) vs. 15.7 yr (SD ± 1.5) Gender - Female: 47 (92%) vs. 55 (95%) - Male: 4 (8%) vs. 3 (5%) Ethnicity - Minority: 50 (46%)	Compared with CBT, the FBT group had higher abstinence rates: 39.4% FBT vs. 19.7% CBT (p=0.040) at the end of treatment; 44% FBT vs. 25.4% CBT (p=0.030) at 6-mo follow-up. Binge Eating, Episodes - Baseline: 17.0/mo (SD ± 22.0) vs. 17.0/mo (SD ± 29.5) - End of Treatment: 4.1/mo (SD ± 7.4) vs. 7.8/mo (SD ± 21.5) - 6-mo follow-up: 7.5/mo (SD ± 16.8) 6.7/mo (SD ± 16.7) - 12-mo follow-up: 6.7/mo (SD ± 19.0) 5.8/mo (SD ± 9.5) Purging, Episodes - Baseline: 28.0/mo (SD ± 28.0) vs. 33.0/mo (SD ± 37.5) - End of Treatment: 7.6/mo (SD ± 10.9) vs. 13.2/mo (SD ± 21.5) - 6-mo follow-up: 10.0/mo (SD ± 16.5) vs. 11.5/mo (SD ± 17.7) - 12-mo follow-up: 7.0/mo (SD ± 11.4) 7.0/mo (SD ± 10.8) More subjects were hospitalized in CBT (N=12, 21%) than in FBT (N=1, 2%) (p=0.015). Attrition at Post-treatment: 17% (9/52) vs. 28% (15/58)	Low

Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EBW=expected body weight; FBT=family-based treatment; mo=month; RCT=randomized controlled trial; SD=standard deviation; SPT=supportive psychotherapy; yr=year

Compared to Cognitive-Behavioral Therapy Guided Self-Help

Schmidt et al. (2007)	Design: RCT Setting: Multi-center, Outpatient: National Health Service Country: United Kingdom Funding: Non-profit	Randomized N=85 CBT Guided Self-Help 10 wk > 6 mo (N=44) Family Therapy 6 mo (N=41) Follow-up: Baseline – 12 mo	Inclusion: 13-20 years of age; BN or EDNOS; at least one close other to accompany them for family treatment Exclusion: BMI below 10th percentile for age and sex; knowledge of English insufficient to understand the treatment; learning disability; severe mental illness; substance dependence	BN or EDNOS: 85 (100%) - BN: 30 (68.2%) vs. 31 (75.6%) - EDNOS: 14 (31.8%) vs. 10 (24.4%) Binge Eating, Objective: 5.2/wk (SD ± 6.4) vs. 5.9/wk (SD ± 6.7) Vomiting, Objective: 9.5/wk (SD ± 11.7) vs. 9.9/wk (SD ± 17.9) BN, Age at Onset: 14.9 yr (SD ± 2.1) vs. 15.2 yr (SD ± 1.8) History of AN: 7 (16%) vs. 8 (20%) Age 13 yr-20 yr: 85 (100%) Age: 17.4 yr (SD ± 1.8) vs. 17.9 yr (SD ± 1.6) Gender - Female: 42 (95.5%) vs. 41 (100%) - Male: 2 (4.5%) vs. 0 (0%) Race, Caucasian: 30 (100%, N=30) vs. 31 (94%, N=33) Ethnicity, Other: 0 (0%, N=30) vs. (6%, N=33)	Binge Eating, Objective, Abstinence - Baseline: 8 (18%) vs. 8 (19.5%) - 6 mo: 13 (41.9%, N=31) vs. 8 (25%, N=32) - 12 mo: 13 (52%, N=25) vs. 16 (55%, N=29) Vomiting, Abstinence - Baseline: 9 (20.5%) vs. 6 (14.6%) - 6 mo: 10 (32.3%, N=31) vs. 9 (28%, N=32) - 12 mo: 14 (56%, N=25) vs. 15 (51.7%, N=29) Binge Eating and Purging, Abstinence - Baseline: 2 (4.5%) vs. 2 (5%) - 6 mo: 6 (19.4%, N=31) vs. 4 (12.5%, N=32) - 12 mo: 9 (36%, N=25) vs. 12 (41.4%, N=29) Hospitalization Costs - Baseline – 12 mo: 481.19 pounds (SD ± 1411.47) vs. 66.28 pounds (SD ± 149.66) Attrition: 30% (13/44) vs. 29% (12/41)	Low

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EDNOS= eating disorder not otherwise specified; mo=month; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Serotonin Reuptake Inhibitors

Fluoxetine

Compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Beumont et al. (1997)	Design: RCT Setting: Multi-center Country: Australia Funding: Industry	Randomized N= 67 Fluoxetine 60 mg + Nutritional Counselling 8 wk (N=34) Placebo + Nutritional Counselling 8 wk (N=33) Follow-up: Baseline – 20 wk	Inclusion: At least 18 years of age; women; BN; BMI between 20 and 25 kg/m^2 Exclusion: Presence of medical illness; psychosis; suicidal ideation; use of other psychotropic medication within 1 wk; use of fluoxetine within the previous 5 weeks	BN: 67 (100%) BMI 20 kg/m²-25 kg/m²: 67 (100%) BMI: 22 kg/m² (SD ± 2) Age >= 18 yr: 67 (100%) Age: 24.2 yr (SD ± 4.5) vs. 25.1 yr (SD ± 5.8) Gender, Female: 67 (100%) Race: NR	Both groups improved during treatment. In the fluoxetine group, there were some improvements on EDE scores and modest weight loss during active treatment, but recurrent symptoms and weight gain occurred post-treatment. Bulimic Episodes – Baseline->8 wk->20 wk: 9.5->1.8->2.2/wk vs. 6.2->1.2-> 1.9 Vomiting – Baseline->8 wk->20 wk: 8.8->1.2->2.5/wk vs. 7.3->2.3->2.3 Vomiting, % Change - Baseline – 8 wk: -86% vs. -69% Binge Eating, Abstinence - 8 wk: 24 (69.6%) vs. 20 (61.5%) - 20 wk: 12 (35.7%) vs. 20 (60.9%) Study Withdrawal - Baseline – 20 wk: 17 (50%) vs. 10 (30%) (p=0.3)	High

					Study Withdrawal, Adverse Events - Baseline – 20 wk: 4 (11.76%) vs. 0 (0%) Adverse Events, Severe - Baseline – 20 wk: 5 (14.71%) vs. 4 (12.12%) Attrition at 8 wk: 32% (11/34) vs. 21% (7/33)	
Fluoxetine Bulimia Nervosa Collaborative Study Group (1992); Goldbloom and Olmsted (1993); Goldstein et al. (1999)	Design: RCT; Post-hoc Analysis; Sub-Group Analysis Setting: Multi-center; outpatient Country: United States; Canada Funding: Industry	Randomized N=387 Fluoxetine 20 mg 8 wk (N=129) Fluoxetine 60 mg 8 wk (N=129) Placebo 8 wk (N=129) Current Analysis (N=382) - 128 vs. 127 vs. 127	Inclusion: Women; BN; at least age 18 years; weigh between 85% and 130% of the midpoint of IBW for height; at least three binge-eating episodes/week for at least 6 months Exclusion: Serious medical illness; psychosis; acute suicidal ideation; used psychoactive medications during the 2 weeks prior to the study; initiated some other treatment for BN during the mo prior to enrollment; initiated psychotherapy or behavior therapy during the mo prior to enrollment	BN: 387 (100%) Binge Eating >= 3 episodes/wk, In the Previous >= 6 mo: 387 (100%) Vomiting, Self-Induced: 320 (83%) %IBW 85%-130%: 387 (100%) BMI <= 21.8 kg/m²: 190 (49.22%, N=386) BMI > 21.8 kg/m²: 186 (48.19%, N=386) BMI: 22.7 kg/m² (SD ± 4.2) vs. 22.4 kg/m² (SD ± 3.2) vs. 22.6 kg/m² (SD ± 3.3) Age >= 18 yr: 387 (100%) Age: 27.4 yr (SD ± 7.2) vs. 26.4 yr (SD ± 6.2) vs. 27.7 yr (SD ± 8) Gender, Female: 387 (100%)	Fluoxetine 60 mg had a greater decrease in weekly binge eating (MD -34%, p<0.001) and vomiting episodes (MD -51%, p<0.001) as well as greater weight reduction (-1.6 kg, p<0.001) vs. placebo. Fluoxetine 20 mg had an intermediate effect. Binge Eating – Baseline: 8/wk (SD ± 5) vs. 11/wk (SD ± 10, N=128) vs. 11/wk (SD ± 8) Binge Eating, % Change - Baseline – 8 wk: -45% (N=128) vs. -67% (N=127) vs. -33% (N=127) - Fluoxetine 60 mg vs. Placebo: MD -34% (p<0.001) - Fluoxetine 60 mg vs. Fluoxetine 20 mg: MD -22% (p=0.003) - Fluoxetine 20 mg vs. Placebo: MD -12% (p=0.538) Binge Eating, Responder, Reduction 50%-100% - Baseline – 8 wk: 63 (49%, N=128) vs. 80 (63%, N=127) vs. 55 (43%, N=127)	High

				Race, Caucasian: 123 (95%) vs. 125 (97%) vs. 126 (98%)	- Fluoxetine 60 mg vs. Placebo: p=0.001 - Fluoxetine 60 mg vs. Fluoxetine 20 mg: p=0.003 - Fluoxetine 20 mg vs. Placebo: p=0.453 Vomiting – Baseline: 9/wk (SD ± 10) vs. 11/wk (SD ± 14, N=128) vs. 11/wk (SD ± 14) Vomiting, % Change - Baseline – 8 wk: -29% (N=128) vs. -56% (N=127) vs. -5% (N=127) - Fluoxetine 60 mg vs. Placebo: MD -51% (p<0.001) - Fluoxetine 60 mg vs. Fluoxetine 20 mg: MD -27% (p=0.014) - Fluoxetine 20 mg vs. Placebo: MD -24% (p=0.04) Vomiting, Responder, Reduction 50%-100% - Baseline – 8 wk: 58 (45%, N=128) vs. 72 (57%, N=127) vs. 33 (26%, N=127) - Fluoxetine 60 mg vs. Placebo: p=0.001 - Fluoxetine 60 mg vs. Fluoxetine 20 mg: p=0.011 - Fluoxetine 20 mg vs. Placebo: p=0.021 More adverse effects were reported with fluoxetine but similar rates of treatment withdrawal for adverse effects were noted. Study Withdrawal - Baseline – 8 wk - Adverse Events: 4 (3.1%) vs. 11 (8.53%) vs. 8 (6.2%)	

					- Lack of Efficacy: 5 (3.88%) vs. 1 (0.78%) vs. 11 (8.53%) Attrition: 23% (30/129) vs. 30% (39/129) vs. 37% (48/129)	
Goldstein et al. (1995, 1999)	Design: RCT; Sub-Group Analysis Setting: Multi-center; outpatient Country: United States Funding: NR	Randomized N=398 Fluoxetine 60 mg 16 wk (N=296) Placebo 16 wk (N=102) ITT (N=390) - 290 vs. 100	Inclusion: Outpatients; at least 18 years old; BN; with at least 3 vomiting episodes/wk after binge eating; a history of BN of at least 6 months Exclusion: Participated in a prior fluoxetine study; taken any fluoxetine within the 5 weeks before enrolment; taken a cumulative lifetime fluoxetine dose of more than 140 mg; psychosis; acute suicidality; organic brain disease; history of seizures; diagnosis of AN; medically unstable condition; allergy to fluoxetine; history of severe allergies; multiple adverse drug reactions; hypertension treated with methyldopa; hypertension treated with clonidine; hypertension treated with reserpine; hypertension treated with guanethidine; patients who had used a MAOI within 2 weeks of enrolment; used psychoactive medications in the wk before enrollment; women who were pregnant or lactating; women who were pregnant not using medically accepted contraception; used any other method of bulimic therapy within 1 mo of entry (visit 1); used lithium in the wk before	BN: 398 (100%) BN, Purging Type > 1: 97 (32.8%) vs. 28 (27.5%) Binge Eating and Purging >= 3 episodes/wk: 398 (100%) BN, Duration >= 6 mo: 398 (100%) Binge Eating: 293 (99%) vs. 102 (100%) Vomiting: 296 (100%) vs. 101 (99%) Laxative Abuse: 49 (16.6%) vs. 12 (11.8%) Diuretics, Abuse: 22 (7.4%) vs. 7 (6.9%) Fasting: 53 (17.9%) vs. 15 (14.7%) Age >= 18 yr: 398 (100%) Age: 27 yr vs. 26 yr Gender - Female: 282 (95.3%) vs. 101 (99%)	There was greater decrease with fluoxetine in binge eating ($p=0.0002$) and vomiting ($p<0.0001$) episodes/wk: 9->5 binges/wk with fluoxetine vs. 9.5->7.5/wk with placebo; 9->5 vomiting/wk with fluoxetine vs. 9->7/wk with placebo. Binge Eating, Remission - Baseline – 16 wk: 53 (18.3%, N=290) vs. 12 (12%, N=100) Vomiting, Remission - Baseline – 16 wk: 55 (19%, N=290) vs. 12 (12%, N=100) Depression was reduced and side effects including anxiety, dizziness, emotional lability, myalgia, reduced libido, nausea, sweating, tremor, and yawning were more likely with fluoxetine than placebo. Treatment Adherence, Completed - Baseline – 16 wk: 176 (59.5%) vs. 49 (48%) ($p=0.045$) Treatment Discontinuation - Baseline – 16 wk - Adverse Events: 32 (10.8%) vs. 6 (5.9%) ($p=0.144$)	High

			enrollment; used tryptophan in the wk before enrollment	- Male: 14 (4.7%) vs. 1 (1%) Race, Caucasian: 385 (96.7%)	- Lack of Efficacy: 23 (7.8%) vs. 26 (25.5%) p<0.001) Overall attrition was 43% (173/398)	
Kanerva et al. (1995)	Design: RCT Setting: Outpatient: Helsinki University Central Hospital Country: Southern Finland Funding: Industry and academic	Randomized N=50 Fluoxetine 60 mg 8 wk (N=24) Placebo 8 wk (N=26)	Inclusion: BN; more than 15 years old; BMI of 16 kg/m^2 or more Exclusion: Pregnancy; lactation; inadequate contraception; major somatic illness; psychiatric illness; previous treatment with fluoxetine; previous treatment with any other concurrent psychiatric treatment; recent drug abuse; recent alcohol abuse; severe depressive features; suicidal features; recent administration of other psychotropic drugs; concurrent administration of other psychotropic drugs; lithium; MAOIs	BN: 50 (100%) BN, Duration: 5.7 yr BN, Age at Onset: 19.6 yr Bulimic Investigatory Test, Edinburgh: 24.3 units (SD ± 2.3) vs. 23.9 units (SD ± 3.5) Purging, Abstinence: 8 (16%) BMI >= 16 kg/m²: 50 (100%) Weight: 62.2 kg (SD ± 15.4) vs. 63 kg (SD ± 17) History of AN: 18 (36%) Age > 15 yr: 50 (100%) Age: 25.2 yr Gender, Female: 50 (100%) Race: NR	The fluoxetine group reported greater fatigue (p=0.023), less anxiety (p=0.0004), a decrease in weight (p=0.023), and less depression (p=0.0062) at 4 wk but not 8 wk. Greater weight loss was reported with fluoxetine at 8 wk: -4 kg (SD ± 3.9, N=22) vs. 1.1 kg (SD ± 3, N=24) (MD -5.1 kg, p=0.023) About half had binge eating reduction >50% in both groups from ~10 binge eating episodes/wk at baseline. Binge Eating - Baseline:9.2/wk vs. 10.5/wk - 8 wk: 5.3/wk (N=22) vs. 5.7/wk (N=24) Study Withdrawal, Adverse Events – Varies: 1 (4.17%) vs. 1 (3.85%) Attrition: 8% (2/24) vs. 8% (2/26)	Moderate
Romano et al. (2002)	Design: RCT Setting: Multi-center Country: United States Funding: Industry	Randomized N=150 Fluoxetine 60 mg 52 wk (N=76) Placebo 52 wk (N=74)	Inclusion: At least 18 years old; psychiatric diagnosis of BN; purging type with self-induced vomiting Exclusion: Schizophrenia; bipolar disorder; mood-congruent psychotic features; mood-incongruent psychotic	BN, Purging Type: 150 (100%) BN, Age at Diagnosis: 25.3 yr (SD ± 7.7) vs. 26.3 yr (SD ± 9.3)	Fluoxetine was associated with a longer time to relapse and greater change in binge eating and vomiting vs. placebo. Binge Eating – Baseline: 3/wk (SD ± 4.8) vs. 3.9/wk (SD ± 5.1)	High

			features; serious suicidal risk; organic brain disease; taken fluoxetine within 5 weeks before enrollment; previously treated with 60 mg/day of fluoxetine for longer than 8 weeks; history of alcohol or other substance abuse disorder within 3 months before enrollment; used psychoactive medications within 4 weeks before enrollment; received CBT within 4 weeks of enrollment	BMI: 22.5 kg/m² (SD ± 3.9) vs. 23 kg/m² (SD ± 3.8) Age >= 18 yr: 150 (100%) Age: 29.5 yr (SD ± 7) vs. 30 yr (SD ± 9.3) Gender - Female: 74 (97.37%) vs. 73 (98.65%) - Male: 2 (2.63%) vs.1 (1.35%) Race - Caucasian: 71 (93.42%) vs. 65 (87.84%) - Non-Caucasian: 5 (6.58%) vs. 9 (12.16%)	Binge Eating, Change - Baseline – 52 wk: 2.47/wk (SD ± 6.58, N=74) vs. 4.11/wk (SD ± 6.7, N=71) (MD -1.64/wk, p<0.02) Vomiting – Baseline: 4.1/wk (SD ± 5.5) vs. 4.5/wk (SD ± 6.1) Vomiting, Change - Baseline – 52 wk: 2.92/wk (SD ± 7.08, N=74) vs. 4.82/wk (SD ± 8.43, N=71) (MD -1.9/wk, p=0.001) Study Withdrawal - Baseline – 52 wk: - Adverse Events: 4 (5.26%) vs. 3 (4.05%) - Symptom Worsening: 17 (22.37%) vs. 22 (29.73%) Attrition: 83% (63/76) vs. 92% (68/74)	
Walsh et al. (2000)	Design: RCT Setting: Multi-center Country: United States Funding: Industry	Randomized N=22 Fluoxetine 60 mg 8 wk (N=13) Placebo 8 wk (N=9)	Inclusion: Women; BN; self-induced vomiting; not responded to, or had relapsed following, a course of CBT or interpersonal psychotherapy Exclusion: NR	BN: 22 (100%) Binge Eating, Duration: 15.6 yr (SD ± 8.9) vs. 13.9 yr (SD ± 9.9) Treatment Failure, Cognitive-Behavioral Therapy or Interpersonal Psychotherapy: 22 (100%) Vomiting, Self-Induced: 22 (100%) BMI: 22 kg/m² (SD ± 4.6) vs. 23 kg/m² (SD ± 3.2)	Binge and purge frequency in prior 28 days decreased with fluoxetine and increased with placebo: 22->4 and 15->18 for binge eating; 30->6 and 15->38 for purging. Binge Eating and Purging, Abstinence - 4 wk – 8 wk: 5 (38%) vs. 0 (0%) (p=0.054) Overall Attrition: 9% (20/22)	High

				History of AN: 2 (15%) vs. 2 (22%) Age: 32 yr (SD ± 7.8) vs. 27.8 yr (SD ± 5.2) Gender, Female: 22 (100%) Race: NR		

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EDE=Eating Disorder Examination; IBW=ideal body weight; ITT=intention-to-treat; MAOI=monoamine oxidase inhibitor; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to another pharmacotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Leombruni et al. (2006)	Design: RCT Setting: Outpatient: Eating Disorder Pilot Center outpatient facility of the Psychiatric Clinic of the University of Turin Country: Italy Funding: NR	Randomized N=37 Citalopram 20-40 mg 12 wk (N=19) Fluoxetine 20-60 mg 12 wk (N=18) Follow-up (N=28) - 14 vs. 14	Inclusion: BN; female Exclusion: Other current Axis I comorbidity; previous pharmacologic treatment in psychiatric specialty centers; previous treatment with citalopram or fluoxetine	BN: 37 (100%) BN, Duration: 7.23 yr (SD ± 6.8) vs. 8.66 yr (SD ± 6.39) BN, Age at Onset: 21.421 yr (SD ± 7.151) vs. 17.88 yr (SD ± 2.63) Weight: 56.929 kg (SD ± 5.063, N=14) vs. 55.614 kg (SD ± 3.062, N=14) Age: 28.684 yr (SD ± 8.246) vs. 26.55 yr (SD ± 6.27) Gender, Female: 37 (100%) Race: NR	Fluoxetine group had more vomiting episodes/wk at baseline (2.75 vs. 4.3) and fewer at 12 weeks (2.44 vs. 1.57). Vomiting, Change - Baseline – 12 wk: -0.312/wk (SD ± 0.77, N=14) vs. -2.72/wk (SD ± 1.48, N=14) Citalopram was associated with a greater decline in BDI scores (14.3->7.8) as compared to fluoxetine (11.6->10.3). Weight change was minimal in both groups and study withdrawal rates were comparable (p<0.926) Attrition: 26% (5/19) vs. 22% (4/18)	High

Milano et al. (2013)	Design: RCT Setting: Outpatient Country: Italy Funding: NR	Randomized N=60 Fluoxetine 60 mg 10 wk (N=20) Fluvoxamine 200 mg 10 wk (N=20) Sertraline 100 mg 10 wk (N=20)	Inclusion: Female; aged 18 to 34 years; BN, binge eating-purging; BN	BN: 60 (100%) BN, Purging Type: 60 (100%) Age 18 yr-34 yr: 60 (100%) Gender, Female: 60 (100%) Race: NR	Fluoxetine and fluvoxamine showed greater percent reduction in binge and vomiting episodes than sertraline: -75% vs. -59% vs. -18% for bulimic episodes; -68% vs. -62% vs. -3.54% for purging Study Withdrawal, Adverse Events, Serious - Baseline – 10 wk: 0 (0%) vs. 0 (0%) vs. 0 (0%) Attrition: NR	High

Abbreviations: BDI=Beck Depression Inventory; BN=bulimia nervosa; hr=hour; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Goldbloom et al. (1997)	Design: RCT Setting: Single Center: Toronto Hospital Country: Canada Funding: Industry	Randomized N=76 CBT 16 wk (N=24) Fluoxetine 60 mg + CBT 16 wk (N=29) Fluoxetine 60 mg 16 wk (N=23) Follow-up: Baseline – 20 wk Current Analysis (N=38) - 14 vs. 12 vs. 12	Inclusion: Females; 18-45 years of age; 85-125% matched population mean weight; BN; binge and vomit frequency of at least twice/wk; minimum 6-mo duration of bulimia Exclusion: Psychosis; ongoing pharmacotherapy or psychotherapy; use of MAOIs within 2 weeks prior to the onset of the study treatment; immediate suicide risk	BN: 76 (100%) BN, Duration >= 6 mo: 76 (100%) Binge Eating and Purging >= 2 episodes/wk: 76 (100%) Percent Average Body Weight, Matched-Population 85%-125%: 76 (100%) Age 18 yr-45 yr: 76 (100%) Age: 25.8 yr (SD ± 5.5, N=38) Gender, Female: 76 (100%) Race: NR	All treatments led to clinically significant improvement with some benefit of CBT + fluoxetine over fluoxetine alone but not over CBT alone. Binge Eating, Objective - Baseline->20 wk: 33.6->7.4/mo vs. 29.6->1.8 vs. 21->10 - Fluoxetine + CBT vs. Fluoxetine at 20 wk: MD -8.2/mo (p<0.03) Vomiting - Baseline->20 wk: 41.8->9/mo vs. 30.9->3.3 vs. 24.6->17.3 - CBT vs. Fluoxetine at 20 wk: MD -8.3/mo (p<0.07) - Fluoxetine + CBT vs. Fluoxetine at 20 wk: MD -14/mo (p<0.03)	High

					Binge Eating or Vomiting, Abstinence - 20 wk: 6 (43%, N=14) vs. 3 (25%, N=12) vs. 2 (17%, N=12) Attrition: 33% (8/24) vs. 55% (16/29) vs. 39% (9/23)	
Jacobi et al. (2002)	Design: RCT Setting: Outpatient: Department of Psychology at the University of Hamburg Country: Germany Funding: Industry	Randomized N=89 Current Analysis N=53 Group CBT 4 mo (N=19) Group CBT + Fluoxetine 60 mg (induction 20-40 mg) 4 mo (N=18) Fluoxetine 60 mg 4 mo (induction 20-40 mg) (N=16) Follow-up: Baseline – 16 mo	Inclusion: Women; 18-65 years of age; BN; minimum of 2 episodes of binge eating and vomiting for at least 6 months prior to the beginning of the study; actual BMI 17.5-25 kg/m² Exclusion: Concurrent severe psychiatric disturbance; concurrent psychosis or depression with suicidal risk; concurrent alcohol or drug abuse; concurrent involvement in other treatment; concurrent use of other medication	BN: 89 (100%) Binge Eating and Purging >= 2 episodes, In the Previous >= 6 mo: 89 (100%) BMI 17.5 kg/m²-25 kg/m²: 89 (100%) BMI: 20.6 kg/m² (SD ± 2, N=53) Age 18 yr-65 yr: 89 (100%) Age: 26 yr (SD ± 5.8, N=53) Gender, Female: 53 (100%) Race: NR	Baseline mean binges/28 days were 54.2 in the fluoxetine group vs. 36.5 and 33.5 in the CBT and fluoxetine + CBT groups respectively. All treatments led to significant improvements in eating disorder symptoms and in other psychological disturbances. Binge eating abstinence rates for completers were highest for CBT at both post-treatment and follow-up: - 4 mo: 5 (26%) vs. 3 (17%) vs. 2 (13%) - 16 mo: 4 (40%, N=10) vs. 1 (11%, N=9) vs. 1 (13%, N=8) At the end of treatment, vomiting abstinence was greater for CBT (37%) than for fluoxetine (6%) (p=0.046) or fluoxetine + CBT (6%) (p=0.041). Drug Discontinuation, Adverse Events - Baseline – 4 mo: 5 (27.78%) vs. 4 (25%) Attrition: 42% (8/19) vs. 33% (6/18) vs. 25% (4/16)	High

Mitchell et al. (2001)	Design: RCT Setting: Single Center: University of Minnesota Hospital Eating Disorder Program Country: United States Funding: Government, industry, and non-profit	Randomized N=91 Fluoxetine 60mg 16 wk (N=26) SH Manual (manual-based CBT) + Placebo 16 wk (N=22) Fluoxetine 60mg + SH Manual (manual-based CBT) 16 wk (N=21) Placebo 16 wk (N=22)	Inclusion: BN; female; at least 18 years of age; at least 85% of IBW; binge eating three times a wk for the last 6 months; self-induced vomiting 3 times a wk for the last 6 months Exclusion: Receiving pharmacotherapy or psychotherapy; current medical condition that would preclude safe outpatient treatment; history of hypersensitivity to fluoxetine; prior exposure to fluoxetine in a total amount greater than 140 mg; prior exposure to fluoxetine within the preceding 5 weeks before entering the study	BN: 91 (100%) Binge Eating 3 episodes/wk, In the Previous 6 mo: 91 (100%) Vomiting, Self-Induced 3 episodes/wk, In the Previous 6 mo: 91 (100%) %IBW >= 85%: 91 (100%) Age >= 18 yr: 91 (100%) Age: 26.6 yr (SD ± 7.1) - 26.6 yr (SD ± 7.1) vs. 26.8 yr (SD ± 6.9) vs. 29.3 yr (SD ± 7.8) vs. 23.8 yr (SD ± 6.1) Gender, Female: 91 (100%) Race, Caucasian: 25 (100%, N=25) vs. 22 (100%) vs. 20 (95.2%) vs. 21 (95.5%)	Active treatments reduced binge eating and vomiting as compared to placebo. Binge Eating – Baseline: 11.58/wk (SD ± 6.74) vs. 11.91/wk (SD ± 10.7) vs. 11.29/wk (SD ± 5.87) vs. 9.45/wk (SD ± 5.34) Binge Eating, % Change - Baseline – 16 wk: -50.3% (SD ± 52.6) vs. -59.7% (SD ± 39.6) vs. -66.8% (SD ± 29.9, N=20) vs. -32.4% (SD ± 66.7, N=21) Vomiting – Baseline: 16.81/wk (SD ± 27.72) vs. 13.86/wk (SD ± 10.81) vs. 12.43/wk (SD ± 6.92) vs. 11.77/wk (SD ± 6.67) Vomiting, % Change - Baseline – 16 wk: -52.8% (SD ± 50.7) vs. -50.2% (SD ± 55) vs. -66.7% (SD ± 31.2, N=20) vs. -22.8% (SD ± 56.1, N=21) Attrition: 4% (1/26) vs. 5% (1/22) vs. 10% (2/21) vs. 18% (4/22)	High
Mitchell et al. (2002)	Design: RCT Setting: Multi-center: Cornell University; University of Minnesota; Rutgers University Country: United States Funding: non-profit	Randomized N=62 CBT 16 wk > Fluoxetine 60 mg 25 wk > Fluoxetine / Desipramine 50-300 mg 60 wk (N=31) CBT 16 wk > IPT 33 wk (N=31)	Inclusion: BN; women; failed to respond to CBT; adult; purging by self-induced vomiting at least 2 times a wk for 3 months Exclusion: Substance dependence in the last 6 months; any history of psychosis; received an adequate trial of antidepressant treatment for BN previously; suicidality	BN: 62 (100%) Treatment Failure, Cognitive-Behavioral Therapy: 62 (100%) Binge Eating, Duration: 10.4 hr/wk (SD ± 7.1) vs. 11 hr/wk (SD ± 6.7)	Outcomes did not differ for the two groups but study withdrawals were greater in the medication group (48% vs. 32%). Binge Eating and Purging, Abstinence - 34 wk: 3 (10%) vs. 5 (16%) - 60 wk: 3 (9.68%) vs. 5 (16.13%)	High

				Purging, Duration: 8.9 hr/wk (SD ± 6.3) vs. 10.7 hr/wk (SD ± 6.7) Binge Eating, Objective: 5/wk vs. 4/wk Vomiting, Self-Induced >= 2 episodes/wk, In the Previous 3 mo: 62 (100%) History of AN: 11 (36%) vs. 9 (29%) Age >= 18 yr: 62 (100%) Age: 27.1 yr (SD ± 6.3) vs. 28 yr (SD ± 7.3) Gender, Female: 62 (100%) Race: NR	Disease Response, Remission - 34 wk: 3 (18.75%, N=16) vs. 4 (19.05%, N=21) Disease Response, Symptomatic - 34 wk: 13 (81.25%, N=16) vs. 17 (80.95%, N=21) Study Withdrawal: 17 wk – 60 wk: 16 (51.61%) vs. 13 (41.94%) Attrition: 48% (15/31) vs. 32% (10/31)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; hr=hour; IBW=ideal body weight; IPT=interpersonal psychotherapy; MAOI=monoamine oxidase inhibitor; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; yr=year

Sertraline

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Milano et al. (2004)	Design: RCT Setting: Outpatient Country: Italy Funding: NR	Randomized N=20 Sertraline 100 mg 12 wk (N=10) Placebo 12 wk (N=10)	Inclusion: Female; BN, purging type; outpatients	BN, Purging Type: 20 (100%) Age: 24 yr – 36 yr Gender, Female: 20 (100%) Race: NR	The sertraline group showed a statistically significant reduction in the number of bulimic (-75% vs. -10% with placebo) and purging episodes (-55% vs. -8% with placebo).	High

					Bulimic Episodes: 14.9->2.93 vs. 12.9->10.9 (at 12 wk, MD -7.97, p<0.01) Purging: 9->4 vs. 9->8 (at 12 wk, MD -4, p<0.01) Study Withdrawal, Adverse Events, Serious - Baseline – 12 wk: 0 (0%) vs. 0 (0%) Attrition: NR	
Milano et al. (2013)	Design: RCT Setting: Outpatient Country: Italy Funding: NR	Randomized N=60 Fluoxetine 60 mg 10 wk (N=20) Fluvoxamine 200 mg 10 wk (N=20) Sertraline 100 mg 10 wk (N=20)	Inclusion: Female; aged 18 to 34 years; BN, binge eating-purging; BN	BN: 60 (100%) BN, Purging Type: 60 (100%) Age 18 yr-34 yr: 60 (100%) Gender, Female: 60 (100%) Race: NR	Fluoxetine and fluvoxamine showed greater percent reduction in binge and vomiting episodes than sertraline: -75% vs. -59% vs. -18% for bulimic episodes; -68% vs. -62% vs. -3.54% for purging Study Withdrawal, Adverse Events, Serious - Baseline – 10 wk: 0 (0%) vs. 0 (0%) vs. 0 (0%) Attrition: NR	High

Abbreviations: BN=bulimia nervosa; MD=mean difference; NR=not reported; RCT=randomized controlled trial; wk=week; yr=year

Fluvoxamine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Fichter et al. (1996, 1997)	Design: RCT Setting: Outpatient; Inpatient: Roseneck Hospital for Behavioural Medicine	Randomized N=81 Psychotherapy + Fluvoxamine 50-300 mg 2 wk (-) Psychotherapy 15 wk (N=37)	Inclusion: 18-50 years of age; BN of at least 6 months duration at admission to the hospital; BN; body weight between 85% and 125% of IBW Exclusion: Pregnant; lactating; displayed serious medical conditions; psychoses; acute	BN: 81 (100%) BN, Duration >= 6 mo: 81 (100%) BN, Age at Onset: 19 yr (SD ± 3) vs. 19 yr (SD ± 4) (N=72)	Benefit of fluvoxamine in preventing relapse was seen using CGI-severity (8.1% relapse with fluvoxamine vs. 31.4% with placebo, p<0.05) but attrition was high.	High

	Country: Germany Funding: NR	Psychotherapy + Placebo 2 wk (-) Psychotherapy 15 wk (N=35) Follow-up: Baseline – 19 wk Current Analysis (N=72)	suicidal ideation; history of seizures; suffered from insulin-dependent diabetes; multiple drug allergies; psychoactive substance dependency; used other psychoactive medication within the 2 weeks before entering the medication part of the study; appetite suppressants or other relevant medication within the 2 weeks before entering the medication part of the study	Binge Eating: 16/wk (SD ± 15) vs. 15/wk (SD ± 15) (N=72) %IBW 85%-125%: 81 (100%) BMI: 20.6 kg/m² (SD ± 4, N=34) vs. 19.9 kg/m² (SD ± 3.3, N=34) Age 18 yr-50 yr: 81 (100%) Age: 25.2 yr (SD ± 4.9) vs. 23.7 yr (SD ± 5.1) Gender, Unknown: 81 (100%) Race: NR	Binge Eating, Change - Baseline – 15 wk: 17.76/wk (SD ± 11.62) vs. 40.5/wk (SD ± 12.38) Binge Eating, % Change - Baseline – 15 wk: 111% vs. 270% (MD -159%, p<0.05) Binge Eating, Abstinence - 15 wk: 25 (67.33%) vs. 12 (34.68%) Study Withdrawal, Adverse Events - Baseline – 19 wk: 8 (22%) vs. 1 (2.86%) Hospitalization - Baseline – 19 wk: 1.7 d (SD ± 3.9) vs. 0.2 d (SD ± 1.3) Attrition: 38% (9/37) vs. 14% (5/35)	
Milano et al. (2013)	Design: RCT Setting: Outpatient Country: Italy Funding: NR	Randomized N=60 Fluoxetine 60 mg 10 wk (N=20) Fluvoxamine 200 mg 10 wk (N=20) Sertraline 100 mg 10 wk (N=20)	Inclusion: Female; aged 18 to 34 years; BN, binge eating-purging; BN	BN: 60 (100%) BN, Purging Type: 60 (100%) Age 18 yr-34 yr: 60 (100%) Gender, Female: 60 (100%) Race: NR	Fluoxetine and fluvoxamine showed greater percent reduction in binge and vomiting episodes than sertraline: -75% vs. -59% vs. -18% for bulimic episodes; -68% vs. -62% vs. -3.54% for purging Study Withdrawal, Adverse Events, Serious - Baseline – 10 wk: 0 (0%) vs. 0 (0%) vs. 0 (0%) Attrition: NR	High
Schmidt et al. (2004)	Design: RCT Setting: Outpatient centers	Randomized N=267 Fluvoxamine 50-300 mg + Psychotherapy Level 1 8 wk > (-)	Inclusion: Female; BN; 18- 50 years of age; weighed between 85% and 115% of IBW Exclusion: Pregnancy; breast-feeding; inadequate contraception; psychosis; active suicidality; any clinically	BN: 267 (100%) Binge Eating < 2 episodes/wk: 0 (0%, N=267)	At 8 wk, there were no differences among groups on a bulimic severity index (BINGE scale) or % reduction of binges.	High

	Country: United Kingdom Funding: Industry	Psychotherapy Level 1 52 wk (N=134) Fluvoxamine 50-300 mg + Psychotherapy Level 1 8 wk > Placebo 52 wk (N=67) Placebo + Psychotherapy Level 1 8 wk > (-) Psychotherapy Level 1 52 wk (N=66) Follow-up: Baseline – 52 mo Follow-up (N=178) - 83 vs. 46 vs. 49	important medical illness; multiple drug allergies; substance dependence; any serious laboratory abnormality; treatment with psychoactive medication; appetite suppressants within 2 weeks of allocation to double-blind drug; less than 2 binges/wk	%IBW 85%-115%: 267 (100%) Age 18 yr-50 yr: 267 (100%) Gender, Female: 267 (100%) Race: NR	19 adverse events occurred in the first 8 weeks, 17 of these in patients on fluvoxamine. In the second phase of the trial (wk 9 to wk 52) there was no difference between the treatment groups in bulimic severity index, ‘‘performance score’’ (related to maintaining of remission) or proportions of patients with good or poor outcome. Disease Response, Remission - 52 wk: 28 (34%) vs. 17 (36%) vs. 16 (33%) Attrition at 8 wk: 36% (72/201) with fluvoxamine vs. 26% (17/66) with placebo Attrition 8 wk – 52 wk: 73% (61/83) vs. 57% (26/46) vs. 53% (26/49)	

Abbreviations: BMI=body mass index; BN=bulimia nervosa; CGI=Clinical Global Impression; d=day; IBW=ideal body weight; MD=mean difference; NR=not reported; RCT=randomized controlled trial; wk=week; yr=year

Citalopram

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Milano et al. (2005)	Design: RCT Setting: NR Country: Italy Funding: NR	Randomized N=20 Citalopram 20 mg 1 wk > 40 mg 8 wk (N=10) Placebo 8 wk (N=10)	Inclusion: Female; 19-28 years of age; BN with purging behaviors	BN, Purging Type: 20 (100%) Age 19 yr-28 yr: 20 (100%) (N=20) Gender, Female: 20 (100%)	The citalopram group showed a significant reduction in the number of binge eating (-65% vs. -12% with placebo) and purging episodes (-56% vs. -7% with placebo).	High

				Race: NR	Study Withdrawal, Adverse Events, Serious - Baseline – 8 wk: 0 (0%) vs. 0 (0%) Attrition: NR	

Abbreviations: BN=bulimia nervosa; NR=not reported; RCT=randomized controlled trial; wk=week; yr=year

Binge-Eating Disorder Studies Supporting Guideline Statements

Cognitive-Behavioral Therapy

Compared to No Treatment/Wait-List Control

Compared to no treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1995)	Design: RCT Setting: NR Country: United States Funding: NR	Randomized N=50 Group CBT 12 wk > Weight Loss Therapy (for responders) / Group IPT (for non-responders) 24 wk (N=39) Assessment Only Control 24 wk (N=11) Current Analysis (N=42) - 31 vs. 11	Inclusion: BED; overweight Exclusion: Current weight loss program; antidepressant medication; medication that might influence weight; abused drugs or alcohol; current major psychiatric condition; psychosis; history of purging within the previous 6 months; BMI below 27 kg/m^2	BED: 50 (100%) Overweight: 50 (100%) Binge Eating: 4.5 d/wk (SD ± 1.7) BES: 33.3 units (SD ± 5.9) vs. 27.2 units (SD ± 6.3) Weight: 107.3 kg (SD ± 25.4) BMI: 37.1 kg/m² (SD ± 7.3) Age: 47.6 yr (SD ± 10.1) Gender - Female: 43 (86%) - Male: 7 (14%) Race: NR	Addition of IPT did not yield further benefits over CBT alone. At 12 wk, 55% of the CBT group was abstinent vs. 9% with no treatment (p<0.008). Greater reductions were also seen with CBT in binge days/wk (-77% vs. -22%): - Baseline: 4.4 d/wk (SD ± 1.8, N=31) vs. 3.7 d/wk (SD ± 1.2) - 24 wk: 1 d/wk (SD ± 1.4, N=31) vs. 2.9 d/wk (SD ± 2) (MD -1.9 d/wk, p=0.0001) Weight loss was significantly better in the active treatment group though clinically modest (2.8 kg difference): - Baseline: 108 kg (SD ± 26.7, N=31) vs. 106.1 kg (SD ± 20.3)	High

					- 24 wk: 107.4 kg (SD ± 28, N=31) vs. 110.2 kg (SD ± 22.8) (MD -2.8 kg, p=0.02) Study withdrawal was greater in the CBT group (18% vs. 9%). Overall Attrition: 16% (8/50)	
Schag et al. (2019)	Design: RCT Setting: Outpatient Country: Germany Funding: Academic	Randomized N=80 Group CBT 8 wk (N=41) Assessment Only Control 24 wk (N=39) Follow-up: Baseline – 20 wk	Inclusion: BED; adults Exclusion: Current suicidality, substance addiction, psychotic disorders, or bipolar I disorder; received current psychotherapy; pregnancy or breastfeeding; somatic conditions which influence eating behavior or body weight (e.g., diabetes, thyroid diseases) and in which medication had been adapted in the last 3 weeks	BED: 80 (100%) BED, Duration: 15.9 yr (SD ± 11.4) vs. 15.5 yr (SD ± 12.2) Age: 40.1 yr (SD ± 12.1) vs. 40.5 (SD ± 13.5) Gender - Female: 67 (83.8%) - Male: 13 (16%) Race: NR	Both groups reduced binge-eating episodes from baseline to the end of treatment. At follow-up, the group difference was significant (p=0.005): Mean Binge-Eating Episodes In The Past 4 weeks - Baseline->End of Treatment->Follow-Up: 13.6->7.5->6.3 vs. 13.1->9.2->11.4. At follow-up (but not at the end of treatment), the group CBT condition showed higher abstinence rate and lower deterioration rate compared to the control condition in the ITT analyses: Abstinence - Baseline->End of Treatment->Follow-Up: 0 (0%)->6 (14.6%)->14 (34.1%) vs. 0(0%)->7 (17.9%)->4 (10.3%) Deterioration - Baseline->End of Treatment->Follow-Up: 0 (0%)->5 (12.2%)->6 (14.6%) vs. 0 (0%)->8 (20.5%)->9 (23.2%) Attrition: 20% (8/41) vs. 10% (439)	Low

Abbreviations: BED=binge-eating disorder; BES=Binge Eating Scale; BMI=body mass index; CBT=cognitive-behavioral therapy; d=day; IPT=interpersonal psychotherapy; ITT=intention-to-treat; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to wait-list control

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Eldredge et al. (1997)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=46 Group CBT 12 wk > BWL Treatment (for those with improvement) / Group CBT (remaining participants) 24 wk (N=36) WLC 24 wk (N=10)	Inclusion: Obese; BMI of 27 or above; BED Exclusion: Additional treatment which might interfere with CBT; concurrent involvement in a weight loss program; concurrent involvement in antidepressant medication; concurrent involvement in any other medication which might influence weight; current drug or alcohol abuse; a history of purging within the prior 6 months; current major medical or psychiatric condition which might interfere with treatment; pregnancy; psychosis; severe suicidality	BED: 46 (100%) Binge Eating: 6.9 d (SD ± 3) Obesity: 46 (100%) Weight: 106.8 kg (SD ± 28.2) BMI >= 27 kg/m²: 46 (100%) BMI: 38.4 kg/m² (SD ± 9.5) Age: 45.2 yr (SD ± 9.8) Gender - Female: 44 (95.65%) - Male: 2 (4.35%) Race: NR	Active treatment was associated with greater percent change in binge eating from baseline: -68.2% vs. -19.8% (MD -48.4%, p=0.046). Half of the active treatment group responded by 12 wk. Binge Eating, Abstinence - 24 wk: 24 (66.67%) vs. NR BMI - Baseline: 36.33 kg/m² vs. 44.58 kg/m² - 12 wk: 36.29 kg/m² vs. 44.73 kg/m² Attrition: 19% (7/36) vs. 20% (2/10)	High
Gorin et al. (2003)	Design: RCT Setting: NR Country: United States Funding: Academic	Randomized N=94 Group CBT 12 wk (Standard Group) (N=32) Group CBT 12 wk (Spouse Involvement) (N=31) WLC 12 wk (N=31)	Inclusion: 18-65 years of age; BED; BMI >=25 kg/m^2; female; overweight Exclusion: Engaged in purging behaviors more than once per mo; AN; BN; EDNOS; receiving concurrent treatment for weight loss; currently taking appetite suppressants; pregnant	BED: 94 (100%) Overweight: 94 (100%) BMI >= 25 kg/m²: 94 (100%) BMI: 39.42 kg/m² (SD ± 7.72) Age 18 yr-65 yr: 94 (100%) Age: 45.2 yr (SD ± 10.03)	Both active treatment groups improved more than WLC but no significant differences between active treatments. Binge Eating - Baseline: 3.81 d/wk (SD ± 1.66) vs. 3.41 d/wk (SD ± 2.09) vs. 3.77 d/wk (SD ± 1.82)	High

		Follow-up: Baseline – 38 wk		Gender, Female: 94 (100%) Race, Caucasian: 81 (86%)	Binge Eating, Change - Baseline – 12 wk: -2 d/wk vs. -2.23 d/wk vs. -0.82 d/wk Binge Eating, Abstinence - 12 wk: 9 (29%) vs. 14 (46%) vs. 3 (9%) - Standard Group CBT vs. Spouse Involvement CBT: p=0.35 - CBT groups vs. WLC: p=0.02 BMI - Baseline: 38.72 kg/m² (SD ± 8.78) vs. 40.51 kg/m² (SD ± 8.29) vs. 39.37 kg/m² (SD ± 7.53) BMI, Change - Baseline – 12 wk: -0.07 kg/m² (SD ± 6.7) vs. -0.14 kg/m² (SD ± 6.44) vs. 0.36 kg/m² (SD ± 5.94) Overall Attrition: 34% (32/94)	
Kristeller et al. (2014)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=140 Psychoeducational CBT 5 mo (N=48) MB-EAT 5 mo (N=50) WLC 5 mo (N=42) With BED (N=35 vs. 31 vs. 31 Follow-up: Baseline – 6 mo	Inclusion: Overweight or obese; BMI >= 28 kg/m^2 Exclusion: Suicidal symptomology; psychiatric symptoms potentially likely to interfere with group participation or follow-up; psychotic symptoms; drug or alcohol abuse; unstable medication use; previous regular meditation practice; concurrent participation in a weight loss program; concurrent psychotherapy focused on weight or eating issues; purging or laxative abuse within 6 months	Overweight or Obesity: 140 (100%) BED: 35 (70%, N=50) vs. 31 (58.49%, N=53) vs. 31 (65.96%, N=47) Weight: 242.7 lbs (N=150) BMI >= 28 kg/m²: 140 (100%) BMI: 40.26 kg/m² (N=150) Age: 46.55 yr (N=150) Gender	Compared to WLC, psychoeducational CBT and MB-EAT showed comparable improvement at 1-mo post-intervention on binge days per mo (15.31->5.23 d/mo CBT vs. 14.84->4.78 MB-EAT vs. 14.04->12.83 WLC). The proportion of individuals with no BED diagnosis at 1-mo post-treatment was 75% with psychoeducational CBT vs. 95% with MB-EAT vs. 48% with WLC but attrition was considerable. BMI – Baseline: 39.04 kg/m² (SD ± 8.61, N=27) vs. 39.63	High

				- Female: 132 (88%, N=150) - Male: 18 (12%, N=150) Race - Black or African American: 20 (13.33%, N=150) - Minority: 21 (14%, N=150) Ethnicity, Hispanic/Latino: 1 (0.67%, N=150)	kg/m² (SD ± 7.99, N=39) vs. 38.14 kg/m² (SD ± 6.42, N=26) BMI, Change - Baseline – 6 mo: -0.11 kg/m² (SD ± 6.83, N=27) vs. 0.42 kg/m² (SD ± 6.76, N=39) vs. 0.28 kg/m² (SD ± 5.01, N=26) Study Withdrawal, Treatment Dissatisfaction - Baseline – 6 mo: 5 (10%, N=50) vs. 0 (0%, N=53) vs. 0 (0%, N=47) Attrition: 43% (21/48) vs. 22% (11/50) vs. 38% (16/42)	
Peterson et al. (1998, 2001)	Design: RCT; Subgroup Follow-up Analysis Setting: NR Country: United States Funding: Non-profit and government	Randomized N=61 Group CBT Therapist-Led 8 wk (N=16) Group CBT Partial SH 8 wk (N=19) Group CBT SH 8 wk (N=15) WLC 8 wk (N=11) Follow-up: Baseline – 60 wk Follow-up (N=51) 16 vs. 19 vs. 16	Inclusion: Female; 18-65 years of age; BED Exclusion: Receiving current psychoactive medication or psychotherapy; met criteria for substance abuse or dependence within the past 6 months; medically unstable at the time of enrollment; at risk of self-injury at the time of enrollment; nonpurging BN; had engaged in any compensatory behavior; self-induced vomiting, abuse of laxatives or diuretics, excessive exercise, or fasting in the past 6 months	BED: 61 (100%) Binge Eating – Baseline - 7.7/wk (SD ± 3.8) vs. 8.2/wk (SD ± 5.9) vs. 6.8/wk (SD ± 2.4) vs. 5.7/wk (SD ± 6) - 9 hr/wk (SD ± 6.7) vs. 13.4 hr/wk (SD ± 13) vs. 9.8 hr/wk (SD ± 5.5) vs. 8.3 hr/wk (SD ± 7.6) BMI: 34.7 kg/m² (SD ± 7.5) Age 18 yr-65 yr: 61 (100%) Age: 42.4 yr (SD ± 10.2) Gender, Female: 61 (100%) Race, Caucasian: 59 (96.5%)	Reductions in binge-eating episodes and associated symptoms were observed for all active treatments at post-treatment, 1-mo, 6-mo, and 1-yr follow-ups, with no significant differences among the three conditions. Binge-eating episodes/wk at 60 wk were 3.5/wk therapist led; 3.1/wk partial SH; 3.3/wk SH. Binge Eating, Change - Baseline – 8 wk - -4.4/wk vs. -5.5/wk vs. -5/wk vs. -0.9/wk - -4.8 hr/wk vs. -10.2 hr/wk vs. 7.5 hr/wk vs. 1.3 hr/wk Abstinence rates were: - at 8 wk: 18.8% vs. 36.8% vs. 53.3% vs. 0% - at 60 wk: 16.7% vs. 46.2% vs. 33.3% vs. NR	High

					Attrition: 13% (2/16) vs. 11% (2/19) vs. 27% (4/15) vs. 18% (2/11)	
Peterson et al. (2009)	Design: RCT Setting: Multi-center Country: United States Funding: Government	Randomized N=259 Group CBT Therapist-Led 20 wk (N=60) Group CBT Therapist-Assisted 20 wk (N=63) Group CBT SH 20 wk (N=67) WLC 20 wk (N=69) Follow-up: Baseline – 72 wk	Inclusion: Adults; BED; BMI >=25 kg/m^2 Exclusion: Pregnancy; lactation; lifetime diagnosis of bipolar disorder; lifetime diagnosis of psychotic disorder; current diagnosis of substance abuse or dependence; medical instability; psychiatric instability; acute suicide risk; current psychotherapy; current participation in a formal weight loss program	BED: 259 (100%) BMI >= 25 kg/m²: 259 (100%) BMI: 39 kg/m² (SD ± 7.8) - 39.2 kg/m² (SD ± 8.3) vs. 40.7 kg/m² (SD ± 8.8) vs. 38.2 kg/m² (SD ± 7.2) vs. 38.1 kg/m² (SD ± 6.9) Age: 47.1 yr (SD ± 10.4) - 48.1 yr (SD ± 9.1) vs. 48.1 yr (SD ± 9.1) vs. 47.1 yr (SD ± 10.4) vs. 47.6 yr (SD ± 10.6) Gender - Female: 60 (100%) vs. 51 (81%) vs. 60 (89.6%) vs. 56 (81.2%) - Male: 0 (0%) vs. 12 (19%) vs. 7 (10.4%) vs. 13 (18.8%) Race, Caucasian: 55 (91.7%) vs. 60 (95.2%) vs. 67 (100%) vs. 67 (97.1%)	At 20 wk, therapist-led and therapist-assisted groups had significantly greater abstinence rates (51.7% therapist-led vs. 33.3% therapist-assisted vs. 17% SH vs. 10.1% WLC) (p<0.008) but abstinence rates were comparable at follow-up (20-27%). Reductions in binge eating were greater with therapist-led (24.6->6.3/mo) vs. 21.9->9.7/mo with therapist-assisted, 22.4->9.7/mo with SH, and 23.1->17.6/mo with WLC. Attrition: 12% (7/60) vs. 32% (20/63) vs. 40% (27/67) vs. 19% (13/69)	High
Schlup et al. (2009)	Design: RCT Setting: University of Basel Country: Switzerland Funding: Non-profit	Randomized N=36 Group CBT 8 wk > Booster Sessions 60 wk (N=18) WLC 8 wk (N=18)	Inclusion: 18-70 years of age; BED Exclusion: Severe mental disorders; major depression with acute suicidal risk; psychosis; bipolar disorder; current substance use disorder; participation in a diet program; participation in psychotherapy; weight loss medication in the past 3 months; previous	BED: 36 (100%) Binge Eating: 3.53/wk BMI: 33.4 kg/m² (SD ± 7.6) - 32.4 kg/m² (SD ± 5.6) vs. 34.3 kg/m² (SD ± 9.1) Age 18 yr-70 yr: 36 (100%) Age: 44.3 yr (SD ± 10.3)	Binge-eating episodes/wk were 3.53 at baseline and decreased by 1.58/wk in the CBT group vs. a 0.35/wk increase in the WLC group (WLC vs. CBT with booster sessions SMD 1.15, p=0.0004).	High

			surgical treatment of obesity; male	- 47.1 yr (SD ± 8.5) vs. 41.2 yr (SD ± 11.1) Gender, Female: 36 (100%) Race: NR	Abstinence was achieved in 39% of CBT subjects vs. 0% of WLC subjects. Treatment was discontinued in 5.6% of CBT subjects vs. 0% of WLC subjects. BMI change was comparable: 0.01 kg/m² vs. 0.42 kg/m². Overall Attrition: 14% (5/36)	
Tasca et al. (2006, 2012)	Design: RCT Setting: NR Country: Canada Funding: Non-profit	Randomized N=135 Group CBT 16 wk (N=47) Group Psychodynamic IPT 16 wk (N=48) WLC 16 wk (N=40) Follow-up: Baseline – 16 mo	Inclusion: BED; a minimum of 2 days of binge eating/wk for at least the previous 6 months Exclusion: Current problems with substance use; bipolar disorder; psychotic disorder; current suicidality; current other medical or psychological treatment for BED; history of an eating disorder other than BED; current purging behavior; age less than 18 years	BED: 135 (100%) Binge Eating >= 2 d/wk, In the Previous >= 6 mo: 135 (100%) BED, Duration: 19.62 yr (SD ± 9.19) BMI: 41.11 kg/m² (SD ± 9.95) Age: 42.75 yr (SD ± 10.76) Gender - Female: 123 (91.11%) - Male: 12 (8.89%) Race, Caucasian: 132 (97.7%)	Binge-eating abstinence at 16 wk was 62.2% CBT, 59.5% IPT, and 9.1% WLC. Abstinence rates at 68 wk were 67.7% CBT vs. 56.8% IPT. Both treatments were noted to reduce interpersonal problem subscale ratings including cold/distant subscale ratings. BMI - Baseline: 42.59 kg/m² (SD ± 12.95, N=37) vs.40.03 kg/m² (SD ± 9.69, N=37) vs. 42.58 kg/m² (SD ± 9.57, N=33) BMI, Change - Baseline – 68 wk: -1.57 kg/m² (SD ± 9.9, N=37) vs. -2.36 kg/m² (SD ± 7.25, N=37) vs. NR (N=33) Attrition: 21% (10/47) vs. 23% (11/48) vs. 18% (7/40)	High
Telch et al. (1990)	Design: RCT Setting: NR Country: United States	Randomized N=44 Group CBT 10 wk (N=23)	Inclusion: BN; an average of 2 or more binge episodes a wk for the past 6 months; recurrent episodes of binge eating; feeling of lack of control or inability to stop eating during the eating binges; persistent	BN: 44 (100%) Binge Eating >= 2/wk, In the Previous 6 mo: 44 (100%) Binge Eating:	Group CBT was associated with greater reductions from baseline on binge episodes/wk (5.3->0.32 vs. 5.29->4.14, MD -3.82/wk, p<0.0001) and binge days/wk	High

	Funding: Government	WLC 10 wk (N=21) Follow-up: Baseline – 20 wk	concern with body shape and weight; female Exclusion: Purging; age below 18 years or above 65; current or history of self-induced vomiting; current or history of laxative use; current or history of other purging behaviors; current use of antidepressant medication; current use of appetite suppressants; concurrent treatment for weight loss; concurrent unipolar affective disorder, bipolar affective disorder or psychosis; concurrent drug abuse; concurrent alcoholism	- 5.3/wk (SD ± 2.98) vs. 5.29/wk (SD ± 3.3) - 4.3 d/wk (SD ± 1.61) vs. 4.14 d/wk (SD ± 1.59) Binge Eating, Duration: 22.9 yr (SD ± 11.9) Age: 42.6 yr (SD ± 8.4) Gender, Female: 44 (100%) Race - Caucasian: 40 (91%) - Black or African American: 1 (2%) - Asian: 1 (2%) Ethnicity, Hispanic/Latino: 2 (5%)	(4.3->0.32 vs. 5.29->3.57, MD -3.25 d/wk, p<0.0001). Numbers of binge-eating abstinence at 10 wk were 15 (79%, N=19) vs. 0 (0%). Weight – Baseline: 86.83 kg (SD ± 13.72) vs. 86.81 kg (SD ± 10.62) Weight, Change - Baseline – 10 wk: 0.31 kg (SD ± 10.91, N=19) vs. 0.92 kg (SD ± 8.66) Attrition: 17% (4/23) vs. 0% (0/21)	
Wagner et al. (2016)	Design: RCT Setting: NR Country: Germany Funding: NR	Randomized N=139 CBT Web 16 wk (N=69) WLC 16 wk (N=70)	Inclusion: BED; 18-65 years of age Exclusion: Current BN or AN; severe major depressive symptoms; acute suicidal ideation; severe substance abuse; severe dependence disorder; type 1 diabetes mellitus; thyroid problems; ongoing psychotherapy; bariatric surgery; serious medical conditions influencing weight; serious medical conditions influencing eating; BDI >26	BED: 139 (100%) BMI: 32.4 kg/m² (SD ± 7.4) Weight: 93.7 kg (SD ± 22.6) BDI: 16.2 units (SD ± 5.8) Age 18 yr-65 yr: 139 (100%) Age: 35.1 yr (SD ± 9.9) - 34.9 yr (SD ± 10.1) vs. 35.3 yr (SD ± 9.7) Gender - Female: 65 (94.2%) vs. 69 (98.6%) - Male: 4 (5.8%) vs. 1 (1.4%) Race: NR	Web CBT group showed fewer binge-eating episodes per mo at 16 wk (16->6.8/mo vs. 17.1->14.9/mo), higher rates of recovery and remission (47.8% vs. 4.3%, p<0.001; 14.6% vs. 0% p<0.001, respectively), and lower ratings of binge eating psychopathology. 27.5% of the web CBT group withdrew from the study vs. 8.6% of the WLC group (p=0.004). Weight – Baseline: 91.9 kg (SD ± 21) vs. 95.4 kg (SD ± 24.1) Weight, Change - Baseline – 16 wk: -1.3 kg (SD ± 16.39) vs. 0.2 kg (SD ± 18.91)	High

					Attrition: 28% (19/69) vs. 9% (6/70)	
Wilfley et al. (1993)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=56 Group CBT 16 wk (N=18) Group IPT 16 wk (N=18) WLC 16 wk (N=20)	Inclusion: Nonpurging BN; female; 18- 65 years of age; average of two or more binge episodes per wk for the past 6 months Exclusion: Age below 18 years or above 65; current self-induced vomiting, laxative use, or purging behaviors; past history of self-induced vomiting, laxative use, or purging behaviors; current use of antidepressant medication; current use of appetite suppressants; concurrent treatment for weight loss; concurrent unipolar disorder, bipolar affective disorder, or psychosis; concurrent drug abuse; concurrent alcoholism	BN, Non-Purging Type: 56 (100%) Binge Eating >= 2/wk, In the Previous 6 mo; 56 (100%) Binge Eating, Duration: 23.7 yr (SD ± 13.4) Binge Eating: 4.2 d/wk (SD ± 1.5) vs. 4.7 d/wk (SD ± 1.8) vs. 4.4 d/wk (SD ± 1.8) Weight: 87.3 kg (SD ± 14.2) Age 18 yr-65 yr: 56 (100%) Age: 44.3 yr (SD ± 8.3) Gender, Female: 56 (100%) Race - Caucasian: 48 (86%) - Black or African American: 3 (5%) - Indian: 1 (2%) - Pacific Islander: 1 (2%) Ethnicity, Hispanic/Latino: 3 (5%)	Abstinence from binge eating at 16 wk was 28% with group CBT vs. 44% with group IPT vs. 0% with WLC. IPT had a greater binge-eating percent change but it was not statistically significant: -48% with CBT vs. -71% with IPT vs. -10% with WLC. Weight, Change - Baseline – 68 wk: 0 kg vs. -3 kg vs. NR Adherence was greater in the IPT group (88% vs. 72% with CBT). Study withdrawal rates were low in all groups (11% CBT, 0% IPT, 5% WLC). Attrition: 33% (6/18) vs. 11% (2/18) vs. NR	High

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; BWL=behavioral weight loss; CBT=cognitive-behavioral therapy; d=day; EDNOS=eating disorder not otherwise specified; hr=hour; IPT=interpersonal psychotherapy; MB-EAT=mindful-based eating awareness training; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; SMD=standardized mean difference; wk=week; WLC=wait-list control; yr=year

Compared to Interpersonal Psychotherapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and	Outcome measures, main results, and overall percent attrition	Risk of bias

		size (N), dose, duration, and follow-up		baseline clinical features (e.g., BMI)		
Tasca et al. (2006, 2012)	Design: RCT Setting: NR Country: Canada Funding: Non-profit	Randomized N=135 Group CBT 16 wk (N=47) Group Psychodynamic IPT 16 wk (N=48) WLC 16 wk (N=40) Follow-up: Baseline – 16 mo	Inclusion: BED; a minimum of 2 days of binge eating/wk for at least the previous 6 months Exclusion: Current problems with substance use; bipolar disorder; psychotic disorder; current suicidality; current other medical or psychological treatment for BED; history of an eating disorder other than BED; current purging behavior; age less than 18 years	BED: 135 (100%) Binge Eating >= 2 d/wk, In the Previous >= 6 mo: 135 (100%) BED, Duration: 19.62 yr (SD ± 9.19) BMI: 41.11 kg/m² (SD ± 9.95) Age: 42.75 yr (SD ± 10.76) Gender - Female: 123 (91.11%) - Male: 12 (8.89%) Race, Caucasian: 132 (97.7%)	Binge-eating abstinence at 16 wk was 62.2% CBT, 59.5% IPT, and 9.1% WLC. Abstinence rates at 68 wk were 67.7% CBT vs. 56.8% IPT. Both treatments were noted to reduce interpersonal problem subscale ratings including cold/distant subscale ratings. BMI - Baseline: 42.59 kg/m² (SD ± 12.95, N=37) vs.40.03 kg/m² (SD ± 9.69, N=37) vs. 42.58 kg/m² (SD ± 9.57, N=33) BMI, Change - Baseline – 68 wk: -1.57 kg/m² (SD ± 9.9, N=37) vs. -2.36 kg/m² (SD ± 7.25, N=37) vs. NR (N=33) Attrition: 21% (10/47) vs. 23% (11/48) vs. 18% (7/40)	High
Wilfley et al. (1993)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=56 Group CBT 16 wk (N=18) Group IPT 16 wk (N=18) WLC 16 wk (N=20)	Inclusion: Nonpurging BN; female; 18- 65 years of age; average of two or more binge episodes per wk for the past 6 months Exclusion: Age below 18 years or above 65; current self-induced vomiting, laxative use, or purging behaviors; past history of self-induced vomiting, laxative use, or purging behaviors; current use of antidepressant medication; current use of appetite suppressants; concurrent treatment for weight loss; concurrent unipolar disorder,	BN, Non-Purging Type: 56 (100%) Binge Eating >= 2/wk, In the Previous 6 mo; 56 (100%) Binge Eating, Duration: 23.7 yr (SD ± 13.4) Binge Eating: 4.2 d/wk (SD ± 1.5) vs. 4.7 d/wk (SD ± 1.8) vs. 4.4 d/wk (SD ± 1.8) Weight: 87.3 kg (SD ± 14.2)	Abstinence from binge eating at 16 wk was 28% with group CBT vs. 44% with group IPT vs. 0% with WLC. IPT had a greater binge-eating percent change but it was not statistically significant: -48% with CBT vs. -71% with IPT vs. -10% with WLC. Weight, Change - Baseline – 68 wk: 0 kg vs. -3 kg vs. NR	High

			bipolar affective disorder, or psychosis; concurrent drug abuse; concurrent alcoholism	Age 18 yr-65 yr: 56 (100%) Age: 44.3 yr (SD ± 8.3) Gender, Female: 56 (100%) Race - Caucasian: 48 (86%) - Black or African American: 3 (5%) - Indian: 1 (2%) - Pacific Islander: 1 (2%) Ethnicity, Hispanic/Latino: 3 (5%)	Adherence was greater in the IPT group (88% vs. 72% with CBT). Study withdrawal rates were low in all groups (11% CBT, 0% IPT, 5% WLC). Attrition: 33% (6/18) vs. 11% (2/18) vs. NR	
Wilfley et al. (2002); Hilbert et al. (2012)	Design: RCT; Follow-up Setting: Multi-center Country: United States Funding: Government	Randomized N=162 Group CBT NR (N=81) Group IPT NR (N=81) Follow-up: Baseline – 4 yr Follow-up (N=90) 45 vs. 45	Inclusion: Overweight; BED; 18-65 years of age; BMI 27-48 kg/m^2; average of >=2 days of binge eating/wk for at least 6 months' duration; marked distress regarding binge eating; at least 3 of 5 behavioral features associated with BED Exclusion: Taking weight-affecting medications; taking psychotropic medications; psychiatric conditions warranting immediate treatment; psychotic symptoms; substance dependence; suicidality	BED: 162 (100%) Binge Eating >= 2 d/wk, Duration 6 mo: 162 (100%) Overweight: 162 (100%) BMI 27 kg/m²-48 kg/m²: 162 (100%) Binge Eating: 17.3 d/mo (SD ± 6.9) vs. 16.3 d/mo (SD ± 7.2) Age 18 yr-65 yr: 162 (100%) Age: 45.6 yr (SD ± 9.6) vs. 44.9 yr (SD ± 9.6) Gender - Female: 67 (82.7%) vs. 67 (82.7%) - Male: 14 (17.3%) vs. 14 (17.3%) Race	Binge-eating recovery rates were equivalent for CBT and IPT at post-treatment (79% vs 73%) and at 1-yr follow-up (59% vs 62%). Persistent recovery was present at 4 yr in 27.3% of the CBT group and 22.2% of the IPT group. Binge days per mo showed similar reductions: 17.3 baseline -> 1.7 at 12 mo with CBT; 16.3-> 1.2 with IPT. Disease Response, Remission - Post-Treatment: 73 (94%, N=78) vs. 72 (90%, N=80) - 12 mo: 56 (84%, N=67) vs. 63 (89%, N=71) - 4 yr: 18 (72%, N=25) vs. 26 (83.9%, N=31)	High

				- Caucasian: 76 (93.9%) vs. 74 (91.4%) - Black or African American: 3 (3.7%) vs. 3 (3.7%) - Native American/Alaska Native: 1 (1.2%) vs. 0 (0%) Ethnicity, Hispanic/Latino: 1 (1.2%) vs. 4 (4.9%)	BMI – Baseline: 37.4 kg/m² (SD ± 5.3) vs. 37.4 kg/m² (SD ± 5.1) BMI, Change - Baseline – 12 mo: -0.2 kg/m² (SD ± 4.03, N=67) vs. -1.1 kg/m² (SD ± 4.08, N=71) Attrition: 11% (9/81) vs. 9% (7/81)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; IPT=interpersonal psychotherapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Compared to Behavioral Weight Loss

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1994b)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=108 CBT 12 wk > Weight Loss Treatment 36 wk (N=36) CBT 12 wk > Weight Loss Therapy + Desipramine 25-300 mg 36 wk (N=36) Weight Loss Treatment 36 wk (N=37) CBT > Weight Loss Therapy +/- Desipramine 25-300 mg 36 wk (pooled) (N=72)	Inclusion: Female; BED; binge eating at least twice a wk for a 6-mo period; overweight Exclusion: Current weight loss program; antidepressant medication; any medication that may affect weight; suicidality; abuse of drugs or alcohol; history of purging in the prior 12 months; BMI below 27 kg/m^2; current BN	BED: 108 (100%) Binge Eating >= 2/wk, Duration 6 mo: 108 (100%) Binge Eating: 4.5 d/wk (SD ± 1.4) - 4.4 d/wk (SD ± 1.4, N=30) vs. 5.1 d/wk (SD ± 1.4, N=27) vs. 4.5 d/wk (SD ± 1.6, N=27) Overweight: 108 (100%) Weight: 104.9 kg (SD ± 18.5) - 102.1 kg (SD ± 15.7, N=30) vs. 111.9 kg (SD ± 17.4, N=27) vs. 102.9 kg (SD ± 15.8, N=27)	At 12 wk, CBT groups had significantly less binge eating (67% reduction vs. 44% with weight loss alone, MD -23 %, p<0.01) and the weight loss group had more weight loss (-2.0 kg) compared to CBT groups (0.7 kg) (MD 2.7 kg, p<0.002). No differences were noted between groups at the end of treatment or follow-up except weight loss (0 kg vs. -4.8 kg vs. -4.15 kg at 48 wk) - CBT > Weight Loss Treatment vs. CBT > Weight Loss Therapy + Desipramine: MD 4.8 kg (p<0.05)	High

		Follow-up: Baseline – 48 wk Current Analysis (N=84) - 30 vs. 27 vs. 27		BMI: 38.6 kg/m² (SD ± 6.6) Age: 45 yr (SD ± 10) Gender, Female: 108 (100%) Race: NR	Binge Eating, Abstinence - 48 wk: 8 (28%, N=30) vs. 9 (32%, N=27) vs. 4 (14%, N=27) BDI – Baseline: 13.5 units (SD ± 7.8, N=30) vs. 13.7 units (SD ± 8.1, N=27) vs. 12.9 units (SD ± 6.5, N=27) BDI, Change - Baseline – 36 wk: -4.6 units (SD ± 10.5, N=30) vs. -5.9 units (SD ± 10.84, N=27) vs. -1.6 units (SD ± 11.79, N=27) Attrition: 17% (11/36) vs. 23% (12/36) vs. 27% (16/37)	
Grilo et al. (2011)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=125 Group CBT 24 wk (N=45) Group CBT + Group BWL Treatment 40 wk (N=35) Group BWL Treatment 24 wk (N=45) Follow-up Period: Baseline – 76 wk for Group CBT or Group BWL Treatment Baseline – 92 wk for Group CBT + Group BWL Treatment	Inclusion: Obese; BED; 18-60 years of age; BMI 30-55 kg/m^2 Exclusion: Concurrent treatment for eating problems or weight problems; psychosis or bipolar disorder requiring alternative treatment	BED: 125 (100%) Obesity: 125 (100%) BMI 30 kg/m²-55 kg/m²: 125 (100%) BMI: 38.8 kg/m² (SD ± 5.8) - 39.3 kg/m² (SD ± 6.1) vs. 39 kg/m² (SD ± 6.1) vs. 38 kg/m² (SD ± 5.3) Weight: 250.1 lbs (SD ± 52.6) vs. 237.2 lbs (SD ± 42.8) vs. 242.7 lbs (SD ± 45.8) Age 18 yr-60 yr: 125 (100%) Age: 44.8 yr (SD ± 9.4) - 45.2 yr (SD ± 8.5) vs. 44.5 yr (SD ± 9.2) vs. 44.6 yr (SD ± 10.5) Gender	At 12-mo follow-up, ITT binge-eating remission rates were 51% with CBT, 40% with CBT + BWL, and 36% with BWL. Binge eating with CBT had greater reductions at 24 wk than BWL (15.6->2.2/mo vs. 14.9->4.6/mo) and these differences were maintained at 50-wk follow-up. At post-treatment, BWL or CBT+ BWL had significantly greater percent BMI reduction than CBT alone: - -0.5% (SD ± 3.5) with CBT vs. -2.6% (SD ± 5.3) with BWL (MD -2.1 %, p=0.03) - -0.5% (SD ± 3.5) with CBT vs. -2.7% (SD ± 6) with CBT+ BWL (MD -2.2 %, p=0.04)	High

				- Female: 28 (64.4%) vs. 28 (80%) vs. 28 (62.2%) - Male: 41 (33%) - 17 (35.6%) vs. 7 (20%) vs. 17 (37.8%) Race - Caucasian: 34 (75.6%) vs. 26 (74.3%) vs. 36 (80%) - Black or African American: 5 (11.1%) vs. 8 (22.9%) vs. 7 (15.6%) - Asian: 2 (4.4%) vs. 1 (2.9%) vs. 0 (0%) - American Indian/Alaskan Native: 1 (2.2%) vs. 0 (0%) vs. 0 (0%) Ethnicity, Hispanic/Latino: 3 (6.7%) vs. 0 (0%) vs. 2 (4.4%)	Attrition: 24% (11/45) vs. 40% (14/35) vs. 31% (14/45)	
Munsch et al. (2007, 2012)	Design: RCT; Follow-Up/Extension Setting: Outpatient: Department of Clinical Psychology and Psychotherapy Country: Switzerland Funding: NR	Randomized N=80 Group CBT 64 wk (N=44) Group BWL Treatment 64 wk (N=36) Follow-up: 323.5 wk (Mean, SD ± 46.9)	Inclusion: 18- 70 years of age; BMI 27-40 kg/m^2; BED; obese Exclusion: Participation in a diet program; other psychotherapy program; treatment with weight loss medication during the last 3 months; suicidal tendency; psychosis; mania; substance use disorder	BED: 80 (100%) Obesity: 80 (100%) BMI 27 kg/m²-40 kg/m²: 80 (100%) Age 18 yr-70 yr: 80 (100%) Age: 44.4 yr (SD ± 11.5) vs. 47.8 yr (SD ± 11.8) Gender - Female: 40 (90.9%) vs. 31 (86.1%) - Male: 4 (9.1%) vs. 5 (13.9%) Race: NR	Self-reported binge eating was less with group CBT at follow-ups: Baseline->16 wk->64 wk: 3.81->0.14->0.52/wk vs. 4.1->1.15->1.5/wk; MD -1.01/wk (p<0.001) at 16 wk; MD -0.98/wk (p<0.045 at 64 wk). Abstinence rates were greater in the BWL group at 16 wk but not 64 wk: 41% CBT vs. 58% BWL (p=0.01) at 16 wk; 52% vs. 50% (p=0.39) respectively at 64 wk. Follow up at 6 years (mean 323.5 wk) showed comparable results of the two treatments for	High

					abstinence rates: 20% CBT vs. 17% BWL (OR 1.12) BMI - Baseline: 33.66 kg/m² (SD ± 4.31, N=42) vs. 34.36 kg/m² (SD ± 3.74, N=33) - 16 wk: 33.62 kg/m² (SD ± 4.7, N=30) vs. 33.08 kg/m² (SD ± 3.69, N=27) (MD 0.54 kg/m², p=0.004) - 64 wk: 32.36 kg/m² (SD ± 5.38, N=23) vs. 33.62 kg/m² (SD ± 3.99, N=21) (MD -1.26 kg/m², p=0.7) Study Withdrawal, Treatment Dissatisfaction - Baseline – 16 wk: 5 (11.36%) vs. 2 (5.56%) Attrition: 36% (22/44) vs. 36% (10/36)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; BWL=behavioral weight loss; CBT=cognitive-behavioral therapy; d=day; ITT=intention-to-treat; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to Cognitive-Behavioral Therapy

Compared to group cognitive-behavioral therapy spouse involvement

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Gorin et al. (2003)	Design: RCT Setting: NR Country: United States Funding: Academic	Randomized N=94 Group CBT 12 wk (Standard Group) (N=32)	Inclusion: 18- 65 years of age; BED; BMI >=25 kg/m^2; female; overweight Exclusion: Engaged in purging behaviors more than once per mo; AN; BN; EDNOS; receiving concurrent treatment for weight	BED: 94 (100%) Overweight: 94 (100%) BMI >= 25 kg/m²: 94 (100%) BMI: 39.42 kg/m² (SD ± 7.72)	Both active treatment groups improved more than WLC but no significant differences between active treatments. Binge Eating - Baseline: 3.81 d/wk (SD ± 1.66) vs. 3.41 d/wk	High

		Group CBT 12 wk (Spouse Involvement) (N=31) WLC 12 wk (N=31) Follow-up: Baseline – 38 wk	loss; currently taking appetite suppressants; pregnant	Age 18 yr-65 yr: 94 (100%) Age: 45.2 yr (SD ± 10.03) Gender, Female: 94 (100%) Race, Caucasian: 81 (86%)	(SD ± 2.09) vs. 3.77 d/wk (SD ± 1.82) Binge Eating, Change - Baseline – 12 wk: -2 d/wk (SD ± 1.43465) vs. -2.23 d/wk (SD ± 1.52182) vs. -0.82 d/wk (SD ± 1.41763) Binge Eating, Abstinence - 12 wk: 9 (29%) vs. 14 (46%) vs. 3 (9%) - Standard Group CBT vs. Spouse Involvement CBT: p=0.35 - CBT groups vs. WLC: p=0.02 BMI - Baseline: 38.72 kg/m² (SD ± 8.78) vs. 40.51 kg/m² (SD ± 8.29) vs. 39.37 kg/m² (SD ± 7.53) BMI, Change - Baseline – 12 wk: -0.07 kg/m² (SD ± 6.7) vs. -0.14 kg/m² (SD ± 6.44) vs. 0.36 kg/m² (SD ± 5.94) Overall Attrition: 34% (32/94)	

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; EDNOS=eating disorder not otherwise specified; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Group cognitive-behavioral therapy with body exposure component compared to group cognitive-behavioral therapy with cognitive restructuring component

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Hilbert and Tuschen-Caffier (2004)	Design: RCT Setting: Outpatient: University of Marburg Country: Germany Funding: Government	Randomized N=28 Group CBT + Body Exposure Component NR (N=14) Group CBT + Cognitive Restructuring Component Focused on Body Image NR (N=14)	Inclusion: Female; patients suffering from full syndrome BED or subclinical binge eaters Exclusion: Pregnancy; presence of psychotic symptoms; substance dependence; suicidality; use of psychoactive medication; use of medication affecting body weight	BED, Full Syndrome or BED, Subclinical: 28 (100%) - Full Syndrome: 10 (83.3%, N=12) vs. 10 (83.3%, N=12 - Subclinical: 2 (16.7%, N=12) vs. 2 (16.7%, N=12) Binge Eating, Duration: 13.5 yr (SD ± 10.7) vs. 17.7 yr (SD ± 13.2) Age: 42.1 yr (SD ± 12.1) vs. 38.6 yr (SD ± 8.5) Gender, Female: 28 (100%) Race: NR	Comparable decreases were seen in binges/wk (2.9->0.6/wk vs. 3.4->1/wk) and proportion with BED at the end of treatment (16.7% vs. 25%). BMI - Baseline: 34 kg/m² (SD ± 10.2, N=12) vs. 36.4 kg/m² (SD ± 10.4, N=12) - End of Treatment: 33.1 kg/m² (SD ± 10.4, N=12) vs. 37.2 kg/m² (SD ± 10.3, N=12) Disease Response, Recovery - Baseline – End of Treatment: 4 (33.3%, N=12) vs. 9 (75%, N=12) (p=0.098) - Baseline – 4 mo: 6 (50%, N=12) vs. 8 (66.7%, N=12) (p=0.408) Attrition: 14% (2/14) vs. 14% (2/14)	High

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to appetite focused cognitive-behavioral therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
McIntosh et al. (2016)	Design: RCT Setting: NR Country: New Zealand Funding: Government	Randomized N=112 CBT12 mo (N=38) Appetite-Focused CBT12 mo (N=36)	Inclusion: Female; 16-65 years of age; BED or BN; the subjective experience of dyscontrol Exclusion: Severe major depression; serious suicidal intent; severe psychoactive	BED or BN: 112 (100%) - BED: 18 (48.2%) vs. 18 (50%) vs. 18 (47.4%) - BN: 20 (51.8%) vs. 18 (50%) vs. 20 (52.6%) Eating Disorder, Duration: 15.2 yr (SD ± 12.7)	Binge-eating abstinence was not statistically different between groups either at 12 mo or 24 mo. - At 24 mo: 53.3% vs. 67.9% vs. 62.1%	High

		Schema Therapy 12 mo (N=38) Follow-up: Baseline – 24 mo	substance dependence; bipolar I disorder; schizophrenia; severe physical illness; severe medical complications of the eating disorder; cognitive impairment; psychotropic medication; an adequate trial of CBT in the past yr; an adequate trial of schema therapy in the past yr; currently underweight	- 14.6 yr (SD ± 13.2) vs. 15.4 yr (SD ± 13.9) vs. 15.7 yr (SD ± 11.4) Weight: 83.2 kg (SD ± 22.4) BMI: 29.9 kg/m² (SD ± 7.8) Age 16 yr-65 yr: 112 (100%) Age: 35.3 yr (SD ± 12.6) - 34.4 yr (SD ± 13) vs. 34.3 yr (SD ± 11.9) vs. 37.1 yr (SD ± 12.9) AN, Lifetime: 4 (10.53%) vs. 2 (5.56%) vs. 2 (5%) Gender, Female: 112 (100%) Race - Caucasian: 19 (17%) - Asian: 4 (4%) - Pacific Islander: 0 (0%) Nationality, New Zealand and Race, Caucasian: 75 (67%) Nationality, New Zealand and Race, Maori: 11 (10%)	Weight – Baseline: 83 kg (SD ± 22.4) vs. 84.7 kg (SD ± 23.8) vs. 82 kg (SD ± 21.5) Weight, Change - Baseline – 12 mo: 1.4 kg (SD ± 5.55) vs. -1 kg (SD ± 5.4) vs. 1.5 kg (SD ± 5.55) - Baseline – 24 mo: 0.79 kg (SD ± 7.67, N=30) vs. -0.056 kg (SD ± 7.41, N=28) vs. 0.8 kg (SD ± 7.54, N=29) Disease Response, Remission - 12 mo: 13 (34.2%) vs. 20 (55.6%) vs. 20 (52.6%) - 24 mo: 16 (53.3%, N=30) vs. 19 (67.9%, N=28) vs. 17 (58.6%, N=29) Attrition: 29% (11/38) vs. 36% (13/36) vs. 24% (9/38)	

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to maintenance group cognitive-behavioral therapy and maintenance group cognitive-behavioral therapy with exercise

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Pendleton et al. (2002)	Design: RCT Setting: NR	Randomized N=114	Inclusion: Females; 25-60 years of age; >30 lbs overweight; binge eating; history of	Binge Eating: 114 (100%) Binge Eating: 4.8 d/wk (SD ± 2) vs. 4.6 d/wk (SD ± 2.1) vs.	Group CBT + Exercise with maintenance was superior on binge eating d/wk to group CBT alone for 4 mo: 4.2->0.6 d/wk	High

	Country: NR Funding: NR	Group CBT 4 mo (N=17) Group CBT + Exercise 4 mo (N=20) Group CBT 4 mo > 10 mo (Maintenance) (N=23) Group CBT 4 mo > + Exercise 10 mo (Maintenance) (N=24) Observational Period: Baseline – 16 mo	sedentary lifestyle and occupation; nonsmoker Exclusion: History of cardiovascular disease, diabetes, metabolic disorder, or gastrointestinal disorder or surgery; history of drug abuse	4.6 d/wk (SD ± 1.9) vs. 4.2 d/wk (SD ± 2.3)Weight: 97.2 kg (SD ± 17.8, N=84) BMI: 36.2 kg/m² (SD ± 6.5, N=84) Age: 45 yr (SD ± 8.3, N=84) Gender, Female: 114 (100%) Race - Caucasian: 64 (76%, N=84) - Black or African American: 11 (13%, N=84) - Mexican American: 7 (8%, N=84) - Other: 3 (3%, N=84)	vs. 4.8->1.9 d/wk (MD -1.3 d/wk, p=0.039). - 10 mo: 0.5 vs. 2 d/wk (MD -1.5 d/wk, p=0.002) - 16 mo: 1 vs. 2.5 d/wk (MD -1.5 d/wk, p=0.007) Binge abstinence at 16 mo was: 18% CBT alone, 65% CBT+Exercise, 39% CBT alone+maintenance, 58% CBT+Exercise+maintenance. BMI was significantly reduced in the subjects in both the exercise and maintenance conditions at 16 mo: 1.33 kg/m² (SD ± 2) vs. -0.75 kg/m² (SD ± 2.4) vs. -0.24 kg/m² (SD ± 3) vs. -2.26 kg/m² (SD ± 3.9) Overall Attrition: 26% (30/114)	

Abbreviations: BMI=body mass index; CBT=cognitive-behavioral therapy; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Group cognitive-behavioral therapy therapist-led compared to group cognitive-behavioral therapy with partial self-help or self-help

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Peterson et al. (1998, 2001)	Design: RCT; Subgroup Follow-up Analysis Setting: NR Country: United States Funding: Non-profit and government	Randomized N=61 Group CBT Therapist-Led 8 wk (N=16) Group CBT Partial SH 8 wk (N=19)	Inclusion: Female; 18-65 years of age; BED Exclusion: Receiving current psychoactive medication or psychotherapy; met criteria for substance abuse or dependence within the past 6 months; medically unstable at the time of enrollment; at risk of self-injury at the time of	BED: 61 (100%) Binge Eating – Baseline - 7.7/wk (SD ± 3.8) vs. 8.2/wk (SD ± 5.9) vs. 6.8/wk (SD ± 2.4) vs. 5.7/wk (SD ± 6) - 9 hr/wk (SD ± 6.7) vs. 13.4 hr/wk (SD ± 13) vs. 9.8 hr/wk (SD ± 5.5) vs. 8.3 hr/wk (SD ± 7.6)	Reductions in binge-eating episodes and associated symptoms were observed for all active treatments at post-treatment, 1-mo, 6-mo, and 1-yr follow-ups, with no significant differences among the three conditions. Binge-eating episodes/wk at 60 wk were	High

		Group CBT SH 8 wk (N=15) WLC 8 wk (N=11) Follow-up: Baseline – 60 wk Follow-up (N=51) - 16 vs. 19 vs. 16	enrollment; non-purging BN; had engaged in any compensatory behavior; self-induced vomiting, abuse of laxatives or diuretics, excessive exercise, or fasting in the past 6 months	BMI: 34.7 kg/m² (SD ± 7.5) Age 18 yr-65 yr: 61 (100%) Age: 42.4 yr (SD ± 10.2) Gender, Female: 61 (100%) Race, Caucasian: 59 (96.5%)	3.5/wk therapist led; 3.1/wk partial SH; 3.3/wk SH. Binge Eating, Change - Baseline – 8 wk - -4.4/wk vs. -5.5/wk vs. -5/wk vs. -0.9/wk - -4.8 hr/wk vs. -10.2 hr/wk vs. 7.5 hr/wk vs. 1.3 hr/wk Abstinence rates were: - at 8 wk: 18.8% vs. 36.8% vs. 53.3% vs. 0% - at 60 wk: 16.7% vs. 46.2% vs. 33.3% vs. NR Attrition: 13% (2/16) vs. 11% (2/19) vs. 27% (4/15) vs. 18% (2/11)	
Peterson et al. (2009)	Design: RCT Setting: Multi-center Country: United States Funding: Government	Randomized N=259 Group CBT Therapist-Led 20 wk (N=60) Group CBT Therapist-Assisted 20 wk (N=63) Group CBT SH 20 wk (N=67) WLC 20 wk (N=69) Follow-up: Baseline – 72 wk	Inclusion: Adults; BED; BMI >=25 kg/m^2 Exclusion: Pregnancy; lactation; lifetime diagnosis of bipolar disorder; lifetime diagnosis of psychotic disorder; current diagnosis of substance abuse or dependence; medical instability; psychiatric instability; acute suicide risk; current psychotherapy; current participation in a formal weight loss program	BED: 259 (100%) BMI >= 25 kg/m²: 259 (100%) BMI: 39 kg/m² (SD ± 7.8) - 39.2 kg/m² (SD ± 8.3) vs. 40.7 kg/m² (SD ± 8.8) vs. 38.2 kg/m² (SD ± 7.2) vs. 38.1 kg/m² (SD ± 6.9) Age: 47.1 yr (SD ± 10.4) - 48.1 yr (SD ± 9.1) vs. 48.1 yr (SD ± 9.1) vs. 47.1 yr (SD ± 10.4) vs. 47.6 yr (SD ± 10.6) Gender - Female: 60 (100%) vs. 51 (81%) vs. 60 (89.6%) vs. 56 (81.2%) - Male: 0 (0%) vs. 12 (19%) vs. 7 (10.4%) vs. 13 (18.8%)	At 20 wk, therapist-led and therapist-assisted groups had significantly greater abstinence rates (51.7% therapist-led vs. 33.3% therapist-assisted vs. 17% SH vs. 10.1% WLC) (p<0.008) but abstinence rates were comparable at follow-up (20-27%). Reductions in binge eating were greater with therapist-led (24.6->6.3/mo) vs. 21.9->9.7/mo with therapist-assisted, 22.4->9.7/mo with SH, and 23.1->17.6/mo with WLC. Attrition: 12% (7/60) vs. 32% (20/63) vs. 40% (27/67) vs. 19% (13/69)	High

				Race, Caucasian: 55 (91.7%) vs. 60 (95.2%) vs. 67 (100%) vs. 67 (97.1%)		

Abbreviations: BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; hr=hour; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; WLC=wait-list control; yr=year

Individual compared to group cognitive-behavioral therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Ricca et al. (2010)	Design: RCT Setting: Outpatient: Outpatient Clinic for Eating Disorders of the Psychiatric Unit of the Department of Neuroscience of the University of Florence Country: Italy Funding: NR	Randomized N=144 Individual CBT 24 wk (N=72) Group CBT 24 wk (N=72) With BED (N=81) - 40 vs. 41 Follow-up: Baseline – 3.5 yr	Inclusion: 18-60 years of age; BED or subthreshold BED; binge eating frequency of at least once a wk for a minimum duration of 6 consecutive months Exclusion: Recurrent severe compensatory behaviors; current severe mental disorders; schizophrenia; bipolar disorder; severe major depression; suicide ideation; psychoactive substance dependence; prior CBTs; psychoactive medications within the past 3 months; previous surgical treatment for obesity	BED or BED, Subclinical: 144 (100%) BED: 40 (56%) vs. 41 (57%) BED, Subclinical: 32 (44.4%) vs. 31 (43.1%) Binge Eating >= 1/wk, Duration 6 mo: 144 (100%) Age 18 yr-60 yr: 144 (100%) - 46.5 yr (SD ± 12.4) vs. 47.4 yr (SD ± 11.9) Gender - Female: 62 (86.1%) vs. 65 (90.3%) - Male: 10 (13.9%) vs. 7 (9.7%) Race, Caucasian: 144 (100%)	The two treatment conditions had similar outcomes at 24 wk and at 3.5-yr follow-up with decrease in binges/mo from 8 to 4 for each. 56% individual CBT and 57% group CBT met BED criteria at baseline vs. about 20% at 24 wk and at 3.5 yr. Rate of recovery was 33.3% with individual CBT vs. 16.7% with group CBT at 24 wk (p=0.02) but not statistically different at 3.5 yr (36.1% vs. 27.8%, OR 1.49, 95% CI 0.72 – 3.03). BMI – Baseline: 38 kg/m² (SD ± 7.78) vs. 38.2 kg/m² (SD ± 6.52) BMI, Change - Baseline – 24 wk: -1.5 kg/m² (SD ± 5.94) vs. -0.8 kg/m² (SD ± 4.87) - Baseline – 3.5 yr: -2 kg/m² (SD ± 6.42) vs. -1.2 kg/m² (SD ± 5.42) Attrition: 4% (3/72) vs. 6% (4/72)	Low

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; CI=confidence interval; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Short-term group compared to long-term

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Schlup et al. (2010)	Design: Non-Randomized Comparison; Short-Term CBT Data from Schlup B (2009); Long-Term CBT Data from Munsch S (2006) Setting: NR Country: Switzerland Funding: NR	Total N=76 Short-Term Group CBT 8 wk> Booster Sessions 60 wk (N=36) Long-Term Group CBT 16 wk> Booster Sessions 68 wk (N=40)	Inclusion: BED; obese Exclusion: NR	BED: 76 (100%) Obesity: 76 (100%) BMI: 33.2 kg/m² (SD ± 6.9) vs. 33.2 kg/m² (SD ± 4.3) Age: 44.4 yr (SD ± 10.2) vs. 44.6 yr (SD ± 11.2) Gender, Female: 76 (100%) Race: NR	Study withdrawal rates were greater with the long-term (16 wk) CBT, 35% over 68 weeks vs. 14% over 60 weeks with 8 wk CBT (p=0.034). Treatment discontinuation rates showed a similar pattern: 1 (2.8%) with 8 wk CBT vs. 12 (30%) with 16 wk CBT (p=0.002) at the end of treatment. Remission rates were significantly greater with 16 wk CBT at the end of treatment: 12 (46%, N=27) vs. 21 (86%, N=24) (OR 0.1351, p=0.008). BMI – Varies: 32.63 kg/m² (N=27) vs. 32.24 kg/m² (N=23) (SMD 0.06, p=0.79) Attrition: 14% (5/36) vs. 35% (14/40)	----------

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; SMD=standardized mean difference; wk=week; yr=year

Compared to Other Psychotherapy

Compared to brief strategic therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Castelnuovo (2011); Jackson (2018) (STRATOB)	Design: RCT Setting: Inpatient: Saint Joseph Hospital - Istituto Auxologico Italiano Country: Italy Funding: Non-profit	Randomized N=60 CBT + Diet Therapy + Physical Activity Counseling 1 mo> CBT 7 mo (N=30) Brief Strategic Therapy+ Diet Therapy + Physical Activity Counseling 1 mo> Brief Strategic Therapy 7 mo (N=30) Follow-up: Baseline – 7 mo	Inclusion: 18- 65 years of age; obesity; BED; BMI >= 30 kg/m^2; female Exclusion: Other severe psychiatric disturbance	BED: 60 (100%) Binge Eating: 2.82/wk (SD ± 0.77) - 2.83/wk (SD ± 0.74) vs. 2.8/wk (SD ± 0.8) Obesity: 60 (100%) Weight: 106.95 kg (SD ± 6.95) BMI >= 30 kg/m²: 60 (100%) Age 18 yr-65 yr: 60 (100%) Age: 46.05 yr (SD ± 10.54) - 46.2 yr (SD ± 10.5) vs. 45.9 yr (SD ± 10.76) Gender, Female: 60 (100%) Race: NR	Lack of remission was greater at 7 mo with CBT (63.3%) vs. Brief Strategic Therapy (20%) (p=0.001). Weight – Baseline: 107.37 kg (SD ± 6.83) vs. 106.53 kg (SD ± 7.14) Weight, % Change - Baseline – 7 mo: -11.92% (SD ± 16.9) vs. -16.93% (SD ± 5.51) (MD 5.01%, p=0.128) Attrition: NR	High

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; STRATOB=Systemic and STRATegic psychotherapy for OBesity; wk=week; yr=year

Compared to web-based guided self-help

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
de Zwaan et al. (2017) (INTERBED)	Design: Prospective RCT Setting: Multi-center; outpatient Country: Germany; Switzerland Funding: Government	Randomized N=178 CBT 4 mo (N=89) GSH Web 4 mo (N=89) Follow-up: Baseline – 22 mo	Inclusion: Overweight or obese; full or subsyndromal BED; age 18 years or older; BMI between 27 and 40 kg/m^2 Exclusion: Ongoing psychotherapy; current BN; current substance abuse; psychotic disorder; current suicidal ideation; current intake	BED, Full Syndrome: 74 (86%) vs. 77 (92.8%) BED, Subclinical: 12 (14%) vs. 6 (7.2%) BED, Duration: 10.4 yr (± 11.1) vs. 7.9 yr (± 9.3)	Per-protocol sample (N=153) failed to show noninferiority of GSH web. In modified ITT analysis, GSH web was inferior to CBT in reducing objective binge-eating episode days at the end of treatment.	High

			of antipsychotic drugs; current intake of weight-affecting drugs	BMI: 34.4 kg/m² (SD ± 3.9) - (N=86) vs. 33.4 kg/m² (SD ± 3.9, N=83) Age: 43.2 yr (SD ± 12.3, N=169) - 42.7 yr (SD ± 12, N=86) vs. 43.7 yr (SD ± 12.7, N=83) Gender - Female: 74 (86%, N=86) vs.74 (89.2%, N=83) - Male: 12 (14%, N=86) vs.9 (10.8%, N=83) Race: NR	CBT was superior to GSH web at 6 mo but not 1.5-yr follow-up. Abstinence at 4 mo was 61.2% CBT vs. 35.5% GSH Web as compared to 22 mo with 46.6% vs. 43.1%, respectively. BMI, Change - Baseline – 4 mo: -0.2 kg/m² (SD ± 3.3, N=85) vs. -0.5 kg/m² (SD ± 3.02, N=77) - Baseline – 10 mo: -0.9 kg/m² (SD ± 3.35, N=80) vs. -0.3 kg/m² (SD ± 3.15, N=70) Attrition: 9% (8/89) vs. 19% (17/89)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; GSH=guided self-help; INTERBED=Internet and Binge-eating disorder; ITT=intention-to-treat; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to mindful-based eating awareness training

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Kristeller et al. (2014)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=140 Psychoeducational CBT 5 mo (N=48) MB-EAT 5 mo (N=50) WLC 5 mo (N=42) With BED (N=35 vs. 31 vs. 31	Inclusion: Overweight or obese; BMI >= 28 kg/m^2 Exclusion: Suicidal symptomology; psychiatric symptoms potentially likely to interfere with group participation or follow-up; psychotic symptoms; drug or alcohol abuse; unstable medication use; previous regular meditation practice; concurrent participation in a weight loss program; concurrent psychotherapy focused on	Overweight or Obesity: 140 (100%) BED: 35 (70%, N=50) vs. 31 (58.49%, N=53) vs. 31 (65.96%, N=47) Weight: 242.7 lbs (N=150) BMI >= 28 kg/m²: 140 (100%) BMI: 40.26 kg/m² (N=150)	Compared to WLC, psychoeducational CBT and MB-EAT showed comparable improvement at 1-mo post-intervention on binge days per mo (15.31->5.23 d/mo CBT vs. 14.84->4.78 MB-EAT vs. 14.04->12.83 WLC). The proportion of individuals with no BED diagnosis at 1-mo post-treatment was 75% with psychoeducational CBT vs. 95%	High

		Follow-up: Baseline – 6 mo	weight or eating issues; purging or laxative abuse within 6 months	Age: 46.55 yr (N=150) Gender - Female: 132 (88%, N=150) - Male: 18 (12%, N=150) Race - Black or African American: 20 (13.33%, N=150) - Minority: 21 (14%, N=150) Ethnicity, Hispanic/Latino: 1 (0.67%, N=150)	with MB-EAT vs. 48% with WLC but attrition was considerable. BMI – Baseline: 39.04 kg/m² (SD ± 8.61, N=27) vs. 39.63 kg/m² (SD ± 7.99, N=39) vs. 38.14 kg/m² (SD ± 6.42, N=26) BMI, Change - Baseline – 6 mo: -0.11 kg/m² (SD ± 6.83, N=27) vs. 0.42 kg/m² (SD ± 6.76, N=39) vs. 0.28 kg/m² (SD ± 5.01, N=26) Study Withdrawal, Treatment Dissatisfaction - Baseline – 6 mo: 5 (10%, N=50) vs. 0 (0%, N=53) vs. 0 (0%, N=47) Attrition: 43% (21/48) vs. 22% (11/50) vs. 38% (16/42)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; d=day; MB-EAT=mindfulness-based eating awareness training; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; WLC=wait-list control; yr=year

Compared to group dialectical behavior therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Lammers et al. (2020)	Design: RCT Setting: Outpatient Country: Netherlands Funding: NR	Randomized N=74 Group CBT 20 wk (N=33) Group DBT 20 wk (N=41) Follow-up: Baseline – 44 wk	Inclusion: Overweight or obese; BMI >= 30 kg/m^2; BED; an above average urge to eat in response to negative emotions (score >= 2.38 on the DEBQ subscale Emotional Eating Exclusion: Previous CBT treatment; current substance abuse, psychosis, suicidality; severe personality disorder;	BED: 74 (100%) Overweight or Obesity: 74 (100%) BED Duration: 15.3 yr (SD ± 10.9) BMI >= 30 kg/m²: 74 (100%)	The CBT group experienced greater reductions in EDE-Q Global score at the end of treatment (p=0.060) and at follow-up (p=0.020): Baseline->End of Treatment->Follow-Up: 3.06 ->1.64->1.61 units vs. 3.48->2.31->2.35 units	High

			obesity caused by physical illness; concurrent treatment for being overweight or for eating disorder	BMI: 39.9 kg/m² (SD ± 5.6) Age: 37.3 yr (SD ± 11.8) Gender - Female: 66 (89.2%) - Male: 8 (10.8%) Race: NR	The CBT group also showed greater reductions in objective binge eating episodes at the end of treatment (p=0.035) but not at follow-up (p=0.095): Baseline->End of Treatment->Follow-Up: 8.27->0.74->1.85 vs. 7.51->1.64->2.75 Attrition: 6% (2/33) vs. 12% (5/41)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; DBT=dialectical behavior therapy; EDE-Q=Eating Disorder Examination Questionnaire; DEBQ=Dutch Eating Behavior Questionnaire; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to with ecological momentary assessment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Le Grange et al. (2002)	Design: RCT Setting: NR Country: United States Funding: Non-profit	Randomized N=41 Group CBT 12 wk (N=22) Group CBT + EMA 12 wk (N=19) Follow-up: Baseline – 64 wk	Inclusion: 18-65 years of age; BED; BMI >=27 kg/m^2; female Exclusion: Purged or self-induced vomiting more than once per mo on average during the preceding 6 months; laxative use or diuretic use as a means of weight control more than once per mo on average during the preceding 6 months; receiving concurrent treatment for weight loss; currently taking appetite suppressants; suffering from any medical condition that may impact weight; pregnancy; diabetes; thyroid conditions	BED: 41 (100%) BED, Duration: 27.7 yr (SD ± 11.7) BMI >= 27 kg/m²: 41 (100%) BMI: 37.9 kg/m² (SD ± 8.2) Age 18 yr-65 yr: 41 (100%) Age: 44.2 yr (SD ± 8.5) Gender, Female: 41 (100%) Race, Caucasian: 38 (93%)	Both groups showed a decrease in binge episodes/wk, presence of BED, and other rating scale measures, without any added benefit of EMA. Binge Eating - Baseline: 4.27/wk (SD ± 2.95) vs. 3.95/wk (SD ± 1.75) Binge Eating, Change - Baseline – 12 wk: -2.16/wk (SD ± 2.11) vs. -2.31/wk (SD ± 1.72) - Baseline – 64 wk: -2.09/wk (SD ± 2.23) vs. -1.67/wk (SD ± 1.69) BED - 12 wk: 13 (59%) vs. 7 (37%) (p=0.15)	High

					- 64 wk: 12 (55%) vs 11 (58%) (p=0.83) BMI – Baseline: 37.77 kg/m² (SD ± 8.21) vs. 35.53 kg/m² (SD ± 7.69) BMI, Change - Baseline – 12 wk: 0.14 kg/m² (SD ± 6.42) vs. 0.62 kg/m² (SD ± 6.39) - Baseline – 64 wk: 2.16 kg/m² (SD ± 7.12) vs. 1.7 kg/m² (SD ± 6.75) Attrition: 27% (6/22) vs. 37% (7/19)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; EMA=ecological momentary assessment; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to schema therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
McIntosh et al. (2016)	Design: RCT Setting: NR Country: New Zealand Funding: Government	Randomized N=112 CBT 12 mo (N=38) Appetite-Focused CBT12 mo (N=36) Schema Therapy 12 mo (N=38) Follow-up: Baseline – 24 mo	Inclusion: Female; 16-65 years of age; BED or BN; the subjective experience of dyscontrol Exclusion: Severe major depression; serious suicidal intent; severe psychoactive substance dependence; bipolar I disorder; schizophrenia; severe physical illness; severe medical complications of the eating disorder; cognitive impairment; psychotropic medication; an adequate trial of CBT in the past year; an adequate trial of schema	BED or BN: 112 (100%) - BED: 18 (48.2%) vs. 18 (50%) vs. 18 (47.4%) - BN: 20 (51.8%) vs. 18 (50%) vs. 20 (52.6%) Eating Disorder, Duration: 15.2 yr (SD ± 12.7) - 14.6 yr (SD ± 13.2) vs. 15.4 yr (SD ± 13.9) vs. 15.7 yr (SD ± 11.4) Weight: 83.2 kg (SD ± 22.4) BMI: 29.9 kg/m² (SD ± 7.8) Age 16 yr-65 yr: 112 (100%)	Binge-eating abstinence was not statistically different between groups either at 12 mo or 24 mo. - At 24 mo: 53.3% vs. 67.9% vs. 62.1% Weight – Baseline: 83 kg (SD ± 22.4) vs. 84.7 kg (SD ± 23.8) vs. 82 kg (SD ± 21.5) Weight, Change - Baseline – 12 mo: 1.4 kg (SD ± 5.55) vs. -1 kg (SD ± 5.4) vs. 1.5 kg (SD ± 5.55) - Baseline – 24 mo: 0.79 kg (SD ± 7.67, N=30) vs. -0.056 kg (SD ± 7.41, N=28)	High

			therapy in the past year; currently underweight	Age: 35.3 yr (SD ± 12.6) - 34.4 yr (SD ± 13) vs. 34.3 yr (SD ± 11.9) vs. 37.1 yr (SD ± 12.9) AN, Lifetime: 4 (10.53%) vs. 2 (5.56%) vs. 2 (5%) Gender, Female: 112 (100%) Race - Caucasian: 19 (17%) - Asian: 4 (4%) - Pacific Islander: 0 (0%) Nationality, New Zealand and Race, Caucasian: 75 (67%) Nationality, New Zealand and Race, Maori: 11 (10%)	vs. 0.8 kg (SD ± 7.54, N=29) Disease Response, Remission - 12 mo: 13 (34.2%) vs. 20 (55.6%) vs. 20 (52.6%) - 24 mo: 16 (53.3%, N=30) vs. 19 (67.9%, N=28) vs. 17 (58.6%, N=29) Attrition: 29% (11/38) vs. 36% (13/36) vs. 24% (9/38)	

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to group behavioral treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Nauta et al. (2000, 2001)	Design: RCT; Follow-up Setting: NR Country: Netherlands Funding: NR	Randomized N=74 Subjects with BED (N=37) - Group Cognitive Treatment 15 wk (N=21) - Group Behavioral Treatment 15 wk (N=16)	Inclusion: Obese; women; between 18 and 50 years of age; BMI of 27 kg/m^2 or higher Exclusion: Participation in a weight-loss program; current alcohol or drug dependence; psychosis; met some criteria for BED, but not all	BED: 37 (50%) BED, Duration: 12.5 yr (SD ± 6.4) Obesity: 74 (100%) BMI >= 27 kg/m²: 74 (100%) Age 18 yr-50 yr: 74 (100%)	Cognitive treatment reduced binge eating more than behavioral treatment at 41 wk (91% vs. 75%), but other comparisons did not differ. Cognitive treatment was noted to be better at 1 yr in shape, weight, and eating concerns. Cognitive treatment showed greater binge-eating abstinence at 67 wk follow-up: 15 (83%,	High

		Follow-up: Baseline – 67 wk		Age: 38.3 yr (SD ± 7.1) Gender, Female: 74 (100%) Race: NR	N=18) vs. 7 (54%, N=13) (p=0.08). Weight loss was minimal (0.3 kg) with cognitive treatment vs. 3 kg with behavioral treatment at 1 yr. Attrition: 14% (3/21) vs. 19% (3/16)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to group cognitive-behavioral therapy with exercise

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Pendleton et al. (2002)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=114 Group CBT 4 mo (N=17) Group CBT + Exercise 4 mo (N=20) Group CBT 4 mo > 10 mo (Maintenance) (N=23) Group CBT 4 mo > + Exercise 10 mo (Maintenance) (N=24) Observational Period: Baseline – 16 mo	Inclusion: Females; 25-60 years of age; >30 lbs overweight; binge eating; history of sedentary lifestyle and occupation; nonsmoker Exclusion: History of cardiovascular disease, diabetes, metabolic disorder, or gastrointestinal disorder or surgery; history of drug abuse	Binge Eating: 114 (100%) Binge Eating: 4.8 d/wk (SD ± 2) vs. 4.6 d/wk (SD ± 2.1) vs. 4.6 d/wk (SD ± 1.9) vs. 4.2 d/wk (SD ± 2.3) Weight: 97.2 kg (SD ± 17.8, N=84) BMI: 36.2 kg/m² (SD ± 6.5, N=84) Age: 45 yr (SD ± 8.3, N=84) Gender, Female: 114 (100%) Race - Caucasian: 64 (76%, N=84) - Black or African American: 11 (13%, N=84)	Group CBT + Exercise with maintenance was superior on binge eating d/wk to group CBT alone for 4 mo: 4.2->0.6 d/wk vs. 4.8->1.9 d/wk (MD -1.3 d/wk, p=0.039). - 10 mo: 0.5 vs. 2 d/wk (MD -1.5 d/wk, p=0.002) - 16 mo: 1 vs. 2.5 d/wk (MD -1.5 d/wk, p=0.007) Binge abstinence at 16 mo was: 18% CBT alone, 65% CBT+Exercise, 39% CBT alone+maintenance, 58% CBT+Exercise+maintenance. BMI was significantly reduced in the subjects in both the exercise and maintenance conditions at 16 mo: 1.33 kg/m² (SD ± 2) vs. -0.75 kg/m² (SD ± 2.4) vs. -0.24	High

				- Mexican American: 7 (8%, N=84) - Other: 3 (3%, N=84)	kg/m² (SD ± 3) vs. -2.26 kg/m² (SD ± 3.9) Overall Attrition: 26% (30/114)	

Abbreviations: BMI=body mass index; CBT=cognitive-behavioral therapy; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to Pharmacotherapy

Compared to fluoxetine and placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Devlin et al. (2005)	Design: RCT Setting: NR Country: United States Funding: Industry and government	Randomized N=116 Placebo 5 mo (N=31) Fluoxetine 60 mg 5 mo (N=32) Individual CBT + Placebo 5 mo (N=25) Individual CBT + Fluoxetine 60 mg 5 mo (N=28) (All received Group Behavioral Weight Control Treatment) Fluoxetine 60 mg +/- Individual CBT 5 mo (pooled) (N=60) Individual CBT + Placebo/Fluoxetine 60	Inclusion: 18-70 years of age; BMI ≥ 27 kg/m^2; maximum weight of 159 kg; BED for at least 6 months; overweight or obese Exclusion: Substance-related disorders within the past yr; acutely suicidal; current psychotic disorder, bipolar disorder, major depressive disorder with melancholic features, or BN; history of AN or psychotic disorder; concurrent eating or weight control treatment; currently taking antidepressants, mood stabilizers, or appetite suppressants; MAOIs within the prior 2 weeks; previously had an adverse reaction to fluoxetine	Overweight or Obesity: 116 (100%) BED, Duration >= 6 mo: 116 (100%) Binge Eating, Duration: 27.3 yr (SD ± 14.7) BMI >= 27 kg/m²: 116 (100%) BMI: 40.9 kg/m² (SD ± 6.9) - 40.3 kg/m² (SD ± 7.1) vs. 40.1 kg/m² (SD ± 6.6) vs. 41.1 kg/m² (SD ± 7.6) vs. 42.1 kg/m² (SD ± 6.9) Weight <= 159 kg: 116 (100%) Age 18 yr-70 yr: 116 (100%) Age: 43 yr (SD ± 12) - 44.1 yr (SD ± 10.2) vs. 45.9 yr (SD ± 13.6) vs. 43.4 yr (SD ± 11.8) vs. 39.4 yr (SD ± 12.1)	No significant difference was noted in weight change: - Baseline: 113.5 kg (SD ± 22.2) vs. 113.8 kg (SD ± 22.8) vs. 116.5 kg (SD ± 22.2) vs. 116.9 kg (SD ± 20.8) - Baseline – 5 mo Change: -2.4 kg (SD ± 5.9) vs. -1.9 kg (SD ± 6.9) vs. -1.9 kg (SD ± 7.1) vs. -4.1 kg (SD ± 6.9) CBT +/- Fluoxetine had more binge-eating abstinence than no CBT at 5 mo: 33 (62%) vs. 21 (33%) (p<0.001) BDI - Baseline: 15.6 units (SD ± 9.3) vs. 14.5 units (SD ± 7.2) vs. 13.9 units (SD ± 10.6) vs. 16.9 units (SD ± 9.1) BDI, Change - Baseline – 5 mo: -5 units (SD ± 7.5) vs. -7 units (SD ± 8.5) vs. -5.5 units (SD ± 8.5) vs. -10.6 units (SD ± 9.3)	High

		mg 5 mo (pooled) (N=53) Placebo/Fluoxetine 60 mg 5 mo (pooled) (N=63)		Gender - Female: 90 (78%) - Male: 26 (22%) Race - Caucasian: 89 (77%) - Black or African American: 14 (12%) - Multiracial or Other: 1 (1%) Ethnicity, Hispanic/Latino: 12 (10%)	Attrition: 48% (15/31) vs. 31% (10/32) vs. 40% (10/25) vs. 25% (7/28)	
Grilo et al. (2005a, 2012b)	Design: RCT; Follow-up Setting: NR Country: United States Funding: Government; product donation by industry	Randomized N=108 Current Analysis (N=81) CBT16 wk (N=28) CBT+ Fluoxetine 60 mg 16 wk (N=26) Fluoxetine 60 mg 16 wk (N=27) Placebo 16 wk (N=27) Follow-up: Baseline – 16 mo	Inclusion: 18- 60 years of age; BED; between 100% and 200% of ideal weight for height Exclusion: Concurrent treatment for eating or weight; concurrent treatment for psychiatric problems; medical conditions that influence weight or eating; diabetes; thyroid problems; hypoglycemia; severe psychiatric conditions requiring different treatments; psychosis or bipolar disorder requiring treatment; lactation; pregnancy; purging behaviors	BED: 108 (100%) BMI: 36.3 kg/m² (SD ± 7.9) - 35 kg/m² (SD ± 6.2) vs. 35.7 kg/m² (SD ± 8.3) vs. 38.9 kg/m² (SD ± 9.5) vs. 35.7 kg/m² (SD ± 7.2) Age 18 yr-60 yr: 108 (100%) Age: 44 yr (SD ± 8.6 - 43.6 yr (SD ± 8.5) vs. 44.7 yr (SD ± 8.1) vs. 44.3 yr (SD ± 9.5) vs. 43.6 yr (SD ± 8.5) Gender - Female: 22 (78.6%) vs. 20 (76.9%) vs. 19 (70.4%) vs. 23 (85.2%) - Male: 6 (21.4%) vs. 6 (23.1%) vs. 8 (29.6%) vs. 4 (14.8%) Race - Caucasian: 26 (92.9%) vs. 23 (88.5%) vs. 27 (100%) vs. 20 (74.1%) - Black or African American: 2 (7.1%) vs. 2	Remission rates at 16 wk were much higher in both CBT groups: 61% vs. 50% vs. 22% vs. 26%. - CBT vs. Placebo: p=0.008 - CBT+ Fluoxetine vs. Placebo: p=0.05 - CBT+ Fluoxetine vs. Fluoxetine: p=0.03 - CBT+ Fluoxetine vs. CBT: p=0.42 - Fluoxetine vs. Placebo: p=0.83 In the 12-mo follow-up study, these conclusions persisted, and the CBT groups were more likely to achieve remission, though the rates were less at the end of treatment: 36% vs. 27% vs. 4% vs. NR - CBT vs. Fluoxetine: p=0.005 - CBT+ Fluoxetine vs. Fluoxetine: p=0.024 - CBT vs. CBT+ Fluoxetine: p=0.57 Weight loss was modest in all treatment groups.	High

				(7.7%) vs. 0 (0%) vs. 5 (18.5%) Ethnicity, Hispanic/Latino: 0 (0%) vs. 1 (3.8%) vs. 0 (0%) vs. 2 (7.4%)	- Change - Baseline – 16 mo: -9.84 lbs vs. -4.13 lbs vs. -1.48 lbs vs. NR BDI Baseline: 16.5 units (SD ± 8.4) vs. 20.2 units (SD ± 12.1) vs. 16.9 units (SD ± 8.4) vs. 18.7 units (SD ± 9.7) 16 wk: 6.5 units (SD ± 6.8) vs. 9.2 units (SD ± 7.3) vs. 11.8 units (SD ± 9.8) vs. 11.7 units (SD ± 10.3) - CBT vs. Placebo: MD -5.2 units (p=0.04) - CBT vs. Fluoxetine: MD -5.3 units (p=0.01) - CBT+ Fluoxetine vs. Fluoxetine: MD -2.6 units (p=0.04) Attrition: 21% (6/28) vs. 23% (6/26) vs. 22% (6/27) vs. 15% (4/27)	

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; MAOI=monoamine oxidase inhibitor; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Compared to fluoxetine and fluvoxamine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Ricca et al. (2001)	Design: RCT Setting: Outpatient: the Outpatient Clinic for Eating Disorders of the Units of Psychiatry and	Randomized N=108 CBT24 wk (N=20) CBT+ Fluoxetine 20-60 mg 24 wk (N=22)	Inclusion: BED; 18-45 years of age Exclusion: Diabetes mellitus; thyroid disorders; any other disease interfering with eating behavior; pregnancy; lactation; heart disease	BED: 108 (100%) Binge Eating, Duration: 5.6 yr (SD ± 5) - 6.4 yr (SD ± 6) vs. 4.9 yr (SD ± 5.1) vs. 4.8 yr (SD ± 4.4) vs. 5.1 yr (SD ± 4.7) vs. 5.3 yr (SD ± 4.8)	BMI scores were significantly reduced at 24 wk in CBT groups (-2.2 kg/m² vs. -3.8 kg/m² vs. NR vs. -0.7 kg/m² vs. NR). Improvements persisted at 1-yr follow-up but with some weight regain with fluoxetine alone (-1.6	High

	Endocrinology of the University of Florence Country: Italy Funding: Reimbursed by government	CBT+ Fluvoxamine 100-300 mg 24 wk (N=23) Fluoxetine 20-60 mg 24 wk (N=21) Fluvoxamine 100-300 mg 24 wk (N=22) Follow-up: Baseline – 76 wk		BMI: 32.3 kg/m² (SD ± 5.8) - 32 kg/m² (SD ± 6) vs. 31.7 kg/m² (SD ± 5.6) vs. 32.5 kg/m² (SD ± 6.1) vs. 32.1 kg/m² (SD ± 3.8) vs. 32.7 kg/m² (SD ± 4.1) Mental Disorder, Other and not BED: 15 (13.89%) Age 18 yr-45 yr: 108 (100%) Age: 25.9 yr (SD ± 6.8) - 26.3 yr (SD ± 6.7) vs. 25.2 yr (SD ± 6.3) vs. 25.1 yr (SD ± 6.9) vs. 25.1 yr (SD ± 6.1) vs. 26.1 yr (SD ± 5.9) Gender - Female: 13 (65%) vs. 13 (59.09%) vs. 13 (56.52%) vs. 12 (57.14%) vs. 13 (59.09%) - Male: 7 (35%) vs. 9 (40.91%) vs. 10 (43.48%) vs. 9 (42.86%) vs. 9 (40.91%) Race: NR	kg/m² vs. -3.3 kg/m² vs. NR vs. 0.5 kg/m² vs. NR). BDI - Baseline: 22 units vs. 16.5 units vs. 22 units vs. 20 units vs. 21 units - Baseline – 24 wk: -8 units (SD ± 9.62, N=17) vs. -6 units (SD ± 12.96, N=16) vs. NR (N=18) vs. -5 units (SD ± 10.21, N=16) vs. NR (N=16) - Baseline – 76 wk: -8 units (SD ± 9.93, N=17) vs. -6 units (SD ± 12.96, N=16) vs. NR (N=18) vs. -4 units (SD ± 10.39, N=16) vs. NR (N=16) Adverse Events - Baseline – 24 wk: 0 (0%) vs. 6 (27.2%) vs. 6 (26.09%) vs. 7 (33.33%) vs. 7 (31.82%) Treatment Discontinuation, Adverse Events - Baseline – 24 wk: 0 (0%) vs. 3 (13.64%) vs. 3 (13.04%) vs. 2 (9.52%) vs. 4 (18.18%) Attrition: 15% (3/20) vs. 27% (6/22) vs. 22% (5/23) vs. 24% (5/21) vs. 27% (6/22)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to sertraline

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Brambilla et al. (2009)	Design: RCT Setting: Single Center; Eating Disorder Center, Sacco Hospital Country: Italy Funding: NR	Randomized N=30 CBT + Nutritional Counseling 6 mo (N=10) Sertraline 50-150 mg + CBT + Diet Therapy 6 mo (N=10) Sertraline 50-150 mg + Topiramate 25-150 mg + CBT + Diet Therapy 6 mo (N=10) (All received CBT)	Inclusion: BED; obese; female Exclusion: DSM-IV Axis I or Axis II disorders; pharmacological treatments in the past 6 months	BED: 30 (100%) Obesity: 30 (100%) Binge Eating, Duration: 13 yr (SD ± 6) vs. 9 yr (SD ± 5) vs. 15 yr (SD ± 10) Age: 46 yr (SD ± 8) vs. 45 yr (SD ± 11) vs. 47 yr (SD ± 8) Gender, Female: 30 (100%) Race: NR	Binge-eating frequency significantly decreased in sertraline + topiramate group: 5->4 episodes/wk vs. 6->5 episodes/wk vs. 5->2 episodes/wk. Weight decreased more in sertraline + topiramate group: 88->87 kg vs. 86->84 kg vs. 105->93 kg. Adverse effects were not reported. Overall Attrition: 0% (0/30)	High

Abbreviations: BED=binge-eating disorder; CBT=cognitive-behavioral therapy; DSM-IV= Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to methylphenidate

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Quilty et al. (2019)	Design: RCT Setting: Outpatient Country: Canada Funding: Non-profit	Randomized N=49 CBT 12 wk (N=27) Long-Acting Methylphenidate 18-72 mg 12 wk (N=22)	Inclusion: BED; BMI above 25; female; 18-50 years of age Exclusion: Current pregnancy or lactation; psychotherapy or behavioral treatment for eating or weight; psychotropic or investigational medication changes; current mental disorders; current severe suicidality or homicidality; current uncontrolled medical conditions; other serious medical illnesses or events; history of seizures or tics; uncontrolled or clinically relevant hypertension	BED: 49 (100%) Age: 32.78 yr (SD ± 8.62) vs. 45 yr (SD ± 11) vs. 47 yr (SD ± 8) Gender, Female: 49 (100%) Race - Caucasian: 77.6% (N=38) - South Asian: 2.0% (N=1) - Black 4.1% (N=2) - Other: 16.3% (N=8)	Both groups experienced fewer objective and subjective binge episodes at post-treatment than at baseline (p<0.001). Binge Eating, Objective – Baseline->Post-Treatment: 2.26 (SD ± 1.89)->0.11 (SD ± 0.32) vs. 2.19 (SD ± 1.47)->0.69 (SD ± 1.49) Binge Eating, Subjective – Baseline->Post-Treatment: 5.59 (SD ± 5.92)->0.26 (SD ± 0.45)	High

			(>140/90), tachycardia (heart rate > 110), arrhythmias or conduction abnormalities		vs. 4.62 (SD ± 4.65)->1.38 (SD ± 3.18) There was a significant difference in BMI for the methylphenidate group, (p< 0.001), but not for the CBT group (p=0.13). BMI - Baseline->Post-Treatment: 39.26 (SD ± 8.80)->40.16 (SD ± 9.45) vs. 36.53 (SD ± 6.55)->34.38 (SD ± 6.22) Attrition: 26% (7/27) vs. 22% (5/22)	

Abbreviations: BED=binge-eating disorder; BMI=BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Adjunctive Cognitive-Behavioral Therapy With Psychotherapy

Cognitive-behavioral therapy with weight loss treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1994b)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=108 CBT 12 wk > Weight Loss Treatment 36 wk (N=36) CBT 12 wk > Weight Loss Therapy + Desipramine 25-300 mg 36 wk (N=36) Weight Loss Treatment 36 wk (N=37)	Inclusion: Female; BED; binge eating at least twice a wk for a 6-mo period; overweight Exclusion: Current weight loss program; antidepressant medication; any medication that may affect weight; suicidality; abuse of drugs or alcohol; history of purging in the prior 12 months; BMI below 27 kg/m^2; current BN	BED: 108 (100%) Binge Eating >= 2/wk, Duration 6 mo: 108 (100%) Binge Eating: 4.5 d/wk (SD ± 1.4) - 4.4 d/wk (SD ± 1.4, N=30) vs. 5.1 d/wk (SD ± 1.4, N=27) vs. 4.5 d/wk (SD ± 1.6, N=27) Overweight: 108 (100%) Weight: 104.9 kg (SD ± 18.5)	At 12 wk, CBT groups had significantly less binge eating (67% reduction vs. 44% with weight loss alone, MD -23 %, p<0.01) and the weight loss group had more weight loss (-2.0 kg) compared to CBT groups (0.7 kg) (MD 2.7 kg, p<0.002). No differences were noted between groups at the end of treatment or follow-up except weight loss (0 kg vs. -4.8 kg vs. -4.15 kg at 48 wk) - CBT > Weight Loss Treatment vs. CBT >	High

		CBT > Weight Loss Therapy +/- Desipramine 25-300 mg 36 wk (pooled) (N=72) Follow-up: Baseline – 48 wk Current Analysis (N=84) - 30 vs. 27 vs. 27		- 102.1 kg (SD ± 15.7, N=30) vs. 111.9 kg (SD ± 17.4, N=27) vs. 102.9 kg (SD ± 15.8, N=27) BMI: 38.6 kg/m² (SD ± 6.6) Age: 45 yr (SD ± 10) Gender, Female: 108 (100%) Race: NR	Weight Loss Therapy + Desipramine: MD 4.8 kg (p<0.05) Binge Eating, Abstinence - 48 wk: 8 (28%, N=30) vs. 9 (32%, N=27) vs. 4 (14%, N=27) BDI – Baseline: 13.5 units (SD ± 7.8, N=30) vs. 13.7 units (SD ± 8.1, N=27) vs. 12.9 units (SD ± 6.5, N=27) BDI, Change - Baseline – 36 wk: -4.6 units (SD ± 10.5, N=30) vs. -5.9 units (SD ± 10.84, N=27) vs. -1.6 units (SD ± 11.79, N=27) Attrition: 17% (11/36) vs. 23% (12/36) vs. 27% (16/37)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Cognitive-behavioral therapy with nutritional counseling or with diet therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Brambilla et al. (2009)	Design: RCT Setting: Single Center; Eating Disorder Center, Sacco Hospital Country: Italy Funding: NR	Randomized N=30 CBT + Nutritional Counseling 6 mo (N=10) Sertraline 50-150 mg + CBT + Diet Therapy 6 mo (N=10)	Inclusion: BED; obese; female Exclusion: DSM-IV Axis I or Axis II disorders; pharmacological treatments in the past 6 months	BED: 30 (100%) Obesity: 30 (100%) Binge Eating, Duration: 13 yr (SD ± 6) vs. 9 yr (SD ± 5) vs. 15 yr (SD ± 10) Age: 46 yr (SD ± 8) vs. 45 yr (SD ± 11) vs. 47 yr (SD ± 8)	Binge-eating frequency significantly decreased in sertraline + topiramate group: 5->4 episodes/wk vs. 6->5 episodes/wk vs. 5->2 episodes/wk. Weight decreased more in sertraline + topiramate group: 88->87 kg vs. 86->84 kg vs. 105->93 kg.	High

		Sertraline 50-150 mg + Topiramate 25-150 mg + CBT + Diet Therapy 6 mo (N=10) (All received CBT)		Gender, Female: 30 (100%) Race: NR	Adverse effects were not reported. Overall Attrition: 0% (0/30)	

Abbreviations: BED=binge-eating disorder; CBT=cognitive-behavioral therapy; DSM-IV= Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Group cognitive-behavioral therapy with very low calorie diet

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
de Zwaan et al. (2005)	Design: Sub-Group Analysis of RCT Setting: NR Country: United States Funding: NR	Randomized N=71 Very Low Calorie Diet (1200 cal/d + low level exercise) 16 wk > (+) Group CBT 26 wk (N=36) Very Low Calorie Diet (1200 cal/d + low level exercise) 26 wk (N=35) Follow-up: Baseline – 18 mo	Inclusion: 18-55 years of age; at least 50 lb above "ideal" body weight; BED; women; obese Exclusion: Current use of any psychotropic medication; current evidence of psychosis; current evidence of suicidality; current evidence of chemical abuse; current psychiatric treatment; current obesity treatment	BED: 71 (100%) Obesity: 71 (100%) BMI: 36.1 kg/m² - 36.1 kg/m² (SD ± 3.7) vs. 35.7 kg/m² (SD ± 4.1) Age 18 yr-55 yr: 71 (100%) Age: 39.3 yr - 40.9 yr (SD ± 7.7) vs. 37.7 yr (SD ± 6.5) Gender, Female: 71 (100%) Race, Caucasian: 69 (97.2%) - 35 (97.2%) vs. 34 (97.1%)	Change in binges/wk was numerically less in the very low calorie diet +CBT group (3.9->2.3/wk vs. 6.2->1.5/wk) at 7 mo. Binge eating abstinence rates at 18 mo were comparable: 11 (33.3%) vs. 10 (32.3%). The mean total weight loss at the end of the very low-calorie diet program was 35.2 lb or 16.1% (SD=8.2) of the original weight with mean weight loss of 5.5% of initial body weight at 1 yr follow-up. Weight - Baseline: 217.3 lbs (SD ± 24.8) vs. 214.9 lbs (SD ± 27.9) Weight, Change - Baseline – 24 wk: -34.2 lbs (SD ± 21.36) vs. -36.3 lbs (SD ± 20.66)	High

					- Baseline – 18 mo: -12.4 lbs (SD ± 22.85, N=31) vs. -12.2 lbs (SD ± 22.65, N=31) Weight, Regain of Lost >= 50 % - 24 wk – 18 mo: 12 (39.2%, N=31) vs. 17 (56.3%, N=31) (p=0.19) Adherence, Sessions Unattended, Diet Therapy - Baseline – 24 wk: 3.1 (SD ± 2.6) vs. 5.3 (SD ± 5) (MD -2.2, p=0.02) Attrition: 6% (2/36) vs. 20% (7/36)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Group cognitive-behavioral therapy with group behavioral weight loss treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Grilo et al. (2011)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=125 Group CBT 24 wk (N=45) Group CBT + Group BWL Treatment 40 wk (N=35) Group BWL Treatment 24 wk (N=45) Follow-up Period:	Inclusion: Obese; BED; 18-60 years of age; BMI 30-55 kg/m^2 Exclusion: Concurrent treatment for eating problems or weight problems; psychosis or bipolar disorder requiring alternative treatment	BED: 125 (100%) Obesity: 125 (100%) BMI 30 kg/m²-55 kg/m²: 125 (100%) BMI: 38.8 kg/m² (SD ± 5.8) - 39.3 kg/m² (SD ± 6.1) vs. 39 kg/m² (SD ± 6.1) vs. 38 kg/m² (SD ± 5.3) Weight: 250.1 lbs (SD ± 52.6) vs. 237.2 lbs (SD ± 42.8) vs. 242.7 lbs (SD ± 45.8)	At 12-mo follow-up, ITT binge-eating remission rates were 51% with CBT, 40% with CBT + BWL, and 36% with BWL. Binge eating with CBT had greater reductions at 24 wk than BWL (15.6->2.2/mo vs. 14.9->4.6/mo) and these differences were maintained at 50-wk follow-up. At post-treatment, BWL or CBT+ BWL had significantly greater percent BMI reduction than CBT alone:	High

		Baseline – 76 wk for Group CBT or Group BWL Treatment Baseline – 92 wk for Group CBT + Group BWL Treatment		Age 18 yr-60 yr: 125 (100%) Age: 44.8 yr (SD ± 9.4) - 45.2 yr (SD ± 8.5) vs. 44.5 yr (SD ± 9.2) vs. 44.6 yr (SD ± 10.5) Gender - Female: 28 (64.4%) vs. 28 (80%) vs. 28 (62.2%) - Male: 17 (35.6%) vs. 7 (20%) vs. 17 (37.8%) Race - Caucasian: 34 (75.6%) vs. 26 (74.3%) vs. 36 (80%) - Black or African American: 5 (11.1%) vs. 8 (22.9%) vs. 7 (15.6%) - Asian: 2 (4.4%) vs. 1 (2.9%) vs. 0 (0%) - American Indian/Alaskan Native: 1 (2.2%) vs. 0 (0%) vs. 0 (0%) Ethnicity, Hispanic/Latino: 3 (6.7%) vs. 0 (0%) vs. 2 (4.4%)	- -0.5% (SD ± 3.5) with CBT vs. -2.6% (SD ± 5.3) with BWL (MD -2.1 %, p=0.03) - -0.5% (SD ± 3.5) with CBT vs. -2.7% (SD ± 6) with CBT+ BWL (MD -2.2 %, p=0.04) Attrition: 24% (11/45) vs. 40% (14/35) vs. 31% (14/45)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; BWL=behavioral weight loss; CBT=cognitive-behavioral therapy; ITT=intention-to-treat; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Group cognitive-behavioral therapy with ecological momentary assessment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Le Grange et al. (2002)	Design: RCT Setting: NR	Randomized N=41 Group CBT 12 wk (N=22)	Inclusion: 18-65 years of age; BED; BMI >=27 kg/m^2; female Exclusion: Purged or self-induced vomiting more than	BED: 41 (100%) BED, Duration: 27.7 yr (SD ± 11.7)	Both groups showed a decrease in binge episodes/wk, presence of BED, and other rating scale measures, without any added benefit of EMA.	High

	Country: United States Funding: Non-profit	Group CBT + EMA 12 wk (N=19) Follow-up: Baseline – 64 wk	once per mo on average during the preceding 6 months; laxative use or diuretic use as a means of weight control more than once per mo on average during the preceding 6 months; receiving concurrent treatment for weight loss; currently taking appetite suppressants; suffering from any medical condition that may impact weight; pregnancy; diabetes; thyroid conditions	BMI >= 27 kg/m²: 41 (100%) BMI: 37.9 kg/m² (SD ± 8.2 Age 18 yr-65 yr: 41 (100%) Age: 44.2 yr (SD ± 8.5) Gender, Female: 41 (100%) Race, Caucasian: 38 (93%)	Binge Eating - Baseline: 4.27/wk (SD ± 2.95) vs. 3.95/wk (SD ± 1.75) Binge Eating, Change - Baseline – 12 wk: -2.16/wk (SD ± 2.11) vs. -2.31/wk (SD ± 1.72) - Baseline – 64 wk: -2.09/wk (SD ± 2.23) vs. -1.67/wk (SD ± 1.69) BED - 12 wk: 13 (59%) vs. 7 (37%) (p=0.15) - 64 wk: 12 (55%) vs 11 (58%) (p=0.83) BMI – Baseline: 37.77 kg/m² (SD ± 8.21) vs. 35.53 kg/m² (SD ± 7.69) BMI, Change - Baseline – 12 wk: 0.14 kg/m² (SD ± 6.42) vs. 0.62 kg/m² (SD ± 6.39) - Baseline – 64 wk: 2.16 kg/m² (SD ± 7.12) vs. 1.7 kg/m² (SD ± 6.75) Attrition: 27% (6/22) vs. 37% (7/19)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; EMA=ecological momentary assessment; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Cognitive-behavioral therapy with general nutrition counseling compared to cognitive-behavioral therapy with low energy-density diet

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias

Masheb et al. (2011)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=50 General Nutrition Counseling + CBT 6 mo (N=25) Low Energy-Density Diet + CBT 6 mo (N=25) Follow-up: Baseline – 12 mo	Inclusion: BED; 21-60 years of age; obese; BMI of 30 kg/m^2 or greater Exclusion: Co-existing psychiatric conditions requiring alternative treatments; co-existing psychiatric conditions requiring hospitalization; current substance dependence; receiving treatment known to affect eating or weight; serious neurologic illness	BED: 50 (100%) Obesity: 50 (100%) BMI >= 30 kg/m²: 50 (100%) BMI: 39.1 kg/m² (SD ± 6.6) - 39 kg/m² (SD ± 6.5) vs. 39.2 kg/m² (SD ± 6.9) Age 21 yr-60 yr: 50 (100%) Age: 45.8 yr (SD ± 7.6) - 43.7 yr (SD ± 6.7) vs. 47.9 yr (SD ± 7.9) Gender - Female: 18 (72%) vs. 20 (80%) - Male: 7 (28%) vs. 5 (20%) Race - Caucasian: 22 (88%) vs. 18 (72%) - Black or African American: 3 (12%) vs. 6 (24%) Ethnicity, Hispanic/Latino: 0 (0%) vs. 1 (4%)	Disease response/remission was comparable (72% nutritional counseling vs. 60% low energy diet, p=0.37) as was the proportion with at least 5% weight decrease at 12 mo (28% vs. 20%, p=0.747). Weight, % Change - Baseline – 6 mo: -1.5% (SD ± 4.2) vs. -3.1% (SD ± 6.2) - Baseline – 12 mo: -1.4% (SD ± 7.6) vs. -2.8% (SD ± 6.1) Adherence, Sessions Completed - Baseline – 6 mo: 19.1 (SD ± 3.1) vs. 16.8 (SD ± 7.4) (MD 2.3, p=0.161) Attrition: 8% (2/25) vs 20% (5/25)	High

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Group cognitive-behavioral therapy with nutritional intervention

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Painot et al. (2001)	Design: RCT	Randomized N=60	Inclusion: BED; obese	BED: 60 (100%)	Scores for depression (p<0.01), anxiety (p<0.01), and eating disorders (p<0.001) are	Moderate

	Setting: NR Country: Switzerland Funding: Non-industry	Group CBT 12 wk (N=35) Group CBT + Nutritional Intervention 12 wk (N=25)	Exclusion: Compensatory behavior in the past 6 months; substance abuse or dependence; concurrent treatment	Obesity: 60 (100%) BMI: 33 kg/m² (SD ± 7.75) Weight: 91 kg (SD ± 11.83) vs. 91 kg (SD ± 15) Age: 42 yr (SD ± 15.49) - 42 yr (SD ± 11.83) vs. 44 yr (SD ± 10 Gender, Female: 60 (100%) Race, Caucasian: 60 (100%)	significantly and similarly improved with both types of treatments although mean weight loss is significant only with the combined approach (-0.5kg vs. -1.9kg, p<0.001). Attrition: NR	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Group cognitive-behavioral therapy with exercise

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Pendleton et al. (2002)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=114 Group CBT 4 mo (N=17) Group CBT + Exercise 4 mo (N=20) Group CBT 4 mo > 10 mo (Maintenance) (N=23) Group CBT 4 mo > + Exercise 10 mo (Maintenance) (N=24)	Inclusion: Females; 25-60 years of age; >30 lbs overweight; binge eating; history of sedentary lifestyle and occupation; nonsmoker Exclusion: History of cardiovascular disease, diabetes, metabolic disorder, or gastrointestinal disorder or surgery; history of drug abuse	Binge Eating: 114 (100%) Binge Eating: 4.8 d/wk (SD ± 2) vs. 4.6 d/wk (SD ± 2.1) vs. 4.6 d/wk (SD ± 1.9) vs. 4.2 d/wk (SD ± 2.3) Weight: 97.2 kg (SD ± 17.8, N=84) BMI: 36.2 kg/m² (SD ± 6.5, N=84) Age: 45 yr (SD ± 8.3, N=84) Gender, Female: 114 (100%)	Group CBT + Exercise with maintenance was superior on binge eating d/wk to group CBT alone for 4 mo: 4.2->0.6 d/wk vs. 4.8->1.9 d/wk (MD -1.3 d/wk, p=0.039). - 10 mo: 0.5 vs. 2 d/wk (MD -1.5 d/wk, p=0.002) - 16 mo: 1 vs. 2.5 d/wk (MD -1.5 d/wk, p=0.007) Binge abstinence at 16 mo was: 18% CBT alone, 65% CBT+Exercise, 39% CBT alone+maintenance, 58% CBT+Exercise+maintenance.	High

		Observational Period: Baseline – 16 mo		Race - Caucasian: 64 (76%, N=84) - Black or African American: 11 (13%, N=84) - Mexican American: 7 (8%, N=84) - Other: 3 (3%, N=84)	BMI was significantly reduced in the subjects in both the exercise and maintenance conditions at 16 mo: 1.33 kg/m² (SD ± 2) vs. -0.75 kg/m² (SD ± 2.4) vs. -0.24 kg/m² (SD ± 3) vs. -2.26 kg/m² (SD ± 3.9) Overall Attrition: 26% (30/114)	

Abbreviations: BMI=body mass index; CBT=cognitive-behavioral therapy; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Adjunctive Cognitive-Behavioral Therapy With Pharmacotherapy

Cognitive-behavioral therapy with desipramine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1994b)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=108 CBT 12 wk > Weight Loss Treatment 36 wk (N=36) CBT 12 wk > Weight Loss Therapy + Desipramine 25-300 mg 36 wk (N=36) Weight Loss Treatment 36 wk (N=37) CBT > Weight Loss Therapy +/- Desipramine 25-300 mg 36 wk (pooled) (N=72)	Inclusion: Female; BED; binge eating at least twice a wk for a 6-mo period; overweight Exclusion: Current weight loss program; antidepressant medication; any medication that may affect weight; suicidality; abuse of drugs or alcohol; history of purging in the prior 12 months; BMI below 27 kg/m^2; current BN	BED: 108 (100%) Binge Eating >= 2/wk, Duration 6 mo: 108 (100%) Binge Eating: 4.5 d/wk (SD ± 1.4) - 4.4 d/wk (SD ± 1.4, N=30) vs. 5.1 d/wk (SD ± 1.4, N=27) vs. 4.5 d/wk (SD ± 1.6, N=27) Overweight: 108 (100%) Weight: 104.9 kg (SD ± 18.5) - 102.1 kg (SD ± 15.7, N=30) vs. 111.9 kg (SD ± 17.4, N=27) vs. 102.9 kg (SD ± 15.8, N=27)	At 12 wk, CBT groups had significantly less binge eating (67% reduction vs. 44% with weight loss alone, MD -23 %, p<0.01) and the weight loss group had more weight loss (-2.0 kg) compared to CBT groups (0.7 kg) (MD 2.7 kg, p<0.002). No differences were noted between groups at the end of treatment or follow-up except weight loss (0 kg vs. -4.8 kg vs. -4.15 kg at 48 wk) - CBT > Weight Loss Treatment vs. CBT > Weight Loss Therapy + Desipramine: MD 4.8 kg (p<0.05)	High

		Follow-up: Baseline – 48 wk Current Analysis (N=84) - 30 vs. 27 vs. 27		BMI: 38.6 kg/m² (SD ± 6.6) Age: 45 yr (SD ± 10) Gender, Female: 108 (100%) Race: NR	Binge Eating, Abstinence - 48 wk: 8 (28%, N=30) vs. 9 (32%, N=27) vs. 4 (14%, N=27) BDI – Baseline: 13.5 units (SD ± 7.8, N=30) vs. 13.7 units (SD ± 8.1, N=27) vs. 12.9 units (SD ± 6.5, N=27) BDI, Change - Baseline – 36 wk: -4.6 units (SD ± 10.5, N=30) vs. -5.9 units (SD ± 10.84, N=27) vs. -1.6 units (SD ± 11.79, N=27) Attrition: 17% (11/36) vs. 23% (12/36) vs. 27% (16/37)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Cognitive-behavioral therapy with sertraline or sertraline and topiramate

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Brambilla et al. (2009)	Design: RCT Setting: Single Center; Eating Disorder Center, Sacco Hospital Country: Italy Funding: NR	Randomized N=30 CBT + Nutritional Counseling 6 mo (N=10) Sertraline 50-150 mg + CBT + Diet Therapy 6 mo (N=10) Sertraline 50-150 mg + Topiramate 25-150 mg	Inclusion: BED; obese; female Exclusion: DSM-IV Axis I or Axis II disorders; pharmacological treatments in the past 6 months	BED: 30 (100%) Obesity: 30 (100%) Binge Eating, Duration: 13 yr (SD ± 6) vs. 9 yr (SD ± 5) vs. 15 yr (SD ± 10) Age: 46 yr (SD ± 8) vs. 45 yr (SD ± 11) vs. 47 yr (SD ± 8) Gender, Female: 30 (100%) Race: NR	Binge-eating frequency significantly decreased in sertraline + topiramate group: 5 ->4 episodes/wk vs. 6->5 episodes/wk vs. 5->2 episodes/wk. Weight decreased more in sertraline + topiramate group: 88->87 kg vs. 86->84 kg vs. 105->93 kg.	High

		+ CBT + Diet Therapy 6 mo (N=10) (All received CBT)			Adverse effects were not reported. Overall Attrition: 0% (0/30)	

Abbreviations: BED=binge-eating disorder; CBT=cognitive-behavioral therapy; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Cognitive-behavioral therapy with topiramate

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Claudino et al. (2007)	Design: RCT Setting: 4 university centers Country: Brazil Funding: Industry	Randomized N=73 Placebo + CBT 21 wk (N=36) Topiramate 200-300mg (up-titrate) + CBT 21 wk (N=37)	Inclusion: Obese; 18-60 years of age; BMI >= 30 kg/m^2; BED; Score of > 17 on the BES Exclusion: Clinically significant schizophrenia, major affective disorders, or alcohol or drug abuse; unstable schizophrenia, major affective disorders, or alcohol or drug abuse; high potential suicide risk; concurrent use of antipsychotics, cyproheptadine, antiepileptics, systemic steroids, or antiobesity agents; psychotherapy for weight loss within 3 months	BED: 73 (100%) BES > 17 units: 73 (100%) Obesity: 73 (100%) BMI >= 30 kg/m²: 73 (100%) BMI: 37.4 kg/m² (SD ± 3.5) vs. 37.4 kg/m² (SD ± 4.9) Weight: 98.4 kg (SD ± 10.9) vs. 96.6 kg (SD ± 16.7) Age 18 yr-60 yr: 73 (100%) Age: 35.4 yr (SD ± 10.7) vs. 41.1 yr (SD ± 9.9) Gender - Female: 34 (94.4%) vs. 36 (97.3%) - Male: 2 (5.6%) vs. 1 (2.7%) Race, Caucasian:19 (52.8%) vs. 23 (62.1%)	Amount and rate of weight reduction was greater with topiramate: -0.9 kg with CBT vs. -6.8 kg with topiramate; 11.5% vs. 36.7% lost more than 10% of body weight (p=0.05). More patients with topiramate achieved remission (61.1% vs. 83.8% (p=0.03)) though reductions in binge frequency did not differ. - Baseline: 3.8/wk (SD ± 1.5) vs. 4.7/wk (SD ± 3.3) - Baseline: 3.4 d/wk (SD ± 1.3) vs. 4.2 d/wk (SD ± 3.4) - % Change - Baseline – 21 wk: -92.9% (SD ± 17.7, N=24) vs. -99.5% (SD ± 2.6, N=29) (MD 6.6 %, p=0.08) BDI - Baseline: 15.9 units (SD ± 9.4) vs. 16.8 units (SD ± 8.3) BDI, Change - Baseline – 21 wk: -6.7 units (SD ± 11.26) vs. -5.9 units (SD ± 10.48) (MD -0.66 units)	High

					Study withdrawal rates did not differ significantly but topiramate had more paresthesia and dysgeusia and placebo had more insomnia. Attrition: 28% (10/36) vs. 19% (7/37)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BES=Binge Eating Scale; BMI=body mass index; CBT=cognitive-behavioral therapy; d=day; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Individual cognitive-behavioral therapy with fluoxetine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Devlin et al. (2005)	Design: RCT Setting: NR Country: United States Funding: Industry and government	Randomized N=116 Placebo 5 mo (N=31) Fluoxetine 60 mg 5 mo (N=32) Individual CBT + Placebo 5 mo (N=25) Individual CBT + Fluoxetine 60 mg 5 mo (N=28) (All received Group Behavioral Weight Control Treatment) Fluoxetine 60 mg +/- Individual CBT 5 mo (pooled) (N=60) Individual CBT + Placebo/Fluoxetine 60	Inclusion: 18-70 years of age; BMI ≥ 27 kg/m^2; maximum weight of 159 kg; BED for at least 6 months; overweight or obese Exclusion: Substance-related disorders within the past yr; acutely suicidal; current psychotic disorder, bipolar disorder, major depressive disorder with melancholic features, or BN; history of AN or psychotic disorder; concurrent eating or weight control treatment; currently taking antidepressants, mood stabilizers, or appetite suppressants; MAOIs within the prior 2 weeks; previously had an adverse reaction to fluoxetine	Overweight or Obesity: 116 (100%) BED, Duration >= 6 mo: 116 (100%) Binge Eating, Duration: 27.3 yr (SD ± 14.7) BMI >= 27 kg/m²: 116 (100%) BMI: 40.9 kg/m² (SD ± 6.9) - 40.3 kg/m² (SD ± 7.1) vs. 40.1 kg/m² (SD ± 6.6) vs. 41.1 kg/m² (SD ± 7.6) vs. 42.1 kg/m² (SD ± 6.9) Weight <= 159 kg: 116 (100%) Age 18 yr-70 yr: 116 (100%) Age: 43 yr (SD ± 12) - 44.1 yr (SD ± 10.2) vs. 45.9 yr (SD ± 13.6) vs.	No significant difference was noted in weight change: - Baseline: 113.5 kg (SD ± 22.2) vs. 113.8 kg (SD ± 22.8) vs. 116.5 kg (SD ± 22.2) vs. 116.9 kg (SD ± 20.8) - Baseline – 5 mo Change: -2.4 kg (SD ± 5.9) vs. -1.9 kg (SD ± 6.9) vs. -1.9 kg (SD ± 7.1) vs. -4.1 kg (SD ± 6.9) CBT +/- Fluoxetine had more binge-eating abstinence than no CBT at 5 mo: 33 (62%) vs. 21 (33%) (p<0.001) BDI - Baseline: 15.6 units (SD ± 9.3) vs. 14.5 units (SD ± 7.2) vs. 13.9 units (SD ± 10.6) vs. 16.9 units (SD ± 9.1) BDI, Change - Baseline – 5 mo: -5 units (SD ± 7.5) vs. -7 units	High

		mg 5 mo (pooled) (N=53) Placebo/Fluoxetine 60 mg 5 mo (pooled) (N=63)		43.4 yr (SD ± 11.8) vs. 39.4 yr (SD ± 12.1) Gender - Female: 90 (78%) - Male: 26 (22%) Race - Caucasian: 89 (77%) - Black or African American: 14 (12%) - Multiracial or Other: 1 (1%) Ethnicity, Hispanic/Latino: 12 (10%)	(SD ± 8.5) vs. -5.5 units (SD ± 8.5) vs. -10.6 units (SD ± 9.3) Attrition: 48% (15/31) vs. 31% (10/32) vs. 40% (10/25) vs. 25% (7/28)	

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; MAOI=monoamine oxidase inhibitor; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Cognitive-behavioral therapy with fluoxetine or with fluvoxamine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Ricca et al. (2001)	Design: RCT Setting: Outpatient: the Outpatient Clinic for Eating Disorders of the Units of Psychiatry and Endocrinology of the University of Florence Country: Italy Funding: Reimbursed by government	Randomized N=108 CBT24 wk (N=20) CBT+ Fluoxetine 20-60 mg 24 wk (N=22) CBT+ Fluvoxamine 100-300 mg 24 wk (N=23) Fluoxetine 20-60 mg 24 wk (N=21)	Inclusion: BED; 18-45 years of age Exclusion: Diabetes mellitus; thyroid disorders; any other disease interfering with eating behavior; pregnancy; lactation; heart disease	BED: 108 (100%) Binge Eating, Duration: 5.6 yr (SD ± 5) - 6.4 yr (SD ± 6) vs. 4.9 yr (SD ± 5.1) vs. 4.8 yr (SD ± 4.4) vs. 5.1 yr (SD ± 4.7) vs. 5.3 yr (SD ± 4.8) BMI: 32.3 kg/m² (SD ± 5.8) - 32 kg/m² (SD ± 6) vs. 31.7 kg/m² (SD ± 5.6) vs. 32.5 kg/m² (SD ± 6.1) vs. 32.1 kg/m² (SD ± 3.8) vs. 32.7 kg/m² (SD ± 4.1) Mental Disorder, Other and not BED: 15 (13.89%)	BMI scores were significantly reduced at 24 wk in CBT groups (-2.2 kg/m² vs. -3.8 kg/m² vs. NR vs. -0.7 kg/m² vs. NR). Improvements persisted at 1-yr follow-up but with some weight regain with fluoxetine alone (-1.6 kg/m² vs. -3.3 kg/m² vs. NR vs. 0.5 kg/m² vs. NR). BDI - Baseline: 22 units vs. 16.5 units vs. 22 units vs. 20 units vs. 21 units - Baseline – 24 wk: -8 units (SD ± 9.62, N=17) vs. -6 units (SD ± 12.96, N=16)	High

		Fluvoxamine 100-300 mg 24 wk (N=22) Follow-up: Baseline – 76 wk		Age 18 yr-45 yr: 108 (100%) Age: 25.9 yr (SD ± 6.8) - 26.3 yr (SD ± 6.7) vs. 25.2 yr (SD ± 6.3) vs. 25.1 yr (SD ± 6.9) vs. 25.1 yr (SD ± 6.1) vs. 26.1 yr (SD ± 5.9) Gender - Female: 13 (65%) vs. 13 (59.09%) vs. 13 (56.52%) vs. 12 (57.14%) vs. 13 (59.09%) - Male: 7 (35%) vs. 9 (40.91%) vs. 10 (43.48%) vs. 9 (42.86%) vs. 9 (40.91%) Race: NR	vs. NR (N=18) vs. -5 units (SD ± 10.21, N=16) vs. NR (N=16) - Baseline – 76 wk: -8 units (SD ± 9.93, N=17) vs. -6 units (SD ± 12.96, N=16) vs. NR (N=18) vs. -4 units (SD ± 10.39, N=16) vs. NR (N=16) Adverse Events - Baseline – 24 wk: 0 (0%) vs. 6 (27.2%) vs. 6 (26.09%) vs. 7 (33.33%) vs. 7 (31.82%) Treatment Discontinuation, Adverse Events - Baseline – 24 wk: 0 (0%) vs. 3 (13.64%) vs. 3 (13.04%) vs. 2 (9.52%) vs. 4 (18.18%) Attrition: 15% (3/20) vs. 27% (6/22) vs. 22% (5/23) vs. 24% (5/21) vs. 27% (6/22)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Cognitive-behavioral therapy with zonisamide

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Ricca et al. (2009)	Design: Non-RCT Setting: Outpatient: Outpatient Clinic for Eating Disorders of the University of Florence Country: Italy	Randomized N=52 CBT + Placebo 24 wk (N=24) CBT+ Zonisamide 0-150 mg 29 wk (N=28)	Inclusion: 18-60 years of age; BED or subthreshold BED Exclusion: Any organic disease interfering with eating behavior; illiteracy and intellectual disability; lifetime history of psychotic disorder, bipolar disorder, or substance abuse disorders; history of seizures;	BED or BED, Subclinical: 52 (100%) BED, Subclinical: 14 (58.33%) vs. 16 (57.14%) Binge Eating: 10 (41.67%) vs. 12 (42.86%)	No statistically significant differences were noted though some numerical differences favoring zonisamide were described at the end of treatment and at 1-yr follow-up. BMI	----------

	Funding: NR	(All received CBT) Binge Eating N=10 vs. 12 BED, Subclinical N=14 vs. 16 Follow-up: Baseline – 18 mo	contraindication to treatment with zonisamide; pregnancy; lactation	Binge Eating, Duration: 7.67 yr (SD ± 3.07) vs. 6.06 yr (SD ± 3.96) Age 18 yr-60 yr: 52 (100%) Age: 34.8 yr (SD ± 11.09) vs. 36.07 yr (SD ± 11.56) Gender - Female: 20 (83.33%) vs. 23 (82.14%) - Male: 4 (16.67%) vs. 5 (17.86%) Race: NR	- Baseline: 39.22 kg/m² (SD ± 7.84) vs. 38.43 kg/m² (SD ± 5.7) - 6 mo: 38.41 kg/m² (SD ± 7.67) vs. 36.77 kg/m² (SD ± 5.84) - 18 mo: 38.99 kg/m² (SD ± 7.02) vs. 36.49 kg/m² (SD ± 5.96) Binge Eating - Baseline:10.4/mo (SD ± 7.35) vs. 9.2/mo (SD ± 6.88) - 6 mo: 4.9/mo (SD ± 5.88) vs. 3.5/mo (SD ± 4.23) - 18 mo: 5.1/mo (SD ± 5.88) vs. 3.6/mo (SD ± 4.23) Study Withdrawal - Baseline – 18 mo - Adverse Events: NR vs. 6 (21.43%) - Lack of Efficacy: NR vs. 1 (3.57%) Attrition: 33% (8/24) vs. 50% (14/28)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Interpersonal Psychotherapy

Compared to Group Cognitive-Behavioral Therapy, Behavioral Weight Loss, and Wait-List Control

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Tasca et al. (2006, 2012)	Design: RCT Setting: NR	Randomized N=135	Inclusion: BED; a minimum of 2 days of binge eating/wk for at least the previous 6 months	BED: 135 (100%)	Binge-eating abstinence at 16 wk was 62.2% CBT, 59.5% IPT, and 9.1% WLC. Abstinence	High

	Country: Canada Funding: Non-profit	Group CBT 16 wk (N=47) Group Psychodynamic IPT 16 wk (N=48) WLC 16 wk (N=40) Follow-up: Baseline – 16 mo	Exclusion: Current problems with substance use; bipolar disorder; psychotic disorder; current suicidality; current other medical or psychological treatment for BED; history of an eating disorder other than BED; current purging behavior; age less than 18 years	Binge Eating >= 2 d/wk, In the Previous >= 6 mo: 135 (100%) BED, Duration: 19.62 yr (SD ± 9.19) BMI: 41.11 kg/m² (SD ± 9.95) Age: 42.75 yr (SD ± 10.76) Gender - Female: 123 (91.11%) - Male: 12 (8.89%) Race, Caucasian: 132 (97.7%)	rates at 68 wk were 67.7% CBT vs. 56.8% IPT. Both treatments were noted to reduce interpersonal problem subscale ratings including cold/distant subscale ratings. BMI - Baseline: 42.59 kg/m² (SD ± 12.95, N=37) vs.40.03 kg/m² (SD ± 9.69, N=37) vs. 42.58 kg/m² (SD ± 9.57, N=33) BMI, Change - Baseline – 68 wk: -1.57 kg/m² (SD ± 9.9, N=37) vs. -2.36 kg/m² (SD ± 7.25, N=37) vs. NR (N=33) Attrition: 21% (10/47) vs. 23% (11/48) vs. 18% (7/40)	
Wilfley et al. (1993)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=56 Group CBT 16 wk (N=18) Group IPT 16 wk (N=18) WLC 16 wk (N=20)	Inclusion: Nonpurging BN; female; 18- 65 years of age; average of two or more binge episodes per wk for the past 6 months Exclusion: Age below 18 years or above 65; current self-induced vomiting, laxative use, or purging behaviors; past history of self-induced vomiting, laxative use, or purging behaviors; current use of antidepressant medication; current use of appetite suppressants; concurrent treatment for weight loss; concurrent unipolar disorder, bipolar affective disorder, or psychosis; concurrent drug abuse; concurrent alcoholism	BN, Non-Purging Type: 56 (100%) Binge Eating >= 2/wk, In the Previous 6 mo; 56 (100%) Binge Eating, Duration: 23.7 yr (SD ± 13.4) Binge Eating: 4.2 d/wk (SD ± 1.5) vs. 4.7 d/wk (SD ± 1.8) vs. 4.4 d/wk (SD ± 1.8) Weight: 87.3 kg (SD ± 14.2) Age 18 yr-65 yr: 56 (100%) Age: 44.3 yr (SD ± 8.3)	Abstinence from binge eating at 16 wk was 28% with group CBT vs. 44% with group IPT vs. 0% with WLC. IPT had a greater binge-eating percent change but it was not statistically significant: -48% with CBT vs. -71% with IPT vs. -10% with WLC. Weight, Change - Baseline – 68 wk: 0 kg vs. -3 kg vs. NR Adherence was greater in the IPT group (88% vs. 72% with CBT).	High

				Gender, Female: 56 (100%) Race - Caucasian: 48 (86%) - Black or African American: 3 (5%) - Indian: 1 (2%) - Pacific Islander: 1 (2%) Ethnicity, Hispanic/Latino: 3 (5%)	Study withdrawal rates were low in all groups (11% CBT, 0% IPT, 5% WLC). Attrition: 33% (6/18) vs. 11% (2/18) vs. NR	
Wilfley et al. (2002); Hilbert et al. (2012)	Design: RCT; Follow-up Setting: Multi-center Country: United States Funding: Government	Randomized N=162 Group CBT NR (N=81) Group IPT NR (N=81) Follow-up: Baseline – 4 yr Follow-up (N=90) - 45 vs. 45	Inclusion: Overweight; BED; 18-65 years of age; BMI 27-48 kg/m^2; average of >=2 days of binge eating/wk for at least 6 months' duration; marked distress regarding binge eating; at least 3 of 5 behavioral features associated with BED Exclusion: Taking weight-affecting medications; taking psychotropic medications; psychiatric conditions warranting immediate treatment; psychotic symptoms; substance dependence; suicidality	BED: 162 (100%) Binge Eating >= 2 d/wk, Duration 6 mo: 162 (100%) Overweight: 162 (100%) BMI 27 kg/m²-48 kg/m²: 162 (100%) Binge Eating: 17.3 d/mo (SD ± 6.9) vs. 16.3 d/mo (SD ± 7.2) Age 18 yr-65 yr: 162 (100%) Age: 45.6 yr (SD ± 9.6) vs. 44.9 yr (SD ± 9.6) Gender - Female: 67 (82.7%) vs. 67 (82.7%) - Male: 14 (17.3%) vs. 14 (17.3%) Race - Caucasian: 76 (93.9%) vs. 74 (91.4%)	Binge-eating recovery rates were equivalent for CBT and IPT at posttreatment (79% vs 73%) and at 1-yr follow-up (59% vs 62%). Persistent recovery was present at 4 yr in 27.3% of the CBT group and 22.2% of the IPT group. Binge days per mo showed similar reductions: 17.3 baseline -> 1.7 at 12 mo with CBT; 16.3-> 1.2 with IPT. Disease Response, Remission - Posttreatment: 73 (94%, N=78) vs. 72 (90%, N=80) - 12 mo: 56 (84%, N=67) vs. 63 (89%, N=71) - 4 yr: 18 (72%, N=25) vs. 26 (83.9%, N=31) BMI – Baseline: 37.4 kg/m² (SD ± 5.3) vs. 37.4 kg/m² (SD ± 5.1) BMI, Change - Baseline – 12 mo: -0.2 kg/m² (SD ± 4.03,	High

				- Black or African American: 3 (3.7%) vs. 3 (3.7%) - Native American/Alaska Native: 1 (1.2%) vs. 0 (0%) Ethnicity, Hispanic/Latino: 1 (1.2%) vs. 4 (4.9%)	N=67) vs. -1.1 kg/m² (SD ± 4.08, N=71) Attrition: 11% (9/81) vs. 9% (7/81)	
Wilson et al. (2010)	Design: RCT Setting: Outpatient: University clinics Country: United States Funding: Government	Randomized N=205 CBT-GSH 6 mo (N= 66) BWL Treatment 6 mo (N= 64) IPT 6 mo (N= 75) Treatment Setting, Rutgers University subgroup (N= 31 vs. 32 vs. 37) Treatment Setting, Washington University in St. Louis subgroup (N= 35 vs. 32 vs. 38) Follow-up: Baseline – 30 mo	Inclusion: Aged 18 years and older; BMI 27-45 kg/m^2; overweight or obese; BED Exclusion: Current psychosis; bipolar disorder; current suicidal state; alcohol or drug dependence within the past 6 months; current participation in a weight-control program; taking medication that would affect weight	BED: 205 (100%) Overweight or Obesity: 205 (100%) BMI 27 kg/m²-45 kg/m²: 205 (100%) Age >= 18 yr: 205 (100%) Age: 50.3 yr (SD ± 13.6) vs. 46.2 yr (SD ± 10.9) vs. 48.7 yr (SD ± 11.2) Gender - Female: 54 (82%) vs. 57 (89%) vs. 64 (85%) - Male: 12 (18%) vs. 7 (11%) vs. 11 (15%) Race - Caucasian: 54 (82%) vs. 56 (88%) vs. 58 (77%) - Black or African American: 7 (11%) vs. 7 (11%) vs. 13 (17%) - Native American/Alaska Native: 0 (0%) vs. 0 (0%) vs. 1 (1%) Ethnicity, Hispanic/Latino: 5 (8%) vs. 1 (2%) vs. 3 (4%)	BMI- Baseline: 36.2 kg/m² (SD ± 4.3) vs. 36.8 kg/m² (SD ± 5.5) vs. 36.3 kg/m² (SD ± 5.1) BMI - 6 mo: 36.1 kg/m² (SD ± 4.4) vs. 35.4 kg/m² (SD ± 5.7) vs. 35.9 kg/m² (SD ± 5.3) - CBT-GSH vs. BWL Treatment: SMD 0.741 - IPT vs. BWL Treatment: SMD 0.48 - CBT-GSH vs. IPT 6 mo: SMD 0.15 BMI - 30 mo: 35.7 kg/m² (SD ± 5) vs. 36.3 kg/m² (SD ± 6.2) vs. 36.1 kg/m² (SD ± 5.5) - BWL Treatment vs. CBT-GSH: SMD 0.52 - BWL Treatment vs. IPT: SMD 0.29 - IPT vs. CBT-GSH: SMD 0.2 Weight – Baseline->6 mo: 100.3->100 kg vs. 103.5->99.8 kg vs. 100.4->99.1 kg Weight, Decrease >= 5 % - Baseline – 6 mo: 10 (15%) vs. 26 (41%) vs. 11 (15%) - BWL Treatment vs. CBT-GSH: OR 3.9 - BWL Treatment vs. IPT: OR 3.9	High

					Disease Response, Remission - 30 mo: 41 (62.1%) vs. 28 (43.9%) vs. 51 (67.9%) - CBT-GSH vs. BWL Treatment: OR 2.3 - IPT vs. BWL Treatment: OR 2.6 - IPT vs. CBT-GSH: OR 1.2 Attrition: 30% (20/66) vs. 28% (18/64) vs. 7% (5/75)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; BWL=behavioral weight loss; CBT=cognitive-behavioral therapy; CBT-GSH=cognitive-behavioral therapy guided self-help; d=day; IPT=interpersonal psychotherapy; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; SMD=standardized mean difference; wk=week; WLC=wait-list control; yr=year

Antidepressants

Fluoxetine

Compared to placebo and individual cognitive-behavioral therapy

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Devlin et al. (2005)	Design: RCT Setting: NR Country: United States Funding: Industry and government	Randomized N=116 Placebo 5 mo (N=31) Fluoxetine 60 mg 5 mo (N=32) Individual CBT + Placebo 5 mo (N=25) Individual CBT + Fluoxetine 60 mg 5 mo (N=28)	Inclusion: 18-70 years of age; BMI ≥ 27 kg/m^2; maximum weight of 159 kg; BED for at least 6 months; overweight or obese Exclusion: Substance-related disorders within the past yr; acutely suicidal; current psychotic disorder, bipolar disorder, major depressive disorder with melancholic features, or BN; history of AN or psychotic disorder; concurrent eating or weight control treatment; currently taking antidepressants, mood stabilizers, or appetite suppressants; MAOIs within the	Overweight or Obesity: 116 (100%) BED, Duration >= 6 mo: 116 (100%) Binge Eating, Duration: 27.3 yr (SD ± 14.7) BMI >= 27 kg/m²: 116 (100%) BMI: 40.9 kg/m² (SD ± 6.9) - 40.3 kg/m² (SD ± 7.1) vs. 40.1 kg/m² (SD ± 6.6) vs. 41.1 kg/m² (SD ± 7.6) vs. 42.1 kg/m² (SD ± 6.9)	No significant difference was noted in weight change: - Baseline: 113.5 kg (SD ± 22.2) vs. 113.8 kg (SD ± 22.8) vs. 116.5 kg (SD ± 22.2) vs. 116.9 kg (SD ± 20.8) - Baseline – 5 mo Change: -2.4 kg (SD ± 5.9) vs. -1.9 kg (SD ± 6.9) vs. -1.9 kg (SD ± 7.1) vs. -4.1 kg (SD ± 6.9) CBT +/- Fluoxetine had more binge-eating abstinence than no CBT at 5 mo: 33 (62%) vs. 21 (33%) (p<0.001)	High

		(All received Group Behavioral Weight Control Treatment) Fluoxetine 60 mg +/- Individual CBT 5 mo (pooled) (N=60) Individual CBT + Placebo/Fluoxetine 60 mg 5 mo (pooled) (N=53) Placebo/Fluoxetine 60 mg 5 mo (pooled) (N=63)	prior 2 weeks; previously had an adverse reaction to fluoxetine	Weight <= 159 kg: 116 (100%) Age 18 yr-70 yr: 116 (100%) Age: 43 yr (SD ± 12) - 44.1 yr (SD ± 10.2) vs. 45.9 yr (SD ± 13.6) vs. 43.4 yr (SD ± 11.8) vs. 39.4 yr (SD ± 12.1) Gender - Female: 90 (78%) - Male: 26 (22%) Race - Caucasian: 89 (77%) - Black or African American: 14 (12%) - Multiracial or Other: 1 (1%) Ethnicity, Hispanic/Latino: 12 (10%)	BDI - Baseline: 15.6 units (SD ± 9.3) vs. 14.5 units (SD ± 7.2) vs. 13.9 units (SD ± 10.6) vs. 16.9 units (SD ± 9.1) BDI, Change - Baseline – 5 mo: -5 units (SD ± 7.5) vs. -7 units (SD ± 8.5) vs. -5.5 units (SD ± 8.5) vs. -10.6 units (SD ± 9.3) Attrition: 48% (15/31) vs. 31% (10/32) vs. 40% (10/25) vs. 25% (7/28)	
Grilo et al. (2005a. 2012b)	Design: RCT; Follow-up Setting: NR Country: United States Funding: Government; product donation by industry	Randomized N=108 Current Analysis (N=81) CBT16 wk (N=28) CBT+ Fluoxetine 60 mg 16 wk (N=26) Fluoxetine 60 mg 16 wk (N=27) Placebo 16 wk (N=27) Follow-up: Baseline – 16 mo	Inclusion: 18- 60 years of age; BED; between 100% and 200% of ideal weight for height Exclusion: Concurrent treatment for eating or weight; concurrent treatment for psychiatric problems; medical conditions that influence weight or eating; diabetes; thyroid problems; hypoglycemia; severe psychiatric conditions requiring different treatments; psychosis or bipolar disorder requiring treatment; lactation; pregnancy; purging behaviors	BED: 108 (100%) BMI: 36.3 kg/m² (SD ± 7.9) - 35 kg/m² (SD ± 6.2) vs. 35.7 kg/m² (SD ± 8.3) vs. 38.9 kg/m² (SD ± 9.5) vs. 35.7 kg/m² (SD ± 7.2) Age 18 yr-60 yr: 108 (100%) Age: 44 yr (SD ± 8.6) - 43.6 yr (SD ± 8.5) vs. 44.7 yr (SD ± 8.1) vs. 44.3 yr (SD ± 9.5) vs. 43.6 yr (SD ± 8.5) Gender - Female: 22 (78.6%) vs. 20 (76.9%) vs. 19 (70.4%) vs. 23 (85.2%)	Remission rates at 16 wk were much higher in both CBT groups: 61% vs. 50% vs. 22% vs. 26%. - CBT vs. Placebo: p=0.008 - CBT+ Fluoxetine vs. Placebo: p=0.05 - CBT+ Fluoxetine vs. Fluoxetine: p=0.03 - CBT+ Fluoxetine vs. CBT: p=0.42 - Fluoxetine vs. Placebo: p=0.83 In the 12-mo follow-up study, these conclusions persisted, and the CBT groups were more likely to achieve remission, though the rates were less at the end of	High

				- Male: 6 (21.4%) vs. 6 (23.1%) vs. 8 (29.6%) vs. 4 (14.8%) Race - Caucasian: 26 (92.9%) vs. 23 (88.5%) vs. 27 (100%) vs. 20 (74.1%) - Black or African American: 2 (7.1%) vs. 2 (7.7%) vs. 0 (0%) vs. 5 (18.5%) Ethnicity, Hispanic/Latino: 0 (0%) vs. 1 (3.8%) vs. 0 (0%) vs. 2 (7.4%)	treatment: 36% vs. 27% vs. 4% vs. NR - CBT vs. Fluoxetine: p=0.005 - CBT+ Fluoxetine vs. Fluoxetine: p=0.024 - CBT vs. CBT+ Fluoxetine: p=0.57 Weight loss was modest in all treatment groups. - Change - Baseline – 16 mo: -9.84 lbs vs. -4.13 lbs vs. -1.48 lbs vs. NR BDI Baseline: 16.5 units (SD ± 8.4) vs. 20.2 units (SD ± 12.1) vs. 16.9 units (SD ± 8.4) vs. 18.7 units (SD ± 9.7) 16 wk: 6.5 units (SD ± 6.8) vs. 9.2 units (SD ± 7.3) vs. 11.8 units (SD ± 9.8) vs. 11.7 units (SD ± 10.3) - CBT vs. Placebo: MD -5.2 units (p=0.04) - CBT vs. Fluoxetine: MD -5.3 units (p=0.01) - CBT+ Fluoxetine vs. Fluoxetine: MD -2.6 units (p=0.04) Attrition: 21% (6/28) vs. 23% (6/26) vs. 22% (6/27) vs. 15% (4/27)	

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; MAOI=monoamine oxidase inhibitor; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to sertraline

Author (year)	Study characteristics, including design,	Interventions, including study arm, co-intervention, sample	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and	Outcome measures, main results, and overall percent attrition	Risk of bias

(trial name)	setting, country, and funding	size (N), dose, duration, and follow-up		baseline clinical features (e.g., BMI)		
Leombruni et al. (2008)	Design: RCT Setting: Outpatient: Eating Disorders Pilot Centre of the Psychiatric Clinic of the University of Turin Country: Italy Funding: NR	Randomized N=42 Fluoxetine 40-80 mg 6 mo (10 mg induction) (N=20) Sertraline 100-200 mg 6 mo (25 mg induction) (N=22) Current Analysis (N=37) - 17 vs. 20	Inclusion: Obese; female; BED; BMI => 30 kg/m^2; 18- 65 years of age Exclusion: BMI < 30 kg/m^2; overweight caused by pharmacologic treatments; overweight condition secondary to metabolic or endocrine disorders; comorbidity of an acute full-syndrome Axis I disorder; comorbid mood disorder or anxiety disorder; comorbidity with BN; male	BED: 42 (100%) Obesity: 42 (100%) BED, Duration: 144 mo (SD ± 46.5) Binge Eating: 4.6/wk (SD ± 3.2) vs. 6.2/wk (SD ± 7.3) Weight: 101.9 kg (SD ± 12.5) vs. 99.6 kg (SD ± 14.5) BMI >= 30 kg/m²: 42 (100%) BMI: 39.3 kg/m² (SD ± 3.5) - 40.2 kg/m² (SD ± 3.9) vs. 38.6 kg/m² (SD ± 3.8) Age 18 yr-65 yr: 42 (100%) Age: 39.6 yr (SD ± 8.5) Gender, Female: 42 (100%) Race: NR	9/17 (52.9%) with fluoxetine vs. 12/20 (60%) with sertraline achieved abstinence from binge eating at 6 mo (p=0.664). Binge Eating, Change - Baseline – 6 mo: -3.7/wk (SD ± 2.55, N=15) vs. -5.1/wk (SD ± 5.52, N=16) About half (47% vs. 45%) lost more than 5% of baseline weight at 3 mo (p=0.768). Both groups showed significant improvements in BES scores. - Baseline: 32.1 units (SD ± 3.5, N=17) vs. 26.1 units (SD ± 8.5, N=20) - Baseline – 6 mo: -12.9 units (SD ± 9.39, N=17) vs. -10.2 units (SD ± 6.47, N=20) BDI - Baseline: 11.1 units (SD ± 4.5) vs. 13.3 units (SD ± 7) - 6 mo: 8.4 units (SD ± 6.2, N=15) vs. 9.9 units (SD ± 5.9, N=16) Adverse Events, Treatment-Related - 2 mo – 6 mo: 2 (11.8%, N=17) vs. 3 (15%, N=20) (p=0.665) Attrition: 25% (5/20) vs. 27% (6/22)	High

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BES=Binge Eating Scale; BMI=body mass index; BN=bulimia nervosa; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to cognitive-behavioral therapy and fluvoxamine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Ricca et al. (2001)	Design: RCT Setting: Outpatient: the Outpatient Clinic for Eating Disorders of the Units of Psychiatry and Endocrinology of the University of Florence Country: Italy Funding: Reimbursed by government	Randomized N=108 CBT24 wk (N=20) CBT+ Fluoxetine 20-60 mg 24 wk (N=22) CBT+ Fluvoxamine 100-300 mg 24 wk (N=23) Fluoxetine 20-60 mg 24 wk (N=21) Fluvoxamine 100-300 mg 24 wk (N=22) Follow-up: Baseline – 76 wk	Inclusion: BED; 18-45 years of age Exclusion: Diabetes mellitus; thyroid disorders; any other disease interfering with eating behavior; pregnancy; lactation; heart disease	BED: 108 (100%) Binge Eating, Duration: 5.6 yr (SD ± 5) - 6.4 yr (SD ± 6) vs. 4.9 yr (SD ± 5.1) vs. 4.8 yr (SD ± 4.4) vs. 5.1 yr (SD ± 4.7) vs. 5.3 yr (SD ± 4.8) BMI: 32.3 kg/m² (SD ± 5.8) - 32 kg/m² (SD ± 6) vs. 31.7 kg/m² (SD ± 5.6) vs. 32.5 kg/m² (SD ± 6.1) vs. 32.1 kg/m² (SD ± 3.8) vs. 32.7 kg/m² (SD ± 4.1) Mental Disorder, Other and not BED: 15 (13.89%) Age 18 yr-45 yr: 108 (100%) Age: 25.9 yr (SD ± 6.8) - 26.3 yr (SD ± 6.7) vs. 25.2 yr (SD ± 6.3) vs. 25.1 yr (SD ± 6.9) vs. 25.1 yr (SD ± 6.1) vs. 26.1 yr (SD ± 5.9) Gender - Female: 13 (65%) vs. 13 (59.09%) vs. 13 (56.52%) vs. 12 (57.14%) vs. 13 (59.09%) - Male: 7 (35%) vs. 9 (40.91%) vs. 10 (43.48%) vs. 9 (42.86%) vs. 9 (40.91%)	BMI scores significantly reduced at 24 wk in CBT groups (-2.2 kg/m² vs. -3.8 kg/m² vs. NR vs. -0.7 kg/m² vs. NR). Improvements persisted at 1-yr follow-up but with some weight regain with fluoxetine alone (-1.6 kg/m² vs. -3.3 kg/m² vs. NR vs. 0.5 kg/m² vs. NR). BDI - Baseline: 22 units vs. 16.5 units vs. 22 units vs. 20 units vs. 21 units - Baseline – 24 wk: -8 units (SD ± 9.62, N=17) vs. -6 units (SD ± 12.96, N=16) vs. NR (N=18) vs. -5 units (SD ± 10.21, N=16) vs. NR (N=16) - Baseline – 76 wk: -8 units (SD ± 9.93, N=17) vs. -6 units (SD ± 12.96, N=16) vs. NR (N=18) vs. -4 units (SD ± 10.39, N=16) vs. NR (N=16) Adverse Events - Baseline – 24 wk: 0 (0%) vs. 6 (27.2%) vs. 6 (26.09%) vs. 7 (33.33%) vs. 7 (31.82%) Treatment Discontinuation, Adverse Events - Baseline – 24 wk: 0 (0%) vs. 3 (13.64%) vs. 3	High

				Race: NR	(13.04%) vs. 2 (9.52%) vs. 4 (18.18%) Attrition: 15% (3/20) vs. 27% (6/22) vs. 22% (5/23) vs. 24% (5/21) vs. 27% (6/22)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Fluvoxamine

Compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Hudson et al. (1998)	Design: RCT Setting: Outpatient: Multi-center Country: United States Funding: Industry	Randomized N=85 Fluvoxamine 50-300 mg 9 wk (N=42) Placebo 9 wk (N=43) ITT (N=83) - 40 vs. 43	Inclusion: BED; at least 3 binge-eating episodes/wk for at least 6 months; 18-60 years of age; weigh over 85% of the midpoint of the IBW for their height Exclusion: Concurrent AN; major depression within 1 year of study entry; OCD within 1 yr of study entry; lifetime substance dependence; psychosis; mania; organic dementia; significant suicide risk; received psychotherapy or behavioral therapy within 3 months of study entry; history of psychosurgery; history of seizures; clinically unstable medical illness; received MAOIs, tricyclics, neuroleptics, lithium, or fluoxetine within 4 weeks of randomization; received investigational medications or depot neuroleptics within 3 months of randomization; history of	BED: 85 (100%) Binge Eating >= 3 episodes/wk, In the Previous >= 6 mo: 85 (100%) BMI: 34.2 kg/m² (SD ± 6) vs. 36.8 kg/m² (SD ± 8.2) Age 18 yr-60 yr: 85 (100%) Age: 41.2 yr (SD ± 9.9) vs. 43 yr (SD ± 9.5) Gender - Female: 39 (93%) vs. 38 (88%) - Male: 3 (7%) vs. 5 (12%) Race, Caucasian: 41 (98%) vs. 41 (95%)	Fluvoxamine was associated with reduction in binge frequency, CGI severity, and BMI and greater response. Binge Eating: 5.4->1.12/wk vs.5.3->2.16/wk BMI, Change - Baseline – 9 wk: MD -0.167 kg/m² (p=0.04) Remission rates were 38% with fluvoxamine vs. 26% with placebo. Weight changes were minimal: 2.7 lbs vs. 0.3 lbs. 5 (12%) of fluvoxamine withdrew due to adverse events vs. 0 (0%) for placebo (p=0.03). Insomnia, nausea, and abnormal dreams were reported	High

			fluvoxamine; fewer than 3 binges in the wk before randomization		significantly more with fluvoxamine vs. placebo: - Insomnia: 18 (44%, N=40) vs. 6 (14%) (p<0.05) - Nausea: 14 (34%, N=40) vs. 5 (12%) (p<0.01) - Dreams, Abnormal: 8 (20%, N=40) vs. 2 (5%) (p<0.01) Fluvoxamine was significantly associated with more study withdrawal (p=0.04). Attrition: 31% (13/42) vs. 12% (5/43)	
Pearlstein et al. (2003)	Design: RCT Setting: NR Country: United States Funding: Industry	Randomized N=20 Fluvoxamine 150 mg 12 wk (N=9) Placebo 12 wk (N=11)	Inclusion: BED	BED: 20 (100%) BDI, Item Average: 0.44 units (SD ± 0.22, N=7) vs. 0.68 units (SD ± 0.57, N=9) BMI: 41.16 kg/m² Weight: 243 lbs (SD ± 85) vs. 258 lbs (SD ± 96) Age: 41 yr Gender - Female: 17 (85%) - Male: 3 (15%) Race, Caucasian: 18 (90%)	Abstinence rates were comparable (50% of each group). Reductions were noted in binge frequency and EDE subscale score but none were significant. Weight, Change - Baseline – 12 wk: -1 lbs (SD ± 64.74) vs. 4 lbs (SD ± 75.57) Sedation - Baseline – 12 wk: 8 (88.89%) vs. 3 (27.27%) Libido, Decrease - Baseline – 12 wk: 3 (33.33%) vs. 0 (0%) Nausea - Baseline – 12 wk: 4 (44.44%) vs. 1 (9.09%) Overall Attrition: 25% (5/20)	Moderate

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CGI=Clinical Global Impression; EDE=Eating Disorder Examination; ITT=intention-to-treat; MAOI=monoamine oxidase inhibitor; MD=mean difference; mo=month; NR=not reported; OCD=obsessive-compulsive disorder; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Compared to cognitive-behavioral therapy and fluvoxamine

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Ricca et al. (2001)	Design: RCT Setting: Outpatient: the Outpatient Clinic for Eating Disorders of the Units of Psychiatry and Endocrinology of the University of Florence Country: Italy Funding: Reimbursed by government	Randomized N=108 CBT24 wk (N=20) CBT+ Fluoxetine 20-60 mg 24 wk (N=22) CBT+ Fluvoxamine 100-300 mg 24 wk (N=23) Fluoxetine 20-60 mg 24 wk (N=21) Fluvoxamine 100-300 mg 24 wk (N=22) Follow-up: Baseline – 76 wk	Inclusion: BED; 18-45 years of age Exclusion: Diabetes mellitus; thyroid disorders; any other disease interfering with eating behavior; pregnancy; lactation; heart disease	BED: 108 (100%) Binge Eating, Duration: 5.6 yr (SD ± 5) - 6.4 yr (SD ± 6) vs. 4.9 yr (SD ± 5.1) vs. 4.8 yr (SD ± 4.4) vs. 5.1 yr (SD ± 4.7) vs. 5.3 yr (SD ± 4.8) BMI: 32.3 kg/m² (SD ± 5.8) - 32 kg/m² (SD ± 6) vs. 31.7 kg/m² (SD ± 5.6) vs. 32.5 kg/m² (SD ± 6.1) vs. 32.1 kg/m² (SD ± 3.8) vs. 32.7 kg/m² (SD ± 4.1) Mental Disorder, Other and not BED: 15 (13.89%) Age 18 yr-45 yr: 108 (100%) Age: 25.9 yr (SD ± 6.8) - 26.3 yr (SD ± 6.7) vs. 25.2 yr (SD ± 6.3) vs. 25.1 yr (SD ± 6.9) vs. 25.1 yr (SD ± 6.1) vs. 26.1 yr (SD ± 5.9) Gender - Female: 13 (65%) vs. 13 (59.09%) vs. 13 (56.52%) vs. 12 (57.14%) vs. 13 (59.09%) - Male: 7 (35%) vs. 9 (40.91%) vs. 10 (43.48%) vs. 9 (42.86%) vs. 9 (40.91%)	BMI scores were significantly reduced at 24 wk in CBT groups (-2.2 kg/m² vs. -3.8 kg/m² vs. NR vs. -0.7 kg/m² vs. NR). Improvements persisted at 1-yr follow-up but with some weight regain with fluoxetine alone (-1.6 kg/m² vs. -3.3 kg/m² vs. NR vs. 0.5 kg/m² vs. NR). BDI - Baseline: 22 units vs. 16.5 units vs. 22 units vs. 20 units vs. 21 units - Baseline – 24 wk: -8 units (SD ± 9.62, N=17) vs. -6 units (SD ± 12.96, N=16) vs. NR (N=18) vs. -5 units (SD ± 10.21, N=16) vs. NR (N=16) - Baseline – 76 wk: -8 units (SD ± 9.93, N=17) vs. -6 units (SD ± 12.96, N=16) vs. NR (N=18) vs. -4 units (SD ± 10.39, N=16) vs. NR (N=16) Adverse Events - Baseline – 24 wk: 0 (0%) vs. 6 (27.2%) vs. 6 (26.09%) vs. 7 (33.33%) vs. 7 (31.82%) Treatment Discontinuation, Adverse Events - Baseline – 24 wk: 0 (0%) vs. 3 (13.64%) vs. 3	High

				Race: NR	(13.04%) vs. 2 (9.52%) vs. 4 (18.18%) Attrition: 15% (3/20) vs. 27% (6/22) vs. 22% (5/23) vs. 24% (5/21) vs. 27% (6/22)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; CBT=cognitive-behavioral therapy; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Escitalopram

Compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Guerdjikova et al. (2008)	Design: RCT Setting: NR Country: NR Funding: Industry	Randomized N=44 Escitalopram 10-30 mg 12 wk (N=21) Placebo 12 wk (N=23) ITT (N=43) - 20 vs. 23	Inclusion: BED; obese; BMI >=30 kg/m^2; 18-60 years of age; at least 2 binge0eating episodes in the wk prior to study entry Exclusion: Concurrent AN or BN; concurrent substance abuse or dependence; substance abuse within 1 yr of study entry; received interpersonal therapy, CBT, or DBT for BED within 3 months; received MAOIs within 4 weeks; previously treated with escitalopram; < 2 binge days in the wk before randomization; lifetime history of psychosis, mania, hypomania, or dementia; significant suicide risk; received psychotropic medication within 2 weeks of randomization; received investigational medications or depot antipsychotics within 3 months	BED: 44 (100%) Obesity: 44 (100%) Binge Eating >= 2 episodes, In the Previous 1 wk: 44 (100%) BMI >= 30 kg/m²: 44 (100%) BMI: 40.1 kg/m² (SD ± 6.8) vs. 40.3 kg/m² (SD ± 4.8) Weight: 113 kg (SD ± 20) vs. 109.2 kg (SD ± 17.2) Age 18 yr-60 yr: 44 (100%) Age: 36.9 yr (SD ± 10) vs. 41 yr (SD ± 10.7) Gender - Female: 21 (100%) vs.22 (95.7%)	Rates of reduction of binge episodes/binge days were comparable: 4.9->0.9 binges/wk vs. 5.1->1.7 binges/wk; 4->0.9 d/wk vs. 4.1->1.6 d/wk. - Binge Episodes, Change - Baseline – 12 wk: MD -0.27/wk (95% CI -0.5 – 0.07) - Binge Days, Change - Baseline – 12 wk: MD -0.28 d/wk (95% CI -0.5 – 0.05) 85% of escitalopram group were “very much improved” vs. 39.1% with placebo (p=0.029). Remission rates did not differ significantly: 50% escitalopram vs. 26% placebo (p=0.088). BMI, Change - Baseline – 12 wk: 0.3 kg/m² (SD ± 5.35, N=20)	High

				- Male: 0 (0%) vs. 1 (4.3%) Race, Caucasian: 16 (76.19%) vs. 17 (73.9%)	vs. 0.2 kg/m² (SD ± 3.8) (MD -0.6 kg/m², 95% CI -1.1 – 0) Weight, Change - Baseline – 12 wk: -1 kg (SD ± 2.6, N=20) vs. 0.6 kg (SD ± 2.4) (MD -1.7 kg, 95% CI -3.2 – 0.1) There were no significant differences between treatment groups in the incidence of adverse events. Attrition: 29% (6/21) vs. 17% (4/23)	

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; CI=confidence interval; d=day; DBT=dialectal behavioral therapy; ITT=intention-to-treat; MAOI=monoamine oxidase inhibitor; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Duloxetine

Compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Guerdjikova et al. (2012)	Design: RCT Setting: Multi-center Country: United States Funding: Industry	Randomized N=40 Duloxetine 30-120 mg 12 wk (N=20) Placebo 12 wk (N=20)	Inclusion: BED; 18-65 years of age (inclusive); a current depressive disorder for a duration of at least 1 mo, including the time preceding and during the screening period; displayed >=2 binge days/wk; had scores of >=25 on the clinician-rated version of the Inventory of Depressive Symptoms Scale at screening and baseline visits Exclusion: A significant risk for suicide; psychotherapy from a mental health professional for	BED: 40 (100%) Binge Eating >= 2 d/wk: 40 (100%) Binge Eating - 4.5/wk (SD ± 2.3) - 4 d/wk (SD ± 1.7) Weight: 114.7 kg (SD ± 23.6) BMI: 40.6 kg/m² (SD ± 7.4)	Duloxetine was superior to placebo on the following: Reduced binge-eating episodes: 4.5->1.1 episodes/wk vs. 4->1.3 episodes/wk (MD -0.62/wk, 95% CI -0.89 – -0.03) Reduced frequency of weekly binge-eating days: 4->1 d/wk vs. 3.5->1.3 d/wk (SD ± 1.2) (MD -0.77 d/wk, 95% CI -1 – -0.17) Reduced weight	High

			treatment of BED or depression, within 3 months before randomization; alcohol or substance abuse within 6 months before randomization; alcohol or substance dependence within 6 months before randomization; BN or AN within 6 months before randomization; lifetime history of a psychotic disorder, bipolar disorder, or dementia; an Axis II disorder which might interfere with study procedures; clinically unstable medical disease; history of seizures; MAOIs, tricyclics, antipsychotics, lithium, or fluoxetine within 4 weeks before randomization	Age 18 yr-65 yr: 40 (100%) Age: 40.1 yr (SD ± 12) - 44.4 yr (SD ± 12.1) vs. 35.7 yr (SD ± 10.4) Gender - Female:16 (80%) vs. 19 (95%) - Male: 4 (20%) vs. 1 (5%) Race, Caucasian: 33 (83%)	- Baseline: 111.1 kg (SD ± 24.1) vs. 118.3 kg (SD ± 23.1) - Baseline – 12 wk: -3.2 kg (SD ± 6.4) vs. -0.3 kg (SD ± 2.2) (MD -2.91 kg, 95% CI -5.74 – -0.09, p=0.04) Binge-eating remission rates were 56% duloxetine vs. 30% placebo (p=0.09). Depression remission rates were 28% duloxetine vs. 20% placebo (p=0.71). Greater side effects were noted with duloxetine than placebo. - Xerostomia: 7 (35%) vs. NR - Hyperhidrosis: 5 (25%) vs. NR - Nausea: 9 (45%) vs. NR - Constipation: 5 (25%) vs. NR Treatment Discontinuation - Baseline – 12 wk - Adverse Events: 3 (15%) vs. 0 (0%) (p=0.19) - Lack of Efficacy: 0 (0%) vs. 1 (5%) Attrition: 35% (7/20) vs. 30% (6/20)	

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; CI=confidence interval; d=day; MAOI=monoamine oxidase inhibitor; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Atomoxetine

Compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and	Outcome measures, main results, and overall percent attrition	Risk of bias

		size (N), dose, duration, and follow-up		baseline clinical features (e.g., BMI)		
McElroy et al. (2007a)	Design: RCT Setting: Outpatient: University of Cincinnati Medical Center Country: United States Funding: Industry	Randomized N=40 Atomoxetine 40-120 mg 10 wk (N=20) Placebo 10 wk (N=20) Follow-up: Baseline – 11 wk ITT (N=39) - 20 vs. 19	Inclusion: 18-65 years of age; BED; weighed >= 85% of the midpoint of IBW for height; had >= 3 binge-eating episodes in the wk before receiving study medication; had >= 2 binge days in the wk before receiving study medication Exclusion: Concurrent AN or BN; substance use disorder within 6 months of study entry; lifetime history of a psychotic disorder, bipolar disorder, dementia, or other cognitive disorder; personality disorder that could interfere with diagnostic assessment, treatment, or compliance; clinically significant suicidality or homicidality; cognitive-behavioral psychotherapy or interpersonal psychotherapy within 3 months of study entry; behavioral weight management for BED within 3 months of study entry; clinically unstable medical illness; history of seizures; required treatment with any drug that might adversely interact with study medication; required treatment with any drug that might obscure the action of the study medication; MAOIs, tricyclics, lithium, antipsychotics, fluoxetine within 4 weeks prior to randomization; other psychoactive medication within 2 weeks of study medication initiation; investigational medications or depot	BED: 40 (100%) Binge Eating >= 3 episodes, In the Previous 1 wk: 40 (100%) Binge Eating >= 2 d, In the Previous 1 wk: 40 (100%) Binge Eating: 3.8 d/wk (SD ± 1.1) vs. 3.9 d/wk (SD ± 1.5) Weight: 106.9 kg (SD ± 20.2) vs. 116.6 kg (SD ± 30.1) BMI: 37.3 kg/m² (SD ± 6.7) vs. 41.4 kg/m² (SD ± 8.5) Age 18 yr-65 yr: 40 (100%) Age: 43.1 yr (SD ± 10.2) vs. 39.2 yr (SD ± 7.7) Gender - Female: 16 (80%) vs. 17 (85%) - Male: 4 (20%) vs. 3 (15%) Race - Caucasian: 17 (85%) vs. 17 (85%) - Black or African American: 5 (12.5%) - Asian: 1 (2.5%)	Atomoxetine was associated with greater reduction in frequency of binge-eating episodes and days (4.2 at baseline->0.23 episodes/wk at 10 wk->0.42 at 11 wk vs. 4.9->1.14->0.98 episodes/wk). Binge Eating, Change - Baseline – 10 wk - Atomoxetine vs. Placebo: MD -0.16/wk (95% CI -0.29 – -0.01) - Atomoxetine vs. Placebo: MD -0.17 d/wk (95% CI -0.3 – -0.03) Atomoxetine was associated with greater reduction in weight and BMI. - Weight, Change - Baseline – 10 wk: -2.7 kg (SD ± 3.7) vs. 0 kg (SD ± 3.2, N=19) (MD -2.69 kg, 95% CI -4.88 – 0.49) - Weight, Change - Baseline – 11 wk: -2.65 kg vs. 0 kg - BMI, Change - Baseline – 10 wk: Atomoxetine vs. Placebo: MD -0.89 kg/m² (95% CI -1.66 – -0.12) Remission rate was greater with atomoxetine (70%) vs. placebo (32%) (p=0.026). Adverse Events - Baseline – 10 wk - Xerostomia: 11 (55%) vs. 4 (20%) (p=0.048)	High

			antipsychotics within 3 months prior to randomization; treated with atomoxetine; pregnant; lactating		Treatment Discontinuation, Adverse Events - Baseline – 10 wk: 3 (15%) vs. 1 (5%) (p=0.6) Study Withdrawal, Lack of Efficacy - Baseline – 10 wk: 0 (0%) vs. 1 (5%) Attrition: 30% (6/20) vs. 45% (9/20)	

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CI=confidence interval; d=day; IBW=ideal body weight; ITT=intention-to-treat; MAOI=monoamine oxidase inhibitor; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Bupropion

Compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
White and Grilo (2013)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=61 Bupropion 300 mg 8 wk (N=31) Placebo 8 wk (N=30)	Inclusion: Overweight or obese; women; BED; BMI 25–50 kg/m^2; 18-65 years of age Exclusion: Diabetes; seizure disorders; uncontrolled hypertension; hypothyroidism; current pregnancy; current breastfeeding; current medications with psychoactive properties; current herbal supplements with psychoactive properties; current treatment for eating or weight; serious psychiatric disorder that warrants a higher level of treatment; bipolar disorder; current substance use disorder; homicidal ideation; suicidal ideation; history of AN or BN	BED: 61 (100%) Overweight or Obesity: 61 (100%) BMI 25 kg/m²-50 kg/m²: 61 (100%) BMI: 35.8 kg/m² (SD ± 6.8) Age 18 yr-65 yr: 61 (100%) Age: 44.1 yr (SD ± 12.5) - 45.2 yr (SD ± 12.1) vs. 43.1 yr (SD ± 13) Gender, Female: 61 (100%) Race - Caucasian: 51 (84%)	Both groups improved, and weight loss was statistically greater with bupropion compared with placebo: 36.2->35.7 kg/m² vs. 35.4->35.2 kg/m² - BMI, % Change - Baseline – 8 wk: Bupropion vs. Placebo: MD 1.2 % (p=0.002) Disease Response, Remission - Baseline – 8 wk: 13 (42%) vs. 8 (27%) (p=0.21) Binge Eating, Objective, Self-Reported - Baseline: 3.3/wk (SD ± 3.3) vs. 3/wk (SD ± 2.6) - 8 wk: 0.8/wk (SD ± 1.2) vs. 1/wk (SD ± 1.5)	High

				- Race, Non-Caucasian: 7 (22.6%) vs. 3 (10%) - Black or African American: 5 (8%) - Other: 2 (3%) Ethnicity, Hispanic/Latino: 3 (5%)	BDI - Baseline: 13.4 units (SD ± 9.8) vs. 10.8 units (SD ± 6.1) BDI, Change - Baseline – 8 wk: -5.4 units (SD ± 12.39) vs. -2.1 units (SD ± 9.11) Adverse Events, Treatment-Related - Baseline – 8 wk: 0 (0%) vs. 0 (0%) Attrition: 13% (4/31) vs. 10% (3/30)	

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Vortioxetine

Compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Grant et al. (2019)	Design: RCT Setting: NR Country: United States Funding: Industry	Randomized N=80 Vortioxetine (10-20 mg 12 wk (N=40) Placebo 12 wk (N=40)	Inclusion: BED; at least 3 binge-eating days/wk for the 2 weeks; 18-65 years of age Exclusion: Unstable medical illness; clinically significant abnormalities on baseline physical examination; an immediate suicide risk; past 12-month psychotic disorder, bipolar disorder, or major depressive disorder; past 6-month alcohol or substance use disorder; illegal substance use; initiation of psychological or weight-loss interventions; use of any other psychotropic medication; current pregnancy or lactation; previous	BED: 80 (100%) BMI: 37.9 kg/m² (SD ± 8.8) - 39.3 kg/m² (SD ± 9.6) vs. 36.5 kg/m² (SD ± 7.9) Weight: 246.61 lb (SD ± 67.89) vs. 228.76 lb (SD ± 54.81) Age 18 yr-65 yr: 80 (100%) Age: 40.0 yr (SD ± 13.1) - 40.3 yr (SD ± 13.2) vs. 39.8 yr (SD ± 13.2) Age at onset: 29.4 (SD ± 14.9) vs. 18.7 (SD ± 9.8)	Both groups had significant reduction in binge-eating frequency but no significant change in weight and BMI. Binge Eating, Days, Baseline: 4.4/wk (SD ± 2.7) vs. 4.3/wk (SD ± 1.6) Binge Eating, Change - Baseline – 12 wk: -2.53/wk (2.88) (p=0.006) vs. -2.17/wk (2.64) (p=0.008) BMI, Change - Baseline – 12 wk: -0.81 (3.56) vs. 0.19 (1.06)	High

			vortioxetine treatment; current OTC weight loss medications; cognitive impairment	Gender - Female: 23 (69.7%) vs. 22 (64.7%) - Male: 10 (30.3%) vs. 12 (35.3%) Race - Black or African American: 36 (53.7%) - Caucasian: 15 (22.4%) - Asian: 2 (3%) - Other: 2 (3%) Ethnicity, Hispanic/Latino: 12 (17.9%)	Weight, Change - Baseline – 12 wk: -5.35 lb (22.93) vs. -4.07 lb (21.44) Adverse Events, Nausea - Baseline – 12 wk: 19 vs. 9 Attrition: 25% (10/40) vs. 25% (10/40)	

Abbreviations: BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; NR=not reported; OTC=over-the-counter; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Tricyclic Antidepressants

Imipramine Compared to Placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Laederach-Hofmann et al. (1999)	Design: RCT Setting: Outpatient: Medical Outpatient Clinic of the University of Berne Country: Switzerland Funding: NR	Randomized N=31 Imipramine 25 mg + Diet Counseling + Psychological Support 8 wk > Diet Counseling + Psychological Support 32 wk (N=15) Placebo + Diet Counseling + Psychological Support 32 wk (N=16)	Inclusion: BED; overweight or obese; BMI >27.5 kg/m^2; 20-60 years of age Exclusion: Psychoactive medication; appetite suppressants; cyclothymia; Schizophrenia; major depression; personality disorders; concomitant psychotherapy; BN; AN; endocrine disorders; diabetes; other eating disorders	BED: 31 (100%) Overweight or Obesity: 31 (100%) Weight: 96 kg (SD ± 14.2) vs. 114.8 kg (SD ± 29.5, N=14) BMI > 27.5 kg/m²: 31 (100%) BMI: 39.5 kg/m² (SD ± 8.6) - 36.1 kg/m² (SD ± 6.3) vs. 43.2 kg/m² (SD ± 9.4, N=14)	Percent change of binge-eating episodes was -42% with imipramine vs. 7% with placebo at 32 wk. - Binge Eating - Baseline: 7.1/wk (SD ± 4.1) vs. 7.1/wk (SD ± 4.9, N=14) - Binge Eating, Change - Baseline – 32 wk: -3.2/wk (SD ± 2.9) vs. 0/wk (SD ± 1.4, N=14) Imipramine group had 5 kg weight loss vs. 2.1 kg weight gain with placebo group.	Moderate

				Age 20 yr-60 yr: 31 (100%) Age: 40.7 yr (SD ± 10.9) vs. 35.7 yr (SD ± 10.3) Gender - Female: 27 (87.1%) - Male: 4 (12.9%) Race: NR	Imipramine group had 34.3% reduction in HAM-D vs. 21.5% increase with placebo group. Treatment Discontinuation, Adverse Events - Baseline – 8 wk: 1 (6.67%) vs. 1 (6.25%) Attrition: 7% (1/15) vs. 6% (1/16)	

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; HAM-D=Hamilton Depression Rating Scale; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Desipramine Compared to Cognitive-Behavioral Therapy and Weight Loss Treatment

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Agras et al. (1994b)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=108 CBT 12 wk > Weight Loss Treatment 36 wk (N=36) CBT 12 wk > Weight Loss Therapy + Desipramine 25-300 mg 36 wk (N=36) Weight Loss Treatment 36 wk (N=37) CBT > Weight Loss Therapy +/- Desipramine 25-300 mg 36 wk (pooled) (N=72)	Inclusion: Female; BED; binge eating at least twice a wk for a 6-mo period; overweight Exclusion: Current weight loss program; antidepressant medication; any medication that may affect weight; suicidality; abuse of drugs or alcohol; history of purging in the prior 12 months; BMI below 27 kg/m^2; current BN	BED: 108 (100%) Binge Eating >= 2/wk, Duration 6 mo: 108 (100%) Binge Eating: 4.5 d/wk (SD ± 1.4) - 4.4 d/wk (SD ± 1.4, N=30) vs. 5.1 d/wk (SD ± 1.4, N=27) vs. 4.5 d/wk (SD ± 1.6, N=27) Overweight: 108 (100%) Weight: 104.9 kg (SD ± 18.5) - 102.1 kg (SD ± 15.7, N=30) vs. 111.9 kg (SD ± 17.4, N=27) vs. 102.9 kg (SD ± 15.8, N=27) BMI: 38.6 kg/m² (SD ± 6.6)	At 12 wk, CBT groups had significantly less binge eating (67% reduction vs. 44% with weight loss alone, MD -23 %, p<0.01) and the weight loss group had more weight loss (-2.0 kg) compared to CBT groups (0.7 kg) (MD 2.7 kg, p<0.002). No differences were noted between groups at the end of treatment or follow-up except weight loss (0 kg vs. -4.8 kg vs. -4.15 kg at 48 wk) - CBT > Weight Loss Treatment vs. CBT > Weight Loss Therapy + Desipramine: MD 4.8 kg (p<0.05)	High

		Follow-up: Baseline – 48 wk Current Analysis (N=84) - 30 vs. 27 vs. 27		Age: 45 yr (SD ± 10) Gender, Female: 108 (100%) Race: NR	Binge Eating, Abstinence - 48 wk: 8 (28%, N=30) vs. 9 (32%, N=27) vs. 4 (14%, N=27) BDI – Baseline: 13.5 units (SD ± 7.8, N=30) vs. 13.7 units (SD ± 8.1, N=27) vs. 12.9 units (SD ± 6.5, N=27) BDI, Change - Baseline – 36 wk: -4.6 units (SD ± 10.5, N=30) vs. -5.9 units (SD ± 10.84, N=27) vs. -1.6 units (SD ± 11.79, N=27) Attrition: 17% (11/36) vs. 23% (12/36) vs. 27% (16/37)	

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Stimulants

Lisdexamfetamine compared to placebo

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition	Risk of bias
Guerdjikova et al. (2016)	Design: RCT Setting: Outpatient: Lindner Center of HOPE Country: United States Funding: Industry	Randomized N=50 LDX 20-70 mg 12 wk (N=25) Placebo 12 wk (N=25) Screening: -2 wk – Baseline Follow-up: Baseline – 13 wk	Inclusion: BED; 18-55 years of age; >=3 binge-eating days/wk for the 2 weeks before receiving study medication Exclusion: Current AN or BN; current suicidal ideation; a suicide attempt within the last year; receipt of a psychological intervention or weight loss intervention for BED that begun within 3 months of study entry; substance use disorder or	BED: 50 (100%) Binge Eating >= 3 d/wk, In the Previous 2 wk: 50 (100%) Binge Eating - 5.1/wk (SD ± 3.1) - 4.2 d/wk (SD ± 1.2) Weight: 111.3 kg (SD ± 26.4) - 113 kg (SD ± 26.6, 95% CI: 103.7 – 122.9) vs.	No change was noted on primary longitudinal analyses but more patients had a weight decrease of >7% with LDX compared to placebo (26% vs. 0%, p=0.02). - Weight, Change - Baseline – 12 wk: -4.3 kg (SD ± 3.4) vs. -0.6 kg (SD ± 3.2) (MD -3.7 kg, 95% CI -5.7 – -1.8) Binge Eating, Baseline	High

			stimulant misuse within 6 months of study entry; lifetime history of psychosis, mania, or hypomania; clinically unstable medical illness; clinically significant laboratory or electrocardiogram abnormalities; receipt of psychotropic medication within 4 weeks prior to randomization; pregnant; lactating	109.5 kg (SD ± 26.7, 95% CI: 100.3 – 118.5) BMI: 39.8 kg/m² (SD ± 9.3) - 40.8 kg/m² (SD ± 9.7) vs. 38.8 kg/m² (SD ± 9.1) Age 18 yr-55 yr: 50 (100%) Age: 37.7 yr (SD ± 8.9) - 39.2 yr (SD ± 9.2) vs. 36.1 yr (SD ± 8.5) Gender - Female: 24 (96%) vs. 22 (88%) - Male: 1 (4%) vs. 3 (12%) Race - Caucasian: 19 (76%) vs. 20 (80%) - Black or African American: 6 (24%) vs. 4 (16%) - Asian: 0 (0%) vs. 1 (4%)	- 5.6/wk (SD ± 3.7) vs. 4.7/wk (SD ± 2.4) - 4.3 d/wk (SD ± 1.3, 95% CI 3.81 – 4.7) vs. 4.1 d/wk (SD ± 1.1, 95% CI 3.68 – 4.4) Binge Eating, Change - Baseline – 12 wk: - -4.7/wk (SD ± 3.1) vs. -2.2/wk (SD ± 2.5) (MD -2.5/wk, 95% CI -4.1 – -0.8) - -3.4 d/wk (SD ± 1.3) vs. -2.3 d/wk (SD ± 1.8) (MD -1 d/wk, 95% CI -1.9 – -0.1) Disease Response, Remission - Baseline – 12 wk: 10 (43%, N=23) vs. 8 (35%, N=23) There were greater rates of jitteriness (28% vs. 0%, p<0.05), xerostomia (48% vs. 0%, p<0.05), and insomnia (44% vs. 8%, p<0.05) with LDX. Treatment Discontinuation, Adverse Events - Baseline – 12 wk: 2 (8%) vs. 2 (8%) Attrition: 48% (12/25) vs. 44% (11/25)	
Hudson et al. (2017)	Design: RCT (withdrawal design) Setting: Multi-center Country: Canada; Germany; Spain; Sweden; United States Funding: Industry	Randomized N=275 LDX 50-70 mg 26 wk (N=137) Placebo 26 wk (N=138) Follow-up: Baseline – 27 wk	Inclusion: 18-55 years of age; history of BED, moderate to severe; BMI 18- 45 kg/m^2; treatment response, LDX; <= 1 binge-eating day/wk for the last 4 consecutive weeks; CGI-S score <= 2 units Exclusion: current AN or BN; psychiatric disorders, comorbid; psychotherapy or weight loss support for BED within 3	BED, Moderate to Severe, History of: 275 (100%) Binge Eating <= 1 d/wk, In the Previous 4 wk: 275 (100%) BMI 18 kg/m²-45 kg/m²: 275 (100%) Binge Eating	Relapse was less frequent with continued LDX (3.7% vs. 32.1% placebo) at 26 wk and LDX was also superior on time to relapse (via log rank test with HR 0.09, 95% CI 0.04 – 0.23, p<0.001). Binge Eating, Change - Baseline – 26 wk: 0.02 d/wk (SD ± 0.62, N=102) vs. 0.63 d/wk (SD ±	High

			months of screening; psychostimulant use for BED within 6 months before screening; MADRS score of 18 or higher at screening; past suicide attempt; current active suicidal ideation; history of stimulant abuse or dependence; substance abuse or dependence within the past 6 months; use of prohibited medications	- -12 wk: 4.8 d/wk (SD ± 1.19, N=136) vs. 4.71 d/wk (SD ± 1.23, N=131) - Baseline: 0.12 d/wk (SD ± 0.262, N=136) vs. 0.13 d/wk (SD ± 0.274, N=131) Binge Eating, 4 wk Cessation: 89 (65.4%, N=136) vs. 86 (65.6%, N=131) Weight - -12 wk: 92.42 kg (SD ± 18.42, N=136) vs. 97.66 kg (SD ± 21.06, N=131) Age 18 yr-55 yr: 275 (100%) Age: 37.3 yr (SD ± 10, N=136) vs. 40.1 yr (SD ± 9.9, N=131) Gender - Female: 122 (89.7%, N=136) vs. 112 (85.5%, N=131) - Male: 14 (10.3%, N=136) vs. 19 (14.5%, N=131) Race - Caucasian: 112 (82.4%, N=136) vs. 113 (86.3%, N=131) - Non-Caucasian: 24 (17.6%, N=136) vs. 18 (13.7%, N=131)	0.54, N=50) (MD -0.61 d/wk, 95% CI -0.81 – -0.42, p<0.001) Weight, Change - Baseline – 26 wk: -8.29 kg (SD ± 7.62, N=136) vs. -4.25 kg (SD ± 5.29, N=131) Adverse Events - Baseline – 27 wk: 82 (60.3%, N=136) vs. 62 (46.3%, N=134) - Treatment-Related: 32 (23.5%, N=136) vs. 19 (14.2%, N=134) - Severe: 4 (2.9%, N=136) vs. 0 (0%, N=134) Treatment Discontinuation, Adverse Events - Baseline – 26 wk: 6 (4.38%) vs. 0 (0%) Attrition: 26% (35/137) vs. 64% (88/138)	
McElroy et al. (2015b, 2016b)	Design: RCT; Secondary Analysis Setting: Multi-center	Randomized N=260 LDX 30 mg 11 wk (N=66)	Inclusion: 18-55 years of age; BED; BMI of at least 25 and no greater than 45 kg/m^2	BED: 260 (100%) BMI 25 kg/m²-45 kg/m²: 260 (100%)	Log-transformed scale was used to measure binge-eating days/wk as primary outcome. Binge-eating days/wk were reduced with 50 mg and 70 mg doses (4.5 d/wk->0.4 d/wk and	High

	Country: United States Funding: Industry	LDX 50 mg 11 wk (30 mg induction) (N=65) LDX 70 mg 11 wk (30 mg induction) (N=65) Placebo 11 wk (N=64) LDX 30/50/70 mg 11 wk (30 mg induction) (pooled) (N=196) Follow-up: Baseline – 12 wk ITT (N=255) - 66 vs. 64 vs. 63 vs. 62	Exclusion: Current AN or BN; current ADHD; current psychiatric disorder; lifetime history of bipolar disorder or psychosis; MADRS score of at least 18; psychological interventions or weight-loss interventions initiated within 3 months of screening; use of a psychostimulant within the prior 6 months; recent history of suspected substance abuse; lifetime history of psychostimulant abuse or dependence; Investigational compounds, sedatives, anxiolytics, antipsychotics, antidepressants, norepinephrine reuptake inhibitors, benzodiazepines or weight-reducing agents within the past 30 days; psychostimulants within the past 60 days; current Axis I disorder	Weight: 98.6 kg (SD ± 17.85, N=259) - 98.5 kg (SD ± 18.65) vs. 100.6 kg (SD ± 18.84) vs. 98.4 kg (SD ± 16.7) vs. 96.8 kg (SD ± 17.28, N=63) - LDX 30/50/70 mg (pooled): 99.2 kg (SD ± 18.03) BMI: 34.9 kg/m² (SD ± 5.3, N=259) - 35 kg/m² (SD ± 5.39) vs. 35.2 kg/m² (SD ± 5.73) vs. 35 kg/m² (SD ± 4.82) vs. 34.3 kg/m² (SD ± 5.31, N=63) - LDX 30/50/70 mg (pooled): 35.1 kg/m² (SD ± 5.3) Age 18 yr-55 yr: 260 (100%) - Age < 40 yr: 134 (51.7%, N=259) - Age >= 40 yr: 125 (48.3%, N=259) Age: 38.7 yr (SD ± 10.17) - (N=259) - 38.4 yr (SD ± 11.14) vs. 39.6 yr (SD ± 9.32) vs. 38.6 yr (SD ± 10.01) vs. 38 yr (SD ± 10.3, N=63) Gender - Female: 57 (86.4%) vs. 50 (76.9%) vs. 55 (84.6%) vs. 49 (77.8%, N=63) - Male: 9 (13.6%) vs. 15 (23.1%) vs. 10 (15.4%) vs. 14 (22.2%, N=63) Race	4.6->0.5 respectively) but not 30 mg and placebo (4.5 ->1 and 4.3->1.1 respectively). Similarly, rates of 4-wk cessation of binge eating were 42.2% 50 mg and 50% 70 mg vs. 34.9% 30 mg and 21.3% placebo. - LDX 30 mg vs. Placebo: p=0.09 - LDX 50 mg vs. Placebo: p=0.01 - LDX 70 mg vs. Placebo: p<0.001 Weight reduction was -3.1 kg 30 mg, -4.9 kg 50 or 70 mg, and -1 kg placebo. Decreased appetite, insomnia, and dry mouth occurred in more than 10% of LDX patients at numerically greater rates than placebo, though, study withdrawal rates were comparable. Appetite, Decrease: 17 (25.8%) vs. 13 (20%) vs. 12 (18.5%) vs. 4 (6.3%, N=63) Insomnia: 7 (10.6%) vs. 10 (15.4%) vs. 9 (13.8%) vs. 1 (1.6%, N=63) Xerostomia: 22 (33.3%) vs. 22 (33.8%) vs. 27 (41.5%) vs. 5 (7.9%, N=63) Adverse Events - Baseline – 11 wk: 57 (86.4%) vs. 56 (86.2%)	

				- Caucasian: 48 (72.7%) vs. 53 (81.5%) vs. 49 (75.4%) vs. 52 (82.5%, N=63) - Black or African American: 15 (22.7%) vs. 10 (15.4%) vs. 12 (18.5%) vs. 9 (14.3%, N=63) - Asian: 1 (1.5%) vs. 0 (0%) vs. 1 (1.5%) vs. 2 (3.2%, N=63) - Other: 2 (3%) vs. 1 (1.5%) vs. 1 (1.5%) vs. 0 (0%, N=63) - Native American/Alaska Native: 0 (0%) vs. 1 (1.5%) vs. 2 (3.1%) vs. 0 (0%, N=63) Ethnicity - Non-Hispanic/Non-Latino: 59 (89.4%) vs. 58 (89.2%) vs. 55 (84.6%) vs. 58 (92.1%, N=63) - Hispanic/Latino: 7 (10.6%) vs. 7 (10.8%) vs. 10 (15.4%) vs. 5 (7.9%, N=63)	vs. 53 (81.5%) vs. 37 (58.7%, N=63) Mortality, All-Cause - Baseline – 11 wk: 0 (0%) vs. 0 (0%) vs. 1 (1.54%) vs. 0 (0%, N=63) Study Withdrawal, Adverse Events: 3 (4.55%) vs. 1 (1.54%) vs. 3 (4.62%) vs. 0 (0%, N=63) Adherence, Completed Treatment - Baseline – 11 wk - 51 (77.27%) vs. 52 (80%) vs. 52 (80%) vs. 47 (74.6%, N=63) - LDX 30/50/70 mg (pooled): 155 (79.08%) Attrition: 23% (15/66) vs. 20% (13/65) vs. 20% (13/65) vs. 27% (17/64)	
McElroy et al. (2016a, 2017)	Design: RCT (Study 1); Post-hoc Analysis Setting: Multi-center Country: Germany; Spain; Sweden; United States Funding: Industry	Randomized N=383 LDX 50-70 mg 12 wk (30 mg induction) (N=192) Placebo 12 wk (N=191) Follow-up: Baseline – 13 wk	Inclusion: 18-55 years of age; moderate to severe BED; BMI >=18 and <=45 kg/m^2; CGI-S score >= 4 Exclusion: Current AN or BN; comorbid current psychiatric disorders; psychotherapy within <= 3 months; weight loss support for BED within <= 3 months; psychostimulants for fasting or dieting <= 6 months before screening; MADRS total score >= 18 at screening; considered a suicide risk by the	BED, Moderate to Severe: 383 (100%) BMI >= 18 kg/m²-<= 45 kg/m²: 383 (100%) Binge Eating: 6.41/wk (SD ± 2.957) vs. 5.96/wk (SD ± 2.535, N=187) Binge Eating: 4.69 d/wk (SD ± 1.237, N=379)	LDX was associated with greater changes in binge days/wk (3.87 fewer vs. 2.51 fewer, MD -1.35 d/wk, 95% CI -1.7 – -1.01), greater rates of binge abstinence (40% vs. 14.1%), and weight reduction (-6.25 kg mean vs. -0.1 kg). LDX was superior at all post-treatment assessments and showed greater rates of patients being much improved or very	High

			investigator; previously made a suicide attempt; currently demonstrating active suicidal ideation; lifetime history of psychosis, mania, hypomania, dementia or ADHD; lifetime amphetamine abuse or dependence; lifetime stimulant abuse or dependence; recent history of substance abuse or dependence	- 4.78 d/wk (SD ± 1.266) vs. 4.59 d/wk (SD ± 1.201, N=187) Weight: 94.3 kg (SD ± 19.732) vs. 92.7 kg (SD ± 19.331, N=187) BMI: 33.68 kg/m² (SD ± 6.292) vs. 33.21 kg/m² (SD ± 6.234, N=187) Age 18 yr-55 yr: 383 (100%) Age: 38.5 yr (SD ± 10.4) vs. 37.6 yr (SD ± 10.21, N=187) Gender - Female: 165 (85.9%) vs. 163 (87.2%, N=187) - Male: 27 (14.1%) vs. 24 (12.8%, N=187) Race - Caucasian: 150 (78.1%) vs. 144 (77%, N=187) - Black or African American: 33 (17.2%) vs. 29 (15.5%, N=187) - Asian: 3 (1.6%) vs. 5 (2.7%, N=187) - Multiracial: 1 (0.5%) vs. 6 (3.2%, N=187) - Native American/Alaska Native: 2 (1%) vs. 2 (1.1%, N=187) - Native Hawaiian/Pacific Islander: 2 (1%) vs. 1 (0.5%, N=187)	much improved on CGI (86% vs. 47%). Severe adverse events were also greater with LDX (8.9% vs. 3.2%). Adverse Events - Baseline – 12 wk: 158 (82.3%) vs. 110 (58.8%) (N=187) - Xerostomia: 76 (39.6%) vs. 16 (8.6%, N=187) - Heart Rate, Increase: 14 (7.3%) vs. 5 (2.7%, N=187) - Hyperhidrosis: 10 (5.2%) vs. 1 (0.5%, N=187) - Decreased Appetite: 17 (8.9%) vs. 6 (3.2%, N=187) - Headache: 26 (13.5%) vs.17 (9.1%, N=187) - Insomnia: 34 (17.7%) vs. 14 (7.5%, N=187) Adverse Events, Treatment-Related - Baseline – 12 wk: 134 (69.8%) vs. 71 (38%, N=187) Treatment Discontinuation - Baseline – 12 wk - Adverse Events: 12 (6.3%) vs. 5 (2.62%) - Lack of Efficacy: 5 (2.6%) vs. 2 (1.05%) Attrition: 18% (34/192) vs. 18% (34/191)	
McElroy et al. (2016a, 2017)	Design: RCT (Study 2); Post-hoc Analysis	Randomized N=390	Inclusion: 18-55 years of age; moderate to severe BED; BMI	BED, Moderate to Severe: 390 (100%)	LDX was associated with greater rates of binge abstinence (36.2% vs. 13.1%)	High

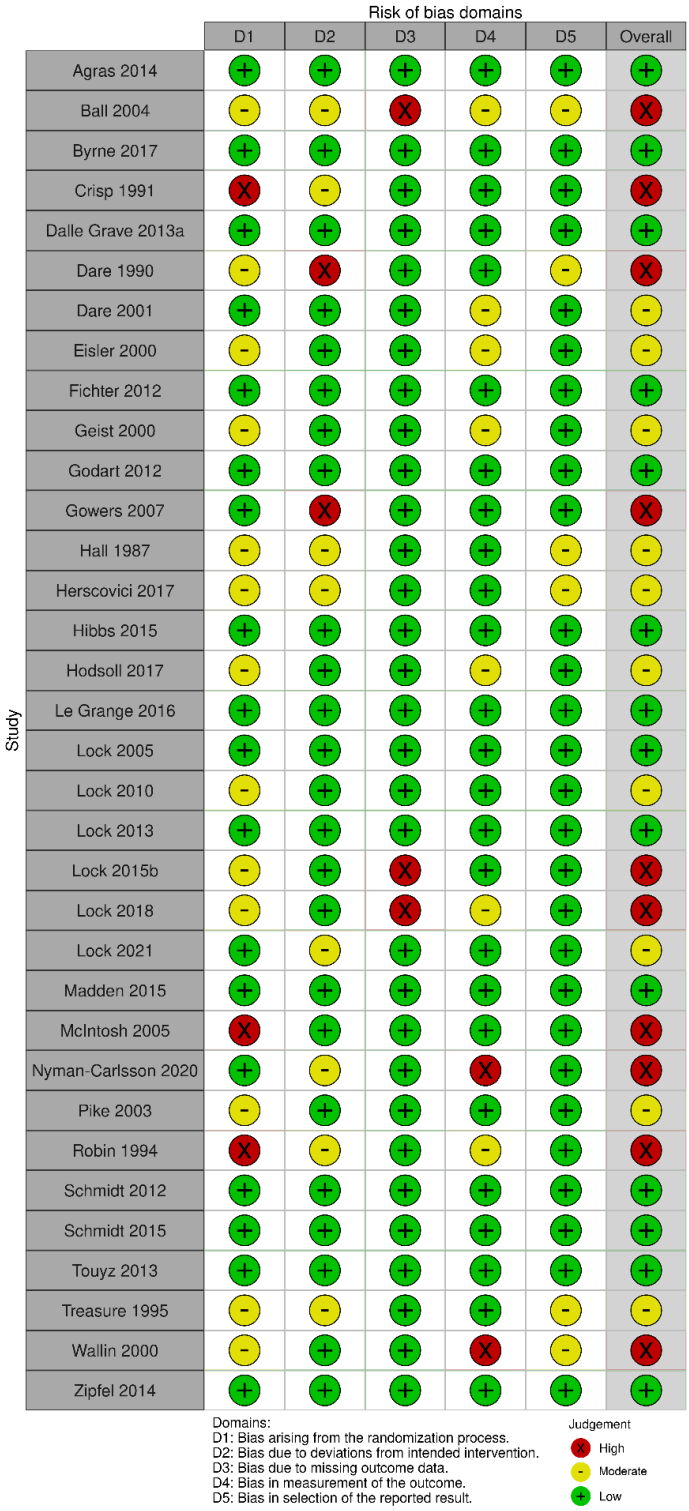
	Setting: Multi-center Country: Germany; United States Funding: Industry	LDX 50-70 mg 12 wk (30 mg induction) (N=195) Placebo 12 wk (N=195) Follow-up: Baseline – 13 wk	>=18 and <=45 kg/m^2; CGI-S score >= 4 Exclusion: Current AN or BN; comorbid current psychiatric disorders; psychotherapy within <= 3 months; weight loss support for BED within <= 3 months; psychostimulants for fasting or dieting <= 6 months before screening; MADRS total score >= 18 at screening; considered a suicide risk by the investigator; previously made a suicide attempt; currently demonstrating active suicidal ideation; lifetime history of psychosis, mania, hypomania, dementia, or ADHD; lifetime amphetamine abuse or dependence; lifetime stimulant abuse or dependence; recent history of substance abuse or dependence	Binge Eating: 6.39/wk (SD ± 3.439, N=181) vs. 6.65/wk (SD ± 3.787, N=185) Binge Eating: 4.75 d/wk (SD ± 1.359, N=366)Weight: 94.75 kg (SD ± 21.745, N=181) vs. 93.05 kg (SD ± 20.33, N=185) BMI >= 18 kg/m²-<= 45 kg/m²: 390 (100%) BMI: 33.85 kg/m² (SD ± 6.202, N=181) vs. 33.2 kg/m² (SD ± 6.341, N=185) Age 18 yr-55 yr: 390 (100%) Age: 37.1 yr (SD ± 10, N=181) vs. 38.7 yr (SD ± 10.01, N=185) Gender - Female: 159 (87.8%, N=181) vs. 153 (82.7%, N=185) - Male: 22 (12.2%, N=181) vs. 32 (17.3%, N=185) Race - Caucasian: 130 (71.8%, N=181) vs. 137 (74.1%, N=185) - Black or African American: 43 (23.8%, N=181 vs. 32 (17.3%, N=185) - Asian: 3 (1.7%, N=181) vs. 4 (2.2%, N=185) - Multiracial: 3 (1.7%, N=181) vs. 8 (4.3%, N=185)	and weight reduction (-5.57 kg mean vs. -0.15 kg). Binge Eating - Baseline: 4.66 d/wk (SD ± 1.28, N=181) vs. 4.85 d/wk (SD ± 1.43, N=185) Binge Eating, Change - Baseline – 12 wk: -3.92 d/wk (SD ± 1.78, N=174) vs. -2.26 d/wk (SD ± 1.82, N=176) (MD -1.66 d/wk, 95% CI -2.04 – -1.28) Post-hoc analysis showed greater rates of patients being much improved or very much improved on CGI with LDX (86% vs. 43%). Treatment Discontinuation, Adverse Events - Baseline – 12 wk: 7 (3.59%) vs. 5 (2.56%) Adverse Events, Serious - Baseline – 12 wk: 1 (0.6%, N=181) vs. 2 (1.1%, N=185) Adverse Events - Baseline – 12 wk: 140 (77.3%, N=181) vs. 94 (50.8%, N=185) Dry mouth, insomnia, and headache were the most commonly reported side effects. - Xerostomia: 60 (33.1%, N=181) vs. 11 (5.9%, N=185) - Insomnia: 19 (10.5%, N=181) vs. 6 (3.2%, N=185) - Headache: 32 (17.7%, N=181) vs. 16 (8.6%, N=185)	

				- Native American/Alaska Native: 0 (0%, N=181) vs. 4 (2.2%, N=185) - Native Hawaiian/Pacific Islander: 2 (1.1%, N=181) vs. 0 (0%, N=185)	Attrition: 25% (48/195) vs. 25% (48/195)	

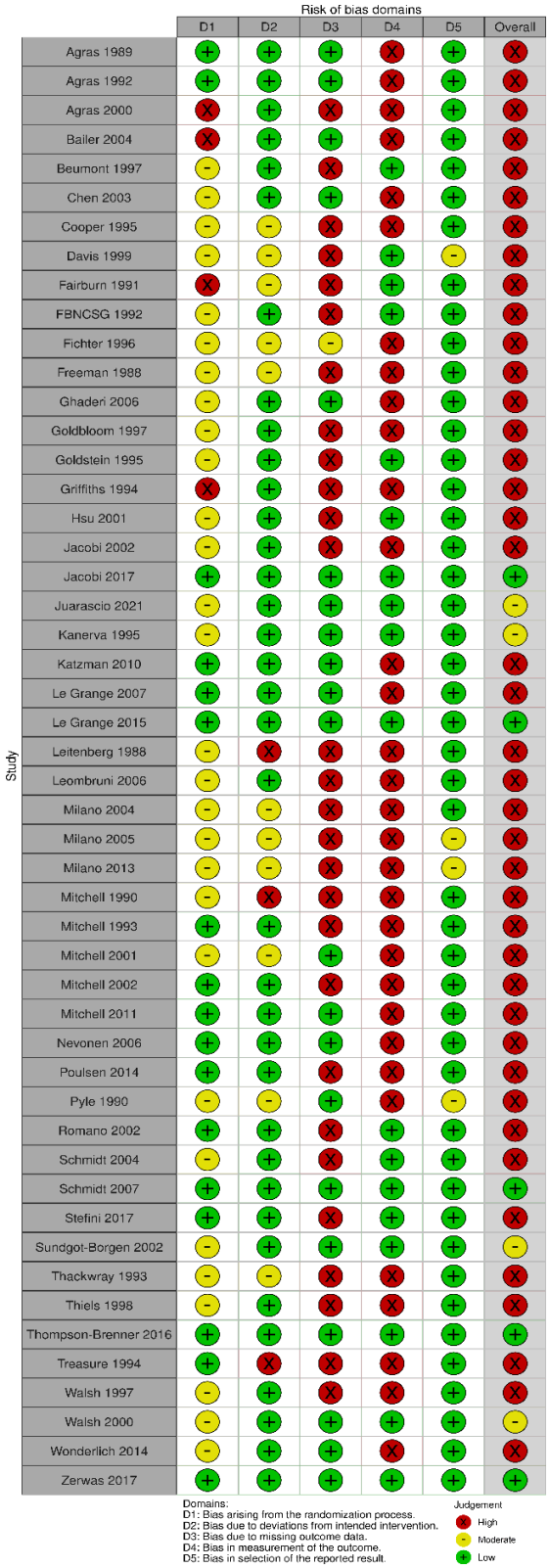
Abbreviations: ADHD=attention-deficit/hyperactivity disorder; AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CGI=Clinical Global Impression; CGI-S=Clinical Global Impression-Severity; CI=confidence interval; d=day; HR=hazard ratio; LDX=lisdexamfetamine; MADRS=Montgomery-Asberg Depression Rating Scale; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Appendix F. Risk of Bias Ratings for Individual Studies Supporting Guideline Statements

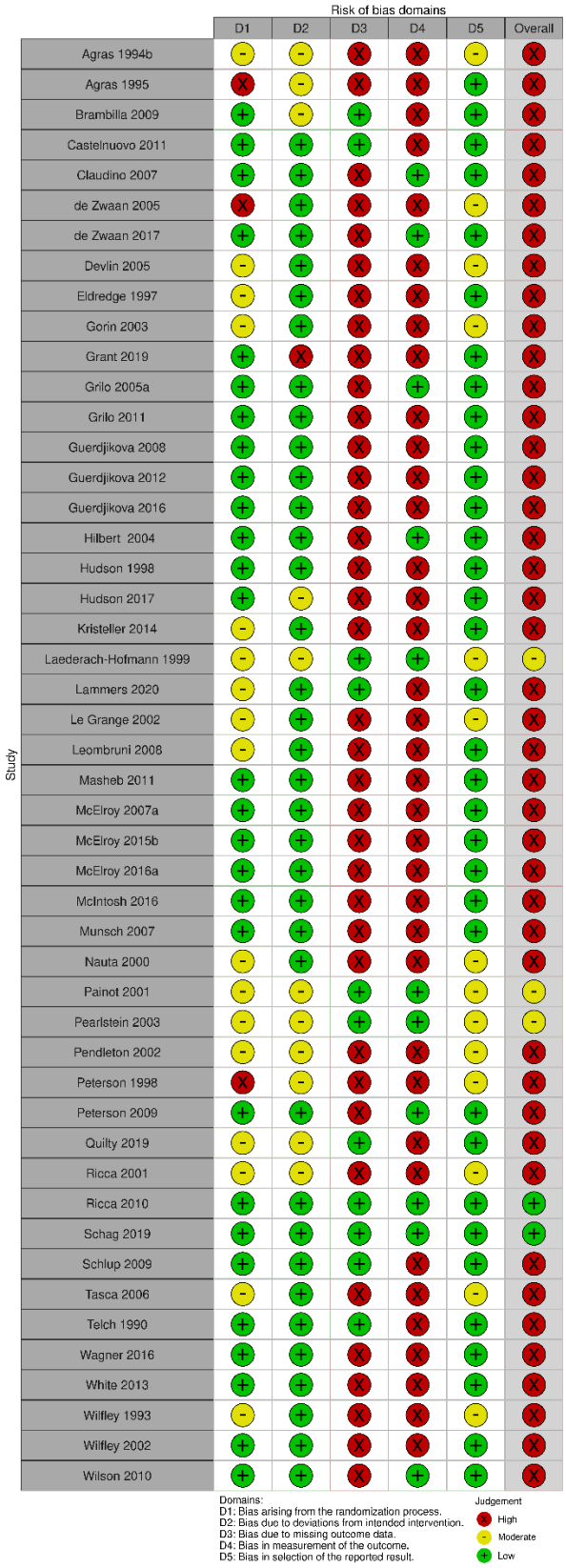
Anorexia Nervosa Studies



Bulimia Nervosa Studies



Binge-Eating Disorder Studies



Appendix G. Balancing of Potential Benefits and Harms in Rating the Strength of the Guideline Statements and Quality Measurement Considerations

Use of Guidelines to Enhance Quality of Care

Clinical practice guidelines can help enhance quality by synthesizing available research evidence and delineating recommendations for care on the basis of the available evidence. In some circumstances, practice guideline recommendations will be appropriate to use in developing quality measures. Guideline statements can also be used in other ways, such as educational activities or electronic decision support, to enhance the quality of care that patients receive. Furthermore, when availability of services is a major barrier to implementing guideline recommendations, improved tracking of service availability and program development initiatives may need to be implemented by health organizations, health insurance plans, federal or state agencies, or other regulatory programs.

Typically, guideline recommendations that are chosen for development into quality measures will advance one or more aims of the Institute of Medicine's report on "Crossing the Quality Chasm" (Institute of Medicine 2001) by facilitating care that is safe, effective, patient-centered, timely, efficient, and equitable. To achieve these aims, quality measures (Watkins et al. 2015) are needed that span the continuum of care (e.g., prevention, screening, assessment, treatment, continuing care), address the different levels of the health system hierarchy (e.g., system-wide, organization, program/department, individual clinicians), and include measures of different types (e.g., process, outcome, patient-centered experience). Emphasis is also needed on factors that influence the dissemination and adoption of evidence-based practices (Drake et al. 2008; Greenhalgh et al. 2004; Horvitz-Lennon et al. 2009).

Often, quality measures will focus on gaps in care or on care processes and outcomes that have significant variability across specialties, health care settings, geographic areas, or patients' demographic characteristics. Administrative databases, registries, and data from electronic health record (EHR) systems can help to identify gaps in care and key domains that would benefit from performance improvements (Acevedo et al. 2015; Patel et al. 2015; Watkins et al. 2016). Nevertheless, for some guideline statements, evidence of practice gaps or variability will be based on anecdotal observations if the typical practices of psychiatrists and other health professionals are unknown. Variability in the use of guideline-recommended approaches may reflect appropriate differences that are tailored to the patient's preferences, treatment of co-occurring illnesses, or other clinical circumstances that may not have been studied in the available research. On the other hand, variability may indicate a need to strengthen clinician knowledge or address other barriers to adoption of best practices (Drake et al. 2008; Greenhalgh et al. 2004; Horvitz-Lennon et al. 2009). When performance is compared among organizations, variability may reflect a need for quality improvement initiatives to improve overall outcomes but could also reflect case-mix differences such as socioeconomic factors or the prevalence of co-occurring illnesses.

Conceptually, quality measures can be developed for purposes of accountability, for internal or health system–based quality improvement, or both. Accountability measures require clinicians to report their rate of performance of a specified process, intermediate outcome, or outcome in a specified group of patients. Because these data are used to determine financial incentives or penalties based on

performance, accountability measures must be scientifically validated, have a strong evidence base, fill gaps in care, and be broadly relevant and meaningful to patients, clinicians, and policy makers. Development of such measures is complex and requires detailed development of specification and pilot testing (Center for Health Policy/Center for Primary Care and Outcomes Research and Battelle Memorial Institute 2011; Fernandes-Taylor and Harris 2012; Iyer et al. 2016; Pincus et al. 2016; Watkins et al. 2011). In contrast, internal or health system–based quality improvement measures are typically designed by and for individual providers, health systems, or payers. They typically focus on measurements that can suggest ways for clinicians or administrators to improve efficiency and delivery of services within a particular setting. Internal or health system–based quality improvement programs may or may not link performance with payment, and, in general, these measures are not subject to strict testing and validation requirements.

Regardless of the purpose of the quality measure, it must be possible to define the applicable patient group (i.e., the denominator) and the clinical action or outcome of interest that is measured (i.e., the numerator) in validated, clear, and quantifiable terms. The measure also needs to be feasible. More specifically, the health system's or clinician's performance on the measure must be readily ascertained from chart review, patient-reported outcome measures, registries, or administrative data. In addition, use of the measure should yield improvements in quality of care to justify any clinician burden (e.g., documentation burden) or related administrative costs (e.g., for manual extraction of data from charts, for modifications of EHRs to capture required data elements).

Documentation of quality measures can be challenging, and, depending on the practice setting, can pose practical barriers to meaningful interpretation of quality measures based on guideline recommendations. For example, when recommendations relate to patient assessment or treatment selection, clinical judgment may need to be used to determine whether the clinician has addressed the factors that merit emphasis for an individual patient. In other circumstances, standardized instruments can facilitate quality measurement reporting, but it is difficult to assess the appropriateness of clinical judgment in a validated, standardized manner. Furthermore, utilization of standardized assessments remains low (Fortney et al. 2017), and clinical findings are not routinely documented in a standardized format. Many clinicians appropriately use free text prose to describe symptoms, response to treatment, discussions with family, plans of treatment, and other aspects of care and clinical decision-making. Reviewing these free text records for measurement purposes would be impractical, and it would be difficult to hold clinicians accountable to such measures without advances in natural language processing technology and further increases in EHR use among mental health professionals.

Possible unintended consequences of any derived measures would also need to be addressed in testing of a fully specified measure in a variety of practice settings. For example, in many health care systems, multiple clinicians are involved in the care of a patient and attributing measure performance to one clinician, or one group of clinicians, can be misleading. As another challenge, highly specified measures may lead to overuse of standardized language that does not accurately reflect what has occurred in practice. If multiple discrete fields are used to capture information, data will be easily retrievable and reportable, but oversimplification is a possible unintended consequence of measurement and documentation burden is likely to be high (Johnson et al. 2021). Just as guideline developers must

balance the benefits and harms of a particular guideline recommendation, developers of performance measures must weigh the potential benefits, burdens, and unintended consequences in optimizing quality measure design and testing.

Assessment and Determination of Treatment Plan

Statement 1 – Screening for Presence of an Eating Disorder

APA recommends (1C) screening for the presence of an eating disorder as part of an initial psychiatric evaluation.

Benefits

Screening for the presence of an eating disorder can identify individuals who require more detailed evaluation to determine whether an eating disorder may be present that requires intervention. Systematic screening is likely to identify individuals who might otherwise have their eating disorder go undetected and may allow earlier intervention, with the potential for better outcomes.

Harms[1]

The harms of screening for an eating disorder include the need to incorporate such questions into clinician workflows, including EHRs, and the time required for asking screening questions that might otherwise be used to inquire about other symptoms or issues that are of greater relevance to a specific patient.

Patient Preferences

Clinical experience suggests that the vast majority of patients are cooperative with and accepting of brief screening questions as part of an initial assessment.

Balancing of Benefits and Harms

The potential benefits of this recommendation were viewed as far outweighing the potential harms. (See Appendix C, Statement 1 for additional discussion of the research evidence.) The level of research evidence for this recommendation is rated as low as there is limited information on whether systematic implementation of screening for an eating disorder is associated with improved detection of these conditions or better outcomes. However, expert opinion suggests that harms of screening are negligible compared with the potential benefit of screening in improving identification of eating disorders.

Differences of Opinion Among Writing Group Members

[1] Harms may include serious adverse events; less serious adverse events that affect tolerability; minor adverse events; negative effects of the intervention on quality of life; barriers and inconveniences associated with treatment; and other negative aspects of the treatment that may influence decision-making by the patient, the clinician, or both.

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other guidelines are consistent in recognizing the importance of screening for and identifying eating disorders, although the scope of individual guidelines' recommendations varies. For example, the American Academy of Child and Adolescent Psychiatry (AACAP) recommends that "mental health clinicians should screen all child and adolescent patients for eating disorders" (Lock et al. 2015a). The American Academy of Pediatrics (AAP) does not make a specific recommendation but notes that annual health visits and pre-sports assessments in pediatrics offer opportunities to screen for eating disorders (Hornberger et al. 2021). Other organizations recommend more targeted screening approaches in at-risk populations (ACOG Committee Opinion 2018; Catalan Agency for Health Technology Assessment and Research 2009; French Haute Autorité de Santé 2010; Hackert et al. 2020; Herpertz et al. 2019; Resmark et al. 2019).

Quality Measurement Considerations

Screening measures can be appropriate for quality measure development, although in this instance further evidence would likely be warranted before incorporating this recommendation into a performance-based measure. However, health plans or local organizations may wish to determine whether screening for an eating disorder is occurring as part of the psychiatric assessment. In addition, incorporation of screening measures for eating disorders into EHRs would be a necessary first step to regular use of such scales. Analysis of data from the EHR would allow initial determinations of screening rates and feedback to practitioners to increase screening for eating disorders. Because several of the available screening measures were developed before BED became part of the DSM-5, information on the reliability, validity, and predictive value for each diagnosis would be important to obtain before use of a screening measure for quality-related purposes. Specific attention is also crucial to identify effects of age, gender, race/ethnicity, culture, language, symptom pattern (e.g., focusing on eating, body shape, muscularity, driven exercise), and setting (e.g., primary care, specialty care) on screening results (see Areas for Future Research). It would also be essential to show an impact of screening on clinical outcomes prior to incorporation of this guideline statement into a formal measure.

Statement 2 – Initial Evaluation of Eating History

APA recommends (1C) that the initial evaluation of a patient with a possible eating disorder include assessment of

- **the patient's height and weight history (e.g., maximum and minimum weight, recent weight changes);**
- **presence of, patterns in, and changes in restrictive eating, food avoidance, binge eating, and other eating-related behaviors (e.g., rumination, regurgitation, chewing and spitting);**
- **patterns and changes in food repertoire (e.g., breadth of food variety, narrowing or elimination of food groups);**

- **presence of, patterns in, and changes in compensatory and other weight control behaviors, including dietary restriction, compulsive or driven exercise, purging behaviors (e.g., laxative use, self-induced vomiting), and use of medication to manipulate weight;**
- **percentage of time preoccupied with food, weight, and body shape;**
- **prior treatment and response to treatment for an eating disorder;**
- **psychosocial impairment secondary to eating or body image concerns or behaviors; and**
- **family history of eating disorders, other psychiatric illnesses, and other medical conditions (e.g., obesity, inflammatory bowel disease, diabetes mellitus).**

Benefits

Assessment of current and prior symptoms as well as previous treatment is beneficial in verifying that an eating disorder is present and in identifying its severity and longitudinal course. Knowledge of the patient's current eating patterns and any compensatory behaviors provides important baseline data for assessing the severity of the clinical presentation and effects of subsequent interventions. Information about family psychiatric history, can help to identify risk factors for the development of an eating disorder, such as a family history of eating disorders and attitudes towards eating, weight, and shape. Obtaining a family history can also identify family-related issues that need to be incorporated into the treatment plan.

Harms

Some individuals may become anxious or frustrated if asked multiple questions, including questions about eating disorders symptoms, during the evaluation. This could interfere with the therapeutic relationship between the patient and the clinician. Another potential consequence is that time used to focus on assessment of eating, shape, and weight concerns could reduce time available to address other issues of importance to the patient or of relevance to diagnosis and treatment planning.

Patient Preferences

Although there is no specific evidence on patient preferences related to assessment in individuals with eating disorders, clinical experience suggests that the majority of patients are cooperative with and accepting of these types of questions as part of an initial assessment.

Balancing of Benefits and Harms

The potential benefits of this recommendation were viewed as far outweighing the potential harms. (See Appendix C, Statement 2 for additional discussion of the research evidence.) The level of research evidence is rated as low because there is minimal research on the benefits and harms of assessing specific eating disorder symptoms and behaviors as part of the initial psychiatric evaluation of an individual with a possible eating disorder. However, expert opinion suggests that conducting such assessments as part of the psychiatric evaluation improves the identification and diagnosis of eating disorders. It is also crucial to treatment planning if an eating disorder is present.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other guidelines are consistent in recognizing the importance of a detailed assessment once a possible eating disorder has been identified (Catalan Agency for Health Technology Assessment and Research 2009; French Haute Autorité de Santé 2010; Hackert et al. 2020; Hay et al. 2014; Herpertz et al. 2019; Lock et al. 2015a; National Guideline Alliance (UK) 2020; Resmark et al. 2019). Although recommendations about specific elements of the assessment vary, typical aspects include questions about weight, body image, diet, eating patterns, and restricting, purging, and exercise behaviors (Catalan Agency for Health Technology Assessment and Research 2009; French Haute Autorité de santé 2010; Hackert et al. 2020; Herpertz et al. 2019; National Guideline Alliance (UK) 2020; Resmark et al. 2019) with some guidelines also recommending questions related to family, personal, and psychosocial history (Catalan Agency for Health Technology Assessment and Research 2009; Couturier et al. 2020; Hackert et al. 2020; Hornberger et al. 2021).

Quality Measurement Considerations

For patients with eating disorders, several components of the initial psychiatric evaluation have potential relevance for quality measure development, although such quality measures do not exist at present. A first step toward development of scientifically sound quality measures is identification of discrete indicators that signal the delivery of high-quality care for individuals with an eating disorder. This step may be challenging to accomplish given the breadth of content within the initial psychiatric assessment and the difficulty in ascertaining evaluation details from chart or administrative data. In addition, many aspects of the initial evaluation of a patient with an eating disorder are already subsumed under good clinical practice. Nevertheless, it may still be possible to use available evidence and expert-recommended consensus to develop and specify electronic and clinical data registry quality measures. Additionally, as discussed in the APA Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd edition (American Psychiatric Association 2016), quality improvement efforts at the local level could assess whether specific aspects of the evaluation were completed while still allowing flexibility in the documentation of findings.

Statement 3 – Quantitative Measures

APA recommends (1C) that the initial psychiatric evaluation of a patient with a possible eating disorder include weighing the patient and quantifying eating and weight control behaviors (e.g., frequency, intensity, or time spent on dietary restriction, binge eating, purging, exercise, and other compensatory behaviors).

Benefits

Use of a quantitative measure as part of the initial evaluation can establish baseline information on the patient's eating disorder symptom severity and associated impairment. As compared with a clinical interview, use of a quantitative measure may improve the consistency with which this information is obtained. When administered through paper-based or electronic self-report, use of quantitative measures may allow routine questions to be asked more efficiently.

Harms

The harms of using a quantitative measure include the time required for administration and review. Overreliance on quantitative measures may lead other aspects of the patient's symptoms and clinical presentation to be overlooked. Patients may also provide inaccurate information about their eating disorder symptoms, such as minimizing symptom severity or frequency, leading to an underestimation of severity of illness. Reliance on inaccurate information can have a negative impact on clinical decision-making, including recommendations for treatment. Some patients may view quantitative measures as impersonal or may feel frustrated by having to complete detailed questionnaires, resulting in possible straining of patient-clinician rapport. Changes in the workflow of clinical practices may be needed to incorporate quantitative measures into routine care. Modification of EHRs or use of other technologies may also be required to facilitate capture of quantitative measure data.

Patient Preferences

Clinical experience suggests that the majority of patients are cooperative with and accepting of quantitative measures as part of an initial assessment.

Balancing of Benefits and Harms

The potential benefits of this recommendation were viewed as far outweighing the potential harms. (See Appendix C, Statement 3 for additional discussion of the research evidence.) This recommendation is also consistent with Guideline VII, "Quantitative Assessment," as part of the APA Practice Guidelines for the Psychiatric Evaluation of Adults (American Psychiatric Association 2016). The level of research evidence for this recommendation is rated as low. Evidence suggests that quantitative measures are reliable methods of assessing disordered eating behaviors and eating-related psychopathology. There is minimal research on the harms of using quantitative measures as part of the psychiatric evaluation as compared with assessment as usual. However, expert opinion suggests that harms of assessment are minimal compared with the benefits of such assessments in improving identification and assessment of eating disorders.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Although many guidelines recommend or discuss the importance of taking a history of weight, eating, purging, exercise, and other behaviors, quantification of symptoms and behaviors is less often mentioned; however, the joint German guideline does discuss use of dimensional approaches such as rating scales in assessing and monitoring patients with an eating disorder (Herpertz et al. 2019; Resmark et al. 2019) and the Danish Health Authority (DHA) notes the value of systematically monitoring eating disorder symptoms in patients with BN (Danish Health Authority 2016b).

Quality Measurement Considerations

Weight and BMI are already available in EHRs and are already incorporated into quality measures for other conditions such as obesity. Consequently, these parameters would be feasible to incorporate into a quality measure that examines whether weights are obtained in individuals with an eating disorder. For individuals with AN, an associated measure could assess whether weight restoration is occurring as treatment proceeds. Information on other weight control behaviors (e.g., frequency, intensity, or time spent on dietary restriction, binge eating, purging, exercise, and other compensatory behaviors) is less commonly available as structured data elements in EHRs, but could be incorporated and serve as a first step for quantitative monitoring of initial symptoms and response to treatment. Additional evidence would be needed to show associations with improved outcomes before adoption of these measures as a performance measure. Also, if such a measure were used for accountability purposes, it should be developed at the level of the organization and not the individual provider level as most care of individuals with AN is delivered using a multidisciplinary team. However, a process-focused internal measure or health system-based measure could be used to support quality improvement efforts to increase the frequency of measuring weight and weight control behaviors in individuals with eating disorders.

Statement 4 – Identification of Co-Occurring Conditions

APA recommends (1C) that the initial psychiatric evaluation of a patient with a possible eating disorder identify co-occurring health conditions, including co-occurring psychiatric disorders.

Benefits

Individuals with eating disorders often have co-occurring health conditions, including co-occurring psychiatric disorders. Some co-occurring health conditions (e.g., diabetes mellitus, celiac disease, Crohn's disease) can exacerbate or increase the likelihood of developing an eating disorder. Other co-occurring health conditions are independent of the eating disorder but add complexity to treatment planning. Co-occurring psychiatric disorders can, similarly, complicate diagnosis and require adjustments to the treatment plan. Alternatively, some medications may be indicated for more than one condition (e.g., an SSRI antidepressant for BN as well as depression). Thus, knowledge of all relevant signs, symptoms, and diagnoses can aid in development of a comprehensive approach to treatment that improves patient outcomes.

Harms

Some individuals may have difficulty concentrating or may become frustrated if asked multiple questions during the evaluation. This could interfere with the therapeutic relationship between the patient and the clinician. Another potential consequence is that time used to focus on assessment of co-occurring disorders could reduce time available to address other issues of importance to the patient or of relevance to diagnosis and treatment planning.

Patient Preferences

Clinical experience suggests that the majority of patients are cooperative with and accepting of assessments for co-occurring conditions that may influence treatment options and health status.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 4 for additional discussion of the research evidence.) This recommendation is also consistent with Guideline I. "Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History" and Guideline VI, "Assessment of Medical Health," as part of the APA Practice Guidelines for the Psychiatric Evaluation of Adults (American Psychiatric Association 2016). The level of research evidence is rated as low because there is minimal research on the benefits and harms of assessing for co-occurring health conditions, including psychiatric conditions, as part of the psychiatric evaluation as compared with not conducting such assessments. However, expert opinion strongly suggests that such assessments improve the identification of other psychiatric disorders, other medical disorders, or complications of eating disorders that can influence treatment planning for an eating disorder. (For additional details, see the APA Practice Guidelines for the Psychiatric Evaluation of Adults; American Psychiatric Association 2016).

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Consistent with APA's recommendation to identify co-occurring health conditions in individuals with a possible eating disorder, other guidelines also recommend evaluating for co-occurring symptoms or disorders as part of eating disorder assessment (French Haute Autorité de Santé 2010; Hackert et al. 2020; Hay et al. 2014; Herpertz et al. 2019; Hornberger et al. 2021; Lock et al. 2015a; National Guideline Alliance (UK) 2020; Resmark et al. 2019).

Quality Measurement Considerations

Although assessing for co-occurring conditions is important to the evaluation of individuals with an eating disorder, it would be challenging to incorporate this recommendation into a highly specified quality measure given the breadth of potential co-occurring conditions and the difficulty in ascertaining evaluation details from chart or administrative data. In addition, assessing co-occurring conditions in a

patient with an eating disorder is already subsumed under good clinical practice. However, quality related efforts at the local level could assess whether EHR templates include prompts for documenting co-occurring conditions and whether such aspects of the evaluation are typically completed, while still allowing flexibility in the documentation of findings.

Statement 5 – Initial Review of Systems

APA recommends (1C) that the initial psychiatric evaluation of a patient with a possible eating disorder include a comprehensive review of systems.

Benefits

Individuals with eating disorders often have other co-occurring medical symptoms and disorders. Co-occurring gastrointestinal, neurological, endocrine, or sexual and reproductive signs and symptoms may emerge as sequelae of eating disorders (Mitchell 2016; Westmoreland et al. 2016). Patients may also have co-occurring medical conditions that can mimic an eating disorder or are independent of the eating disorder but add complexity to treatment planning (e.g., addressing restrictive eating behavior in patients with multiple dietary restrictions related to celiac disease or diabetes mellitus). Knowledge of all relevant signs, symptoms, and diagnoses can aid in development of a comprehensive treatment plan.

Harms

Some individuals may have difficulty concentrating or may become frustrated if asked multiple questions during the evaluation. This could interfere with the therapeutic relationship between the patient and the clinician. Another potential consequence is that time used to focus on assessment of co-occurring disorders could reduce time available to address other issues of importance to the patient or of relevance to diagnosis and treatment planning.

Patient Preferences

Clinical experience suggests that the majority of patients are cooperative with and accepting of assessments for co-occurring medical conditions or complications that may influence treatment options and health status.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 5 for additional discussion of the research evidence.) This recommendation is also consistent with Guideline VI, “Assessment of Medical Health,” as part of the APA Practice Guidelines for the Psychiatric Evaluation of Adults (American Psychiatric Association 2016). The level of research evidence is rated as low because there is minimal research on the benefits and harms of assessing for co-occurring medical conditions as part of the psychiatric evaluation as compared with not conducting such assessments. However, expert opinion strongly suggests that such assessments improve the identification of other medical disorders or complications of eating disorders that can

influence treatment planning for an eating disorder. (For additional details, see the APA Practice Guidelines for the Psychiatric Evaluation of Adults; American Psychiatric Association 2016).

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other guidelines do not specifically mention the importance of a review of systems, but they do describe the value of assessing for symptoms and conditions that could be identified through a review of systems (French Haute Autorité de Santé 2010; Hackert et al. 2020; Hay et al. 2014; Hornberger et al. 2021; Herpertz et al. 2019; Lock et al. 2015a; National Guideline Alliance (UK) 2020; Resmark et al. 2019).

Quality Measurement Considerations

Although the review of systems is an important aspect in the assessment of individuals with an eating disorder, it would be challenging to incorporate this recommendation into a highly specified quality measure given the breadth of symptoms that are relevant to assess and the difficulty in ascertaining evaluation details from chart or administrative data. In addition, many aspects of the review of systems of a patient with an eating disorder are already subsumed under good clinical practice. Nevertheless, quality related efforts at the local level could assess whether EHR templates include prompts for documenting a review of systems and whether such aspects of the evaluation are typically completed while still allowing flexibility in the documentation of findings.

Statement 6 – Initial Physical Examination

APA recommends (1C) that the initial physical examination of a patient with a possible eating disorder include assessment of vital signs, including temperature, resting heart rate, blood pressure, orthostatic pulse, and orthostatic blood pressure; height, weight, and BMI (or percent median BMI, BMI percentile, or BMI Z-score for children and adolescents); and physical appearance, including signs of malnutrition or purging behaviors.

Benefits

Including a comprehensive physical exam in the initial assessment of a patient with a possible eating disorder can establish baseline information about specific symptoms and signs that require intervention or influence decision-making as part of treatment planning. Identifying co-occurring medical conditions, if present, is also important in developing a treatment plan that improves prognosis and can reduce associated symptoms, morbidity, and mortality.

Harms

Some individuals may feel uncomfortable or anxious if asked to participate in a physical examination. This could interfere with the therapeutic relationship between the patient and the clinician.

Patient Preferences

Although there is no specific evidence on patient preferences related to physical examination in individuals with a possible eating disorder, clinical experience suggests that the majority of patients are cooperative with and accepting of a physical examination as part of an initial assessment.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 6 for additional discussion of the research evidence.) This recommendation is also consistent with Guideline VI, "Assessment of Medical Health," as part of the APA Practice Guidelines for the Psychiatric Evaluation of Adults (American Psychiatric Association 2016). The level of research evidence is rated as low because there is minimal research on the benefits and harms of a physical examination as part of the psychiatric evaluation in individuals with a possible eating disorder as compared with not conducting an examination. However, expert opinion suggests that a physical examination is important for diagnosis, evaluation of illness severity, and treatment planning. (For additional details, see the APA Practice Guidelines for the Psychiatric Evaluation of Adults; American Psychiatric Association 2016). In addition, potential effects on the therapeutic relationship can be minimized by having the physical examination conducted by another medically trained clinician who is familiar with common findings in patients with eating disorders.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Consistent with the APA recommendation, a number of other guidelines recommend an initial physical examination (ACOG Committee Opinion 2018; Hackert et al. 2020; Hay et al. 2014; Herpertz et al. 2019; Resmark et al. 2019). Other guidelines describe specific physical findings of eating disorders that warrant evaluation implying that these findings would be identified via a physical examination (Hornberger et al. 2021; Lock et al. 2015a; National Guideline Alliance (UK) 2020). For guidelines that do identify findings of interest, there is some variability in the specific elements mentioned, but these typically relate to weight, height, vital signs, and physical evidence of malnutrition, purging, or common complications of eating disorders.

Quality Measurement Considerations

Although the physical examination is an important aspect in the assessment of individuals with an eating disorder, it would be challenging to incorporate this recommendation into a highly specified quality measure given the breadth of examination elements that are relevant to assess and the difficulty in ascertaining evaluation details from chart or administrative data. In addition, many aspects of the physical examination of a patient with an eating disorder are already subsumed under good clinical practice. Nevertheless, quality related efforts at the local level could assess whether EHR templates

include prompts for documenting relevant physical examination elements and whether such aspects of the evaluation are typically completed, while still allowing flexibility in the documentation of findings.

Statement 7 – Initial Laboratory Assessment

APA recommends (1C) that the laboratory assessment of a patient with a possible eating disorder include a complete blood count and a comprehensive metabolic panel, including electrolytes, liver enzymes, and renal function tests.

Benefits

Laboratory assessment of a patient with a potential eating disorder will help identify laboratory findings that require intervention or influence decision-making as part of treatment planning. For example, electrolyte abnormalities such as hypokalemia may indicate ongoing purging behavior, whereas hypophosphatemia may suggest the onset of refeeding syndrome and indicate the need for supplementation. Obtaining a comprehensive metabolic panel and complete blood count can also establish a baseline for future monitoring of medical stability and treatment response (e.g., normalization of serum electrolytes, resolution of anemia). Additionally, electrolyte abnormalities can be associated with life-threatening complications (e.g., hyponatremia can result in seizures; hypokalemia or hyperkalemia can lead to fatal cardiac arrhythmias).

Harms

Some patients may be fearful of needles or anxious about having bloodwork completed. There are minor risks related to blood drawing such as hematoma at the puncture site.

Patient Preferences

Although there is no specific evidence on patient preferences related to laboratory assessment in individuals with a possible eating disorder, clinical experience suggests that most patients are accepting of this medical intervention.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 7 for additional discussion of the research evidence.) For other laboratory tests such as serum magnesium and phosphorus levels, the potential benefits were viewed as outweighing the potential harms for many patient subgroups. This recommendation is also consistent with Guideline VI, “Assessment of Medical Health,” as part of the APA Practice Guidelines for the Psychiatric Evaluation of Adults (American Psychiatric Association 2016). The level of research evidence is rated as low because there is minimal research on the benefits and harms of conducting a laboratory assessment as part of the psychiatric evaluation as compared with not conducting such assessments. However, expert opinion strongly suggests that such laboratory assessments of patients with possible eating disorders improve the identification of laboratory findings such as electrolyte disturbances or anemia that can help guide treatment planning.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Many other guidelines recommend obtaining laboratory testing as part of assessing individuals with an eating disorder. Some guidelines frame the need for laboratory testing in general terms, such as identifying laboratory abnormalities (National Guideline Alliance [UK] 2020), determining the current level of medical risk (Hay et al. 2014), or reviewing data relevant to nutritional status (Hackert et al. 2020). Among guidelines that make specific recommendations for laboratory testing, a complete blood count, electrolytes, and liver enzymes are most often recommended (ACOG Committee Opinion 2018; French Haute Autorité de Santé 2010; Herpertz et al. 2019; Hornberger et al. 2021; Lock et al. 2015a; Resmark et al. 2019). Other tests that are sometimes recommended include other blood chemistries such as calcium, magnesium, and phosphorus (ACOG Committee Opinion 2018; Hornberger et al. 2021); renal function tests (French Haute Autorité de Santé 2010; Herpertz et al. 2019; Resmark et al. 2019); thyroid stimulating hormone (ACOG Committee Opinion 2018; Herpertz et al. 2019; Hornberger et al. 2021; Resmark et al. 2019); urinalysis (ACOG Committee Opinion 2018; Herpertz et al. 2019; Hornberger et al. 2021; Resmark et al. 2019); amylase and lipase (Herpertz et al. 2019; Resmark et al. 2019); albumin/pre-albumin (French Haute Autorité de Santé 2010); and C-reactive protein (French Haute Autorité de Santé 2010; Herpertz et al. 2019; Resmark et al. 2019). The American College of Obstetricians and Gynecologists (ACOG) guideline also notes that the laboratory assessment should include a urine pregnancy test, serum estradiol, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and prolactin levels in a patient with an eating disorder who presents with oligomenorrhea or amenorrhea (ACOG Committee Opinion 2018).

Quality Measurement Considerations

Administrative data or EHR data could be used to determine whether initial laboratory assessments are occurring in individuals with an eating disorder. Further study would be needed to determine whether there are existing gaps in adherence with this recommendation and whether outcomes would be sufficiently improved to warrant development of a fully specified measure related to laboratory testing. Alternatively, adherence with this recommendation could be assessed on a local level as part of quality improvement initiatives.

Statement 8 – Initial Electrocardiogram

APA recommends (1C) that an electrocardiogram be done in patients with a restrictive eating disorder, patients with severe purging behavior, and patients who are taking medications that are known to prolong QTc intervals.

Benefits

Obtaining an ECG can identify structural and functional cardiac abnormalities which may develop as a complication of low weight, purging behavior, and/or treatment with multiple medications that prolong QTc intervals. Identification of cardiac sequelae resulting from an eating disorder and/or co-occurring medical comorbidities is important in guiding treatment planning to reduce associated symptoms, morbidity, and mortality.

Harms

The potential harms of obtaining an ECG in a patient with a possible eating disorder are minimal and relate primarily to the associated cost of the test.

Patient Preferences

Although there is no specific evidence on patient preferences related to obtaining an ECG in individuals with an eating disorder, clinical experience suggests that most patients are accepting of this intervention as a part of initial assessment and regular monitoring.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 8 for additional discussion of the research evidence.) This recommendation is also consistent with Guideline VI, "Assessment of Medical Health," in the APA Practice Guidelines for the Psychiatric Evaluation of Adults (American Psychiatric Association 2016). The level of research evidence is rated as low because there is minimal research on the benefits and harms of obtaining an ECG as part of the psychiatric evaluation as compared with not obtaining an ECG. However, research does suggest that cardiac conduction abnormalities and other cardiac effects are common in individuals with an eating disorder and expert opinion suggests that obtaining an ECG as part of evaluation can help identify cardiac abnormalities and guide planning of further evaluation and treatment.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other guidelines support obtaining an ECG in individuals with an eating disorder; however, the specific context in which an ECG is recommended varies. The ACOG guideline provide a general recommendation for an ECG (ACOG Committee Opinion 2018), whereas other guidelines recommend an ECG in individuals with AN (French Haute Autorité de Santé 2010; Hay et al. 2014) or under some clinical circumstances (Lock et al. 2015a; National Guideline Alliance (UK) 2020).

Quality Measurement Considerations

Administrative data or EHR data could be used to determine whether an initial ECG is obtained in individuals with an eating disorder. Identifying whether someone meets the precise inclusion criteria for

this recommendation (i.e., AN, patients with severe purging behavior, patients who are taking medications that are known to prolong QTc intervals) may be more challenging from typical administrative or EHR data. Further study would be needed to determine whether there are existing gaps in adherence with this recommendation and whether outcomes would be sufficiently improved to warrant development of a fully specified measure related to obtaining an ECG. Alternatively, adherence with this recommendation could be assessed on a local level as part of quality improvement initiatives.

Statement 9 – Treatment Plan, Including Level of Care

APA recommends (1C) that patients with an eating disorder have a documented, comprehensive, culturally appropriate, and person-centered treatment plan that incorporates medical, psychiatric, psychological, and nutritional expertise, commonly via a coordinated multidisciplinary team.

Benefits

Development and documentation of a comprehensive treatment plan assures that the clinician has considered the available nonpharmacological and pharmacological options for treatment and has identified those treatments that are best suited to the needs of the individual patient, with a goal of improving overall outcomes. It may also assist in forming a therapeutic relationship, eliciting patient preferences, permitting education about possible treatments, setting expectations for treatment, and establishing a framework for shared decision-making. Documentation of a treatment plan promotes accurate communication among all those caring for the patient and can serve as a reminder of prior discussions about treatment.

Harms

The only identifiable harm from this recommendation relates to the time spent in discussion and documentation that may reduce the opportunity to focus on other aspects of the evaluation.

Patient Preferences

Clinical experience suggests that patients are cooperative with and accepting of efforts to establish treatment plans.

Balancing of Benefits and Harms

The potential benefits of this recommendation were viewed as far outweighing the potential harms. (See Appendix C, Statement 9 for additional discussion of the research evidence.) The level of research evidence is low because there is minimal research on the benefits and harms of such an approach. There is also minimal research on whether developing and documenting a specific treatment plan improves outcomes as compared with assessment and documentation as usual.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Many guidelines emphasize the importance of a multidisciplinary approach to treatment (ACOG Committee Opinion 2018; French Haute Autorité de Santé 2010; Hay et al. 2014; Herpertz et al. 2019; Lock et al. 2015a; National Guideline Alliance [UK] 2020; Resmark et al. 2019) with the inclusion of nutrition input from a registered dietician (Hackert et al. 2020; Hornberger et al. 2021; Ozier et al. 2011) and good communication among treatment team members (Hackert et al. 2020; Herpertz et al. 2019; Hornberger et al. 2021; National Guideline Alliance (UK) 2020; Resmark et al. 2019). Other recommended aspects of treatment planning include use of a person-centered (Hay et al. 2014), developmentally aware (Lock et al. 2015a), and culturally-informed (Hackert et al. 2020; Hay et al. 2014) approach to care.

Many guidelines note that most individuals with an eating disorder can be managed in an outpatient setting, particularly as an initial intervention (Catalan Agency for Health Technology Assessment and Research 2009; Couturier et al. 2020; French Haute Autorité de Santé 2010; Golden et al. 2015a; Hay et al. 2014; Herpertz et al. 2019; Lock et al. 2015a; National Guideline Alliance (UK) 2020; Resmark et al. 2019). Other guidelines note that there are multiple factors that can contribute to decisions to hospitalize an individual with an eating disorder, although the specific factors that are mentioned vary (Catalan Agency for Health Technology Assessment and Research 2009; French Haute Autorité de Santé 2010; Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance (UK) 2020; The Royal Colleges of Psychiatrists 2014; Resmark et al. 2019).

Quality Measurement Considerations

It is not known whether psychiatrists and other mental health professionals typically document a comprehensive, culturally appropriate, and person-centered treatment plan that incorporates medical, psychiatric, psychological, and nutritional expertise, but there is likely to be variability in the extent to which this occurs. Although a quality measure could be developed to assess for the implementation of an evidence-based treatment plan that meets consensus-based features of person-centered care, clinical judgment would still be needed to determine whether a documented treatment plan is comprehensive and adapted to individual needs and preferences. Manual review of charts to evaluate for the presence of such a person-centered treatment plan would be burdensome and time-consuming to implement. A quality measure could assess the presence or absence of text in the medical record that would reflect treatment planning; however, when considering the development of such quality measures, there should be a thorough examination of the potential for unintended negative consequences, such as increased documentation burden or overuse of standardized language that meets the quality measure criteria but would inaccurately reflect what occurred in practice. For these reasons, incorporating this recommendation into a highly-specified quality measure is not advised. Nevertheless, EHR note templates could include prompts to foster documentation of a patient-centered treatment plan and local initiatives could engage in quality-related initiatives to improve aspects of treatment planning.

Anorexia Nervosa

Statement 10 – Medical Stabilization, Nutritional Rehabilitation, and Weight Restoration for Patients with Anorexia Nervosa

APA recommends (1C) that patients with anorexia nervosa who require nutritional rehabilitation and weight restoration have individualized goals set for weekly weight gain and target weight.

Benefits

Setting individualized goals for weekly weight gain and target weight in the treatment of anorexia can enhance weight gain during in the treatment course, which in turn improves long-term prognosis (low strength of research evidence). Use of individualized goals allows modifications to be made based on factors such as weight history, co-occurring conditions, likelihood of refeeding syndrome, and treatment setting. If weekly weight gain goals are not being met, adjustments can be made to the patient's estimated caloric intake needs and other revisions to the treatment plan can be implemented if indicated.

Harms

The harms of establishing individualized goals for weekly weight gain and for target weight are unclear (low strength of research evidence). However, for many individuals with AN, the need to gain weight and setting weight-related goals will be associated with anxiety.

Patient Preferences

There is no specific evidence on preferences of individuals with AN related to weekly weight gain goals or target weights. Many individuals will be ambivalent or anxious about weight gain as part of treatment; however, clinical experience suggests that most patients are willing to have their weight assessed as part of treatment if this is approached with sensitivity to their concerns. For example, some individuals may prefer to be told that their goal was met rather than the specific amount of weight gain that occurred.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 10 for additional discussion of the research evidence.) In the expert survey responses, there was significant support for calculating prescribed kcal/day based on initial and target weights and on anticipated and recommended rates of weight gain (Appendix D). The writing group members also noted that long-term outcomes are better when weight gain is monitored and caloric intake is adjusted to meet appropriate weekly goals. Studies also suggest that, with appropriate monitoring, relatively rapid weight gain can occur without a significant risk of complications. Although individuals with anorexia may be ambivalent or anxious about setting weight related goals, the potential harms were viewed as small as compared to the benefits of individualized goal setting and monitoring of weekly weight gain.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Guidelines generally note that weight normalization is a key aspect of AN treatment and that establishing specific weight targets is important (Danish Health Authority 2016a; French Haute Autorité de Santé 2010; Golden et al. 2015a; Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance [UK] 2020; Resmark et al. 2019).

Recommended rates of weekly weight gain vary among guidelines and are typically lower for outpatients than for inpatients (Hilbert et al. 2017). For example, the German guideline recommends weight gains of 0.4 to 1.1 lbs (0.2-0.5 kg) per week for outpatients and 1.1 to 2.2 lbs (0.5-1 kg) per week for inpatients (Herpertz et al. 2019; Resmark et al. 2019). The AAP notes that higher weekly weight gains are being considered of up to 3 to 4.5 lbs (1.4 to 2 kg) per week (Hornberger et al. 2021). The Society for Adolescent Health and Medicine also comments that more aggressive inpatient refeeding protocols can be used for adolescents and young adults with AN as compared to prior recommendations (Golden et al. 2015a). Recommendations for energy intake are also quite variable (Catalan Agency for Health Technology Assessment and Research 2009; Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance [UK] 2020) and caloric intake is typically adjusted on an individual basis to support weekly weight gain targets (Herpertz et al. 2019; Resmark et al. 2019).

Other guidelines are consistent in noting that medications should not be used as sole treatment for AN or for weight gain alone (Catalan Agency for Health Technology Assessment and Research 2009; Herpertz et al. 2019; National Guideline Alliance [UK] 2020; Resmark et al. 2019). However, medications can be used to treat co-occurring conditions (French Haute Autorité de Santé 2010; Lock et al. 2015a) and olanzapine is mentioned in several guidelines as possible to consider in select clinical circumstances (Couturier et al. 2020; Hay et al. 2014).

Other guidelines are also consistent in recommending weight restoration as the best approach to low BMD in individuals with AN (ACOG Committee Opinion 2018; French Haute Autorité de Santé 2010; Hornberger et al. 2021; National Guideline Alliance [UK] 2020). The NICE guideline does describe specific circumstances in which hormonal therapies or bisphosphonates might be considered (National Guideline Alliance [UK] 2020); however, ACOG recommends against the use of combined oral contraceptive pills solely for the treatment of amenorrhea associated with eating disorders (ACOG Committee Opinion 2018).

Quality Measurement Considerations

Weight and BMI are already available in EHRs and are already incorporated into quality measures for other conditions such as obesity. Structured data fields for target weight would be less commonly available and determining this parameter might require more time-consuming manual chart review that would offset the potential value of measurement. Nevertheless, for individuals with AN, it would be possible to develop and test a quality measure to show whether weight restoration is occurring as

treatment proceeds. If such a measure were used for accountability purposes, it should be developed at the level of the organization and not the individual practitioner since most care of individuals with AN is delivered using a multidisciplinary team.

Statement 11 – Psychotherapy in Adults With Anorexia Nervosa

APA recommends (1B) that adults with anorexia nervosa be treated with an eating disorder-focused psychotherapy, which should include normalizing eating and weight control behaviors, restoring weight, and addressing psychological aspects of the disorder (e.g., fear of weight gain, body image disturbance).

Benefits

Use of psychotherapy in the treatment of AN in adults can improve weight related outcomes including change in BMI, change in weight, or %IBW attained (moderate strength of research evidence). CBT, FPT, IPT, MANTRA, and SSCM appear to be associated with modest statistically significant improvements as compared to no treatment, whereas CBT, FPT, some other forms of individual therapy, and family therapy appear to have modest benefits as compared to TAU.

Harms

The harms of psychotherapy in the treatment of AN are not well reported in the literature. However, the harms of using an effective psychotherapy appear to be small. In contrast, use of a psychotherapy that lacks demonstrated benefits in AN could prevent individuals from receiving effective psychotherapy in a timely fashion, thereby influencing prognosis.

Patient Preferences

Clinical experience suggests that most patients are accepting of psychotherapy as part of a treatment plan. However, patients also may have concerns about treatment cost or geographic availability that would influence their choice of psychotherapeutic approaches. In addition, some patients may also prefer one type of psychotherapy over another, based on personal experiences or knowledge about a specific approach. Other patient and clinician factors may affect the therapeutic relationship and may also influence patient preferences.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 11 for additional discussion of the research evidence.) It was recognized that a number of psychotherapies have demonstrated modest efficacy in AN and the harms of these treatments seem small though not well studied. However, there is no single psychotherapy that can be recommended over the other effective psychotherapies in adults with AN. In addition, efficacies overlap among treatments and the effects of treatment vary for different outcomes. Furthermore, patient preferences for specific therapies may differ and additional research evidence may influence our knowledge of effective psychotherapies for this condition. Thus, in balancing of benefits and harms, the

guideline statement focuses on use of an effective eating-disorder focused psychotherapy rather than a specific psychotherapeutic modality.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Guidelines are consistent in recommending psychotherapy for the treatment of individuals with AN (Danish Health Authority 2016a; French Haute Autorité de Santé 2010; Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance [UK] 2020; Resmark et al. 2019) with group formats as well as individual formats mentioned as appropriate (Danish Health Authority 2016a; French Haute Autorité de Santé 2010). CBT-E, MANTRA, and SSCM are most commonly recommended as therapeutic approaches (French Haute Autorité de Santé 2010; Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance [UK] 2020; Resmark et al. 2019) but psychodynamic (French Haute Autorité de Santé 2010; Herpertz et al. 2019; National Guideline Alliance (UK) 2020) and systemic/strategic therapies (French Haute Autorité de Santé 2010) are also mentioned.

Quality Measurement Considerations

This guideline statement may not be appropriate for a performance-based quality measure because of the diversity of effective psychotherapeutic approaches and variations in the availability of psychotherapies. Measurement of psychotherapy utilization using structured EHR data or claims data would require codes for specific types of therapy, but Current Procedural Terminology (CPT) codes refer to psychotherapy in general terms. In addition, patients may be receiving psychotherapies that include a mix of effective elements rather than rigid adherence to a specific psychotherapeutic approach, which would make it hard to specify use of a single modality. For these same reasons, reminders about psychotherapy would be difficult to incorporate into an EHR. In addition, most individuals with AN are receiving some form of psychotherapy and a gap in quality would need to be documented before pursuing additional quality measure development. Nevertheless, individual organizations and health plans may wish to implement programs to assure that effective psychotherapies are being used to treat individuals with AN.

Statement 12 – Family-Based Treatment in Adolescents and Emerging Adults With Anorexia Nervosa

APA recommends (1B) that adolescents and emerging adults with anorexia nervosa who have an involved caregiver be treated with eating disorder-focused family-based treatment, which should include caregiver education aimed at normalizing eating and weight control behaviors and restoring weight.

Benefits

In individuals with AN, use of FBT can improve weight related outcomes including change in BMI or %IBW attained in adolescents as well as in emerging adults, 18 to 26 years of age (moderate strength of research evidence for adolescents; low strength of evidence for emerging adults). Benefits of FBT are apparent when compared to no treatment or TAU.

Harms

The harms of FBT in the treatment of AN are not well reported in the literature but appear to be small. Depending upon family relationships, however, it is possible that greater conflict may occur among family members during treatment.

Patient Preferences

Clinical experience suggests that most patients and family members are accepting of psychotherapy as part of a treatment plan but some patients, particularly emerging adults, may not be willing to receive FBT. Furthermore, in some circumstances, family members or other care partners may not be available or may not wish to participate in FBT. In addition, some patients or families may also prefer one type of psychotherapy over another, based on personal experiences or knowledge about a specific approach. Patients and family may also have concerns about treatment cost or geographic availability that would influence their choice of psychotherapeutic approaches. Other patient, family, and clinician factors may affect the therapeutic relationship and may also influence patient and family preferences.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 12 for additional discussion of the research evidence.) In adolescents, there were statistically significant benefits of FBT on weight-related outcomes. In emerging adults, evidence was limited but FBT was still viewed as beneficial in this age group. For adolescents as well as emerging adults, the harms of FBT seemed small though not well studied.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

For children and adolescents, guidelines consistently recommend psychotherapy, with FBT either recommended or highlighted as having the greatest evidence for efficacy (Couturier et al. 2020; Danish Health Authority 2016a; French Haute Autorité de Santé 2010; Golden et al. 2015a; Hay et al. 2014; Herpertz et al. 2019; Hornberger et al. 2021; Lock et al. 2015a; National Guideline Alliance [UK] 2020; Resmark et al. 2019). Other approaches that are mentioned if FBT is not possible include adolescent focused therapy (Couturier et al. 2020; Lock et al. 2015a), parent-focused therapy (Hornberger et al. 2021), and multi-family therapy (Couturier et al. 2020), and CBT (Couturier et al. 2020).

Quality Measurement Considerations

This guideline statement may not be appropriate for a performance-based quality measure. Measurement of psychotherapy utilization using structured EHR data or claims data would require codes for specific types of therapy but current CPT codes refer to psychotherapy in general terms. Use of a CPT code for family psychotherapy, however, would assure family involvement though not specifying the type of therapy being delivered. For these same reasons, reminders about a FBT would be difficult to incorporate into an EHR. Nevertheless, individual organizations and health plans may wish to implement programs to assure that eating disorder-focused FBT is being used to treat adolescents and emerging adults with AN.

Bulimia Nervosa

Statement 13 – Cognitive-Behavioral Therapy and Serotonin Reuptake Inhibitor Treatment for Adults With Bulimia Nervosa

APA recommends (1C) that adults with bulimia nervosa be treated with eating disorder-focused cognitive-behavioral therapy and that a serotonin reuptake inhibitor (e.g., 60 mgfluoxetine daily) also be prescribed, either initially or if there is minimal or no response to psychotherapy alone by 6 weeks of treatment.

Benefits

Use of CBT in the treatment of BN in adults can promote binge-eating and purging abstinence (moderate strength of research evidence) and can also lead to reductions in the frequencies of binge eating and purging (low strength of research evidence). In comparison with placebo, antidepressants as a group were associated with reductions in binge-eating frequency and a greater rate of binge-eating abstinence in adults (low strength of research evidence). CBT in combination with antidepressant medications also showed efficacy on these outcomes and on depression measures.

Harms

The harms of CBT in the treatment of BN in adults are not well studied but appear to be small. SSRI antidepressant medications have side effects that vary with the specific medication but can include GI effects, headache, insomnia, dry mouth, tremor, weight gain, and sexually related side effects. In individuals with a co-occurring bipolar disorder, use of an SSRI antidepressant may increase the risk of experiencing an episode of mania or hypomania. In clinical trials involving children, adolescents, and young adults up to age 24, SSRIs have been associated with increases in suicidal ideation, hostility, and psychomotor agitation. Clinical trials in individuals with BN also showed a risk of study withdrawal that was higher with antidepressants as compared to no treatment. Despite these potential side effects of SSRIs, evidence suggests that most adults are able to tolerate these medications relatively well, particularly when benefits of treatment are present.

Patient Preferences

Clinical experience suggests that most adults with BN are accepting of treatment with CBT or with antidepressants, although some may prefer one of these treatments over the other based on factors such as prior experiences, treatment availability, or costs.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 13 for additional discussion of the research evidence.) With CBT, the improvements in patient outcomes related to binge eating and purging were significant and the harms appeared small, although not well studied. For this reason, CBT was recommended as an initial treatment. SSRIs alone did not show the same extent of benefits as CBT but combined treatment was associated with benefit. In addition, many individuals with BN will have another condition for which an SSRI may be indicated. For this reason, initial treatment with both CBT and an SSRI was viewed as appropriate for some patients. In addition, some patients may prefer initial treatment with both treatment modalities. Because initial response to treatment is associated with better long-term outcomes, addition of an SSRI at 6 weeks was also viewed as having much greater potential benefits than harms in individuals whose symptoms of BN had not yet responded to CBT alone.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

In terms of psychotherapy for adults with BN, CBT is recommended most often (Catalan Agency for Health Technology Assessment and Research 2009; Danish Health Authority 2016b; Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance (UK) 2020; Resmark et al. 2019). NICE recommends initiating treatment with CBT based guided self-help, with a change to individual CBT after 4 weeks if guided self-help is not feasible or is ineffective (National Guideline Alliance [UK] 2020). The German guidelines also note that IPT and psychodynamic psychotherapy can be reasonable alternatives to CBT (Herpertz et al. 2019; Resmark et al. 2019).

In terms of medication, other guidelines concur in noting that medication not be offered initially as a sole treatment for BN in adults (Danish Health Authority 2016b; Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance [UK] 2020; Resmark et al. 2019); however, fluoxetine (typically in doses of 60 mg daily) is a first-choice treatment when a medication is used to treat BN (Catalan Agency for Health Technology Assessment and Research 2009; Hay et al. 2014; Herpertz et al. 2019).

Quality Measurement Considerations

This guideline statement may not be appropriate for a performance-based quality measure. Measurement of CBT utilization using structured EHR data or claims data would require a method for specifying the type of psychotherapy being delivered but current CPT codes refer to psychotherapy in general terms. For these same reasons, reminders about psychotherapy would be difficult to

incorporate into an EHR. Electronic decision support using passive alerts may be able to prompt clinicians to consider an SSRI in individuals with BN, but such prompts would depend on accurate information about delivered psychotherapy being available in the same EHR system along with measures of outcomes such as binge eating or purging frequencies. Alternatively, individual organizations and health plans may wish to implement programs to assure that effective interventions are being used to treat individuals with BN and that SSRIs be added to CBT if initial treatment response is incomplete.

Statement 14 – Family-Based Treatment in Adolescents and Emerging Adults With Bulimia Nervosa

APA suggests (2C) that adolescents and emerging adults with bulimia nervosa who have an involved caregiver be treated with eating disorder-focused family-based treatment.

Benefits

In adolescents and emerging adults with BN, use of FBT can improve outcomes including binge eating and purging behaviors (low strength of research evidence).

Harms

The harms of FBT in the treatment of BN are not well reported in the literature but appear to be small. Depending upon family relationships, however, it is possible that greater conflict may occur among family members during treatment.

Patient Preferences

Clinical experience suggests that most patients and family members are accepting of psychotherapy as part of a treatment plan. However, patients and family may have concerns about treatment cost or geographic availability that would influence their choice of psychotherapeutic approaches. In addition, some patients or families may also prefer one type of psychotherapy over another, based on personal experiences or knowledge about a specific approach. Other patient, family, and clinician factors may affect the therapeutic relationship and may also influence patient and family preferences.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as likely outweighing the potential harms. (See Appendix C, Statement 14 for additional discussion of the research evidence.) In adolescents and in emerging adults, there were statistically significant benefits of FBT on binge-eating and purging outcomes. Harms of treatment seemed small though not well studied.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

In children and adolescents, FBT and CBT are each mentioned as possible treatment approaches (Couturier et al. 2020; Herpertz et al. 2019; Hornberger et al. 2021; Lock et al. 2015a; National Guideline Alliance (UK) 2020; Resmark et al. 2019). DHA suggests using FBT for moderate and severe BN (Danish Health Authority 2016b). NICE recommends initiating treatment with family therapy with a change to CBT after 4 weeks if family therapy is ineffective (National Guideline Alliance (UK) 2020). The Canadian practice guidelines strongly recommend FBT but note that CBT can be a reasonable alternative (Couturier et al. 2020), whereas the German guidelines recommend CBT as a first-line treatment with FBT as an alternative approach (Herpertz et al. 2019; Resmark et al. 2019). Thus, there is some difference of opinion as to whether CBT or FBT is best as a first-line treatment in children and adolescents with BN.

Quality Measurement Considerations

As a suggestion, this guideline statement is not appropriate for use as a performance-based quality measure or for incorporation into electronic decision support.

Binge-Eating Disorder

Statement 15 – Psychotherapy in Patients With Binge-Eating Disorder

APA recommends (1C) that patients with binge-eating disorder be treated with eating disorder-focused cognitive-behavioral therapy or interpersonal therapy, in either individual or group formats.

Benefits

Use of CBT in the treatment of BED can improve the likelihood of binge-eating abstinence or remission of BED and can also lead to reductions in the frequency of binge eating (low strength of research evidence). Use of IPT can also improve the likelihood of binge-eating abstinence and reduce the frequency of binge eating (low strength of research evidence).

Harms

The harms of treatment with either CBT or IPT in the treatment of BED are not well studied but appear to be small.

Patient Preferences

Clinical experience suggests that most patients are accepting of psychotherapy as part of a treatment plan. However, patients may have concerns about treatment cost or geographic availability that would influence their choice of psychotherapeutic approaches. In addition, some patients may also prefer one type of psychotherapy over another or may prefer individual or group therapy, based on personal experiences or knowledge about a specific approach. Other patient and clinician factors may affect the therapeutic relationship and may also influence patient preferences.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as far outweighing the potential harms. (See Appendix C, Statement 15 for additional discussion of the research evidence.) Although patient preferences may differ in choice of a specific approach to psychotherapy, both CBT and IPT offer therapeutic benefits in BED and the potential for harm appears to be small.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

In adults with BED, guidelines typically recommend treatment with CBT for BED (Catalan Agency for Health Technology Assessment and Research 2009; Hay et al. 2014; Herpertz et al. 2019; Resmark et al. 2019). NICE recommends initiating treatment with CBT-based GSH for BED with a transition to group CBT for BED after 4 weeks if GSH is not effective (National Guideline Alliance [UK] 2020). Alternatively, group CBT for BED could be used if GSH is not feasible or individual CBT could be used if neither group CBT nor GSH is appropriate (National Guideline Alliance (UK) 2020). The Catalan guideline also comments that GSH CBT could be used initially (Catalan Agency for Health Technology Assessment and Research 2009). In addition, the German guidelines recommend IPT as an alternative to CBT in patients with BED (Herpertz et al. 2019; Resmark et al. 2019). For children and adolescents, AAP notes that some evidence supports the use of CBT in patients with BED (Hornberger et al. 2021).

Quality Measurement Considerations

This guideline statement may not be appropriate for a performance-based quality measure. Measurement of psychotherapy utilization using structured EHR data or claims data would require specific codes for CBT and for IPT but current CPT codes refer to psychotherapy in general terms. For these same reasons, reminders about treatment with CBT or IPT would be difficult to incorporate into an EHR. Nevertheless, individual organizations and health plans may wish to implement programs to assure that CBT or IPT are available and are being used to treat individuals with BED.

Statement 16 – Medications in Adults With Binge-Eating Disorder

APA suggests (2C) that adults with binge-eating disorder who prefer medication or have not responded to psychotherapy alone be treated with either an antidepressant medication or lisdexamfetamine.

Benefits

Use of antidepressants in the treatment of BED can increase the likelihood of clinical improvement and enhance remission from BED (low strength of research evidence). Treatment with lisdexamfetamine is also associated with an increased likelihood of clinical improvement and reductions in binge-eating episodes (low strength of research evidence). However, these benefits were modest and, with lisdexamfetamine, the findings may not be generalizable to individuals with BED seen in specialty care.

Harms

Antidepressant medications have side effects that vary with the specific medication but can include GI effects, headache, insomnia, dry mouth, tremor, weight gain, and sexually related side effects. In individuals with a co-occurring bipolar disorder, use of an antidepressant may increase the risk of experiencing an episode of mania or hypomania. In clinical trials involving children, adolescents, and young adults up to age 24, antidepressants have been associated with increases in suicidal ideation, hostility, and psychomotor agitation. Common side effects of lisdexamfetamine include insomnia, reduced appetite, and dry mouth, but increases in heart rate, blood pressure, anxiety, or jitteriness can also occur. Medication misuse and dependence is also possible and individuals with psychotic symptoms or bipolar disorder (or risk factors for these conditions) may experience a worsening of symptoms with stimulant treatment. Despite these potential side effects of antidepressants and lisdexamfetamine, evidence suggests that most individuals are able to tolerate these medications relatively well, particularly when benefits of treatment are present.

Patient Preferences

Clinical experience suggests that most patients are accepting of one of these medications as part of a treatment plan. However, some patients may have concerns about possible side effects or costs of treatment that would influence their choice of pharmacotherapy. Some patients may also prefer one medication over another one on the basis of prior experiences with treatment, whereas other patients may prefer medication over psychotherapy on the basis of factors such as treatment availability, cost, or time constraints.

Balancing of Benefits and Harms

The potential benefits of this statement were viewed as likely outweighing the potential harms. (See Appendix C, Statement 16 for additional discussion of the research evidence.) The benefits of antidepressants and lisdexamfetamine were more modest than benefits of psychotherapy in individuals with BED. In addition, the potential for side effects of treatment and other harms (e.g., misuse of lisdexamfetamine) is greater with pharmacotherapy of BED than with CBT or IPT. Consequently, these treatments are suggested for use if one or both of the psychotherapeutic treatments for BED is ineffective or if a patient prefers medication treatment.

Differences of Opinion Among Writing Group Members

There were no differences of opinion. The writing group voted unanimously in favor of this recommendation.

Review of Available Guidelines from Other Organizations

Other guidelines recommend that medications generally not be used as sole treatment for BED (Hay et al. 2014; Herpertz et al. 2019; National Guideline Alliance (UK) 2020; Resmark et al. 2019), but use of an SSRI (Hay et al. 2014) or other second-generation antidepressant (Herpertz et al. 2019; Resmark et al. 2019) can be considered if psychotherapy is ineffective or unavailable. The Catalan guideline also notes

that an SSRI antidepressant can be offered to a patient with BED although response rates are modest (Catalan Agency for Health Technology Assessment and Research 2009). Similarly, lisdexamfetamine is not generally recommended for use as sole treatment for BED (National Guideline Alliance (UK) 2020), but could be considered if psychotherapy is ineffective or is not desired by the patient (Herpertz et al. 2019; Resmark et al. 2019).

Quality Measurement Considerations

As a suggestion, this guideline statement is not appropriate for use as a performance-based quality measure or for incorporation into electronic decision support.

Appendix H. Evidence Tables for Additional Studies Reviewed

The studies included in this section were reviewed and discussed by the guideline writing group and, where appropriate, incorporated into the network meta-analysis, but did not provide supporting evidence for one of the guideline statements.

Anorexia Nervosa Studies

Psychotherapies

Author (year) (trial name)	Study characteristics, including design, setting, country, and funding	Interventions, including study arm, co-intervention, sample size (N), dose, duration, and follow-up	Main study inclusion and exclusion criteria	Sample demographics including diagnosis, duration, age, gender, and race, and baseline clinical features (e.g., BMI)	Outcome measures, main results, and overall percent attrition
Biney et al. (2021a)	Design: RCT Setting: Inpatient Country: United Kingdom Funding: Profit Organization	Randomized N=40 Practical Body Image therapy (based on CBT) with mirror exposure + TAU 10 wk (N=20) TAU 10 wk (N=20)	Inclusion: AN; 11-18 years of age; current inpatient treatment Exclusion: Previous practical body image therapy; a primary diagnosis other than AN; severe learning difficulty; active psychosis or detainment under the mental health act	AN: 40 (100%) Age 11 yr-18 yr: 40 (100%) Age: 14.2 (SD ± 1.6) Gender, Female: 40 (100%) Race: NR	BMI – Baseline->End of Treatment: 17.91->19.19 kg/m² vs. 17.95->19.33 kg/m² EDE-Q, Weight Concern – Baseline->End of Treatment: 3.93 (SD ± 1.89)->3.17 units (SD ± 2.26) vs. 4.04 (SD ± 1.67)->2.78 units (SD ± 1.78) EDE-Q, Shape Concern – Baseline->End of Treatment: 4.61 (SD ± 1.80)->3.98 units (SD ± 2.10) vs. 4.75 (SD ± 1.40)->3.47 units (SD ± 1.65) Attrition: 18% (5/20) vs. 18% (4/20)
Biney et al. (2021b)	Design: RCT Setting: Inpatient Country: United Kingdom Funding: Profit Organization	Randomized N=50 Self-Esteem Group (based on CBT) + TAU 6 wk (N=25) TAU 6 wk (N=25) Follow-up: Baseline – 10 wk	Inclusion: AN; 11-18 years of age; current inpatient treatment Exclusion: Previous self-esteem group therapy; moderate and severe learning difficulty; aged under 10 years	AN: 50 (100%) Age 11 yr-18 yr: 50 (100%) Age: 15.2 yr (SD ± 1.62) Gender, Female: 50 (100%) Race: NR	BMI – Baseline->End of Treatment: 18.20->19.46 kg/m² vs. 18.26->19.40 kg/m² RSE – Baseline->End of Treatment: 18.80 (SD ± 4.56)->20.05 units (SD ± 5.58) vs. 18.29 (SD ± 4.03)->21.19 units (SD ± 5.01) Attrition: 40% (10/25) vs. 56% (14/25)
del Valle et al. (2010)	Design: RCT Setting: Single Center:	Randomized N=22	Inclusion: Spanish; Caucasian; restrictive AN; age <16 years; BMI >14.0 kg/m^2; undergoing intrahospital psychotherapy in	AN, Restricting: 22 (100%)	BMI – Baseline->12 wk: 18.7->18.2 kg/m² vs. 18.2->18.3 kg/m²

	Children's Hospital Nino Jesus Country: Spain Funding: NR	Resistance Training + Inpatient Psychotherapy 12 wk (N=11) Inpatient Psychotherapy 12 wk (N=11)	the hospital; undergoing dietary counseling in the hospital Exclusion: NR	BMI > 14 kg/m²: 22 (100%) Age < 16 yr: 22 (100%) Age: 14.7 yr (SD ± 0.6) vs. 14.2 yr (SD ± 1.2) Gender - Female: 10 (90.91%) vs. 10 (90.91%) - Male: 1 (9.09%) vs. 1 (9.09%) Race, Caucasian: 22 (100%)	Weight – Baseline->12 wk: 48.2->47 kg vs. 46.6->47.2 kg Weight, Change – Varies: 5.9 kg (SD ± 3.2) vs. 8.4 kg (SD ± 3.4) Adverse Events, Major - Baseline – 12 wk: 0 (0%) vs. 0 (0%) Attrition: 0% (0/22)
Fernandez-del-Valle et al. (2014)	Design: RCT Setting: Single Center: Nino Jesus Hospital Country: Spain Funding: Unknown	Randomized N=44 High-Resistance Training + Psychological Therapy 8 wk (N=22) Psychological Therapy 8 wk (N=22) Follow-up: Baseline – 12 wk Current Analysis (N=36) - 18 vs. 18	Inclusion: Restricting type AN; age <= 16 years; BMI > 14.0 kg/m^2; female; daily life tracing; diet between 1,800 and 2,500 kcal/day Exclusion: Excessive exercisers; contraindications to performing physical activity	AN, Restricting: 44 (100%) BMI > 14 kg/m²: 44 (100%) BMI: kg/m² (SD ± 2.55, N=18) vs. kg/m² (SD ± 2.11, N=18) Age <= 16 yr: 44 (100%) Age: 12.61 yr (SD ± 0.59, N=18) vs. 13 yr (SD ± 0.6, N=18) Gender, Female: 36 (100%) Race: NR	Weight – Baseline->8 wk->12 wk: 43.14->44.55->44.77 kg (N=18) vs. 46.56->48.14->48.89 kg (N=18) BMI – Baseline->8 wk->12 wk: 17.28->17.82->17.61 kg/m² (N=18) vs. 18.12->18.5->18.92 kg/m² (N=18) Adverse Events - Baseline – 12 wk: 0 (0%, N=18) vs. 0 (0%, N=18) Attrition: 18% (4/22) vs. 18% (4/22)
Herpertz-Dahlmann et al. (2014) (ANDI)	Design: RCT Setting: Multi-center Country: Germany Funding: Government	Randomized N=172 Multidisciplinary Day-Patient Treatment > Multidisciplinary Outpatient Treatment 16.5 wk (Mean, SD ± 7) (N=87) Multidisciplinary Inpatient Treatment > Multidisciplinary	Inclusion: Female; 11-18 years of age; diagnosis of AN; BMI below the tenth percentile; first hospital admission for AN Exclusion: Organic brain disease; psychotic disorder; bipolar disorder; substance dependence or abuse; serious	AN: 172 (100%) Hospitalization, AN, First: 172 (100%) AN, Duration: 48 wk (SD ± 36.8)	AN - 12 mo: 17 (23%, N=74) vs. 17 (24%, N=70) Readmission - 12 mo: 13 (15.1%, N=86) vs. 19 (25.3%, N=75) (RD -10.2 %, 95% CI -22.7 – 2.2) Menstruation - 12 mo - Regular: 16 (20%, N=81) vs. 12 (16%, N=75)

		Outpatient Treatment 14.6 wk (Mean, SD ± 6) (N=85) Follow-up: Baseline – 12 mo Modified, ITT (N=161) - 86 vs. 75	self-injurious behavior; intelligence quotient below 85	- 42.4 wk (SD ± 33.1, N=86) vs. 53.7 wk (SD ± 39.6) BMI < 10 percentile: 172 (100%) BMI: 15 kg/m² (SD ± 1.3) Age 11 yr-18 yr: 172 (100%) Age: 15.2 yr (SD ± 1.5) - 15.3 yr (SD ± 1.5) vs. 15.2 yr (SD ± 1.5) Gender, Female: 172 (100%) Race: NR	- Irregular: 11 (14%, N=81) vs. 19 (25%, N=75) BMI - Baseline: 14.9 kg/m² vs. 15.1 kg/m² - 12 mo: 18.1 kg/m² (N=86) vs. 17.8 kg/m² (N=75) (MD 0.46 kg/m², 95% CI -0.11 – 1.02) %EBW – Baseline->12 mo: 74.4%->88% (N=86) vs. 75.4%->86.8% (N=75) Adverse Events, Serious - Baseline – 12 mo: 7 (8.05%) vs. 8 (9.41%) Study Withdrawal - Baseline – 12 mo: 1 (1.1%) vs. 10 (11.8%) Attrition: 29% (25/87) vs. 12% (10/85)
Mountford et al. (2015)	Design: Non-RCT Setting: NR Country: United Kingdom Funding: NR	Total N=90 Body Wise 8 wk (N=50) TAU 8 wk (N=40)	Inclusion: Adults; AN; undergoing inpatient or day-patient treatment Exclusion: NR	AN: 90 (100%) AN, Binge-Eating and Purging:10 (20%) vs. 9 (22.5%) Age >= 18 yr: 90 (100%) Age: 27.5 yr (SD ± 9.16) vs. 25.2 yr (SD ± 9.15) Gender - Female: 88 (97.78%) - Male: 2 (2.22%) Race: NR	BMI - Baseline: 15.65 kg/m² (SD ± 1.7) vs. 15.37 kg/m² (SD ± 1.82) - 16 wk: 17.23 kg/m² (SD ± 1.93) vs. 16.71 kg/m² (SD ± 1.68) Attrition: 22% (11/50) vs. 43% (17/40)
Neumayr et al. (2019)	Design: RCT Setting: Outpatient	Randomized N=40 Therapist Guided Smartphone App Recovery Record + TAU 8 wk (N=20) TAU 8 wk (N=20)	Inclusion: Female; AN; completion of inpatient treatment; age <= 13 years; owner of a smartphone Exclusion: Major depression; suicidal tendency; very high level of care after inpatient treatment	AN: 40 (100%) - Binge-eating and purging type:1 (5%) vs. 5 (25%) - Restricting type: 14 (70%) vs. 15 (75%) - Atypical AN: 5 (25%) vs. 0 (0%)	BMI – Baseline: 19.05 kg/m² (SD ± 1.91) vs. 18.57 kg/m² (SD ± 1.04) BMI, Change - Baseline−End of Treatment: -0.01 (SD ± 0.97, N=19) vs. -0.30 (SD ± 1.42, N=16) (p=0.47) Attrition: 15% (3/20) vs. 20% (4/20)

	Country: Germany Funding: Industry	Follow-up: Baseline – 6 mo		Inpatient Duration: 103.55 d (SD ± 48.03) 107.10 d (SD ± 27.85) Age >= 13 yr: 90 (100%) Age: 20.8 yr (SD ± 6.4) vs. 18.0 yr (SD ± 3.73) Gender, Female: 40 100%) Race: NR	
Peters et al. (2021)	Design: Non-RCT Setting: Inpatient Country: Germany Funding: Industry	Randomized N=304 Interval Treatment (Inpatient Stays) (Duration Varied) (N=20) No Treatment (Duration Varied) (N=20) Follow-up, Mean: 25.1 mo (SD ± 14.0)	Inclusion: Female; AN; at least 21days inpatient stay; 18-55 years of age Exclusion: NR	AN: 304 (100%) - Binge-eating and purging type:115 (37.8%) - Restricting type: 166 (54.6%) - Atypical AN: 23 (7.6%) Inpatient During Study Period: 2.98 (SD ± 1.61) vs. 1.69 (SD ± 1.30) (p<0.001) Inpatient Duration During Study Period: 169 d (SD ± 87.0) vs. 97.0 d (SD ± 90.3) (p<0.001) Age 18 yr-55 yr: 304 (100%) Age: 25.6 yr (SD ± 7.46) vs. 27.5 yr (SD ± 9.60) Gender, Female: 304 100%) Race: NR	BMI – Baseline: 17.9 kg/m² (SD ± 1.56) vs. 17.5 kg/m² (SD ± 2.16) BMI – Follow-Up: 19.2 kg/m² (SD ± 2.66) vs. 18.5 kg/m² (SD ± 2.96) (N=222) Overall Attrition: 26% (79/304)
Ziser et al. (2021)	Design: RCT Setting: Inpatient Country: Germany	Randomized N=22 MANNA (based on motivational interviewing) 10 wk (N=11)	Inclusion: AN; age <= 18 years Exclusion: BMI < 12 kg/m^2; schizophrenia spectrum disorders, bipolar disorder, or substance abuse	AN: 22 (100%) - Binge-eating and purging type:6 (54.5%) vs. 7 (63.6%) - Restricting type: 5 (45.5%) vs. 4 (36.4%) Age: 31.5 yr (SD ± 9.5) vs. 31.9 yr (SD ± 12.6)	BMI – Baseline: 15.6 kg/m² (SD ± 1.3) vs. 15.3 kg/m² (SD ± 1.5) BMI, Change – End of Treatment: 1.79 kg/m² (SD ± 0.9) vs. 1.26 kg/m² (SD ± 0.8)

	Funding: Non-Profit	TAU10 wk (N=11)		Gender, Female: 22 (100%) Race: NR	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; CI=confidence interval; d=day; EBW=expected body weight; EDE-Q=Eating Disorders Examination Questionnaire; ITT=intention-to-treat; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; RD=risk difference; RSE=Rosenberg Self-Esteem scale; SD=standard deviation; TAU=treatment as usual; wk=week; yr=year

Pharmacotherapy

Fluoxetine

Barbarich et al. (2004)	Design: RCT Setting: Multi-center Country: Canada and United States Funding: NR	Randomized N=26 Fluoxetine 20-60 mg 26 wk (N=11) Fluoxetine 20-60 mg + Nutritional Supplements (Tryptophan + Docosahexanoic Acid + Arachadonic Acid) 26 wk (N=15)	Inclusion: AN Exclusion: NR	AN: 26 (100%) AN, Restricting +/- AN, Purging: 6 (23.08%) AN, Duration: 8.4 yr (SD ± 8.1) (N=9) Age: 23 yr (SD ± 6.3) Gender, Unknown: 26 (100%) Race: NR	Weight, Change - Baseline – 26 wk: 0.1 kg/wk (SD ± 0.1, N=2) vs. 0.27 kg/wk (SD ± 0.3, N=7) Study Withdrawal - Baseline – 18 wk - AN, Binge Eating AND Purging subgroup: 2 vs. 4 - AN, Restricting +/- Purging subgroup: 8 vs. 3 Attrition: 82% (9/11) vs. 53% (8/15)
Halmi et al. (2005)	Design: RCT Setting: Multi-center Country: United States Funding: Non-profit	Randomized N=122 CBT + Medical Management 12 mo (N=42) Fluoxetine (maximum of 60 mg) + Medical Management 12 mo (N=41) CBT + Fluoxetine + Medical Management 12 mo (N=39) (All received Medical Management)	Inclusion: AN; within 75% of a target weight; 14-50 years of age Exclusion: NR	AN: 122 (100%) %IBW > 75 %: 122 (100%) Age 14 yr-50 yr: 122 (100%) Age > 18 yr: 109 (89%) Gender, Unknown: 122 (100%) Race: NR	Study Withdrawal - Baseline – 12 mo: 7 (17%) vs. 7 (17%) vs. 7 (18%) Attrition: 57% (24/42) vs. 73% (30/41) vs. 59% (23/39)

Kaye et al. (2001)	Design: RCT Setting: Inpatient: University of Pittsburgh Medical Center Country: United States Funding: Government and industry	Randomized N=39 Fluoxetine 20 mg every other day-60 mg/day 52 wk (N=19) Placebo 52 wk (N=20) Current Analysis (N=35) - 16 vs. 19	Inclusion: Restricting type AN with or without purging behavior; female Exclusion: Concurrent severe medical illness; concurrent neurologic illness; schizophrenic illness; recent alcohol or substance dependence disorder within the last 12 months; psychotropic medication a month before study entry	Percent Average Body Weight: 88% (SD ± 7, N=10) vs. 89% (SD ± 12, N=3) Age: 23 yr (SD ± 9, N=16) vs. 22 yr (SD ± 6, N=19) Gender, Female: 35 (100%) Race: NR	Percent Average Body Weight, Change - Baseline – 52 wk: 5.3% (SD ± 5.3, N=10) vs. 11.2% (SD ± 11.9, N=3) Study Withdrawal, Symptom Worsening - Baseline – 52 wk: 6 (37.5%, N=16) vs. 16 (84.21%, N=19) Attrition: 47% (9/19) vs. 85% (17/20)
Ruggiero et al. (2001)	Design: RCT Setting: Inpatient: Milan University Hospital Country: Italy Funding: NR	Randomized N=35 Fluoxetine + Weight Restoration 3 mo (N=10) Amisulpride + Weight Restoration 3 mo (N=12) Clomipramine + Weight Restoration 3 mo (N=13)	Inclusion: AN, restricting type; severe underweight condition needing urgent weight restoration Exclusion: Younger than 17 years old; clear psychiatric comorbidity; delusional body image related thinking; depression; anxiety; obsessive-compulsive disorder	Amenorrhea: 7 (70%) vs. 11 (91.66%) vs. 11 (84.61%) Underweight, Severe, Requiring Therapy: 35 (100%) Age: 24.5 yr (SD ± 5.06) vs. 24.33 yr (SD ± 5.76) vs. 23.69 yr (SD ± 4.57) Gender, Unknown: 10 (100%) vs. 12 (100%) vs. 13 (100%) Race: NR	Binge eating – Baseline->3 mo: 0 (0%)->4 (40%) vs. 0 (0%)->3 (25%) vs. 0 (0%)->0 (0%) Purging – Baseline->3 mo: 0 (0%)->3 (30%) vs. 0 (0%)->3 (25%) vs. 0 (0%)->0 (0%) Weight - Baseline->3 mo: 40.9->42.75 kg vs. 42.66->38.42 kg vs. 37.62->38.84 kg Weight, % Change - Baseline – 3 mo: 4.52% (SD ± 5.89) vs. 11.04% (SD ± 13.57) vs. 3.26% (SD ± 6.48) BMI - Baseline->3 mo: 15.97->16.7 kg/m² vs. 14.44->16.03 kg/m² vs. 14.69->15.17 kg/m² Attrition: NR
Ruggiero et al. (2003)	Design: Non-RCT Setting: Inpatient: Endocrinology Department of the Istituto Auxologico Italiano	Total N=95 Fluoxetine 30 ± 9.35 mg (Mean) + Nutritional Management 12 mo (N=21) Nutritional Management 12 mo (N=74)	Inclusion: AN Exclusion: Younger than 15 years old	AN: 95 (100%) AN, Duration: 1.3 yr (SD ± 0.3) vs. 3.2 yr (SD ± 0.45) Age: NR Gender - Female: 91 (95.79%) - Male: 4 (4.21%)	Significantly greater BMI increase with fluoxetine was reported compared to nutritional management alone (p<0.0001). BMI - Baseline: 14.83 kg/m² vs. 14.29 kg/m² - 3 mo: 19.06 kg/m² (SD ± 4.3) vs. 15.15 kg/m² (SD ± 2.69) (MD 3.91 kg/m², p<0.0001)

	Country: Italy Funding: NR			Race: NR	- 12 mo: 19.72 kg/m² (SD ± 4.15) vs. 16.52 kg/m² (SD ± 3.27) (MD 3.2 kg/m², p<0.001) Attrition: NR
Walsh et al. (2006)	Design: RCT Setting: Multi-center Country: United States and Canada Funding: government	Randomized N=93 Fluoxetine 20-80 mg/d + CBT 1 yr (N=49) Placebo + CBT 1 yr (N=44)	Inclusion: Female; 16-45 years of age; AN; successfully completed treatment at 1 of the study sites in an inpatient or day-program setting; BMI reached at least 19 kg/m^2 and was maintained for 2 weeks Exclusion: At imminent risk for suicide; serious medical illness aside from AN; medications	AN: 93 (100%) AN, Duration: 56.5 mo (SD ± 44.7) - 4.05 yr (SD ± 3.12, N=43) vs. 4.92 yr (SD ± 4.06, N=38) BMI >= 19 kg/m², Duration >= 2 wk: 93 (100%) Age 16 yr-45 yr: 93 (100%) Age: 23 yr (SD ± 4.6) - 22.4 yr (SD ± 4.46) vs. 24.2 yr (SD ± 4.52) Gender, Female: 93 (100%) Race: NR	BMI – Baseline->52 wk: 20.16->19.08 kg/m² vs. 20.45->18.36 kg/m² BMI >= 18.5 kg/m² - 52 wk: 26.5% (N=21) vs. 31.5% (N=19) (p=0.57) Weight, Change - Baseline – 1 yr: -1.94 kg/mo vs. -2.14 kg/mo (MD 0.2 kg/mo, p=0.75) Disease Response - Baseline – 1 yr - Complete Response: 7 (14.29%) vs. 4 (9.09%) (p=0.32) - Fair or Improved: 32 (65.31%) vs. 25 (56.82%) (p=0.32) Attrition: 57% (28/49) vs. 57% (25/44)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Citalopram

Fassino et al. (2002)	Design: RCT Setting: Outpatient: Centre for Eating Disorders, Turin University Country: Italy Funding: NR	Randomized N=52 Citalopram 10-20 mg/d 12 wk (N=26) WLC 12 wk (N=26)	Inclusion: Restricting-type AN; 16-35 years of age; female Exclusion: Psychiatric comorbidity; under psychopharmacological therapy or estrogen-progesterone therapy during the mo preceding the beginning of the study	Amenorrhea, Duration: 15.807 mo (SD ± 14.827) vs. 20.115 mo (SD ± 25.346) Age 16 yr-35 yr: 52 (100%) Age: 24.346 yr (SD ± 5.381) vs. 25.23 yr (SD ± 8.645) Gender, Female: 52 (100%) Race: NR	Weight - Baseline: 43.48 kg (SD ± 3.93) vs. 42.48 kg (SD ± 4.6) Weight, Change - Baseline – 3 mo: 2.99 kg (N=19) vs. 1.44 kg (N=20) BMI - Baseline: 16.19 kg/m² (SD ± 0.81) vs. 15.62 kg/m² (SD ± 1.42) BMI, Change - Baseline – 3 mo: 1.28 kg/m² (N=19) vs. 0.71 kg/m² (N=20) Attrition: 27% (7/26) vs. 23% (6/26)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; d=day; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Olanzapine

Attia et al. (2011)	Design: RCT Setting: Multi-center, outpatient Country: Canada and United States Funding: Government and industry	Randomized N=23 Olanzapine 2.5-10 mg 8 wk (N=11) Placebo 8 wk (N=12)	Inclusion: >= 16 years of age; AN; BMI 14-19 kg/m^2 Exclusion: Medical or psychiatric problem requiring urgent care; having a clinical symptom or condition inconsistent with the risk profile of olanzapine; diabetes; hyperglycemia; hyperlipidemia; orthostatic hypotension; comorbid schizophrenia, schizophreniform, or bipolar disorder	AN: 23 (100%) BMI 14 kg/m²-19 kg/m²: 23 (100%) BMI: 17.1 kg/m² (SD ± 1.3) Age >= 16 yr: 23 (100%) Age: 27.7 yr (SD ± 9.1) Gender - Female: 10 (90.91%) vs. 12 (100%) - Male: 1 (9.09%) vs. 0 (0%) Race: NR	Greater BMI increase was reported with olanzapine compared with placebo BMI - Baseline: 16.7 kg/m² (SD ± 1.5) vs. 17.4 kg/m² (SD ± 1) - 8 wk: 17.8 kg/m² (SD ± 2.3) vs. 17.6 kg/m² (SD ± 1.3) (MD 0.2 kg/m², p=0.02) Attrition: 27% (3/11) vs. 25% (3/12)
Bissada et al. (2008)	Design: RCT Setting: Single center: The Ottawa Hospital Country: Canada Funding: Industry	Randomized N=34 Olanzapine+ Day Hospital Program (N=16) Placebo + Day Hospital Program (N=18) Treatment: 2 wk – 12 wk Follow-up: Baseline – 13 wk	Inclusion: AN; AN, restricting or binge/purge subtype; women; BMI <=17.5 kg/m^2; attend the day hospital program for eating disorders; free from psychotropic medication for a 2-wk period prior to beginning the study medication Exclusion: Active suicidal intent; comorbid substance abuse disorder; bipolar disorder; schizophrenia; any psychotic disorder; organic brain syndromes; dissociative disorders; pregnancy; failure to use effective contraception if sexually active	AN: 34 (100%) BMI <= 17.5 kg/m²: 34 (100%) BMI: 16.39 kg/m² (SD ± 1.13) vs. 15.93 kg/m² (SD ± 1.39) Age: 23.61 yr (SD ± 6.5) vs. 29.67 yr (SD ± 11.59) Gender, Female: 34 (100%) Race: NR	Olanzapine group had significantly lesser percent of BMI < 18.5 kg/m² at 13 wk (9.4% vs. 30.1%, p=0.02) - 6 wk->12 wk->13 wk: 88.7%->18.8%->9.4% vs. 88.7%->61.3%->30.1% BMI >= 18.5 kg/m² - 13 wk: 14 (87.5%) vs. 10 (55.6%) BMI - 6 wk->13 wk: 18.17 kg/m² (N=15)-> 20.3 kg/m² (N=14) vs. 17.26 kg/m² (N=16)->19.66 kg/m² (N=12) Attrition: 13% (2/16) vs. 22% (4/18)

Brambilla et al. (2007)	Design: RCT Setting: Multi-center, outpatient Country: Italy Funding: NR	Randomized N=35 Current Analysis (N=30) Olanzapine + CBT 3 mo (N=15) - 2.5 mg for 1 mo->5 mg for 2 mo Placebo + CBT 3 mo (N=15)	Inclusion: 18 years of age; AN; female; outpatients Exclusion: General medical impairments; any type of endocrine, metabolic, or immune alterations; cerebral trauma; epilepsy	AN: 35 (100%) AN, Duration: 6.3 yr (SD ± 5) vs. 4.4 yr (SD ± 3) (N=30) Age >= 18 yr: 35 (100%) Age: 23.7 yr (SD ± 4.8) vs. 26.3 yr (SD ± 8.5) Gender, Female: 30 (100%) Race: NR	BMI - Baseline: 15.5 kg/m² (SD ± 1.9) vs. 15.8 kg/m² (SD ± 1.1) - 1 mo: 15.9 kg/m² (SD ± 0.8) vs. 16.2 kg/m² (SD ± 1) - 2 mo: 16.9 kg/m² (SD ± 1.8) vs. 16.5 kg/m² (SD ± 1.3) - 3 mo: 17.2 kg/m² (SD ± 2) vs. 16.9 kg/m² (SD ± 1.2) Overall Attrition: 14% (5/35)
Kafantaris et al. (2011)	Design: RCT Setting: Single Center: Schneider Children's Hospital of the North Shore-Long Island Jewish Health System Country: United States Funding: Industry	Randomized N=20 Olanzapine 10 wk (N=10) - 2.5 mg/d for 1 wk->increased to 10 mg/d by wk 4 Placebo 10 wk (N=10)	Inclusion: Females; 12-21 years of age; AN-restricting type; participating in a comprehensive eating disorders treatment program; underweight Exclusion: Past or current purge type AN; past or current binge type AN; judged to be a serious suicidal risk; prior treatment with olanzapine; were not on a stable medication regimen for 8 weeks prior to study entry	Underweight: 20 (100%) Amenorrhea: 14 (70%) Weight: 94.77 pounds (SD ± 8.66) vs. 92.2 pounds (SD ± 8.11) BMI: 16.4 kg/m² (SD ± 1.2) Age 12 yr-21 yr: 20 (100%) Age: 16.41 yr (SD ± 2.2) vs. 18.1 yr (SD ± 2.04) Gender, Female: 20 (100%) Race - Caucasian: 16 (80%) - Asian: 2 (10%) - Black or African American: 1 (5%) Ethnicity, Hispanic/Latino: 1 (5%)	BMI - Baseline: 16.9 kg/m² (SD ± 0.6) vs. 16 kg/m² (SD ± 1.5) - 5 wk: 17.8 kg/m² (SD ± 1.4, N=7) vs. 17 kg/m² (SD ± 1.7, N=9) - 10 wk: 18.1 kg/m² (SD ± 2, N=7) vs. 17.7 kg/m² (SD ± 1.8, N=8) Hospitalization - Baseline – 10 wk: 3 d (SD ± 5) vs. 5.9 d (SD ± 11.1) Attrition: 30% (3/10) vs. 20% (2/10)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; CBT=cognitive-behavioral therapy; d=day; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Quetiapine

Court et al. (2010)	Design: RCT Setting: Single Center: Orygen Youth Health Research Centre Country: Australia Funding: Industry	Randomized N=33 Quetiapine 100-400 mg/d + TAU 12 wk (N=15) TAU 12 wk (N=18) ITT (N=27) - 13 vs. 14 Follow-up: Baseline – 12 mo	Inclusion: AN Exclusion: Previously received an atypical antipsychotic for more than 1 wk; comorbid psychotic illness; history of brain infarction or surgery; diabetes	AN: 33 (100%) AN, Duration: 65.4 mo (SD ± 96.2, N=10) vs. 30.3 mo (SD ± 37.3, N=11) Weight: 46.4 kg (N=10) vs. 45.9 kg (N=11) Age: 23.8 yr (SD ± 9.4, N=10) vs. 21 yr (SD ± 3.3, N=11) Gender - Female: 14 (93.33%) vs. 18 (100%) - Male: 1 (6.67%) vs. 0 (0%) Race: NR	Weight, Change - Baseline – 12 wk: 5 kg (SD ± 3.5, N=10) vs. 4.5 kg (SD ± 4, N=11) BMI - Baseline: 16.9 kg/m² (N=10) vs. 16.3 kg/m² (N=11) - 12 wk: 18.6 kg/m² (N=9) vs. 18.1 kg/m² (N=9) - 52 wk: 18.9 kg/m² (N=7) vs. 16.7 kg/m² (N=5) Study Withdrawal - Baseline – 12 wk - Adverse Events:1 (7.69%, N=13) vs. NR (N=14) - Lack of Efficacy: 3 (23.08%, N=13) vs. 3 (21.43%, N=14) Attrition: 33% (5/15) vs. 39% (7/18)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; ITT=intention-to-treat; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; TAU=treatment as usual; wk=week; yr=year

Risperidone

Hagman et al. (2011)	Design: RCT Setting: Single Center: Children's Hospital Colorado Country: United States Funding: Industry	Randomized N=40 Risperidone 2.5 mg (Mean) 8.6 wk (Mean) (N=18) Placebo 9.3 wk (N=22)	Inclusion: AN; female; 12-21 years of age Exclusion: NR	AN: 40 (100%) Age 12 yr-21 yr: 40 (100%) Age: 16.2 yr (SD ± 2.5) vs. 15.8 yr (SD ± 2.3) Gender, Female: 40 (100%) Race: NR	BMI - Baseline: 15.9 kg/m² vs. 16.1 kg/m² - 7 wk: 18 kg/m² vs. 18 kg/m² - 15 wk: 18 kg/m² vs. 19 kg/m² - 17 wk: 18 kg/m² vs. 18 kg/m² %IBW - Baseline: 77.4% vs. 79.1% - 7 wk: 88% vs. 89% - 15 wk: 88% vs. 91% - 17 wk: 86% vs. 91% Adverse Events, Significant - Varies: 0 (0%) vs. 0 (0%) Attrition: 11% (2/18) vs. 0% (0/22)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; IBW=ideal body weight; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Cisapride

Szmukler et al. (1995)	Design: RCT Setting: Inpatient: Specialist treatment center for AN Funding: Industry	Randomized N=34 Follow-up (N=29) Cisapride 10 mg + Refeeding Program 8 wk (N=16) Placebo + Refeeding Program 8 wk (N=13)	Inclusion: AN; 18-40 years of age; hospitalized Exclusion: Concurrent illness effecting gastric emptying	AN: 34 (100%) AN, Duration: 39.5 mo (SD ± 45.6) vs. 23.5 mo (SD ± 17.31) Hospitalization: 34 (100%) BN: 8 (27.59%, N=29) Age 18 yr-40 yr: 34 (100%) Age: 21.5 yr (SD ± 3.2) vs. 22.5 yr (SD ± 7.21) Gender, Unknown: 29 (100%) Race: NR	Weight - Baseline: 40.5 kg (SD ± 6.8 vs. 41.6 kg (SD ± 6.49) Weight, Change - Baseline – 8 wk: 5.1 kg (SD ± 2) vs. 5.7 kg (SD ± 2.16) (MD -0.6 kg, p>0.2) Adverse Events - Baseline – 8 wk: 1 (6.25%) vs. 0 (0%) Diarrhea - Baseline – 8 wk: 1 (6.25%) vs. 0 (0%) Overall Attrition: 15% (5/34)

Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Recombinant Human Growth Hormone/Estrogen Replacement

Fazeli et al. (2010)	Design: RCT Setting: Single Center: Massachusetts General Hospital Clinical Research Center Country: United States Funding: Government, industry, and academic	Randomized N=21 Teriparatide +/- Calcium +/- Vitamin D 12 wk (N=10) Placebo +/- Calcium +/- Vitamin D 12 wk (N=11) (All received Calcium +/- Vitamin D)	Inclusion: AN; female; 18-45 years of age; hematocrit greater than 30%; potassium level greater than 3 mMol/L Exclusion: Any condition known to affect bone metabolism; thyroid dysfunction; Cushing's syndrome; diabetes mellitus; renal failure; premature ovarian failure; ingestion of any medication known to affect bone metabolism in the 3 months preceding the study; oral contraceptives in the 3 months preceding the study	AN: 21 (100%) AN, Duration: 3.5 yr (SD ± 4.44) vs. 2 yr (SD ± 2.96) Amenorrhea: 7 (70%) vs. 4 (44.44%, N=9) Age 18 yr-45 yr: 21 (100%) Age: 28 yr (SD ± 6.64) vs. 29.2 yr (SD ± 8.62) Gender, Female: 21 (100%) Race: NR	Weight - Baseline: 46.6 kg (SD ± 3.79) vs. 46.2 kg (SD ± 6.63) Weight, Change - Baseline – 12 wk: 0.3 kg (SD ± 2.67) vs. 0.85 kg (SD ± 1.8) (MD -0.55 kg, p=0.3) %IBW - Baseline: 78.4% (SD ± 6.32vs. 77.7% (SD ± 7.63) %IBW, Change - Baseline – 12 wk: 0.5% (SD ± 4.59) vs. 1.4% (SD ± 2.92) (MD -0.9 %, p=0.3) Attrition: 0% (0/10) vs. 19% (2/11)

Grinspoon et al. (2001)	Design: RCT Setting: Outpatient: Massachusetts General Hospital Country: United States Funding: Government, non-profit, and academic	Randomized N=27 Calcium + Multivitamin Supplement 9 mo (N=13) OCP + Calcium + Multivitamin Supplement 9 mo (N=14)	Inclusion: AN; female; amenorrheic Exclusion: Received estrogen within 6 months of the beginning of the study	AN: 27 (100%) AN: 13 (100%) vs. 14 (100%) Amenorrhea: 13 (100%) vs. 14 (100%) BMI: 16.1 kg/m² (SD ± 1.56) Age: 26.6 yr (SD ± 6.24) Gender, Female: 27 (100%) Race: NR	BMI, Change - Baseline – 9 mo: 1 kg/m² (SD ± 0.95, N=10) vs. 1.7 kg/m² (SD ± 1.9, N=10) (MD -0.7 kg/m², p=0.608) Weight, Increased - Baseline – 9 mo: 20 (74.07%) Attrition: NR
Misra et al. (2013)	Design: Sub-Group Analysis of RCT (Misra et al. (2011)) Setting: Multi-center Country: United States and Canada Funding: Government	Randomized N=72 17B-Estradiol 100 mcg + Progesterone 2.5 mg 10d/mt 18 mo (N=38) Placebo 18 mo (N=34) Follow-up (N=37) - 20 vs. 17	Inclusion: AN; 13-18 years of age; female Exclusion: Active suicidality; psychosis; substance abuse; hematocrit <30 %; potassium<3.0 mMol/L; glucose <50 mg/dl; use of prescription medications within 3 months of study participation known to affect bone metabolism; other diseases known to affect bone metabolism	AN: 72 (100%) Age 13 yr-18 yr: 72 (100%) Age: 16.9 yr (SD ± 1.23) vs. 16.6 yr (SD ± 1.17) Gender, Female: 72 (100%) Race: NR	BMI - Baseline: 17.2 kg/m² vs. 17.5 kg/m² BMI, Change - Baseline – 18 mo: 1.36 kg/m² (N=20) vs. 1.19 kg/m² (N=17) (MD 0.17 kg/m², p=0.79) Weight, Change - Baseline – 18 mo: 3.8 kg vs. 3.3 kg (MD 0.5 kg, p=0.73) Testosterone, Change - Baseline – 18 mo: -15.2 ng/dL (N=20) vs. -16.6 ng/dL (N=17) (MD 1.4 ng/dL, p=0.93) Attrition: 47% (18/38) vs. 50% (17/34)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; OCP=oral contraceptive pill; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Refeeding

High Calorie Intake

Garber et al. (2013)	Design: Prospective Cohort Study Setting: University of California San Francisco (UCSF)	Total N=56 Higher-Calorie Intake 14d (N=28)	Inclusion: 9-20 years of age; AN Exclusion: Previous admissions for AN; pregnancy; BN; thought	AN: 56 (100%) BMI: 16.1 kg/m² (SD ± 2.24) - 16.6 kg/m² (SD ± 2.12) vs. 15.8 kg/m² (SD ± 2.65)	Greater changes in weight and median percent BMI were reported with higher calorie intake. Weight, Change - Baseline – 3 d: 250 g vs. -270 g (MD 520 g, p=0.001)

	Benioff Children's Hospital Country: United States Funding: NR	Lower-Calorie Intake 14d (N=28) Follow-up: 14.9 d (Mean, 14.9 d)	disorders; schizophrenia; other psychosis	BMI, Median Percent: 79.2% (SD ± 11.22) Age 9 yr-20 yr: 56 (100%) Age: 16.2 yr (SD ± 2.24) - 16.1 yr (SD ± 2.12) vs. 16.2 yr (SD ± 2.12) Gender - Female: 55 (98%) - Male: 1 (2%) Ethnicity, Other: 11 (19%)	BMI, Median Percent - Baseline->4 d: 81.9->83.2% vs. 77.6->81.9% (MD 1.3 %, p=0.006 at 4 d) BMI, Median Percent, Change - Varies: 0.46 %/d vs. 0.26 %/d (MD 0.2 %/d, p<0.001) Blood Phosphorous Decreased - Baseline – 14 d: 12 (42.86%) vs. 8 (28.57%) (p=0.273) Hospitalization, Duration - Baseline – 14 d: 11.9 d (SD ± 5.29) vs. 17.6 d (SD ± 6.35) Attrition: 0% (0/56)
Golden et al. (2013)	Design: Retrospective Cohort Study Setting: Inpatient: Lucile Packard Children's Hospital Country: United States Funding: NR	Total N=310 Higher-Calorie Intake 13 d (Mean, SD ± 7.3) (N=222) Lower-Calorie Intake 16.6 d (Mean, SD ± 9) (N=88)	Inclusion: Adolescents; AN Exclusion: BN; EDNOS	AN: 310 (100%) AN, Duration: 1 yr (SD ± 0.9) vs. 1.39 yr (SD ± 1.3) BMI, Median Percent < 70 %: 31 (13.96%) vs. 18 (20.45%) Weight: 42.9 kg (SD ± 7.5) vs. 41.8 kg (SD ± 6.5) Adolescent: 310 (100%) Age: 16.1 yr (SD ± 2.3) vs. 16.2 yr (SD ± 2.4) Gender, Unknown: 222 (100%) vs. 88 (100%) Race, Caucasian: 257 (82.9%)	Significantly reduced length of hospital stay was reported with higher calorie intake: 13 d vs. 16.6 d (MD -3.6 d, p<0.0001). Weight, Change: 2.9 kg vs. 3.6 kg (MD -0.7 kg, p=0.01) BMI – Baseline->Discharge: 16.1->17.1 kg/m² vs. 15.9->17.2 kg/m² (MD -0.1 kg/m², p=0.63) BMI, Median Percent – Baseline->Discharge: 78.7->83.7% vs. 77.9->84.3% (MD -0.6 %, p=0.54 at discharge) Hypomagnesemia - Varies: 34 vs. 13 (p=1) Refeeding Syndrome – Varies: 0 (0%) vs. 0 (0%) Attrition: NR
Imbierowicz et al. (2002)	Design: Retrospective Cohort Study Setting: Inpatient: Bonn University	Total N=84 Bonn University (N=42) - High-Caloric Supplement 10.7	Inclusion: AN Exclusion: NR	AN: 84 (100%) BMI: 14.5 kg/m² (SD ± 1.3) vs. 14.6 kg/m² (SD ± 1.4, N=29)	Significantly greater weight gain was reported with high-caloric supplement. Weight, Change – Varies: 0.5 kg/wk (SD ± 0.5) vs. 0.3 kg/wk (SD ± 0.3, N=29) (MD 0.2 kg/wk, p=0.02)

	Hospital; Klinik am Korso Country: Germany Funding: NR	wk (Mean, SD ± 4.8) (N=29) Klinik am Korso (N=42) - No High-Caloric Supplement 12.3 wk (Mean, SD ± 1.9) (N=42) Subgroups: - BMI < 14 kg/m² (N=11 vs. 11) - BMI >= 14 kg/m² (N=18 vs. 18) - Anorexia, Binge-Eating and Purging (N=14 vs. 14) - Anorexia, Restricting (N=13 vs. 13)		BMI - Anorexia, Binge-Eating and Purging subgroup: 15.1 kg/m² (SD ± 1) vs. 15.1 kg/m² (SD ± 1.1) - Anorexia, Restricting subgroup: 14.1 kg/m² (SD ± 1.5) vs. 14.2 kg/m² (SD ± 1.5) - BMI < 14 kg/m² subgroup: 13.2 kg/m² (SD ± 0.5) vs. 13.4 kg/m² - BMI >= 14 kg/m² subgroup: 15.3 kg/m² (SD ± 1) vs. 15.5 kg/m² (SD ± 1) Gender, Unknown: 71 (100%) Race: NR	- BMI < 14 kg/m² subgroup: 0.5 kg/wk vs. 0.3 kg/wk (MD 0.2 kg/wk, p=0.07) - BMI >= 14 kg/m² subgroup: 0.6 kg/wk vs. 0.2 kg/wk (MD 0.4 kg/wk, p=0.004) - Anorexia, Binge-Eating and Purging subgroup: 0.5 kg/wk vs. 0.3 kg/wk (MD 0.2 kg/wk, p=0.03) - Anorexia, Restricting subgroup: 0.4 kg/wk vs. 0.2 kg/wk (MD 0.2 kg/wk, p=0.02) BMI - Discharge: 17 kg/m² vs. 15.7 kg/m² (N=29) (MD 1.3 kg/m², p=0.004) - 2.5 yr: 17.4 kg/m² (N=18) vs. 18.1 kg/m² (N=18) (MD -0.7 kg/m², p=0.28) Attrition: NR
O'Connor et al. (2016)	Design: RCT Setting: Inpatient Country: United Kingdom Funding: NR	Randomized N=36 Higher-Calorie Intake 10d (N=18) Low-Calorie Intake 10d (N=18)	Inclusion: Adolescents; AN; <78% median BMI; on a weight-losing trajectory; hospitalization; low weight at hospital admission; 10-16 years of age Exclusion: Taking atypical antipsychotic or antidepressant medication; type 1 diabetes mellitus; malabsorption disorders	AN: 36 (100%) Hospitalization: 36 (100%) BMI, Median Percent: 70.8% (SD ± 5.9) vs. 69% (SD ± 4.3) Adolescent: 36 (100%) Age 10 yr-16 yr: 36 (100%) Age: 13.7 yr (SD ± 1.8) vs.14.1 yr (SD ± 1.8) Gender - Female: 17 (94.44%) vs. 17 (94.44%) - Male: 1 (5.56%) vs. 1 (5.56%)	Weight – Baseline: 32.9 kg (SD ± 7) vs. 34.6 kg (SD ± 5) Weight, Change - Baseline – 4 d: 0.3 kg (SD ± 0.73) vs. -0.2 kg (SD ± 0.8) (MD 0.4 kg, 95% CI -0.1 – 1) - Baseline – 10 d: 1.1 kg (SD ± 1.09) vs. 0.64 kg (SD ± 0.69) (MD 0.47 kg, 95% CI -0.2 – 1.1) BMI - Baseline: 13.6 kg/m² (SD ± 1.3) vs. 13.5 kg/m² (SD ± 1) BMI, Change - Baseline – 4 d: 0.13 kg/m² (SD ± 0.32) vs. -0.02 kg/m² (SD ± 0.34) (MD 0.15 kg/m², 95% CI -0.06 – 0.38) - Baseline – 10 d: 0.5 kg/m² (SD ± 0.4) vs. 0.3 kg/m² (SD ± 0.3) (MD 0.2 kg/m², 95% CI 0 – 0.5)

				Race: NR	Overall Attrition: 0% (0/36)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; d=day; EDNOS=eating disorder not otherwise specified; MD=mean difference; NR=not reported; SD=standard deviation; wk=week; yr=year

Tube Feeding

Agostino et al. (2013)	Design: Retrospective Cohort Study Setting: Inpatient: Montreal Children's Hospital Country: Canada Funding: Academic	Total N=165 cNG Tube 10 d (N=31) No cNG Tube 55.9 d (Mean) (N=134)	Inclusion: 10-18 years of age; met criteria for AN or a restrictive form of EDNOS Exclusion: BN; admitted for a reason other than nutritional rehabilitation for their eating disorder; admitted for depression or suicide; previously admitted for eating disorder treatment prior to the study period	Anorexia Restricting or EDNOS: 165 (100%) %IBW: 82% (SD ± 10) vs. 85% (SD ± 13) BMI: 16.6 kg/m² (SD ± 2.2) vs. 16.7 kg/m² (SD ± 2.3) Age 10 yr-18 yr: 165 (100%) Age: 14.9 yr (SD ± 2.1) vs. 14.9 yr (SD ± 1.7) Gender - Female: 29 (94%) vs. 129 (96%) - Male: 2 (6%) vs. 5 (4%) Race: NR	Significantly reduced length of hospital stay was reported with cNG tube: 33.8 d vs. 50.9 d (MD -17.1 d, p=0.0002) Significantly greater weight gain was reported with cNG tube - Baseline – 7 d: 1.22 kg/wk vs. 0.08 kg/wk (MD 1.14 kg/wk, p=0.0001) - Baseline – 14 d: 1.06 kg/wk) vs. 0.69 kg/wk (MD 0.37 kg/wk, p=0.004) Rehospitalizations – Baseline – 6 mo: 4 (12.9%) vs. 31 (23%) (p=0.32) Constipation – Baseline – 2 wk: 3 (9.6%) vs. 5 (3.7%) (p=0.17) Nausea – Baseline – 2 wk: 1 (3.2%) vs. 4 (2.9%) (p=1) Refeeding Syndrome – Baseline – 2 wk: 0 (0%) vs. 0 (0%) Overall Attrition: 0% (0/165)
Rigaud et al. (2007a)	Design: RCT Setting: Inpatient: Nutrition Unit Country: France Funding: NR	Randomized N=81 Tube Feeding (Cyclic Enteral Nutrition) + Multidisciplinary Therapy 2 mo (N=41) Multidisciplinary Therapy 2 mo (N=40)	Inclusion: AN; malnourished; adult Exclusion: BMI lower than 11 kg/m^2	AN: 81 (100%) Malnourished: 81 (100%) AN, Duration: 4.5 yr vs. 3.2 yr Amenorrhea: 41 (100%) vs. 39 (97.5%)	Significantly greater fat-free mass and weight gain was reported with tube feeding at 2 mo: - Fat-free mass: 109 g/d vs. 61 g/d (MD 48 g/d, p<0.01) - Weight, Change: 194 g/d vs. 126 g/d (MD 68 g/d, p<0.01) At discharge, tube feeding group had more subjects with BMI >= 18.5 kg/m²: 16 (39%) vs. 3 (8%) (p<0.02)

		Anorexia, Binge-Eating and Purging subgroup (N=12 vs. 13) Anorexia, Restricting subgroup (N=29 vs. 27) Follow-up: Baseline – 1 yr		Age >= 18 yr: 81 (100%) Age: 22.5 yr (SD ± 4.5) vs. 24.2 yr (SD ± 3.8) Gender - Female: 40 (97%) vs. 39 (97.5%) - Male: 1 (3%) vs. 1 (2.5%) Race: NR	Among binge-eating and purging type, decrease in vomiting and binge-eating episodes was reported with tube feeding - Baseline: 13/ wk vs. 10/wk - 2 mo: 1.35/wk vs. 5.31/wk (MD -3.96/ wk, p<0.01) Among binge-eating and purging type, more remission was reported with tube feeding at 1 wk: 8 (80%, N=10) vs. 4 (50%, N=8) (p<0.01) Treatment-Related – Baseline – 2 mo - Sinusitis: 2 (4.88%) vs. 0 (0%) - Epistaxis: 1 (2.44%) vs. 0 (0%) Constipation, Requiring Laxative – Baseline – 2 mo: 2 (4.88%) vs. 3 (7.5%) Attrition: 2% (1/41) vs. 0% (0/40)
Robb et al. (2002)	Design: Retrospective Cohort Study Setting: NR Country: United States Funding: NR	Total N=100 Nocturnal NG Refeeding + Oral Refeeding + Multidisciplinary Inpatient Therapy (N=52) Oral Refeeding + Multidisciplinary Inpatient Therapy (N=48)	Inclusion: AN; adolescent; female; hospitalization; Caucasian Exclusion: BN	AN: 100 (100%) Hospitalization: 100 (100%) Amenorrhea - Primary: 21 (40.4%) vs. 9 (18.8%) - Secondary: 31 (59.6%) vs. 39 (81.3%) Weight: 41.1 kg (SD ± 4.7) vs. 42.5 kg (SD ± 7.6) Weight, Maximum: 51.7 kg (SD ± 8) vs. 53.7 kg (SD ± 11.9) BMI: 15.5 kg/m² (SD ± 1.7) vs. 16 kg/m² (SD ± 1.8) Adolescent: 100 (100%)	Greater weight gain and BMI change was reported with supplemental NG refeeding: - Weight, Change – Varies: 5.4 kg vs. 2.4 kg (MD 3 kg, p<0.05) - BMI, Change – Varies: 2.03 kg/m² vs. 0.9 kg/m² (MD 1.13 kg/m², p<0.05) Hospitalization, Duration – Baseline – discharge: 22.3 d (SD ± 13.5) vs. 22.1 d (SD ± 9.4) (MD 0.2 d, p=1) Of NG refeeding group: - Epistaxis: 6 (11.5%) Nasal Irritation: 15 (28.8%) No Refeeding Syndrome or Aspiration Pneumonia Attrition: 0% (0/100)

				Age: 14.8 yr (SD ± 1.9) vs. 15 yr (SD ± 1.8) Gender, Female: 100 (100%) Race, Caucasian: 100 (100%)	

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; cNG=continuous nasogastric; d=day; EDNOS=eating disorder not otherwise specified; IBW=ideal body weight; MD=mean difference; mo=month; NG=nasogastric; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Bone Density

Dehydroepiandrosterone

Bloch et al. (2012)	Design: RCT Setting: Single Center: Rambam Medical Center Eating Disorder Unit Country: Israel Funding: Non-profit	Randomized N=26 DHEA 100mg + Calcium 600 mg + Vitamin D3 200 IU 6 mo (N=15) Placebo + Calcium 600 mg + Vitamin D3 200 IU 6 mo (N=11) Current Analysis (N=21) - 13 vs. 8	Inclusion: Premenopausal; female; AN Exclusion: Any serious acute medical conditions; any chronic medical conditions; taking psychotropic medications; taking oral contraceptives	AN: 26 (100%) AN, Duration: 10.5 yr (SD ± 4.4) BMD, Total Body: 1.03 g/cm² (SD ± 0.81) vs. 1.06 g/cm² (SD ± 0.36) BMD, Z-Score - Total Body: 0.27 units (SD ± 0.81) - Femoral Neck: 0.17 units (SD ± 1.07) - Lumbar Spine:.04 units (SD ± 1.25) Weight: 45.5 kg (SD ± 4.8) BMI: 17.7 kg/m² Age: 26.9 yr (SD ± 8.2) - 26.6 yr (SD ± 8.9) vs. 27.4 yr (SD ± 7.4) Gender, Female: 26 (100%) Race: NR	BMD, % Change- Baseline – 6 mo - Total Body: 0% vs. 0% (MD 0%, p=0.6) - Femur: 0% vs. 0% (MD 0%, p=0.96) - Neck: 0% vs. 0% (MD 0%, p=0.99) - Lumbar Spine: 0% vs. 1% (MD -1%, p=0.45) Body Fat, Total - Baseline – 6 mo: 42% (SD ± 43) vs. 6% (SD ± 19) (MD 36 %, p=0.52) Weight – Baseline->3 mo: 46.2->48.3 kg vs. 44.7->46 kg BMI – Baseline->3 mo->6 mo: 17.75->18.65->18.94 kg/m² vs. 17.76->17.75->18.25 kg/m² Adverse Events, Treatment-Related - Baseline – 6 mo: 0% vs. 0% Attrition: 13% (2/15) vs. 27% (3/11)
Divasta et al. (2012, 2014)	Design: RCT	Randomized N=94	Inclusion: Females; AN; 13-27 years of age; amenorrhea; fear of	AN: 94 (100%)	BMD – Baseline->18 mo - Total Body: 1.07->1.078 g/cm² vs. 1.05->1.042 g/cm²

	Setting: Outpatient: Children's Hospital Boston Country: United States Funding: Government	DHEA 50 mg + (Conjugated Equine Estrogens 0.3 mg 3 mo > [Ethinyl Estradiol 20 µg + Levonorgestrel 0.1 mg]) 18 mo (N=47) Placebo 18 mo (N=47) Current Analysis (N=80) - 43 vs. 37	weight gain; malnutrition; body weight <= 85% median body weight for age and sex Exclusion: Celiac disease; diabetes; glucocorticoids	AN, Duration: 12 mo (SD ± 25.19, N=43) vs. 9 mo (SD ± 10.37, N=37) Amenorrhea: 94 (100%) Amenorrhea, Duration: 11 mo (SD ± 11.11, N=80) BMI: 18 kg/m² (SD ± 1.5N=80) - 18.1 kg/m² (SD ± 1.5, N=43) vs. 17.8 kg/m² (SD ± 1.5, N=37) Weight, Median Percent <= 85%: 94 (100%) Age 13 yr-27 yr: 94 (100%) Age: 18.1 yr (SD ± 2.7, N=80) - 18 yr (SD ± 2.5) vs. 18.3 yr (SD ± 2.8) Gender, Female: 80 (100%) Race, Caucasian: 39 (91%) vs. 32 (86%) Ethnicity, Hispanic/Latino: 1 (2%) vs. 1 (3%)	- Total Body, Z-Score: 0.19->0.09 units vs. -0.06->-0.39 units - Hip: 0.89->0.908 g/cm² vs. 0.89->0.882 g/cm² - Hip, Z-Score: -0.37->-0.34 units vs. -0.35->-0.44 units - Lumbar Spine: 0.89->0.919 g/cm² vs. 0.88->0.87 g/cm² - Lumbar Spine, Z-Score: -0.84->-0.86 units vs. -0.98->-0.99 units - Femoral Shaft, Mean Percent, Change: 0 %/yr (SD ± 2.78, N=31) vs. -1.1 %/yr (SD ± 2.69, N=29) (MD 1.1 %/yr, p=0.12) Weight - Baseline: 49.1 kg (SD ± 5.9, N=43) vs. 48 kg (SD ± 5.6, N=37) - Change: 5.9 kg/yr (SD ± 6.56) vs. 5.2 kg/yr (SD ± 6.69) (MD 0.7 kg/yr, p=0.52) Menstruation, Resumed: NR vs. 22 (76%, N=29) Study Withdrawal, All-Cause: 12 (27.91%, N=43) vs. 8 (21.62%, N=37) Attrition: 34% (16/47) vs. 38% (18/47)
Gordon et al. (2002)	Design: RCT Setting: Single center, outpatient: Children's Hospital Boston; Suburban Adolescent Medicine Practice Country: United States Funding: government	Randomized N=61 DHEA 50 mg + Psychotherapy 12 mo (N=31) Conventional hormonal replacement therapy (Ethinyl Estradiol 20 µg+ Levonorgestrel 0.1 mg) + Psychotherapy 12 mo (N=30)	Inclusion: AN; women; 14-28 years of age; post-menarchal Exclusion: Medications known to affect BMD	AN: 61 (100%) Age 14 yr-28 yr: 61 (100%) Age: 17.8 yr (SD ± 2.9) Gender, Female: 61 (100%) Race: NR	Greater weight change was reported with DHEA at 12 mo: 6.8 kg vs. 5.9 kg (MD 0.9 kg, p<0.001) BMD - Hip - Baseline: 0.86 g/cm² (SD ± 0.11) vs. 0.87 g/cm² (SD ± 0.11) - Hip, Change - Baseline – 12 mo: 0.0168 g/cm² (SD ± 0.04) vs. 0.0179 g/cm² (SD ± 0.04) - Lumbar Spine – Baseline: 0.889 g/cm² (SD ± 0.11) vs. 0.886 g/cm² (SD ± 0.08)

					- Lumbar Spine, Change - Baseline – 12 mo: 0.0045 g/cm² (SD ± 0.05) vs. 0.0095 g/cm² (SD ± 0.05) - Total Body, % Change: -0.3% vs. 0.6% Menstruation - Resumed: 18 (58%) vs. NR - Light, Irregular: 1 (3.23%) vs. NR Attrition: 13% (4/31) vs. 20% (6/30)

Abbreviations: AN=anorexia nervosa; BMD=bone mineral density; BMI=body mass index; DHEA=dehydroepiandrosterone; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Estrogen Replacement

Faje et al. (2012)	Design: Sub-Group Analysis of RCT (Misra et al. (2011)) Setting: Multi-center Country: United States; Canada Funding: government	Randomized N=22 Estradiol 100 mcg + Medroxyprogesterone 2.5 mg 10d/mt + Calcium Carbonate 1200 mg + Vitamin D 400 IU 12 mo (N=13) Placebo + Calcium Carbonate 1200 mg + Vitamin D 400 IU 12 mo (N=9)	Inclusion: Girls; 13-18 years of age Exclusion: Diseases affecting bone metabolism; diseases affecting suicidality or psychosis; history of substance abuse; medications affecting psychosis or suicidality; medications affecting bone metabolism	AN:13 (100%) vs. 9 (100%) Amenorrhea, Duration: 0.9 yr (SD ± 0.69) vs. 0.84 yr (SD ± 0.39) BMD - Lumbar Spine, Z-Score < -0.5 units - 13 (100%) vs. 9 (100%) - Lumbar Spine, Z-Score: -1.33 units vs. -1.53 units - Lumbar Spine: 0.847 g/cm² vs. 0.821 g/cm² BMI: 17.4 kg/m² (SD ± 1.44) vs. 16.7 kg/m² (SD ± 1.2) Weight: 47.5 kg (SD ± 4.69) vs. 45.2 kg (SD ± 6.9) Age 13 yr-18 yr: 22 (100%) Age: 17.2 yr (SD ± 1.08) vs. 16.8 yr (SD ± 1.2) Gender, Female: 22 (100%) Race: NR	Significant improvement was reported with transdermal estradiol on lumbar bone density. BMD - Baseline – 12 mo - Lumbar Spine, Change: 0.035 g/cm² vs. 0.005 g/cm² (MD 0.03 g/cm², p=0.02) - Lumbar Spine, % Change: 4.38% vs. -0.46% (MD 4.84 %, p=0.02) - Lumbar Spine, Z-Score, Change: 0.21 units vs. -0.19 units (MD 0.4 units, p=0.01) Attrition: NR

Gordon et al. (2002)	Design: RCT Setting: Single center, outpatient: Children's Hospital Boston; Suburban Adolescent Medicine Practice Country: United States Funding: government	Randomized N=61 DHEA 50 mg + Psychotherapy 12 mo (N=31) Conventional hormonal replacement therapy (Ethinyl Estradiol 20 µg+ Levonorgestrel 0.1 mg) + Psychotherapy 12 mo (N=30)	Inclusion: AN; women; 14-28 years of age; post-menarchal Exclusion: Medications known to affect BMD	AN: 61 (100%) Age 14 yr-28 yr: 61 (100%) Age: 17.8 yr (SD ± 2.9) Gender, Female: 61 (100%) Race: NR	Greater weight change was reported with DHEA at 12 mo: 6.8 kg vs. 5.9 kg (MD 0.9 kg, p<0.001) BMD - Hip - Baseline: 0.86 g/cm² (SD ± 0.11) vs. 0.87 g/cm² (SD ± 0.11) - Hip, Change - Baseline – 12 mo: 0.0168 g/cm² (SD ± 0.04) vs. 0.0179 g/cm² (SD ± 0.04) - Lumbar Spine – Baseline: 0.889 g/cm² (SD ± 0.11) vs. 0.886 g/cm² (SD ± 0.08) - Lumbar Spine, Change - Baseline – 12 mo: 0.0045 g/cm² (SD ± 0.05) vs. 0.0095 g/cm² (SD ± 0.05) - Total Body, % Change: -0.3% vs. 0.6% Menstruation - Resumed: 18 (58%) vs. NR - Light, Irregular: 1 (3.23%) vs. NR Attrition: 13% (4/31) vs. 20% (6/30)
Golden et al. (2002)	Design: Prospective Cohort Study Setting: Single Center: Eating Disorders Center of Schneider Children's Hospital of Long Island Jewish Medical Center Country: United States Funding: NR	Total N=50 Standard Treatment (Nutritional Intervention) + Calcium + Psychological Therapy NR (N=28) Estrogen-Progestin + Standard Treatment + Calcium + Psychological Therapy NR (N=22) - Estrogen + Progestin Given After 1 yr Standard Treatment (N=4) Follow-up: Baseline – 36 mo	Inclusion: Females; 13-21 years of age; AN; primary amenorrhea or secondary amenorrhea of greater than 6 months duration Exclusion: Receiving hormonal therapy; coexistent medical condition that could contribute to the osteopenia; medical condition that precluded the administration of estrogen or progestin; receiving steroids; receiving injectable contraception; receiving oral contraception	AN: 50 (100%) AN, Duration: 21.9 mo (SD ± 20.6) - 15.7 mo (SD ± 7.1) vs. 29.8 mo (SD ± 28.4) Amenorrhea, Primary or Amenorrhea, Secondary >= 6 mo: 50 (100%) - Primary: 6 (12%) - Secondary: 44 (88%) Amenorrhea, Secondary, Duration: 16 mo (SD ± 8.8) (N=44) %IBW: 79.1% (SD ± 7) vs. 79.9% (SD ± 8.4) Weight: 43.9 kg (SD ± 4.7) Age 13 yr-21 yr: 50 (100%)	BMD Femoral Neck - 1 yr - 0.713->0.723 g/cm² vs. 0.7->0.694 g/cm² - 0.017->0.015 g/cm²/kg vs. 0.016->0.014 g/cm²/kg - 0.515->0.45 g/cm²/m vs. 0.513->0.426 g/cm²/m Lumbar Spine - 1 yr - 0.825->0.819 g/cm² vs. 0.834->0.833 g/cm² - 0.019->0.017 g/cm²/kg vs. 0.019->0.017 g/cm²/kg - 0.515->0.51 g/cm²/m vs. 0.513->0.513 g/cm²/m Menstruation, Resumed – NR: 11 (44%, N=25) vs. NR

				Age: 16.8 yr (SD ± 2.3) - 16.3 yr (SD ± 1.9) vs. 17.5 yr (SD ± 2.5) Gender, Female: 50 (100%) Race: NR	Weight – Baseline->1 yr: 42.9->47.5 kg vs. 45.1->48.8 kg (SD ± 5.1) Weight, % Change - Baseline – 1 yr: 9.8% (SD ± 11.2, N=25) vs. 7.3% (SD ± 12, N=18) Attrition: 11% (3/28) vs. 18% (4/22)
Klibanski et al. (1995)	Design: RCT Setting: NR Country: United States Funding: government and non-profit	Randomized N=48 Ethinyl Estradiol 35 µg / (Estrogen 0.625 mg + Progestin 5 mg) + Calcium Carbonate 1500 mg NR (N=22) Calcium Carbonate 1500 mg NR (N=26) %IBW < 70 % (N=6 vs. 10) %IBW > 70 % (N=13 vs. 15) Follow-up: 1.57 yr (Mean, SD ± 0.89) vs. 1.41 yr (Mean, SD ± 0.69)	Inclusion: Amenorrhea that occurred in close temporal association with the onset of anorexia; women; AN Exclusion: Any other illness known to affect bone density; taking any medication known to affect Bone Density including thyroid hormone; taking any medication known to affect Bone Density, including antiseizure medications; taking any medication known to affect bone density, including glucocorticoids	AN: 48 (100%) Amenorrhea: 48 (100%) Amenorrhea, Duration: 3.3 yr (SD ± 3.1) vs. 4.6 yr (SD ± 5.1) BMD, Spinal: 130 mg/cm² (SD ± 27, Total: 56 – 185) %IBW: 72% (SD ± 9) vs. 72% (SD ± 8) Weight: 43.03 kg (SD ± 7.3) vs. 41 kg (SD ± 5.6) Age: 24.9 yr (SD ± 6.9) - 23.7 yr (SD ± 7.2, vs. 25.8 yr (SD ± 6.6) Gender, Female: 48 (100%) Race: NR	BMD, Spinal – Baseline->Final Visit: 124->128 mg/cm² vs. 134->132 mg/cm² BMD, Spinal, % Change – Varies: 2.8% (SD ± 11, N=19) vs. -5.4% (SD ± 22.6, N=25) - %IBW > 70 % subgroup – Varies: 2.2% (SD ± 12) vs. 4.3% (SD ± 21.2) - %IBW < 70 % subgroup – Varies: 4% (SD ± 8.8) vs. -20.1% (SD ± 16.2) Menstruation, Resumed – Varies: 2 (9.09%) vs. 6 (23.08%) Study Withdrawal, Adverse Events – Varies: 2 (9.09%) vs. NR Attrition: 14% (3/22) vs. 4% (1/26)
Misra et al. (2011)	Design: RCT Setting: Multi-center Country: United States; Canada Funding: government and academic	Total N=150 Normal Weight Controls N=40 Those with AN Randomized N=110 (17-beta Estradiol 100 µg +	Inclusion: Adolescent; girl Exclusion: Other diseases affecting bone metabolism; untreated thyroid disease; premature ovarian failure; diabetes; cancer; pituitary disease; renal disease; bone fracture within the past 6 months; use of	AN: 110 (73.33%) Amenorrhea, Duration: 0.9 yr (SD ± 0.84, SE ± 0.08) Amenorrhea > 3 mo: 110 (73.33%) Those With AN: - BMD, Hip: 0.887 g/cm² (SD ± 0.12)	Greater increases with estrogen replacement was reported in BMD Z-scores at the spine and hip. Lumbar Spine, Z-Score, Change - Primary Endpoint - Baseline – 6 mo: 0.043 units (N=40) vs. -0.155 units (N=46) (MD 0.2 units, p=0.0002)

		Medroxyprogesterone 2.5 mg 10d/mo) / (Ethinyl Estradiol 3.75 µg 6 mo > 7.5 µg 12 mo > 11.25 µg) + Calcium Carbonate 1200 mg + Vitamin D 400 IU + Behavioral Therapy 18 mo (N=55) Placebo + Calcium Carbonate 1200 mg + Vitamin D 400 IU + Behavioral Therapy 18 mo (N=55)	prescription medications affecting bone metabolism within three months; suicidality; psychosis; substance abuse; hematocrit <30 %; potassium <3.0 mMol/L; glucose <50 mg/dL	- BMD, Hip, Z-Score: -0.644 units (SD ± 1.03) - BMD, Lumbar Spine: 0.907 g/cm² (SD ± 0.1) - BMD, Lumbar Spine, Z-Score: -0.623 units (SD ± 1.03) - %IBW: 84.6% (SD ± 6.29) - Weight: 47.2 kg (SD ± 5.24) - BMI: 17.4 kg/m² (SD ± 1.05) - Age: 16.5 yr (SD ± 2.1) Gender, Female: 150 (100%) Race: NR	- Baseline – 18 mo: -0.026 units (N=31) vs. -0.236 units (N=30) (MD 0.21 units, p=0.03) Hip, Z-Score, Change - Baseline – 12 mo: -0.08 units (N=34) vs. -0.193 units (N=39) (MD 0.11 units, p=0.04) - Baseline – 18 mo: -0.177 units (N=31) vs. -0.331 units (N=30) (MD 0.15 units, p=0.049) Hip, Change - Baseline – 12 mo: 0.005 g/cm² (SD ± 0.05, N=34) vs. -0.004 g/cm² (SD ± 0.04, N=39) (MD 0.01 g/cm², p=0.04) - Baseline – 18 mo: -0.001 g/cm² (SD ± 0.04, N=31) vs. -0.013 g/cm² (SD ± 0.06, N=30) (MD 0.01 g/cm² (p=0.04) Lumbar Spine, Change - Baseline – 6 mo: 0.015 g/cm² (N=40) vs. -0.006 g/cm² (N=46) (MD 0.02 g/cm², p=0.0003) - Baseline – 12 mo: 0.02 g/cm² (N=34) vs. -0.002 g/cm² (N=39) (MD 0.02 g/cm², p=0.0004) - Baseline – 18 mo: 0.021 g/cm² (N=31) vs. 0.002 g/cm² (N=30) (MD 0.02 g/cm², p=0.02) Lumbar Spine, Apparent (a height adjusted measure of spine BMD), Change - Baseline – 6 mo: 0.003 g/cm³ (N=40) vs. -0.001 g/cm³ (N=46) (MD 0 g/cm³, p=0.002) - Baseline – 12 mo: 0.003 g/cm³ (N=34) vs. 0 g/cm³ (N=39) (MD 0 g/cm³, p=0.005) - Baseline – 18 mo: 0.003 g/cm³ (N=31) vs. 0 g/cm³ (N=30) (MD 0 g/cm³, p=0.004) Menstruation, Resumed - Baseline – 18 mo: 0 (0%) vs. 5 (9.09%)

					Attrition: 44% vs. 45%

Abbreviations: AN=anorexia nervosa; BMD=bone mineral density; BMI=body mass index; DHEA=dehydroepiandrosterone; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Recombinant Human Growth Hormone

Fazeli et al. (2014)	Design: RCT Setting: Single Center: Massachusetts General Hospital Clinical Research Center Country: United States Funding: Academic and non-profit	Randomized N=21 Teriparatide 20 µg + Vitamin D 400IU + Calcium- Phosphate Binder- 1200 mg 6 mo (N=10) Placebo + Vitamin D 400IU + Calcium-Phosphate Binder- 1200 mg 6 mo (N=11)	Inclusion: Women; adult; AN; T-score of <= -2.5 at any site Exclusion: Abnormal thyroid function tests; chronic diseases known to affect BMD; diabetes mellitus; oral bisphosphonates within 12 months of the study; intravenous bisphosphonates within 3 years of initiating the study; medications known to affect bone metabolism in the 3 months preceding the study	AN: 21 (100%) AN, Duration: 20.4 yr (SD ± 11.7) vs. 18 yr (SD ± 14.26) Amenorrhea: 8 (80%) vs. 7 (63.64%) BMD - Femoral Neck: 0.6 g/cm² vs. 0.6 g/cm² - Lateral Spine: 0.53 g/cm² vs. 0.57 g/cm² - Posteroanterior Spine: 0.77 g/cm² vs. 0.81 g/cm² - Total Hip: 0.71 g/cm² vs. 0.69 g/cm² %IBW - 80.1% (SD ± 6.32) vs. 74.7% (SD ± 5.97) BMI: 17.6 kg/m² (SD ± 1.26) vs. 16.6 kg/m² (SD ± 1.33) Weight: 47.2 kg (SD ± 6.64) vs. 45.4 kg (SD ± 4.64) Age >= 18 yr: 21 (100%) Age: 47 yr (SD ± 8.54) vs. 47.1 yr (SD ± 7.63) Gender, Female: 21 (100%) Race: NR	Significantly more improvement was reported with teriparatide on spinal BMD at 6 mo. Lateral Spine - Change: 0.05 g/cm² vs. -0.003 g/cm² (MD 0.05 g/cm², p<0.01) - % Change: 10.5% vs. -0.6% (MD 11.1 %, p<0.01) Posteroanterior Spine - Change: 0.05 g/cm² vs. 0.002 g/cm² (MD 0.05 g/cm², p<0.01) - % Change: 6% vs. 0.2% (MD 5.8 %, p<0.01) Femoral Neck, Change: 0.001 g/cm² vs. 0.01 g/cm² (MD -0.01 g/cm², p>0.2) Total Hip, Change: -0.003 g/cm² vs. -0.001 g/cm² (MD 0 g/cm², p=0.8) Weight, % Change: -2.4% (SD ± 5.38) vs. 1.8% (SD ± 5.31) (MD -4.2 %, p=0.09) Study Withdrawal, All-Cause: 0 (0%) vs. 0 (0%) Overall Attrition: 0% (0/21)

Grinspoon et al. (2002)	Design: RCT Setting: Single Center: Massachusetts General Hospital Country: United States Funding: government, non-profit, and academic	Randomized N=60 Recombinant Human IGF-I 30 µg/kg + Calcium 1500 mg 9 mo (N=14) [Ethinyl Estradiol 35 µg + norethindrone 0.4 mg] + Calcium 1500 mg 9 mo (N=15) Recombinant Human IGF-I 30 µg/kg + [Ethinyl Estradiol 35 µg + norethindrone 0.4 mg] + Calcium - Phosphate Binder 1500 mg 9 mo (N=16) Placebo + Calcium-Phosphate Binder- 1500 mg 9 mo (N=15)	Inclusion: Confirmed AN; weighed less than 85% of IBW; amenorrheic for at least 3 months before the study; women; osteopenic at the anteroposterior spine Exclusion: Received estrogen within 6 months of the study.; received estrogen-related hormones known to affect bone density within 6 months of the study.; received estrogen-related hormones known to affect bone turnover within 6 months of the study.; previously received bisphosphonate therapy	AN: 60 (100%) Amenorrhea >= 3 mo: 60 (100%) BMD - Anteroposterior Spine: 0.828 g/cm² (SD ± 0.06) vs. 0.838 g/cm² (SD ± 0.09) vs. 0.814 g/cm² (SD ± 0.09vs. 0.793 g/cm² (SD ± 0.07) - Radial: 0.665 g/cm² (SD ± 0.04) vs. 0.705 g/cm² (SD ± 0.04) vs. 0.685 g/cm² (SD ± 0.06) vs. 0.68 g/cm² (SD ± 0.05) - Total Body: 1.01 g/cm² (SD ± 0.06) vs. 1.019 g/cm² (SD ± 0.09) vs. 0.995 g/cm² (SD ± 0.07) vs. 1.021 g/cm² (SD ± 0.09) - Hip: 0.809 g/cm² (SD ± 0.1) vs. 0.762 g/cm² (SD ± 0.12) vs. 0.765 g/cm² (SD ± 0.13) vs. 0.731 g/cm² (SD ± 0.1) %IBW < 85 %: 60 (100%) BMI: 17.8 kg/m² (SD ± 2.32) Weight: 45.3 kg (SD ± 5.61,) vs. 45.9 kg (SD ± 5.81vs. 45.6 kg (SD ± 5.6) vs. 42 kg (SD ± 7.36) Age: 25.2 yr (SD ± 5.42) - 23 yr (SD ± 4.12) vs. 27.6 yr (SD ± 6.2) vs. 24.2 yr (SD ± 6.4) vs. 26.3 yr (SD ± 5.81) Gender, Female: 60 (100%) Race: NR	Significant increase on anteroposterior spine BMD was reported with the combination of rhIGF-I and oral contraceptive compared to placebo (MD 2.8 %, p<0.05) BMD Anteroposterior Spine, % Change - Baseline – 9 mo: 0.3% (SD ± 2.24) vs. -0.2% (SD ± 3.1) vs. 1.8% (SD ± 3.2) vs. -1% (SD ± 5.03) Radial, Change - Baseline – 9 mo: -0.008 g/cm² (SD ± 0.02) vs. -0.005 g/cm² (SD ± 0.02) vs. 0.001 g/cm² (SD ± 0.02) vs. -0.01 g/cm² (SD ± 0.02) Total Body, Change - Baseline – 9 mo: -0.017 g/cm² (SD ± 0.04) vs. -0.03 g/cm² (SD ± 0.03) vs. -0.005 g/cm² (SD ± 0.03) vs. -0.018 g/cm² (SD ± 0.04) Hip, Change - Baseline – 9 mo: 0.007 g/cm² (SD ± 0.04) vs. -0.003 g/cm² (SD ± 0.03) vs. 0.008 g/cm² (SD ± 0.03) vs. 0.004 g/cm² (SD ± 0.02) Weight, Change - Baseline – 9 mo: 3 kg (SD ± 5.24) vs. 3.5 kg (SD ± 5.81) vs. 3.7 kg (SD ± 3.6) vs. 2.7 kg (SD ± 3.1) Attrition: 29% (4/14) vs. 0% (0/15) vs. 13% (2/16) vs. 7% (1/15)
Haines et al. (2021)	Design: RCT	Randomized N=90	Inclusion: AN or atypical AN; osteopenia; 18-45 years of age;	AN: 82 (100%) - Atypical: 42 (51%)	At 12 mo, mean postero-anterior lumbar spine arealBMD in the rhIGF-

	Setting: Single Center: Massachusetts General Hospital Country: United States Funding: Government; product donated by industry	Current Analysis N=82 Recombinant Human IGF-I 30 µg/kg + Placebo 6 mo > Risedronate 35 mg 12 mo (N=33) Placebo Injection + Risedronate 35 mg 6 mo > Risedronate 35 mg 12 mo (N=33) Placebo 12 mo (N=16)	areal BMD z-score or T-score <-1.0; estrogen replete or taking systemic estrogen therapy; normal thyroid function tests and serum 25OH vitamin D (≥20 ng/mL) and calcium levels Exclusion: contraindications to risedronate; binge-eating/purging subtype of AN with regular vomiting and significant periodontal disease; invasive dental procedure; any other disorder or medication known to affect bone or bone metabolism excluding exogenous estrogen; serum potassium <3.0 meq/L, alanine aminotransferase >3x upper limit of normal, or estimated glomerular filtration rate <30 mL/min; pregnant or breastfeeding; diabetes mellitus; active substance abuse; suicidality; malignancy or thromboembolic disorders	AN, Duration: 12.2 yr (SD ± 6.1) vs. 12.6 yr (SD ± 8.9) vs. 7.7 yr (SD ± 6.4) Amenorrhea: 10 (30%) vs. 12 (36%) vs. 4 (25%) Areal BMD - Postero-anterior spine (g/cm^2): 0.88 (SD ± 0.12) vs. 0.86 (SD ± 0.12) vs. 0.88 (SD ± 0.12) - Postero-anterior spine (Z-score): -1.4 (SD ± 1.1) vs. -1.6 (SD ± 1.0) vs. -1.3 (SD ± 1.0) - Lateral spine (g/cm^2): 0.67 (SD ± 0.09) vs. 0.66 (SD ± 0.09) vs. 0.67 (SD ± 0.07) - Lateral spine (Z-score): -1.6 (SD ± 1.1) vs. -1.7 (SD ± 1.1) vs. -1.6 (SD ± 0.9) - Total hip (g/cm^2): 0.84 (SD ± 0.11) vs. 0.79 (SD ± 0.12) vs. 0.87 (SD ± 0.11) - Total hip (Z-score): -0.8 (SD ± 0.9) vs. -1.2 (SD ± 1.0) vs. -0.6 (SD ± 0.9) - Femoral neck (g/cm^2): 0.72 (SD ± 0.10) vs. 0.68 (SD ± 0.13) vs. 0.75 (SD ± 0.11) - Femoral neck (Z-score): -1.1 (SD ± 0.9) vs. -1.4 (SD ± 1.1) vs. -0.8 (SD ± 1.0) - Total radius (g/cm^2): 0.55 (SD ± 0.05) vs. 0.53 (SD ± 0.05) vs. 0.55 (SD ± 0.05) - Total radius (Z-score): -0.3 (SD ± 0.9) vs. -0.8 (SD ± 0.9) vs. -0.5 (SD ± 0.9) Age: 28 yr (SD ± 6) vs. 28 yr (SD ± 7) vs. 25 yr (SD ± 6)	1/risedronate group was significantly higher than the placebo group (0.89 vs, 0.87 g/cm^2) (p=0.03) and was statistically similar compared with the risedronate alone group (0.89 g/cm^2) (p = 0.61). At 12 mo, mean lateral lumbar spine arealBMD in the rhIGF-1/risedronate group was significantly higher compared with both the risedronate group (p=0.04) and the placebo group (p=0.002) (0.69 vs. 0.68 vs. 0.66 g/cm^2). At 12 mo, mean total hip arealBMD did not differ among the groups. At 12 mo, mean femoral neck and total wrist areal BMD did not differ among the groups. Attrition: 30% (10/33) vs. 27% (9/33) vs. 13% (2/16)

				Gender, Female: 82 (100%) Race: NR	
Singhal et al. (2021)	Design: RCT Setting: Single Center: Massachusetts General Hospital Country: United States Funding: Government; product donated by industry	Randomized N=75 Recombinant Human IGF-1 30-46.88 µg/kg + 17-beta Estradiol 0.1 mg/day + Progesterone 100 mg 12 mo (N=38) Placebo 12 mo (N=37)	Inclusion: AN or atypical AN; 14-22 years of age Exclusion: Contraindications to estrogen therapy; history of conditions known to impact bone metabolism; bone fracture; past or current use of medications known to affect bone metabolism; pregnancy; suicidality; substance abuse; psychosis; hematocrit below 30% (indicative of anemia); potassium below 3.0 mMol/L; blood glucose below 50 mg/dL; other causes of hypoestrogenism	AN: 75 (100%) Amenorrhea, Duration: 3 mo vs. 4 mo BMD - Lumbar spine (g/cm^2): 0.90 vs. 0.86 - Lumbar spine (Z-score): -1.08 vs. -1.31 - Total hip (g/cm^2): 0.88 vs. 0.79 (SD ± 0.12) - Total hip (Z-score): -0.61 vs. -0.79 - Femoral neck (g/cm^2): 0.78 vs. 0.76 - Femoral neck (Z-score): -1.09 vs. -1.24 Age: 19.4 yr (SD ± 0.3) vs. 19.4 yr (SD ± 0.4) Gender, Female: 75 (100%) Race - Caucasian: 34 (89.5%) vs. 33 (89.2%) - Black or African American: 1 (2.6%) vs. 0 (0%) Ethnicity, Hispanic/Latino:1 (2.6%) vs. 4 (10.8%)	Over 12 mo, lumbar areal BMD increased in the placebo group compared to IGF-1 group (p=0.004). IGF-1 demonstrated no improvement in areal BMD in the setting of variable compliance to estrogen treatment. BMD, Change – Baseline – 12 mo - Lumbar spine (g/cm^2): 0.010 (N=12) vs. 0.043 (N=21) (p=0.004) - Lumbar spine (Z-score): 0.045 (N=12) vs. 0.280 (N=21) (p=0.028) - Total hip (g/cm^2): 0.016 (N=12) vs. 0.024 (N=21) (p=0.487) - Total hip (Z-score): 0.091 (N=12) vs. 0.155 (N=21) (p=0.555) - Femoral neck (g/cm^2): 0.014 (N=12) vs. 0.011 (N=21) (p=0.849) - Femoral neck (Z-score): 0.101 (N=12) vs. -0.016 (N=21) (p=0.470) More participants in the placebo group experienced irregular menses than in the IGF-1 group, but groups did not differ in incidence of other adverse events. Attrition: 30% (3/38) vs. 27% (1/37)

Abbreviations: AN=anorexia nervosa; BMD=bone mineral density; BMI=body mass index; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Risedronate/Testosterone

Haines et al. (2021	Design: RCT	Randomized N=90	Inclusion: AN or atypical AN; osteopenia; 18-45 years of age;	AN: 82 (100%) - Atypical: 42 (51%)	At 12 mo, mean postero-anterior lumbar spine

	Setting: Single Center: Massachusetts General Hospital Country: United States Funding: Government; product donated by industry	Current Analysis N=82 Recombinant Human IGF-I 30 µg/kg + Placebo 6 mo > Risedronate 35 mg 12 mo (N=33) Placebo Injection + Risedronate 35 mg 6 mo > Risedronate 35 mg 12 mo (N=33) Placebo 12 mo (N=16)	areal BMD z-score or T-score <-1.0; estrogen replete or taking systemic estrogen therapy; normal thyroid function tests and serum 25OH vitamin D (≥20 ng/mL) and calcium levels Exclusion: contraindications to risedronate; binge-eating/purging subtype of AN with regular vomiting and significant periodontal disease; invasive dental procedure; any other disorder or medication known to affect bone or bone metabolism excluding exogenous estrogen; serum potassium <3.0 meq/L, alanine aminotransferase >3x upper limit of normal, or estimated glomerular filtration rate <30 mL/min; pregnant or breastfeeding; diabetes mellitus; active substance abuse; suicidality; malignancy or thromboembolic disorders	AN, Duration: 12.2 yr (SD ± 6.1) vs. 12.6 yr (SD ± 8.9) vs. 7.7 yr (SD ± 6.4) Amenorrhea: 10 (30%) vs. 12 (36%) vs. 4 (25%) Areal BMD - Postero-anterior spine (g/cm^2): 0.88 (SD ± 0.12) vs. 0.86 (SD ± 0.12) vs. 0.88 (SD ± 0.12) - Postero-anterior spine (Z-score): -1.4 (SD ± 1.1) vs. -1.6 (SD ± 1.0) vs. -1.3 (SD ± 1.0) - Lateral spine (g/cm^2): 0.67 (SD ± 0.09) vs. 0.66 (SD ± 0.09) vs. 0.67 (SD ± 0.07) - Lateral spine (Z-score): -1.6 (SD ± 1.1) vs. -1.7 (SD ± 1.1) vs. -1.6 (SD ± 0.9) - Total hip (g/cm^2): 0.84 (SD ± 0.11) vs. 0.79 (SD ± 0.12) vs. 0.87 (SD ± 0.11) - Total hip (Z-score): -0.8 (SD ± 0.9) vs. -1.2 (SD ± 1.0) vs. -0.6 (SD ± 0.9) - Femoral neck (g/cm^2): 0.72 (SD ± 0.10) vs. 0.68 (SD ± 0.13) vs. 0.75 (SD ± 0.11) - Femoral neck (Z-score): -1.1 (SD ± 0.9) vs. -1.4 (SD ± 1.1) vs. -0.8 (SD ± 1.0) - Total radius (g/cm^2): 0.55 (SD ± 0.05) vs. 0.53 (SD ± 0.05) vs. 0.55 (SD ± 0.05) - Total radius (Z-score): -0.3 (SD ± 0.9) vs. -0.8 (SD ± 0.9) vs. -0.5 (SD ± 0.9) Age: 28 yr (SD ± 6) vs. 28 yr (SD ± 7) vs. 25 yr (SD ± 6) Gender, Female: 82 (100%) Race: NR	arealBMD in the rhIGF-1/risedronate group was significantly higher than the placebo group (0.89 vs, 0.87 g/cm^2) (p=0.03) and was statistically similar compared with the risedronate alone group (0.89 g/cm^2) (p = 0.61). At 12 mo, mean lateral lumbar spine arealBMD in the rhIGF-1/risedronate group was significantly higher compared with both the risedronate group (p=0.04) and the placebo group (p=0.002) (0.69 vs. 0.68 vs. 0.66 g/cm^2). At 12 mo, mean total hip arealBMD did not differ among the groups. At 12 mo, mean femoral neck and total wrist areal BMD did not differ among the groups. Attrition: 30% (10/33) vs. 27% (9/33) vs. 13% (2/16)
Miller et al. (2011)	Design: RCT Setting: Single Center: Massachusetts General Hospital Clinical Research Center	Randomized N=77 Risedronate 35 mg +/- Calcium 12 mo (N=20)	Inclusion: AN; women; BMD Z-scores below 1.0 in at least one skeletal site Exclusion: Conditions known to affect bone metabolism; use of medications (other than oral	AN: 77 (100%) AN, Binge-Eating and Purging: 10 (50%) vs. 5 (26%) vs. 8 (40%) vs. 7 (39%)	BMD Posteroanterior Spine, Z-Score – Risedronate vs. No Risedronate (pooled)

	Country: United States Funding: Government; product donated by industry	Testosterone 150-300 µg (titrate) +/- Calcium 12 mo (N=19) Risedronate 35 mg + Testosterone 150-300 µg (titrate) +/- Calcium 12 mo (N=20) Placebo +/- Calcium 12 mo (N=18) Risedronate 35 mg qw +/- Testosterone 150-300 µg qd (titrate) (pooled - Risedronate) (N=40) Testosterone 150-300 µg (titrate) +/- Risedronate 35 mg (pooled - Testosterone) (N=39) Placebo +/- Testosterone 150-300 µg (titrate) (pooled - No Risedronate) (N=37) Placebo +/- Risedronate 35 mg (pooled - No Testosterone) (N=38)	contraceptives) known to affect bone metabolism; active substance abuse	AN, Duration: 5.1 yr (SD ± 5.8) vs. 6.6 yr (SD ± 5.5) vs. 6.3 yr (SD ± 6.8) vs. 5.2 yr (SD ± 4.3) Amenorrhea: 6 (29%) vs. 4 (21%) vs. 5 (26%) vs. 2 (13%) BMD - Hip, Z-Score: -1.2 units (SD ± 1.1) vs. -1.1 units (SD ± 1) vs. -1.4 units (SD ± 0.7) vs. -1.4 units (SD ± 0.6) - Lateral Spine, Z-Score: -1.4 units (SD ± 0.6) vs. -1.2 units (SD ± 0.9) vs. -1.7 units (SD ± 1.1) vs. -1.7 units (SD ± 1) - Posteroanterior Spine, Z-Score: -1.6 units (SD ± 0.8) vs. -1.5 units (SD ± 1.2) vs. -1.5 units (SD ± 0.9) vs. -1.7 units (SD ± 0.8) - Radius, Z-Score: -0.6 units (SD ± 1) vs. -0.5 units (SD ± 0.9) vs. -0.4 units (SD ± 0.8) vs. -0.6 units (SD ± 1.1) %IBW: 78.8% (SD ± 5.7) vs. 78.7% (SD ± 7.5) vs. 78.7% (SD ± 7.1) vs. 78.9% (SD ± 5.7) BMI: 17.6 kg/m² (SD ± 1.2) vs. 17.5 kg/m² (SD ± 1.8) vs. 17.8 kg/m² (SD ± 1.4) vs. 17.9 kg/m² (SD ± 1.2) Weight: 47.2 kg (SD ± 5.6) vs. 47.8 kg (SD ± 5.5) vs. 47.9 kg (SD ± 5.4) vs. 47.6 kg (SD ± 5.2) Age: 25.3 yr (SD ± 6.3) vs. 27.1 yr (SD ± 7.3) vs. 25.2 yr (SD ± 6.2) vs. 26.9 yr (SD ± 7.2) Gender, Female: 77 (100%) Race: NR	6 months: -1.33 units (SD ± 1.01) vs. -1.55 units (SD ± 1.22) 9 months: -1.35 units (SD ± 0.95) vs. -1.53 units (SD ± 1.22) 12 months: -1.4 units (SD ± 0.95) vs. -1.55 units (SD ± 1.4) Study Withdrawal, Adverse Events - Baseline – 12 mo – Risedronate vs. No Risedronate vs. Testosterone vs. No Testosterone (pooled: 0 (0%) vs. 0 (0%) vs. 0 (0%) vs. 0 (0%) Overall Attrition: 23% (18/77)

Abbreviations: AN=anorexia nervosa; BMD=bone mineral density; BMI=body mass index; IBW=ideal body weight; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Alendronate

Golden et al. (2005)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=32 Alendronate 10 mg + Vitamin D 400 IU + Calcium- Phosphate Binder- 1200 mg + Multidisciplinary Therapy 1 yr (N=15) Placebo + Vitamin D 400 IU + Calcium-Phosphate Binder- 1200 mg + Multidisciplinary Therapy 1 yr (N=17) Amenorrhea, Secondary (N=13 vs. 14)	Inclusion: AN; osteopenia; adolescents; 12-21 years of age; primary amenorrhea or secondary amenorrhea of greater than 6 months duration; lumbar vertebral spine BMD more than 1 SD below the age-matched mean (z-score, <-1.0) Exclusion: Pregnancy; already receiving hormone therapy or steroids; already receiving injectable or oral contraceptives; receipt of hormone therapy or steroids within 90 days of enrollment; receipt of injectable or oral contraceptives within 90 days of enrollment; history of self-induced vomiting; coexistent medical condition that could contribute to the osteopenia; medical condition that precluded the administration of alendronate; subjects with primary amenorrhea who had a bone age less than 13.0 years	AN: 32 (100%) AN, Duration: 25.7 mo (SD ± 14.6) vs. 34.7 mo (SD ± 28) Amenorrhea, Primary or Amenorrhea, Secondary >= 6 mo: 32 (100%) Amenorrhea - Primary: 2 (13.33%) vs. 3 (17.65%) - Secondary - 13 (86.67%) vs.14 (82.35%) Amenorrhea, Duration {Amenorrhea, Secondary}: 20.1 mo (SD ± 17.5) vs. 19.9 mo (SD ± 17.3) BMD - Femoral Neck: 0.725 g/cm² (SD ± 0.09, N=14) vs. 0.672 g/cm² (SD ± 0.09, N=15) - Volumetric, Femoral Neck: 0.152 g/cm³ (SD ± 0.02) vs. 0.146 g/cm³ (SD ± 0.03) - Femoral Neck, Z-Score: -1.4 units (SD ± 0.87, N=14) vs. -1.8 units (SD ± 0.62, N=15) - Lumbar Spine: 0.795 g/cm² (SD ± 0.09, N=14) vs. 0.78 g/cm² (SD ± 0.07, N=15) - Volumetric, Lumbar Spine: 0.11 g/cm³ (SD ± 0.009) vs. 0.146 g/cm³ (SD ± 0.03) - Lumbar Spine, Z-Score: -1.9 units (SD ± 0.81, N=14) vs. -2 units (SD ± 0.69, N=15) - Total Hip: 0.783 g/cm² (SD ± 0.11, N=14) vs. 0.735 g/cm² (SD ± 0.1, N=15)	Increased femoral neck and lumbar spine BMDs were reported with alendronate. BMD - Baseline – 1 yr Femoral Neck, % Change: 4.4% (N=14) vs. 2.3% (N=15) (MD 2.1 % (p=0.41) Femoral Neck, Z-Score, Change: 0.16 units (N=14) vs. -0.01 units (N=15) Femoral Neck, Z-Score > -1 units: 5 (35.71%, N=14) vs. 4 (26.67%, N=15) Volumetric, Femoral Neck, Change: 4.4 g/cm³ (SD ± 6.4) vs. 0.004 g/cm³ (SD ± 0.02) Lumbar Spine, % Change: 3.5% (SD ± 4.6, N=14) vs. 2.2% (SD ± 6.1, N=15) (MD 1.3 %, p=0.53) Lumbar Spine, Z-Score, Change: 0.14 units (SD ± 0.35, N=14) vs. 0.04 units (SD ± 0.5, N=15) Lumbar Spine, Z-Score > -1 units: 3 (21.43%, N=14) vs. 2 (13.33%, N=15) Volumetric, Lumbar Spine, Change: -0.007 g/cm³ (SD ± 0.03) vs. 0.004 g/cm³ (SD ± 0.02)

				- Trochanter: 0.621 g/cm² (SD ± 0.07, N=14) vs. 0.569 g/cm² (SD ± 0.08, N=15) - Ward's Triangle: 0.698 g/cm² (SD ± 0.12, N=14) vs. 0.64 g/cm² (SD ± 0.11, N=15) %IBW: 76.9% (SD ± 7.1) vs. 77.3% (SD ± 6.2) Weight: 42.6 kg (SD ± 4.28) vs. 42.9 kg (SD ± 3.94) Age 12 yr-21 yr: 32 (100%) Age: 16.9 yr (SD ± 1.6) vs. 16.9 yr (SD ± 2.2) Gender, Female: 32 (100%) Race, Caucasian: 31 (96.88%) Ethnicity, Hispanic/Latino: 1 (3.13%)	Hip, Increased >= 4 % - 1 yr: 9 (64.3%, N=14) vs. 5 (33.3%, N=15) Total Hip, % Change - 3.6% (SD ± 8.5, N=14) vs. 1.6% (SD ± 8.7, N=15) Trochanter, % Change: 2.67% (SD ± 4.4, N=14) vs. 0.01% (SD ± 2.7, N=15) Ward's Triangle, % Change: 6.6% (SD ± 6.9, N=14) vs. 1.9% (SD ± 11.4, N=15) Weight, % Change: 13.5% (SD ± 9.9) vs. 16.2% (SD ± 16.4) (MD -2.7 %, p=0.59) Study Withdrawal, Dyspepsia, Treatment-Related: 0 (0%) vs. 1 (5.88%) Attrition: 7% (1/15) vs. 12% (2/17)

Abbreviations: AN=anorexia nervosa; BMD=bone mineral density; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Etidronate

Nakahara et al. (2006)	Design: RCT Setting: Outpatient: Kagoshima University Hospital Country: Japan Funding: Government	Randomized N=41 Etidronate 200 mg 3 mo (N=14) Calcium L-aspartate 600 mg qd + Alfacalcidol 1 µg 3 mo (N=15) Placebo 3 mo (N=12)	Inclusion: Women; restricting type of AN; Japanese; secondary amenorrhea for at least 3 months before examination Exclusion: AN patients with binge-eating and purging behavior	AN, Restricting: 41 (100%) AN, Duration: 57.3 mo (SD ± 27.7) vs. 57.7 mo (SD ± 9.17) vs. 41.1 mo (SD ± 15) Amenorrhea, Secondary >= 3 mo: 41 (100%) Amenorrhea, Duration: 48.3 mo (SD ± 25.4) vs. 49.3 mo (SD ± 76.6) vs. 36 mo (SD ± 14.9)	BMI - Baseline: 14.4 kg/m² (SD ± 1.7) vs. 14.7 kg/m² (SD ± 2) vs. 14.2 kg/m² (SD ± 2.3) - 3 mo: 15.8 kg/m² (SD ± 1.6) vs. 15.5 kg/m² (SD ± 2.1) vs.15.4 kg/m² (SD ± 1.7) Attrition: NR

				Age: 24.7 yr (SD ± 6.3) vs. 25.6 yr (SD ± 6.8) vs. 23.9 yr (SD ± 3.3) Gender, Female: 41 (100%) Race: NR	

Abbreviations: AN=anorexia nervosa; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; yr=year

Bulimia Nervosa Studies

Tricyclic Antidepressants

Desipramine

McCann et al. (1990)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=30 Desipramine HCl 25-300 mg (titrate) 12 wk (N=15) Placebo 12 wk (N=15) Follow-up: Baseline – 16 wk	Inclusion: Women; nonpurging bulimia; average of at least 2 binge-eating episodes/wk for a minimum of 1 yr Exclusion: Regularly purged by vomiting or laxative abuse; taking psychotropic medications; suffered from a psychotic condition; suffer from current drug abuse	BN, Non-Purging Type: 30 (100%) Binge Eating >= 2 episodes/wk, In the Previous 1 yr: 30 (100%) Vomiting, Self-Induced: 0 (0%, N=30) Laxative Abuse: 0 (0%, N=30) BMI: 31.7 kg/m² (SD ± 4.7) - (N=10) vs. 30.3 kg/m² (SD ± 5) - (N=13) Gender, Female: 30 (100%) Race: NR	Significantly less binge eating was reported with desipramine at 12 wk. Binge Eating, Episodes - 12 wk: 1.4/wk (N=10) vs. 3.7/wk (N=13) (MD -2.3/wk, p<0.04) Binge Eating – Baseline->12 wk: 3.8->1.4 d/wk (N=10) vs. 2.5->2.9 d/wk (N=13) Binge Eating, % Change - Baseline – 12 wk: - episodes/wk: -63% (N=10) vs. 6% (N=13) - d/wk: -63% (N=10) vs. 16% (N=13) Weight - Baseline: 91.6 kg (SD ± 20.1, N=10) vs. 89.1 kg (SD ± 16, N=13) Weight, Change - Baseline – 12 wk: -3.5 kg (SD ± 15.15, N=10) vs. -1.2 kg (SD ± 12.01, N=13) Attrition: 33% (5/15) vs. 30% (2/15)
Agras et al. (1992, (1994a)	Design: RCT; Follow-up Setting: NR	Randomized N=71	Inclusion: Women; 18-65 years of age; BN Exclusion: Concurrent medical condition that would preclude	BN: 71 (100%) Binge Eating: 5.5/wk (SD ± 4.6) vs. 5.9/wk (SD ± 5.1) vs. 7.5/wk (SD ± 3.4) vs. 9.3/wk	At 16 wk, both CBT and combined treatment were superior to medication given for 16 weeks in reducing binge eating and purging. Binge Eating, % Change - Baseline – 16 wk: -34% vs. -40% vs. -67% vs. -79% vs. -81.7%

	Location: NR Funding: Government	Desipramine HCl 25-350 mg (titrate) 16 wk (N=12) Desipramine HCl 25-350 mg (titrate) 24 wk (N=12) Desipramine HCl 25-350 mg (titrate) + CBT 16 wk (N=12) Desipramine HCl 25-350 mg (titrate) + CBT 16 wk > (-) CBT 24 wk (N=12) CBT 24 wk (N=23) Follow-up: Baseline – 72 wk	the use of antidepressants; evidence of conduction disturbance on electrocardiography; current AN; drug or abuse; psychosis; depression with suicidal risk of sufficient severity to preclude the use of antidepressants on an outpatient basis	(SD ± 5.8) vs. 8.7/wk (SD ± 7.2) Purging: 9.7/wk (SD ± 9.4) vs. 6.3/wk (SD ± 4.9) vs. 8.3/wk (SD ± 4.3) vs. 11.7/wk (SD ± 5.9) vs. 10.1/wk (SD ± 7.7) Age 18 yr-65 yr: 71 (100%) Gender, Female: 71 (100%) Race: NR	- CBT vs. Desipramine 16 wk/24 wk (pooled) (MD -42.9%, p<0.005) - Desipramine + CBT 16 wk > (+/-) Desipramine 24 wk (pooled) vs. Desipramine 16 wk/24 wk (pooled) (MD -43.8%, p<0.004) Purging, % Change - Baseline – 16 wk: -52% vs. -38% vs. -69% vs. -89% vs. -82.6% - CBT vs. Desipramine 16 wk/24 wk (pooled) (MD -39.9%, p<0.004) - Desipramine 16 wk/24 wk (pooled) vs. Desipramine + CBT 16 wk > (+/-) Desipramine 24 wk (pooled) (MD 38.2%, p<0.003) At 32 wk, only combined 24-wk treatment was superior to medication given for 16 wks (-35% vs. -45% vs. -60% vs. -90% vs. -78%). Continuing CBT appeared to prevent relapse in patients withdrawn from medication at 16 weeks. At 1-yr follow-up, combined 24-wk treatment and CBT alone were significantly superior in reducing binge eating to desipramine given for 16 wks: -22% (N=11) vs. -67% (N=9) vs. -55% (N=10) vs. -95% (N=9) vs. -72% (N=22). Only 18% (2 of 11) of those receiving 16 weeks of desipramine were free of binge eating and purging at follow-up compared with 78% (7 of 9) of those receiving the combined 24-wk treatment: 2 (18%, N=11) vs. 6 (67%, N=9) vs. 4 (40%, N=10) vs. 7 (78%, N=9) vs. 12 (54%, N=22) Attrition: 8% (1/12) vs. 25% (3/12) vs. 17% (2/12) vs. 25% (3/12) vs. 4% (1/23)
Walsh et al. (1997); Wilson et al. (1999)	Design: RCT	Randomized N=120	Inclusion: BN; women; 18-45 years of age; self-induced vomiting as a primary method of compensating for binge eating;	BN: 120 (100%) BN, Duration: 7.91 yr (SD ± 4.7)	Greater reductions in binge eating and vomiting were reported with CBT compared with supportive psychotherapy. CBT plus meds was significantly better than medication

	Setting: Outpatient Country: NR Funding: Government	CBT + Placebo 16 wk (N=25) CBT + Desipramine NR-300 mg 10 wk > Desipramine 200-300 mg / Fluoxetine 60 mg 16 wk (N=23) Supportive Psychotherapy + Placebo 16 wk (N=22) Supportive Psychotherapy + Desipramine NR-300 mg 10 wk > Desipramine / Fluoxetine 60 mg 16 wk (N=22) Desipramine NR-300 mg 10 wk > Desipramine 200-300 mg/ Fluoxetine 60 mg 16 wk (N=28) CBT 4 mo (pooled) (N=32) Desipramine 200-300 mg 10 wk > Fluoxetine 60 mg 4 mo (pooled) (N=32) Supportive Psychotherapy 4 mo (pooled) (N=35)	weight was between 80% and 120% of IBW Exclusion: Medically ill; evidence of cardiac conduction disease; pregnant; abused drugs or alcohol within the past yr; acutely suicidal; previous adverse reaction to desipramine or fluoxetine	- 8 yr vs. 7.26 yr vs. 7.55 yr vs. 9.55 yr vs. 7.36 yr Vomiting, Self-Induced: 120 (100%) %IBW 80 %-120 %: 120 (100%) Weight: 130 lbs (SD ± 15) BMI: 21.9 kg/m² (SD ± 2.2) History of AN: 9 (36%) vs. 6 (27%) vs. 6 (27%) vs. 7 (32%) vs. 9 (32%) Age 18 yr-45 yr: 120 (100%) Age: 26.1 yr (SD ± 4.9) - 25.8 yr vs. 26.1 yr vs. 26.9 yr vs. 28 yr vs. 24.3 yr Gender, Female: 120 (100%) Race - Caucasian: 100 (83%) - Black or African American: 7 (6%) - Asian: 6 (5%) Ethnicity, Hispanic/Latino: 7 (6%)	alone. Greater improvement in binge eating was reported with medication than placebo plus psychological treatment. Binge Eating – Baseline-> 16 wk: 7.22->2.56/wk vs. 7.29->0.95/wk vs. 6.18->3.32/wk vs. 7.92->3.57/wk vs. 8.32->2.59/wk - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine at 16 wk: MD -1.64/wk (p = 0.04) Vomiting, Diary – Baseline-> 16 wk: 10.8->5.6/wk vs. 10.8->1.1/wk vs. 11.9->7.5/wk vs. 10.6->5.5/wk vs. 10.5->3.7/wk - CBT + Desipramine/Fluoxetine vs. Desipramine/Fluoxetine16 wk: MD -2.6/wk (p = 0.01) Treatment Adherence, Treatment Sessions, Fulfilled – Baseline – 16 wk: 16.5 (SD ± 5) vs. 16.8 (SD ± 5.2) vs. 17.7 (SD ± 4.6) vs. 17.8 (SD ± 4.3) vs. 11.5 (SD ± 4.5) - CBT / Supportive Psychotherapy +/- Desipramine / Fluoxetine (pooled) vs. Desipramine / Fluoxetine: MD 5.7 (p = 0.0001) Attrition: 36% (9/25) vs. 35% (8/23) vs. 27% (6/22) vs. 27% (6/22) vs. 43% (12/28)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; HCl=hydrochloride; IBW=ideal body weight; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Imipramine

Agras et al. (1987)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=22 Imipramine HCl 50-300 mg (up-titrate) 16 wk (N=10) Placebo 16 wk (N=12)	Inclusion: BN; age >=18 years; 2 or more episodes of binge eating followed by self-induced vomiting within 1 wk prior to study entry; laxative use within 1 wk prior to study entry; women Exclusion: Diagnosis of concurrent AN; alcoholism; drug addiction; psychosis; significant suicidal ideation; previous history of the use of antidepressants for bulimia	BN: 22 (100%) BN, Duration: 8.7 yr (N=20) - 9.6 yr vs. 7.8 yr (N=10) Binge Eating and Purging >= 2 episodes, In the Previous 1 wk: 22 (100%) Purging: 11.8/wk (Total: 2 – 25, N=20) Laxative Abuse, In the Previous 1 wk: 22 (100%) AN: 0 (0%, N=22) Age >= 18 yr: 22 (100%) Age: 30.9 yr (N=20) - 30.3 yr vs. 31.5 yr (N=10) Gender, Female: 20 (100%) Race: NR	Significantly greater binge eating and purging percent changes were reported with imipramine at 16 wk: -72% vs. -43% (N=10) (MD -29 %, p<0.05); -72% vs. -35% (N=10) (MD -37 %, p<0.05), respectively. Binge Eating - Baseline: 11.6/wk (SD ± 6.35) vs. 13/wk (SD ± 8.4, N=10) Binge Eating, Change - Baseline – 16 wk: -8.4/wk vs. -5.6/wk (N=10) Purging – Baseline: 10.7/wk (SD ± 5.89) vs. 12.6/wk (SD ± 7.23, N=10) Purging, Change - Baseline – 16 wk: -7.7/wk vs. -4.4/wk (N=10) Purging, Abstinence - 16 wk: 3 (30%) vs. 1 (10%, N=10) Vomiting, Abstinence, Sum - 16 wk: 19.7 wk vs. 5.7 wk (N=10) Attrition: 0% (0/10) vs. 17% (2/12)
Mitchell et al.1990; Keel et al. (2002)	Design: RCT; Follow-up/Extension Setting: Outpatient: Eating Disorders Clinic; University of Minnesota Country: United States Funding: Government and non-profit	Randomized N=171 Intensive Group Therapy + Placebo 10 wk (N=34) Intensive Group Therapy + Imipramine HCl 200-300 mg 10 wk (50 mg induction) (up-titrate) (N=52) Imipramine HCl 200-300 mg 10 wk (50 mg	Inclusion: 18-40 years of age; female; IBW 80% to 120%; BN, binge eating and purging Exclusion: Current involvement in psychotherapy or pharmacotherapy for BN; concurrent medical condition that would preclude safe outpatient therapy with an antidepressant; active abuse of alcohol or drugs in the past 6 months	BN, Purging Type: 171 (100%) BN, Duration: 6.2 yr (SD ± 4) vs. 7 yr (SD ± 4.9) vs. 6.5 yr (SD ± 2.9) vs. 6.4 yr (SD ± 3.3) History of Laxative Abuse or Laxative Abuse: 62 (40%) (N=155) - 8 (24%, N=33) vs. 22 (46%, N=48) vs. 20 (44%, N=45) vs. 12 (41%, N=29)	All three active treatments led to significant reductions in binge eating and purging and improvement in mood relative to placebo. Intensive group psychotherapy had more improvement than Imipramine alone, with no benefit of combination treatment on eating behaviors (though Imipramine did help depression and anxiety.) Binge Eating - Baseline - 9.2/wk (N=33) vs. 8.4/wk (N=48) vs. 7.3/wk (N=45) vs. 8/wk (N=29) - 11.9 hr/wk (N=33) vs. 10.8 hr/wk (N=48) vs. 10.3 hr/wk (N=45) vs. 10.1 hr/wk (N=29)

		induction) (up-titrate) (N=54) Placebo 10 wk (N=31) Imipramine HCl 200-300 mg / (Intensive Group Therapy + Imipramine HCl 200-300 mg) 10 wk (pooled) (N=106) Intensive Group Therapy / (Intensive Group Therapy + Imipramine HCl 200-300 mg) 10 wk (pooled) (N=86) Placebo / Imipramine HCl 200-300 mg 10 wk (pooled) (N=85) Placebo / Intensive Group Therapy 10 wk (pooled) (N=65) Current Analysis (N=155) - 33 vs. 48 vs. 45 vs. 29 Follow-up: Baseline – 10 yr Follow-up (N=101)		%IBW 80%-120%: 171 (100%) %IBW: 97.7% (SD ± 10.2) vs. 108.2% (SD ± 12.4) vs. 106.5% (SD ± 12.8) vs. 107.6% (SD ± 11.3) History of AN: 25 (16.13%, N=155) - 10 (30%, N=33) vs. 5 (10%, N=48) vs. 8 (18%, N=45) vs. 2 (7%, N=29) Age 18 yr-40 yr: 171 (100%) Age: 22.8 yr (SD ± 4.3) vs. 24.3 yr (SD ± 5.7) vs. 24.1 yr (SD ± 4.4) vs. 24.4 yr (SD ± 5.2) Gender, Female: 171 (100%) Race: NR	Binge Eating, Change - Baseline – 10 wk: -8.2/wk vs. -7.7/wk vs. -3.6/wk vs. -0.2/wk - Intensive Group Therapy vs. Imipramine: MD -4.6/wk, p=0.0001 -10.6 hr/wk vs. -9.7 hr/wk vs. -5.3 hr/wk vs. -1.7 hr/wk - Intensive Group Therapy vs. Imipramine: MD -5.3 hr/wk (p=0.0001) Purging – Baseline: 13.2/wk (N=33) vs. 9.6/wk (N=48) vs. 8.6/wk (N=45) vs. 10/wk (N=29) Purging, Change - Baseline – 10 wk: -11.2/wk vs. -8.6/wk vs. -3.9/wk vs. -1.2/wk - Intensive Group Therapy vs. Imipramine: MD -7.3/wk (p=0.0001) Binge Eating – Baseline->10 yr: 6.3->2.4/d vs. 5.9->2.5/d vs. 5.9->2.5/d vs. 5.6->3.4/d Vomiting– Baseline->10 yr: 6.4->2.3/d vs. 5.4->2.6/d vs. 5.7->2.4/d vs. 5.9->3.4/d Laxative Abuse – Baseline->10 yr: 1.3->1/d vs. 2->1.2/d vs. 2.1->1.4/d vs. 1.9->1.3/d Attrition: 15% (5/34) vs. 25% (13/52) vs. 43% (23/54) vs. 16% (5/31)
Pyle et al. (1990)	Design: Follow-up of RCT (Mitchell et al. 1990)	Randomized N=68	Inclusion: BN; history of binge eating at least 3 times a wk for 6 months; women; 18-40 years of age; responded to intensive	BN: 68 (100%)	Although overall 30% relapsed by 6 mo, initial treatment with intensive group psychotherapy plus placebo or imipramine was associated with a lower relapse rate than imipramine

	Setting: NR Country: NR Funding: Government	Imipramine 200-300 mg 12 wk (N=3) Imipramine 200-300 mg + Intensive Support Group (Group CBT + Nutritional Counseling) 12 wk (N=19) Intensive Support Group (Group CBT + Nutritional Counseling) 12 wk (N=25) Placebo + Intensive Support Group (Group CBT + Nutritional Counseling) 12 wk (N=15) Placebo 12 wk (N=6) Follow-up: Baseline – 6 mo Follow-up (N=61) -3 vs. 18 vs. 21 vs. 13 vs. 6	group psychotherapy plus imipramine or placebo or to imipramine alone; history of self-induced vomiting or laxative abuse at least 3 times a wk for 6 months Exclusion: NR	Binge Eating >= 3 episodes/wk, In the Previous 6 mo: 68 (100%) Vomiting, Self-Induced >= 3 episodes/wk, Duration 6 mo or Laxative Abuse >= 3 episodes/wk, Duration 6 mo: 68 (100%) Age 18 yr-40 yr: 68 (100%) Gender, Female: 68 (100%) Race: NR	alone: 2 (67%) vs. 4 (22%, N=18) vs. 3 (14%, N=21) vs. 4 (31%, N=13) vs. 5 (83%) Binge Eating, % Change - -10 wk – 6 mo: -100% vs. -88% (N=18) vs. -92% (N=21) vs. -94% (N=13) vs. -95% Bulimic Episodes, Abstinence - 6 mo: 1 (33%) vs. 11 (61%, N=18) vs. 13 (62%, N=21) vs. 5 (38%, N=13) vs. 1 (17%) Disease Response, Remission - 6 mo: 1 (33%) vs. 13 (72%, N=18) vs. 17 (81%, N=21) vs. 7 (54%, N=13) vs. 1 (17%) Attrition: 0% (0/3) vs. 6% (1/19) vs. 19% (4/25) vs. 15% (2/15) vs. 0% (0/6)

Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; HCl=hydrochloride; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Amitriptyline

Mitchell and Groat (1984)	Design: RCT Setting: Outpatient: Eating disorders clinic in a university hospital	Randomized N=32 Amitriptyline 150mg + Behavioral Treatment Program 8 wk (N=16)	Inclusion: Bulimia; 18-45 years of age; bulimia for at least 6 months duration Exclusion: Current use of other psychotropic medications; significant medical illness which	BN: 32 (100%) BN, Duration >= 6 mo: 32 (100%) BN, Duration: 5 yr vs. 6 yr %IBW	Binge Eating: - 10.4/wk vs. 7.1/wk - 5.4 d/wk vs. 4.4 d/wk - 13 hr/wk vs. 10.4 hr/wk Binge Eating, % Change - Baseline – 8 wk - episodes/wk: -72.1% vs. -51.8% - d/wk: -63.6% vs. -47.8%

	Country: NR Funding: NR	Placebo + Behavioral Treatment Program 8 wk (N=16) Hamilton Depression Rating Scale >= 20 units subgroup (N=8 vs. 8) Depressive Disorder, None subgroup (N=8 vs. 8)	would preclude safe use of tricyclic antidepressants	- >= 75 %-<= 89 %: 4 (25%) vs. 6 (37.5%) - >= 90 %-<= 110 %: 8 (50%) vs. 8 (50%) - >= 111 %-<= 125 %: 3 (18.75%) vs. 2 (12.5%) - > 125 %: 1 (6.25%) vs. 0 (0%) Weight: 63.3 kg vs. 56.6 kg Age 18 yr-45 yr: 32 (100%) Age: 26 yr vs. 24.5 yr Gender, Female: 32 (100%) Race: NR	- hr/wk: -76.7% vs. -51.9% Vomiting, % Change - Baseline – 8 wk: -78.6% vs. -53.1% Attrition: 31% (5/15) vs. 6% (1/16)

Abbreviations: BN=bulimia nervosa; d=day; hr=hour; IBW=ideal body weight; mo=month; NR=not reported; RCT=randomized controlled trial; wk=week; yr=year

Other Antidepressants

Phenelzine

Walsh et al. (1984)	Design: RCT Setting: NR Country: NR Funding: Government	Randomized N=25 Phenelzine Sulfate 60-90mg 8 wk (N=12) Placebo 8 wk (N=13) Current Analysis (N=20) - 9 vs. 11 Follow-up: Baseline – 3 mo	Inclusion: Bulimia; bulimia for at least one year; currently binge eating at least three times weekly; 18-45 years of age; women; weighed between 80% and 120% of IBW Exclusion: Acute or chronic medical problems; hypokalemia; taking other psychotropic medications; acutely suicidal; history of suicide attempts; history of drug or alcohol abuse; unwilling to follow a tyramine-free diet; unable to follow a tyramine-free diet	BN: 25 (100%) Binge Eating >= 3 episodes/wk: 25 (100%) BN, Duration >= 1 yr: 25 (100%) BN, Duration: 8.1 yr (SD ± 4.8, N=9) vs. 9.4 yr (SD ± 4.7, N=11) Vomiting: 20 (100%, N=20) Laxative Abuse: 2 (10%, N=20)	Significantly reduced binge eating was reported at 8 wk with phenelzine: - Baseline: 10.8/wk (SD ± 6.5, N=9) vs. 11.1/wk (SD ± 6.1, N=11) - 8 wk: 2.6/wk (SD ± 4.3, N=9) vs. 10.5/wk (SD ± 5.9, N=11) (MD -7.9/wk, p<0.01) Binge Eating, Reduction >= 50 % - Baseline – 8 wk: 4 (44.44%, N=9) vs. 2 (18.18%, N=11) Binge Eating, Abstinence - 8 wk: 5 (55.56%, N=9) vs. 0 (0%, N=11) Attrition: 25% (3/12) vs. 15% (2/13)

				History of AN: 3 (33.33%, N=9) vs. 2 (18.18%, N=11) %IBW 80 %-120 %: 25 (100%) Age 18 yr-45 yr: 25 (100%) Age: 26.9 yr (SD ± 5.1, N=9) vs. 26 yr (SD ± 4.5, N=11) Gender, Female: 20 (100%, N=20) Race: NR	
Walsh et al. (1985)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=38 Phenelzine 60-90 mg 8 wk (N=20) Placebo 8 wk (N=18) Major Depressive Disorder subgroup (N=9 vs. 9) Depression, None subgroup (N=5 vs. 7)	Inclusion: 18-45 years of age; normal body weight; had been bulimic for at least 1 year; currently binge eating at least 3 times a wk; women Exclusion: Judged to be at significant risk of attempting suicide; abusing alcohol; abusing other drugs; pre-existing medical illness	BN, Duration: 9.4 yr (SD ± 4.9, N=14) vs. 10.5 yr (SD ± 6.1, N=16) Age 18 yr-45 yr: 38 (100%) Age: 27.8 yr (SD ± 4.7, N=14) vs. 27.2 yr (SD ± 5.3, N=16) Gender, Female: 38 (100%) Race: NR	Significantly greater reduction in binge eating and greater remission were reported with phenelzine. Binge Eating - Baseline: 9.9/wk (N=14) vs. 9.7/wk (N=16) - 8 wk: 3.4/wk (N=14) vs. 9.1/wk (N=16) (MD -5.7/wk, p<0.01) - Major Depressive Disorder subgroup: 4.6/wk vs. 10.7/wk (MD -6.1/wk, p<0.05) - Depression, None subgroup: 1.4/wk vs. 7/wk (MD -5.6/wk, p<0.05) Disease Response, Remission - Baseline – 8 wk: 6 (43%, N=14) vs. 0 (0%, N=16) (p<0.01) - Major Depressive Disorder subgroup: 3 (33.33%) vs. 0 (0%) - Depression, None subgroup: 3 (60%) vs. 0 (0%) (p<0.05) Attrition: 50% (10/20) vs. 28% (5/18)
Walsh et al. (1988)	Design: RCT Setting: NR Country: NR	Randomized N=62 Phenelzine 60-90 mg (up-titrate) 8 wk (N=31)	Inclusion: BN; bulimia for at least one year; currently binge eating at least three times weekly; 18-45 years of age;	BN: 62 (100%) Binge Eating >= 3 episodes/wk: 62 (100%)	Significantly greater binge eating reduction was reported with phenelzine. Binge Eating - Baseline: 11.9/wk (N=23) vs. 9.2/wk (N=27)

	Funding: Government, industry, and non-profit	Placebo 8 wk (N=31) Major Depressive Disorder subgroup (N=15 vs. 13) Depression, None subgroup (N=8 vs. 14)	women; weighed between 80% and 120% of IBW Exclusion: Taking psychotropic medications; acute medical problems; chronic medical problems; acutely suicidal; recent histories of drug or alcohol abuse; unwilling or unable to follow a tyramine-free diet	BN, Duration >= 1 yr: 62 (100%) BN, Moderate to Severe, Duration: 9 yr (SD ± 4.4, N=23) vs. 9.8 yr (SD ± 5.5, N=27) %IBW 80 %-120 %: 62 (100%) History of AN: 5 (21.74%, N=23) vs. 7 (25.93%, N=27) Age 18 yr-45 yr: 62 (100%) Age: 27 yr (N=50) - 26.9 yr (SD ± 4.3, N=23) vs. 27.1 yr (SD ± 5.2, N=27) Gender, Female: 62 (100%) Race: NR	- 8 wk: 5.4/wk (N=23) vs. 8.4/wk (N=27) (MD -3/wk, p<0.01) Baseline->8 wk - Major Depressive Disorder subgroup: 11.9->6.7/wk vs. 9.9->9.9/wk - Depression, None subgroup: 12->3/wk vs. 8.6->6.9/wk Binge Eating, % Change - Baseline – 8 wk: -64% (N=23) vs. -5% (N=27) (MD -59 %, p = 0.001) Study Withdrawal, Adverse Events - Baseline – 8 wk: 9 (29.03%) vs. NR Attrition: 26% (8/31) vs. 13% (4/31)

Abbreviations: AN=anorexia nervosa; BN=bulimia nervosa; IBW=ideal body weight; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Bupropion

Horne et al. (1988)	Design: RCT Setting: Multi-center: Carrier Foundation; La Jolla Eating Disorders Clinic; McLean Hospital Country: United States Funding: NR	Randomized N=81 Bupropion 450 mg 8 wk (75 mg induction) (N=55) Placebo 8 wk > +/- Bupropion NR (N=26)	Inclusion: Bulimia; women; 18-55 years of age; bulimic symptoms for at least 1 year and less than 15 years; body weight between 80% and 130% of desirable body weight Exclusion: Current use of any psychiatric medication; current use of medication with possible psychiatric effects; current use of propranolol; current use of Inderal; current use of reserpine; current use of Serpasil; current use of stimulants; serious neurological	BN: 81 (100%) BN, Duration 1 yr-< 15 yr: 81 (100%) BN, Duration: 6.5 yr (SD ± 3.5) vs. 6.6 yr (SD ± 4.9) Binge Eating: 10.8/wk %IBW 80 %-130 %: 81 (100%)	Significantly greater binge eating and purging percent changes were reported with bupropion. Binge Eating - Baseline: 11.9/wk (SD ± 7.3) vs. 8.5/wk (SD ± 6.9) Binge Eating, % Change - Baseline – 8 wk: -69.2% (SD ± 37, N=37) vs. -1.8% (SD ± 77, N=12) (MD -67.4 %, p=0.0061) Purging - Baseline: 5.9% (N=39) vs. 4.5% (N=17)

			condition; seizures; significant suicidal ideation; psychotic symptoms; manic symptoms; current major depression; current alcohol or drug abuse	PIBW: 106% (SD ± 10) vs. 109% (SD ± 11) Age 18 yr-55 yr: 81 (100%) Age: 26.1 yr (SD ± 6.6) vs. 26.9 yr (SD ± 8.2) Gender, Female: 81 (100%) Race: NR	Purging, % Change - Baseline – 8 wk: -53.5% (SD ± 49, N=34) vs. -2.5% (SD ± 45, N=10) (MD -51 %, p=0.0033) Weight - Baseline: 56.9 kg (SD ± 7.3) vs. 60 kg (SD ± 8.4) Weight, Change - Baseline – 8 wk: -1.2 kg (N=37) vs. 0.4 kg (N=12) (MD -1.6 kg, p<0.05) Headache - Baseline – 8 wk: 8 (14.5%) vs. 11 (42.3%) (p=0.015) Seizures - Baseline – 8 wk: 3 (5.5%) vs. 0 (0%) Attrition: 33% (18/55) vs. 54% (14/26)

Abbreviations: BN=bulimia nervosa; IBW=ideal body weight; MD=mean difference; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Other Pharmacotherapy

Topiramate

Hoopes Scott et al. (2003); Hedges et al. (2003)	Design: RCT Setting: Outpatient: University of Utah Health Sciences Center; Mountain West Clinical Trials Country: United States Funding: Industry	Randomized N=69 Topiramate 25–400 mg (titrate) 10 wk (25 mg induction) (N=35) Placebo 10 wk (N=34) Follow-up: Baseline – 11 wk ITT (N=64) - 31 vs. 33	Inclusion: 16-50 years of age; BN for at least 6 months Exclusion: Recent history of clinically significant suicidality; recent history of clinically significant substance abuse, bipolar disorder I, bipolar disorder II, major depressive disorder, or anxiety disorder; any personality disorder that could have interfered with assessments; history of nephrolithiasis; pregnant; lactating; taken psychoactive medications within 2 weeks prior to the study; diagnosis of AN; BMI <=17.5 kg/m^2; serum potassium level of <3.0 mMol/L	BN, Duration >= 6 mo: 69 (100%) Vomiting, Self-Induced: 64 (100%) Laxative Abuse: 13 (20.3%, N=64) Diuretics: 5 (7.8%, N=64) Weight: 61.5 kg (N=34) vs. 67.4 kg BMI <= 17.5 kg/m^2: 0 (0%, N=69)	Topiramate was associated with significantly greater decreases in binge eating and purging and greater improvements in binge and purging symptoms. Binge Eating - Baseline - 10.8/wk (N=31) vs. 11.3/wk (N=33) - 4.8 d/wk (N=31) vs. 4.7 d/wk (N=33) Binge Eating, % Change, d/wk - Baseline – 10 wk: -48.2% (N=31) vs. -17.7% (N=33) (MD -30.5 %, p=0.015) Purging - Baseline - 13.3/wk (N=30) vs. 12.4/wk (N=33) - 4.8 d/wk (N=31) vs. 4.8 d/wk (N=33)

				Age 16 yr-50 yr: 69 (100%) Age: 29 yr (N=31) vs. 29.6 yr (N=33) Gender - Female: 30 (96.77%, N=31) vs. 33 (100%, N=33) - Male: 1 (3.23%, N=31) vs. 0 (0%, N=33) Race: NR	Purging, % Change, d/wk - Baseline – 10 wk: -43.4% (N=31) vs. -16.6% (N=33) (MD -26.8 %, p=0.016) Binge Eating and/or Purging - Baseline: 5 d/wk (SD ± 1.6, N=31) vs. 5.1 d/wk (SD ± 1.5, N=33) Binge Eating and /or Purging, % Change, d/wk - Baseline – 10 wk - Primary Efficacy Outcome: -44.8% (N=31) vs. -10.7% (N=33) (MD -34.1 %, p=0.004) Binge Eating, Remission and/or Purging, Remission - 10 wk: 7 (22.6%, N=31) vs. 2 (6.1%, N=33) (p=0.012) Weight, Change - Baseline – 10 wk: -1.8 kg vs. 0.2 kg (MD -2 kg, p=0.004) Study Withdrawal - Baseline – 10 wk: - Adverse Events: 1 (2.94%, N=34) vs. 2 (5.88%) - Lack of Efficacy: 0 (0%) vs. 2 (5.88%) Attrition: 38% (13/35) vs. 47% (16/34)
Nickel et al. (2005)	Design: RCT Setting: NR Country: Germany Funding: NR	Randomized N=60 Topiramate 250 mg 10 wk (25 mg induction) (N=30) Placebo 10 wk (N=30)	Inclusion: Women; BN; 18 years or older; suffering from BN for at least 12 months Exclusion: Presence of psychotic or bipolar disease; current use of topiramate; severe somatic illness; currently suicidal; abusing alcohol; psychotic disease; bipolar disease; abusing drugs	BN: 60 (100%) BN, Duration >= 1 yr: 60 (100%) Vomiting, Self-Induced: 60 (100%) Exercise, Excessive: 14 (46.67%) vs. 15 (50%) Laxative Abuse and/or Enema, Abuse and/or Diuretics, Abuse: 11 (36.67%) vs. 10 (33.33%)	Topiramate was associated with significantly greater reduction in weight, binge eating, and purging. Weight - Baseline: 64.9 kg (SD ± 5.8) vs. 64.5 kg (SD ± 6.1) - 10 wk: 60.9 kg (SD ± 5.5) vs. 64.2 kg (SD ± 6) (MD -3.8 kg, 95% CI -54 – -2.1, p<0.001) Binge Eating and Purging - Baseline: 8/wk (SD ± 3) vs. 8/wk (SD ± 2.8) - 10 wk: 4.6/wk (SD ± 2.2) vs. 7.9/wk (SD ± 2.7) (MD -3.3/wk, 95% CI -4.3 – -2.1, p<0.001)

				Fasting: 7 (23.33%) vs. 6 (20%) Age >= 18 yr: 60 (100%) Age: 21.5 yr (SD ± 3.1) vs. 21.1 yr (SD ± 2.6) Gender, Female: 60 (100%) Race: NR	Binge Eating and Purging - Reduction > 50.1 % - Baseline – 10 wk: 11 (36.7%) vs. 1 (3.3%) (p<0.001) Adverse Events, Serious, Other - Baseline – 10 wk: 0 (0%) vs. 0 (0%) Attrition: 17% (5/30) vs. 20% (6/30)

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CI=confidence interval; d=day; ITT=intention-to-treat; MD=mean difference; NR=not reported; SD=standard deviation; wk=week; yr=year

Lithium

Hsu et al. (1991)	Design: RCT Setting: Outpatient: Western Psychiatric Institute and Clinic Country: United States Funding: Government	Randomized N=91 Lithium Carbonate 300-600 mg 8 wk (N=47) Placebo 8 wk (N=44) Major Depressive Disorder subgroup (N=17 vs. 13) Depression, None subgroup (N=30 vs. 31) Current Analysis (N=69) - 39 vs. 30 Major Depressive Disorder subgroup (N =12 vs. 7) Depression, None subgroup (N=27 vs. 23)	Inclusion: Female; bulimia; self-inducing vomiting and/or abusing laxatives for the purpose of weight loss; binge eating at least 2 times a wk in the last 6 months; maintaining body weight at between 85% to 125% of average for age, sex, and height Exclusion: Using any psychotropic medication for at least 4 weeks; evidence of schizophrenia or bipolar disorder; evidence of concurrent alcohol or substance dependence; evidence of organic metal illness; evidence of intellectual disability; having hypokalemia; significant medical illnesses; history of gastroplasty for obesity	BN: 91 (100%) Binge Eating >= 2 episodes/wk, In the Previous 6 mo: 91 (100%) Vomiting, Self-Induced or Laxative Abuse: 91 (100%) %ABW 85 %-125 %: 91 (100%) Binge Eating: 12.6/wk (SD ± 9.9, N=68) Age: 25.4 yr (SD ± 7, N=68) Gender, Female: 91 (100%) Race: NR	Binge Eating – Baseline->8 wk - Major Depressive Disorder subgroup: 6.7->4/wk vs. 14.6->6/wk - Depression, None subgroup: 7.5->4.6/wk vs. 8.1->2.9/wk Vomiting - Baseline->8 wk - Major Depressive Disorder subgroup: 8.6->4.1/wk vs. 13.6->6.1/wk - Depression, None subgroup: 7.7->3.8/wk vs. 10.3->3.9/wk Study Withdrawal, Adverse Events - Baseline – 8 wk: 0 (0%) vs. 0 (0%) Attrition: 20% (9/47) vs. 32% (14/44)

Abbreviations: ABW=average body weight; BN=bulimia nervosa; mo=month; NR=not reported; SD=standard deviation; wk=week; yr=year

Guided Self-Help/Self-Help

Banasiak et al. (2005)	Design: RCT Setting: NR Country: Australia Funding: Government	Randomized N=109 GSH 17 wk (N=54) Delayed Treatment Control 17 wk (N=55) Follow-up: Baseline – 10 mo ITT (N=73) - 43 vs. 30	Inclusion: Female; 18 years or older; BN; Caucasian Exclusion: Receiving psychological treatment; receiving pharmacological treatment; BMI below 18; AN; comorbid severe major depressive episode; substance dependence; psychotic disorder; high suicide risk; current or recent pregnancy; serious medical condition that interfered with eating; serious medical condition that interfered with weight	BN: 109 (100%) - Purging Type: 45 (83.3%) vs. 48 (87.3%) BN, Duration: 9.17 yr (SD ± 6.95) vs. 8.48 yr (SD ± 6.08) Binge Eating - Objective: 25 per 28 days (SD ± 26.33) vs. 23.78 per 28 days (SD ± 22.13) - Subjective: 28.77 per 28 days (SD ± 39.58) vs. 17.55 per 28 days (SD ± 21.18) Purging: 46.4 per 28 days (SD ± 49.15) vs. 40.08 per 28 days (SD ± 34.65) Vomiting: 53.65 per 28 days (SD ± 51.34) vs. 39.67 per 28 days (SD ± 35.78) Laxative Abuse: 8.77 per 28 days (SD ± 8.3) vs. 18.4 per 28 days (SD ± 18.22) History of AN: 12 (22.2%) vs. 14 (25.5%) BMI: 22.6 kg/m² (SD ± 3.58) vs. 23.1 kg/m² (SD ± 3.56) Age >= 18 yr: 109 (100%) Age: 29.5 yr (SD ± 8.72) vs. 28.3 yr (SD ± 8.22)	GSH was associated with greater abstinence from binge eating, purging, and compensatory behaviors. Binge Eating, Objective, Abstinence - 17 wk: 22 (61%, N=36) vs. 7 (18%, N=39) (p<0.001) Purging, Abstinence - 17 wk: 14 (47%, N=30) vs. 6 (17%, N=36) (p<0.01) Compensatory Behaviors, Abstinence - 17 wk: 18 (50%, N=36) vs. 6 (15%, N=39) (p<0.01) Binge Eating or Compensatory Behaviors, Abstinence - 17 wk: 14 (39%, N=36) vs. 6 (15%, N=39) (p<0.05) Binge Eating, Objective, Remission, Absolute - 17 wk: 25 (46%) vs. 7 (13%) (p<0.001) Purging, Remission, Absolute - 17 wk: 15 (33%, N=45) vs. 6 (12%, N=48) (p<0.05) Compensatory Behaviors, Remission, Absolute - 17 wk: 19 (35%) vs. 6 (11%) (p<0.01) Binge Eating and Compensatory Behaviors, Remission, Absolute - 17 wk: 15 (28%) vs. 6 (11%) (p<0.05) Attrition: 33% (18/54) vs. 29% (16/55)

				Gender, Female: 109 (100%) Race, Caucasian: 109 (100%)	
Carter et al. (2003)	Design: RCT Setting: NR Country: NR Funding: Academic	Randomized N=85 Cognitive Behavior SH 8 wk (N=28) Nonspecific SH 8 wk (N=28) WLC (N=29)	Inclusion: BN; seeking specialized treatment for the first time; women; age >=17 years Exclusion: Pregnant; medical illness known to influence eating or weight; treatment known to influence eating or weight; diabetes mellitus; current specialist treatment for an eating disorder; previous specialist treatment for an eating disorder; BMI <18 kg/m^2; Absence of binge eating symptoms; binge eating less than once weekly; episodes of overeating not objectively large; purging symptoms less than once weekly	BN: 85 (100%) - Purging Type: 79 (93%) BN, Duration: 7 yr (SD ± 6) BMI: 23 kg/m² (SD ± 5) Age >= 17 yr: 85 (100%) Age: 27 yr (SD ± 8) Gender, Female: 85 (100%) Race - Caucasian: 71 (83%) - Asian: 6 (7%) - African Caribbean: 2 (2%) Ethnicity, Other: 7 (8%)	Binge Eating, Objective - Baseline->8 wk: 24.5->10 per 28 days vs. 18.5->11.5 per 28 days vs. 28->27 per 28 days Purging - Baseline->8 wk: 26->22.5 per 28 days vs. 27.5->16.5 per 28 days vs. 46.5->32 per 28 days Disease Response, Responder - 8 wk: 15 (53.6%) vs. 14 (50%) vs. 9 (31%) Attrition: 18% (5/28) vs. 25% (7/28) vs. 28% (8/29)
Durand and King (2003)	Design: RCT Setting: Outpatient Country: United Kingdom Funding: NR	Randomized N=68 General Practice-Based SH 9 mo (N=34) Specialist Clinic Treatment 9 mo (N=34)	Inclusion: BN; age >=18 years; female; English-speaking Exclusion: BN but requiring an urgent clinic assessment; pregnancy; medical disorders; diabetes; substance misuse; alcohol misuse; serious suicidal intent	BN: 68 (100%) BN, Duration: 7.7 yr (SD ± 4.6) vs. 5.9 yr (SD ± 3.9) Age >= 18 yr: 68 (100%) Age: 28.3 yr (SD ± 6.5) vs. 24.5 yr (SD ± 5.2) Gender, Female: 68 (100%) Race - Caucasian: 29 (85%) vs. 30 (88%) - Black or African American: 3 (9%) vs. 3 (9%)	Bulimic Episodes - Baseline: 19 per 28 days (SD ± 15.2) vs. 20.4 per 28 days (SD ± 19.6) - 9 mo: 15 per 28 days (SD ± 17.4) vs. 14.9 per 28 days (SD ± 18.9) Vomiting - Baseline: 35.1 per 28 days (SD ± 31, N=28) vs. 37.8 per 28 days (SD ± 33.9, N=20) - 9 mo: 20.3 per 28 days (SD ± 27, N=28) vs. 20.5 per 28 days (SD ± 23.9, N=20) Attrition: 23% (8/34) vs. 18% (6/34)

				Ethnicity, Other: 1 (3%) vs. 1 (3%) Ethnicity, Missing Data: 1 (3%) vs. 0 (0%)	
Fernández-Aranda et al. (2009)	Design: Non-Randomized Controlled Trial Setting: NR Country: Spain Funding: Government	Total N=62 Internet-Based CBT-GSH 16 wk (N=31) WLC 12 wk (N=31)	Inclusion: Female; BN, purging subtype; BN Exclusion: NR	BN, Purging Type: 62 (100%) BN, Duration: 6 yr (SD ± 4.2) BMI: 22.58 kg/m² vs. 22.5 kg/m² Age: 23.7 yr (SD ± 3.6) Gender, Female: 62 (100%) Race: NR	Online CBT-GSH was associated with more abstinence from binge eating and vomiting. Binge Eating, Abstinence - 16 wk: 10 (32.3%) vs. 1 (3.2%) (p = 0.003) Vomiting, Abstinence - 16 wk: 10 (32.3%) vs. 0 (0%) (p = 0.001) Binge Eating and Vomiting, Abstinence - 16 wk: 7 (22.6%) vs. 0 (0%) (p = 0.005) Binge Eating – Baseline->Varies: 5.48->1.79 per 2 weeks vs. 7.35->6.94 per 2 weeks Vomiting – Baseline->Varies: 6.16->1.42 per 2 weeks vs. 7.61->7.61 per 2 weeks Overall attrition at 8 wk: 35.5% (11/31) vs. NR
Huon (1985)	Design: RCT Setting: NR Country: Australia; New Zealand Funding: NR	Randomized N=120 SH Program 7 mo (N=30) SH Program + Supportive Psychotherapy (With a Cured Bulimic Patient) 7 mo (N=30) SH Program + Supportive Psychotherapy (With a Improved Bulimic Patient) 7 mo (N=30)	Inclusion: BN; binge eating and vomiting and/or purging at least once/wk; women Exclusion: Already being treated for bulimia	BN: 120 (100%) Binge Eating and Purging >= 1 episodes/wk: 120 (100%) Binge Eating: 10.5/wk Binge Eating, Duration - <= 1 yr: 19 (15.8%) - > 1 yr-< 2 yr: 34 (28.3%) - > 2 yr-< 3 yr: 17 (14%) - > 3 yr-< 5 yr: 19 (16%) - > 5 yr-< 10 yr: 17 (14.2%) - >= 10 yr: 14 (11.7%) Dieting: 76 (63.6%)	Binge Eating and Purging – Baseline->7 mo: 10.39->7.2/wk vs. 10.77->4.65/wk vs. 10.92->5.15/wk vs. 11.15->12.8/wk Disease Response, Deteriorated - 7 mo: 2 (6.6%) vs. 0 (0%) vs. 1 (3.3%) vs. 10 (33.3%) Disease Response, Improvement - 7 mo: 19 (63.3%) vs. 20 (66.6%) vs. 22 (73.3%) vs. 5 (16.6%) Disease Response, Maintained - 7 mo: 4 (13.3%) vs. 3 (10%) vs. 2 (6.6%) vs. 15 (50%) Disease Response, Abstinence - 7 mo: 5 (16.6%) vs. 7 (23.3%) vs. 5 (16.6%) vs. 0 (0%)

		WLC 7 mo (N=30) SH Program + Supportive Psychotherapy (With a Cured/Improved Bulimic Patient) 7 mo (pooled) (N=60) SH Program / (SH Program + Supportive Psychotherapy (With a Cured/Improved Bulimic Patient)) 7 mo (pooled) (N=90) Follow-up: Baseline – 13 mo		Laxative Abuse: 82 (68%) Diuretics: 36 (30%) Weight-Reducing Drug: 44 (36.7%) Age: 22.5 yr Gender, Female: 120 (100%) Race: NR	Attrition: NR
Schmidt et al. (2007)	Design: RCT Setting: Multi-center, Outpatient: National Health Service Country: United Kingdom Funding: Non-profit	Randomized N=85 CBT Guided Self-Help 10 wk > 6 mo (N=44) Family Therapy 6 mo (N=41) Follow-up: Baseline – 12 mo	Inclusion: 13-20 years of age; BN or EDNOS; at least one close other to accompany them for family treatment Exclusion: BMI below 10th percentile for age and sex; knowledge of English insufficient to understand the treatment; learning disability; severe mental illness; substance dependence	BN or EDNOS: 85 (100%) - BN: 30 (68.2%) vs. 31 (75.6%) - EDNOS: 14 (31.8%) vs. 10 (24.4%) Binge Eating, Objective: 5.2/wk (SD ± 6.4) vs. 5.9/wk (SD ± 6.7) Vomiting, Objective: 9.5/wk (SD ± 11.7) vs. 9.9/wk (SD ± 17.9) BN, Age at Onset: 14.9 yr (SD ± 2.1) vs. 15.2 yr (SD ± 1.8) History of AN: 7 (16%) vs. 8 (20%) Age 13 yr-20 yr: 85 (100%)	Binge Eating, Objective, Abstinence - Baseline: 8 (18%) vs. 8 (19.5%) - 6 mo: 13 (41.9%, N=31) vs. 8 (25%, N=32) - 12 mo: 13 (52%, N=25) vs. 16 (55%, N=29) Vomiting, Abstinence - Baseline: 9 (20.5%) vs. 6 (14.6%) - 6 mo: 10 (32.3%, N=31) vs. 9 (28%, N=32) - 12 mo: 14 (56%, N=25) vs. 15 (51.7%, N=29) Binge Eating and Purging, Abstinence - Baseline: 2 (4.5%) vs. 2 (5%) - 6 mo: 6 (19.4%, N=31) vs. 4 (12.5%, N=32) - 12 mo: 9 (36%, N=25) vs. 12 (41.4%, N=29) Hospitalization Costs - Baseline – 12 mo: 481.19 pounds (SD ± 1411.47) vs. 66.28 pounds (SD ± 149.66)

				Age: 17.4 yr (SD ± 1.8) vs. 17.9 yr (SD ± 1.6) Gender - Female: 42 (95.5%) vs. 41 (100%) - Male: 2 (4.5%) vs. 0 (0%) Race, Caucasian: 30 (100%, N=30) vs. 31 (94%, N=33) Ethnicity, Other: 0 (0%, N=30) vs. (6%, N=33)	Attrition: 30% (13/44) vs. 29% (12/41)
Wagner et al. (2013)	Design: RCT Setting: Department of Child and Adolescent Psychiatry at the Medical University of Vienna; Eating Disorders Department at the Parklandklinik Country: Austria; Germany Funding: Government	Randomized N=155 Online CBT-GSH 4-7 mo (N=83) Manual CBT-GSH NR (N=72) Treatment Compliance, Reachable (N=55 vs. 32) Follow-up: Baseline – 18 mo Current Analysis (N=126) - 70 vs. 56	Inclusion: 16-35 years of age; female; BN purging type; EDNOS; binge eating or purging behavior between once and twice a wk or for less than 3 months; BMI above 18 Exclusion: Acute suicidality; severe depression; mental disorders affecting cognition; current drug misuse; current participation in cognitive-behavioral therapy	BN, Purging Type: 155 (100%) BN: 113 (90%, N=126) EDNOS: 13 (10%, N=126) Binge Eating and Purging 1 episodes/wk-2 episodes/wk or Binge Eating and Purging, In the Previous <= 3 mo: 155 (100%) BN, Duration: 8.21 yr (SD ± 5.19, N=70) vs. 8.82 yr (SD ± 4.6, N=56) BMI > 18 kg/m²: 155 (100%) BMI: 20.61 kg/m² (SD ± 2.12, N=70) vs. 20.72 kg/m² (SD ± 2.91, N=56) Age 16 yr-35 yr: 155 (100%)	Binge Eating, Objective - Baseline: 32.49/mo (SD ± 36.38, N=70) vs. 33.42/mo (SD ± 36.41, N=56) Binge Eating, Objective, Change - Baseline – 18 mo: -16.52/mo (SD ± 26.33, N=70) vs. -13.71/mo (SD ± 26.02, N=56) Vomiting - Baseline: 49.17/mo (SD ± 76.14, N=70) vs. 34.21/mo (SD ± 38, N=56) Vomiting, Change - Baseline – 18 mo: -32.74/mo (SD ± 58.78, N=70) vs. -16.72/mo (SD ± 27.42, N=56) Fasting – Baseline->18 mo: 6.81->0.74/mo (N=70) vs. 4.25->1.73/mo (N=56) Exercise, Excessive - Baseline->18 mo: 5.57->2.71/mo (N=70) vs. 5.9->2.75/mo (N=56) Laxative Abuse – Baseline->18 mo: 2.06->0.96/mo (N=70) vs. 0.77->0.37/mo (N=56) Binge Eating and Compensatory Behaviors, Abstinence OR Remission - 18 mo: 28 (58.3%, N=48) vs. 18 (64.3%, N=28)

				Age: 24.17 yr (SD ± 4.46, N=70) vs. 25.02 yr (SD ± 3.84, N=56) Gender, Female: 155 (100%) Race: NR	Study Withdrawal - Baseline – 18 mo: 15 (21.4%, N=70) vs. 24 (42.9%, N=56) Attrition: 34% vs. 56%

Abbreviations: AN=anorexia nervosa; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; EDNOS=eating disorder not otherwise specified; GSH=guided self-help; ITT=intention-to-treat; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SH=self-help; wk=week; WLC=wait-list control; yr=year

Other Interventions

Bachar et al. (1999)	Design: RCT Setting: NR Country: Israel Funding: NR	Randomized N=44 Nutritional Counseling 6 mo (N=10) Cognitive Orientation + Nutritional Counseling 1 yr (N=17) Self-Psychology + Nutritional Counseling 1 yr (N=17) BN, Purging Type (N=10 vs. 11 vs. 10) Current Analysis (N=36) - 10 vs. 12 vs. 14 - BN, Purging Type (N=7 vs. 10 vs. 8) Follow-up: Baseline – 2 yr	Inclusion: Female; bulimic or anorexic Exclusion: NR	AN or BN: 44 (100%) AN, Restricting: 8 (24.24%, N=33) BN, Purging Type: 25 (75.76%, N=33) - BN, Duration: 6.1 yr (SD ± 1.4, N=25) - Binge Eating: 2.4 per day (SD ± 1.5, N=25) - Vomiting: 1.01 per day (SD ± 0.54, N=25) - Age: 24.1 yr (SD ± 3.3, N=25) Gender, Female: 44 (100%) Race: NR	Self-psychology group had more remission than the other two groups. Disease Response, Remission - 1 yr Self-Psychology vs. Cognitive Orientation: 9 (64%, N=14) vs. 2 (17%, N=12) (p<0.02) BN, Purging Type - Self-Psychology vs. Nutritional Counseling.: 4 (50%, N=8) vs. 1 (14%, N=7) (p<0.05) Study Withdrawal - Baseline – 6 mo - Self-Psychology: 1 (5.88%) - Cognitive Orientation: 3 (17.65%) Attrition: 30% (3/10) vs. 29% (5/17) vs. 18% (3/17)
Bauer et al. (2012)	Design: RCT Setting: NR	Randomized N=165 Text Messaging Intervention 4 mo (N=82)	Inclusion: Female; age above 18 years; BN or EDNOS; met level 3 criteria according to the Longitudinal Interval Follow-Up Evaluation for either BN or EDNOS; full or subthreshold	BN or EDNOS: 165 (100%) - BN: 46 (56.1%) vs. 51 (61.4%) - EDNOS: 36 (43.9%) vs. 32 (38.6%)	Greater abstinence of binge eating or compensatory behaviors was reported with text messaging intervention at 8 wk: 31 (37.8%) vs. 15 (18.1%) (p<0.01)

	Country: Germany Funding: Government	TAU 8 mo (N=83) Outpatient Treatment, Utilizers subgroup (N=38 vs. 36) Outpatient Treatment, Non-utilizers subgroup (N=33 vs. 33)	BN; minimum of two binge-eating episodes/wk for a minimum duration of 1 mo Exclusion: NR	Binge Eating >= 2 episodes/wk, In the Previous 1 mo: 165 (100%) BN, Duration - < 1 yr: 6 (8.1%, N=76) vs. 3 (3.9%, N=78) - 1 yr-2 yr: 7 (9.5%, N=76) vs. 9 (11.7%, N=78) - 3 yr-5 yr: 16 (21.6%, N=76) vs. 20 (26%, N=78) - 6 yr-10 yr: 18 (24.3%, N=76) vs. 19 (24.7%, N=78) - > 10 yr: 28 (36.5%, N=76) vs. 26 (33.8%, N=78) Binge Eating: 6.8/wk (SD ± 7.36) vs. 7.43/wk (SD ± 7.24) Vomiting: 58 (70.7%) vs. 61 (73.5%) Laxative Abuse and/or Diuretics, Abuse: 18 (22%) vs. 20 (24.1%) Fasting: 18 (22.5%) vs. 18 (21.7%) Exercise, Excessive: 31 (37.8%) vs. 28 (33.7%) Age > 18 yr: 165 (100%) Age: 29.87 yr (SD ± 7.91) vs. 30.04 yr (SD ± 9.58) Gender, Female: 165 (100%)	Greater remission rate was reported with text messaging intervention at 8 wk, especially among those who did not utilize outpatient treatment. Disease Response, Remission - 8 mo: 42 (51.2%) vs. 30 (36.1%) (p = 0.05) - Outpatient Treatment, Utilizers: 24 (63.2%) vs. 20 (55.6%) (p = 0.51) - Outpatient Treatment, Non-utilizers subgroup: 18 (54.5%) vs. 10 (30.3%) (p = 0.046) Attrition: 13% (11/82) vs. 17% (14/83)

				Race: NR	
Brennan et al. (2020)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=72 Yoga 8 wk (N=36) WLC 8 wk (N=36) Current Analysis (Completers; N=53) 26 vs. 27	Inclusion: Age above 18 years; BN or BED; no or limited yoga experience Exclusion: Suicidal ideation, psychosis, or substance abuse, or a pre-existing diagnosis of BPD	BN or BED: 53 (100%) - BN: 40 (75.5%) - BED: 13 (24.5%) BMI - Overweight: 8 (15.1%) - Obese: 17 (32.1%) Age > 18 yr: 53 (100%) Gender, Female: 53 (100%) Race - Caucasian: 38 (72%) - Asian: 9 (17%) Other: 6 (11%)	Compared to WLC, yoga decreased binge eating frequency: Binge Eating Episodes: 11.46->5.11/28d vs. 12.92->12.11/28d Binge Eating Days: 11.63->4.58 d/28d vs. 11.70->10.60 d/28d Attrition: 28% (10/36) vs. 25% (9/36)
Habibzadeh and Daneshmandi (2010)	Design: RCT Setting: NR Country: Iran Funding: Academic	Randomized N=20 Exercise 2 mo (N=10) Control 2 mo (N=10)	Inclusion: BN; obese; BMI > 30 kg/m^2; women; sedentary life style Exclusion: Cardiovascular diseases; musculoskeletal diseases; respiratory diseases; other chronic diseases; menstrual irregularities; medication; beta-blockers; dieting; apparent occupational responsibilities that could impede participation; apparent leisure time responsibilities that could impede participation	BN: 20 (100%) Obese: 20 (100%) BMI > 30 kg/m²: 20 (100%) Age: 22 yr (SD ± 1.5) - 22.22 yr (SD ± 1.98) vs. 22.67 yr (SD ± 1.5) Gender, Female: 20 (100%) Race: NR	Significantly greater changes were reported with exercise in weight, BMI, fat mass, body fat, and lean mass. Weight – Baseline->2 mo: 74.98->73.27 kg vs. 78.11->78.06 kg (MD -4.79 kg (p<0.001 at 2 mo) BMI – Baseline->2 mo: 30.2->28.88 kg/m² vs. 30.93->30.41 kg/m² (MD -1.53 kg/m² (p<0.001 at 2 mo) Body Composition, Fat Mass – Baseline->2 mo: 29.11->27.17 kg vs. 31.16->31.42 kg (MD -4.25 kg (p<0.001 at 2 mo) Body Fat – Baseline->2 mo: 38.8->36.35% vs. 39.97->39% (MD -2.65 % (p<0.001 at 2 mo) Lean Mass – Baseline->2 mo: 43.27->44.38 kg vs. 43.86->43.25 kg (MD 1.13 kg (p<0.001 at 2 mo)

					Overall Attrition: 0%
Hill et al. (2011)	Design: RCT Setting: NR Country: United States Funding: NR	Randomized N=32 Appetite Focused DBT 12 wk (N=18) WLC 6 wk > Appetite Focused DBT 12 wk (N=14) Appetite Focused DBT/WLC 6 wk (pooled) (N=26) ITT (N=32) - 18 vs. 14	Inclusion: An average of at least one binge eating episode per wk over the previous 3 months; one vomiting episode per wk over the previous 3 months; used vomiting as their primary compensatory behavior; women; BN Exclusion: Age <18 years; current diagnosis of AN; current diagnosis of BED; concurrent psychotherapy focused on eating issues; current suicidal ideation; substance dependence at the level deemed to interfere with treatment; cognitive impairment at the level deemed to interfere with treatment; past and present psychosis	BN: 32 (100%) - BN: 14 (77.78%) vs. 12 (85.71%) - BN, Subclinical: 4 (22.22%) vs. 2 (14.29%) Binge Eating, Objective: 16.5 per 28 days Vomiting: 16.5 per 28 days BMI: 22.6 kg/m² - 23.23 kg/m² (SD ± 5.2) vs. 21.65 kg/m² (SD ± 2.15) Age: 22 yr - 22.67 yr (SD ± 5.86) vs. 21.08 yr (SD ± 2.93) Gender, Female: 32 (100%) Race - Caucasian: 30 (93.8%) - Black or African American: 1 (3.13%) - Asian American: 1 (3.13%)	Binge Eating, Objective – Baseline->6 wk: 15.5->4/28 days vs. 18->19.5 /28 days - Appetite Focused DBT/WLC 6 wk (pooled): 16.5->4.5 /28 days Binge Eating, Objective - 12 wk - Appetite Focused DBT/WLC 6 wk (pooled): 1.5 /28 days Vomiting – Baseline->6 wk: 15.5->2.5 /28 days vs. 23.5->12.5 /28 days - Appetite Focused DBT/WLC 6 wk (pooled): 15.5->2.5 /28 days Vomiting - 12 wk - Appetite Focused DBT/WLC 6 wk (pooled): 2 /28 days Attrition at 6 wk: 11% (2/18) vs. 14% (2/14)
Jäger et al. (1996)	Design: RCT Setting: NR Country: Germany Funding: Non-profit	Randomized N=83 Analytic Inpatient Therapy 38 mo (N=37) Systemic Outpatient Therapy 38 mo (N=46) Current Analysis (N=71)	Inclusion: BN; women Exclusion: NR	BN: 83 (100%) BN, Symptomatic, Duration: 4.7 yr Binge Eating < 2 episodes/wk: 0 (0%, N=32) vs. 0 (0%, N=39) Laxative Abuse: 20 (24%)	Binge Eating – Baseline->38 mo: 12.4->2.1/wk (N=32) vs. 10.6->2.8/wk (N=39) Vomiting, Self-Induced - Baseline->38 mo: 12.2->1.6/wk (N=32) vs. 10.7->2.9/wk (N=39) Binge Eating and Purging, Abstinence - Baseline – 38 mo: 0 (0%)->21 (65.6%) (N=32) vs. 0 (0%)->17 (43.6%) (N=39) (p<0.1)

		- 32 vs. 39		History of AN: 28 (39%, N=71) Age: 23.8 yr Gender, Female: 83 (100%) Race: NR	Weight - Baseline – 38 mo: 59.4->62 kg (N=28) vs. 59.4->59.2 kg (N=37) Study Withdrawal - Baseline – 38 mo: NR vs. 9 (19.57%) Attrition: 19% (7/37) vs. 20% (9/46)
Le Grange et al. (2007)	Design: RCT Setting: Outpatient Country: United States Funding: Government	Randomized N=80 FBT 6 mo (N=41) SPT 6 mo (N=39) Follow-up: Baseline – 12 mo	Inclusion: Adolescent; 12-19 years of age; BN or partial BN; Exclusion: Physical or psychiatric disorder necessitating hospitalization; insufficient knowledge of English; current physical dependence on drugs or alcohol; current low body weight (BMI =< 17.5); current treatment for the eating disorder or current use of medication known to affect eating or weight; and physical conditions (e.g., diabetes mellitus or pregnancy) or treatments known to influence eating or weight; 50 mg or more of fluoxetine	BN: 80 (100%) - BN: 18 (43.9%) vs. 19 (48.7%) - Partial BN: 23 (56.1%) vs. 20 (51.3%) BN, Duration: 22.3 mo (SD ± 20.4) vs. 20.1 mo (SD ± 24.4) BMI: 21.8 kg/m² (SD ± 2.5) vs. 22.4 kg/m² (SD ± 3.4) Age 12 yr-19 yr: 80 (100%) Age: 16 yr (SD ± 1.7) vs. 16.1 yr (SD ± 1.6) Gender - Female: 40 (97.6%) vs. 38 (97.4%) - Male: 1 (2.4%) vs. 1 (2.6%) Race - Caucasian: 31 (75.6%) vs. 20 (51.2%) - African American: 4 (9.8%) vs. 5 (12.8%) - Other: 0 (0%) vs. 4 (10.3%) Ethnicity - Hispanic: 6 (14.6%) vs. 10 (25.6%)	Compared with SPT, remission rates were significantly higher for FBT: 16 (39%) vs. 7 (18%) (p=0.049) at post-treatment; 12 (29%) vs. 4 (10%) (p=0.05) at 6-mo follow-up. Binge Eating, Objective - Baseline: 18.4/mo (SD ± 28.1) vs. 18.9/mo (SD ± 22.3) - Post-treatment: 4.1/mo (SD ± 14.8) vs. 3.2/mo (SD ± 5.1) - 12 mo: 2.5/mo (SD ± 6.8) 5.4/mo (SD ± 13.7) Binge Eating, Subjective - Baseline: 9.9/mo (SD ± 16.6) vs. 7.6/mo (SD ± 10.1) - Post-treatment: 4.5/mo (SD ± 13.3) vs. 4.6/mo (SD ± 8.6) - 12 mo: 2.8/mo (SD ± 6.9) vs. 2.4/mo (SD ± 5.2) Vomiting - Baseline: 34.5/mo (SD ± 31.0) vs. 33.2/mo (SD ± 33.5) - Post-treatment: 4.8/mo (SD ± 9.4) vs. 17.4/mo (SD ± 26.0) - 12 mo: 10.1/mo (SD ± 21.8) vs. 14.5/mo (SD ± 27.7) Attrition: 12% (5/41) vs. 10% (4/39)

Russell et al. (1987)	Design: RCT Setting: Outpatient: Maudsley Hospital Country: United Kingdom Funding: Government	Randomized N=80 Family Therapy 1 yr (N=41) Individual Therapy 1 yr (N=39) BN subgroup (N=12 vs. 11) Follow-up: Baseline – 5 yr	Inclusion: AN or BN	AN or BN: 80 (100%) - BN: 12 (29.27%) vs. 11 (28.21%) BN, Duration - BN subgroup: 4.9 yr (SD ± 3.7) %ABW: 69.6% (SD ± 13) BN, Age at Onset: 17.9 yr (SD ± 6.4) Age: 21.8 yr (SD ± 7.1) Age - BN subgroup: 24 yr (SD ± 8.4) Gender - Female: 73 (91.25%) - Male: 7 (8.75%) Race: NR	Among BN subgroup Disease Response - 1 yr - Poor: 8 (88.89%, N=9) vs. 7 (70%, N=10) - Intermediate: 1 (11.11%, N=9) vs. 2 (20%, N=10) - Good: 0 (0%, N=9) vs. 1 (10%, N=10) PABW – Baseline->1 yr: 77->89% (N=9) vs. 78.7->86.2% (N=10) Study Withdrawal - Early, Baseline – 3 mo: 2 (16.67%) vs.1 (9.09%) - Late - 3 mo – 12 mo: 2 (16.67%) vs. 0 (0%) Treatment Discontinuation - Baseline – 1 yr: 7 (58.33%) vs. 2 (18.18%) Attrition: 37% (15/41) vs. 33% (13/39)
Zeeck et al. (2009a, 2009b)	Design: RCT; Follow-Up Setting: Outpatient Country: Germany Funding: NR	Randomized N=55 Day Clinic 12 wk (N=28) Inpatient Care 12 wk (N=27) Follow-up: Baseline – 12 mo Current Analysis (N=43) - 22 vs. 21	Inclusion: BN; >=18 years of age; reach the clinic in 1 hour or less; fulfilled 1 or more of the following criteria: (1) failed to improve in outpatient psychotherapy (2) severe bulimic symptoms (3) chronic course of their illness and/or (4) severe comorbidity Exclusion: Serious instable medical conditions; current suicidal ideation; current severe substance dependence; current severe psychotic disorder	BN: 55 (100%) BN, Duration: 10.5 yr (SD ± 7.6, N=22) vs. 7 yr (SD ± 6.5, N=21) History of AN: 9 (40.9%, N=22) vs. 7 (33.3%, N=21) BMI: 21.4 kg/m² (SD ± 2.5, N=22) vs. 21.5 kg/m² (SD ± 2.2, N=21) Age >= 18 yr: 55 (100%) Age: 26.2 yr (SD ± 7.2, N=22) vs. 24 yr (SD ± 7.6, N=21) Gender	Disease Response, Complete Remission - Baseline – 12 wk: 4 (18.2%, N=22) vs. 7 (33.3%, N=21) - Baseline – 6 mo: 3 (13.7%, N=22) vs. 4 (19.1%, N=21) - Baseline – 12 mo: 6 (27.27%) vs. 3 (20%, N=15) Disease Response, Relapse - 3 mo – 6 mo - Primary Efficacy Outcome: 0 (0%, N=18) vs. 4 (22.22%, N=18) Study Withdrawal - Baseline – 6 mo: 0 (0%, N=22) vs. 4 (19.05%, N=21) Attrition: 36% (10/28) vs. 33% (9/27)

				- Female: 21 (95.5%, N=22) vs. 19 (90.5%, N=21) - Male: 1 (4.5%, N=22) vs. 2 (9.5%, N=21) Race: NR	

Abbreviations: ABW=average body weight; AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; DBT=dialectical behavior therapy; EDNOS=eating disorder not otherwise specified; ITT=intention-to-treat; MD=mean difference; mo=month; NR=not reported; SD=standard deviation; SPT=supportive psychotherapy; TAU=treatment as usual; wk=week; WLC=wait-list control; yr=year

Binge-Eating Disorder Studies

Psychotherapies

Guided Self-Help/Self-Help

Cachelin et al. (2019)	Design: RCT Setting: NR Country: United States Funding: NR	Randomized N=40 CBT-GSH 12 wk (N=21) WLC 12 wk (N=19)	Inclusion: Female; Latina; 18-55 years of age; BED; BMI >= 18 kg/m^2; overeating Exclusion: Current pregnancy; medical condition or medication that significantly affects weight or eating (e.g., hypothyroidism); brain injury or impairment; serious medical condition or medical risk	BED: 40 (100%) BMI: 30.4 kg/m² (SD ± 7.1) vs. 27.7 kg/m² (SD ± 6.3) Age 18 yr-55 yr: 40 (100%) Age: 26.3 yr (SD ± 6.6) vs. 27.2 yr (SD ± 10.1) Gender, Female: 40 (100%) Ethnicity, Hispanic/Latino: 40 (100%)	CBT-GSH showed significant reductions in frequency of binge eating compared to WLC: Baseline->End of Treatment: 13.3->3.0 (p=0.005) vs. 13.1->11.4 The rate of abstinence from binge eating was much higher for the CBT-GSH group at post-assessment than for the WLC group (47.6% vs. 5.3%). BMI – Baseline->End of Treatment: 29.9->30.4 kg/m² vs. 28.9->29.6 kg/m² Attrition: 29% (6/21) vs. 26% (5/19)
Carrard et al. (2011)	Design: RCT Setting: Education System: University Hospitals of Geneva (Switzerland)	Randomized N=74 Internet CBT-GSH 6 mo (N=37) WLC 6 mo (N=37) Follow-up: Baseline – 12 mo	Inclusion: Women; 18-60 years of age; full or subthreshold BED Exclusion: Recent suicide attempt; past obesity surgery	BED or BED, Subclinical: 74 (100%) - BED: 20 (54.1%) vs. 23 (62.2%) - BED, Subclinical: 17 (45.9%) vs. 14 (37.8%) BMI < 30 kg/m²: 44 (59.5%) Age 18 yr-60 yr: 74 (100%) Age: 36 yr (SD ± 11.4) - 34.4 yr (SD ± 11) vs. 37.8 yr (SD ± 11.8)	BMI – Baseline: 29.8 kg/m² (SD ± 5.9) vs. 27.7 kg/m² (SD ± 5.5) BMI, Change - Baseline – 6 mo: -0.6 kg/m² (SD ± 4.61) vs. 0.2 kg/m² (SD ± 4.22) Attrition: 24% (9/37) vs. 11% (4/37)

	Country: Switzerland Funding: Non-profit			Gender, Female: 74 (100%) Race: NR	
Carter and Fairburn (1998)	Design: RCT Setting: Education System: Oxford University Country: United Kingdom of Great Britain and Northern Ireland Funding: Non-profit	Randomized N=72 CBT-GSH 12 wk (N=24) Pure CBT-SH 12 wk (N=24) WLC 12 wk (Re-randomized to either GSH or pure SH after 12 wk) (N=24) Current Analysis (N=69) - CBT-GSH 12 wk (pooled for re-randomization) (N=34) Pure CBT-SH 12 wk (pooled for re-randomization) (N=35) Follow-up: Baseline – 38 wk	Inclusion: BED; binge frequency of at least 1/wk; women Exclusion: Vomited in the previous 3 months; fasted in the previous 3 months; laxatives in the previous 3 months; diuretics in the previous 3 months; BN; AN; age below 18 years; age above 65 years; current psychiatric treatment; previous treatment for a binge eating problem	BED: 72 (100%) Binge Eating >= 1/wk: 72 (100%) Weight: 85.8 kg (SD ± 19.7) BMI, Obesity > 30 kg/m²: 43 (60%) Age: 39.7 yr (SD ± 10) Gender, Female: 72 (100%) Race - Caucasian: 70 (97.22%) - Afro-Caribbean: 1 (1.39%) - Asian: 1 (1.39%)	BMI - Baseline: 32.2 kg/m² (SD ± 6.4) vs. 30.6 kg/m² (SD ± 6.6) vs. 31.5 kg/m² (SD ± 6.6) - 12 wk: 31.7 kg/m² (SD ± 6.1) vs. 30.7 kg/m² (SD ± 6.6) vs. 31.9 kg/m² (SD ± 7.4) - 38 wk: 31.6 kg/m² (SD ± 6.2) vs. 30.4 kg/m² (SD ± 6.5) vs. NR Treatment Discontinuation - Baseline – 12 wk: 8 (23.53%) vs. NR vs. NR Overall Attrition: 13% (9/72)
Grilo and Masheb (2005)	Design: RCT Setting: NR Country: United States Funding: Non-profit	Randomized N=90 CBT-GSH 12 wk (N=37) BWL 12 wk (N=38) Self-Monitoring 12 wk (N=15)	Inclusion: 18-60 years of age; BMI >= 27 kg/m^2; BED; overweight Exclusion: Concurrent treatment for eating disorder or weight disorder; concurrent treatment for psychiatric illness; psychosis; bipolar disorder; current substance use dependence	BED: 90 (100%) Overweight: 90 (100%) BMI >= 27 kg/m²: 90 (100%) BMI: 35.5 kg/m² (SD ± 6.7) - 33.4 kg/m² (SD ± 5.7) vs. 36 kg/m² (SD ± 6.6) vs. 36.2 kg/m² (SD ± 6.6) Age 18 yr-60 yr: 90 (100%)	By ITT, significantly higher remission rate was reported with CBT-GSH: 46% vs. 18% vs. 13% - GSH vs. BWL: p=0.01 - GSH vs. Self-Monitoring: p=0.03 - BWL vs. Self-Monitoring: p=0.66 Significantly less self-reported binge episodes were reported with CBT-GSH at 12 wk: 3.8/mo vs. 7.3/mo vs. 6.8/mo - GSH vs. BWL: MD -3.5/mo (p=0.016)

				Age: 46.3 yr (SD ± 9) - 46 yr (SD ± 9.2) vs. 46 yr (SD ± 9.2) vs. 48 yr (SD ± 8.2) Gender - Female: 32 (86.5%) vs. 29 (76.3%) vs. 10 (66.7%) - Male: 5 (13.5%) vs. 9 (23.7%) vs. 5 (33.3%) Race - Caucasian: 32 (86.5%) vs. 23 (60.5%) vs. 14 (93.3%) - Black or African American: 2 (5.4%) vs. 6 (15.8%) vs. 1 (6.7%) - Other: 0 (0%) vs. 2 (5.3%) vs. 0 (0%) Ethnicity, Hispanic/Latino: 3 (8.1%) vs. 7 (18.4%) vs. 0 (0%)	- GSH vs. Self-Monitoring: MD -3/mo (p=0.019) Significantly greater treatment adherence was reported with CBT-GSH compared with BWL (p=0.036): 32 (87%) vs. 25 (66%) vs. 13 (87%) Attrition: 14% (5/37) vs. 34% (13/38) vs. 13% (2/15)
Grilo et al. (2013)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=48 CBT-SH + Usual Care 4 mo (N=24) Usual Care 4 mo (N=24)	Inclusion: BED; obese; BMI >=30 kg/m^2 Exclusion: BMI>=50 kg/m^2; over 65 years of age; current antidepressant therapy; current weight loss treatment; schizophrenia; bipolar disorder; current substance use disorder	BED: 48 (100%) - BED: 34 (70.83%) - Subclinical: 7 (29.17%) vs. 7 (29.17%) Obesity: 48 (100%) BMI >= 30 kg/m²: 48 (100%) BMI: 37.62 kg/m² (SD ± 4.79) Age: 45.8 yr (SD ± 11) - 45 yr (SD ± 11.8) vs. 46.5 yr (SD ± 10.2) Gender - Female: 21 (87.5%) vs. 17 (70.8%) - Male: 3 (12.5%) vs. 7 (29.2%) Race - Caucasian: 11 (45.8%) vs. 11 (45.8%)	Disease Response, Remission - 4 mo: 6 (25%) vs. 2 (8.3%) (OR 3.67, 95% CI 0.66 – 20.42, p=0.24) BMI - Baseline: 38.01 kg/m² (SD ± 5.36) vs. 37.22 kg/m² (SD ± 4.22) - 4 mo: 37.45 kg/m² (SD ± 5.34) vs. 37.42 kg/m² (SD ± 4.44) (MD 0.03 kg/m², p=0.4) BDI - Baseline: 14.57 units (SD ± 8.48) vs. 16.09 units (SD ± 8.61) BDI, Change - Baseline – 4 mo: -5.69 units (SD ± 11.02) vs. -4.13 units (SD ± 10.94) Overall Attrition: 0%

				- Black or African American: 6 (25%) vs. 11 (45.8%) Ethnicity - Hispanic/Latino: 2 (8.4%) vs. 1 (4.2%) - Other: 5 (20.8%) vs. 1 (4.2%)	
Grilo et al. (2014)	Design: RCT Setting: Outpatient: University-based medical health-care center in an urban setting Country: United States Funding: Government	Randomized N=104 CBT-SH 4 mo (N=25) Sibutramine 15 mg + CBT-SH 4 mo (N=26) Sibutramine 15 mg 4 mo (N=26) Placebo 4 mo (N=27) Follow-up: Baseline – 16 mo	Inclusion: 18-65 years of age; obese; BMI ≥ 30 and < 50 kg/m^2; BED Exclusion: Current use of antidepressant medication; current use of medication known to influence eating or weight; schizophrenia; bipolar disorder; current substance use disorder	BED: 104 (100%) Obesity: 104 (100%) BMI >= 30 kg/m²-< 50 kg/m²: 104 (100%) BMI: 38.3 kg/m² (SD ± 5.6) Age 18 yr-65 yr: 104 (100%) Age: 43.9 yr (SD ± 11.2) - 45.7 yr vs. 45.6 yr vs. 41.2 yr vs. 43.2 yr Gender - Female: 20 (80%) vs. 16 (61.5%) vs. 19 (73.1%) vs. 18 (66.7%) - Male: 5 (20%) vs. 10 (38.5%) vs. 7 (26.9%) vs. 9 (33.3%) Race - Caucasian: 12 (48%) vs. 10 (38.5%) vs. 13 (50%) vs. 12 (44.4%) - Black or African American: 6 (24%) vs. 10 (38.5%) vs. 8 (30.8%) vs. 12 (44.4%) - Other: 3 (12%) vs. 2 (7.7%) vs. 1 (3.8%) vs. 1 (3.7%) Ethnicity, Hispanic/Latino: 4 (16%) vs. 4 (15.4%) vs. 4 (15.4%) vs. 2 (7.4%)	Disease Response, Remission - 4 mo: 6 (24%) vs. 6 (23.1%) vs. 10 (38.5%) vs. 8 (29.6%) - 16 mo: 10 (40%) vs. 11 (42.3%) vs. 5 (19.2%) vs. 10 (37%) BMI – Baseline: 36.5 kg/m² (SD ± 5.3) vs. 37.8 kg/m² (SD ± 4.6) vs. 39.4 kg/m² (SD ± 6.6) vs. 39.3 kg/m² (SD ± 5.5) BMI, Change - Baseline – 16 mo: -1.1 kg/m² (SD ± 4.37, N=21) vs. -1.3 kg/m² (SD ± 3.89, N=23) vs. -0.1 kg/m² (SD ± 5.63, N=18) vs. 0.2 kg/m² (SD ± 4.43, N=23) BDI - Baseline: 17 units (SD ± 11.6) vs. 14 units (SD ± 7.2) vs. 12.8 units (SD ± 8.1) vs. 13.6 units (SD ± 11.2) BDI, Change - Baseline – 16 mo: -6.5 units (SD ± 14.23, N=21) vs. -3.7 units (SD ± 9.89, N=23) vs. -2.7 units (SD ± 11.18, N=18) vs. -6.1 units (SD ± 12.74, N=23) Attrition: 16% (4/25) vs. 15% (4/26) vs. 23% (6/26) vs. 48% (13/27)
Grilo et al. (2020a, 2020b)	Design: RCT; Follow-Up	Randomized N=191	Inclusion: BED; 18-60 years of age; BMI 30-55 kg/m^2	BED: 191 (100%)	Although stepped care and BWL did not significantly differ in binge-eating remission, there were significant

	Setting: NR Country: United States Funding: Government	BWL (Standard) 24 wk (N=39) Stepped Care 24 wk (N=152) Within stepped care, - BWL 4 wk > BWL + sibutramine or orlistat 20 wk for rapid responders (N=47) - sibutramine (N=22) - orlistat (N=25) - BWL 4 wk > BWL + placebo 20 wk for rapid responders (N=46) BWL 4 wk > CBT-GSH + sibutramine or orlistat 20 wk for non-rapid responders (N=24) - sibutramine (N=16) - orlistat (N=8) BWL 4 wk > BWL + placebo 20 wk for non-rapid responders (N=25) Follow-Up Period: Baseline – 76 wk	Exclusion: Concurrent treatment for eating/weight problems; taking contraindicated medications; uncontrolled medical conditions; pregnancy	BMI 30 kg/m²-55 kg/m²: 191 (100%) BMI: 37.5kg/m² (SD ± 5.7) vs. 39.4 kg/m² (SD ± 6.0) Weight: 103.5 kg (SD ± 21.4) vs. 111.4 kg (SD ± 22.7) Age 18 yr-60 yr: 191 (100%) Age: 48.4 yr (SD ± 9.5) Gender - Female: 32 (82.1%) vs. 104 (68.4%) - Male: 7 (17.9%) vs. 48 (31.6%) Race - Caucasian: 30 ((76.9%) vs. 120 (78.9%) - Black or African American: 5 (12.8%) vs. 23 (15.1%) - Asian: 1 (2.6%) vs. 1 (0.7%) - Other: 0 (0%) vs. 3 (2.0%) Ethnicity, Hispanic/Latino: 3 (7.7%) vs. 5 (3.3%)	differences within the stepped care conditions: - significantly higher remission rate (80.3%) in the medication groups than the placebo groups (80% vs. 58%, p=0.004) - among non-rapid responders, significantly higher remission rates with medications than placebo (p=0.002) Significant improvement in binge-eating frequency was reported across conditions: - Binge eating, Baseline->24 wk: 17.8->1.7/mo vs. 20.2->2.7/mo -within stepped care, significantly greater reduction in binge-eating frequency with medications than placebo (p=0.01) -among non-rapid responders, significantly greater reductions in binge eating with medications than placebo (p=0.004) ITT analyses showed remission rates between BWL and stepped care did not differ significantly at posttreatment (74.4% vs. 66.5%), 6-mo follow-up (38.2% vs. 33.3%), or 12-mo follow-up (44.7% vs. 41.0%). Attrition: 2.6% (1/39) vs. 12.5% (19/152)
Hildebrandt et al. (2017)	Design: RCT Setting: NR Country: United States	Randomized N=66 CBT-GSH 12 wk (N=33) CBT-GSH + Noom Monitor 12 wk (N=33)	Inclusion: BN or BED; binge eating and/or purging at least once weekly; over the age of 18 Exclusion: Co-occurring substance dependence; bipolar disorder; psychotic condition; psychotropic medication initiated 4	BED or BN: 66 (100%) BED: 19 (57.6%) vs. 18 (54.5%) BN: 14 (42.4%) vs. 15 (45.5%)	Objective Binge Eating Episodes - Baseline – 12 wk - 12.7->4.2/28 days vs. 12.4->1.96/28 days Treatment Discontinuation - Baseline – 12 wk

	Funding: Government	Those with BN N=12 vs. 12 Follow-up: Baseline – 36 wk	weeks prior to the screening visit; psychotropic medication dosage change 2 weeks prior to the baseline visit	Binge Eating >= 1/wk and/or Purging >= 1/wk: 66 (100%) AN, Lifetime: 16 (24.24%) BMI: 27.53 kg/m² (SD ± 5.61) Age > 18 yr: 66 (100%) Age: 32.11 yr (SD ± 10.82) - 33.88 yr (SD ± 11.97) vs. 30.33 yr (SD ± 9.39) Gender - Female: 28 (84.8%) vs. 27 (81.8%) - Male: 5 (15.2%) vs. 6 (18.2%) Race, Non-Caucasian: 10 (30.3%) vs. 15 (45.5%) Ethnicity, Hispanic/Latino: 6 (18.2%) vs. 5 (15.2%)	- Need for Additional Intervention: 0 (0%) vs. 2 (6.06%) - Dissatisfaction With Treatment: 0 (0%) vs. 1 (3.03%) Adverse Events - Baseline – 12 wk: 0 (0%) vs. 0 (0%) Attrition: 19% (6/33) vs. 19% (6/33)
Loeb et al. (2000)	Design: RCT Setting: Single Center: Rutgers Eating Disorders Clinic Country: United States Funding: NR	Randomized N=40 GSH 10 wk (N= 20) Unguided SH 10 wk (N= 20)	Inclusion: Female; BMI greater than or equal to 18 kg/m^2; at least one binge eating episode/wk over the past month Exclusion: More than one purging episode/wk over the past three months; actively suicidal; actively depressed; substance abuse; psychosis; BN, purging, threshold; psychotropic medication	BED: 33 (82.5%) BED, Subclinical: 3 (7.5%) BN, Purging, Subclinical: 2 (5%) BN, Purging, None: 2 (5%) Binge Eating >= 1/wk, In the Previous 1 mo: 40 (100%) BDI: 18.8 units (SD ± 8.22) BMI >= 18 kg/m²: 40 (100%) BMI < 25 kg/m²: 4 (10%)	Significantly greater percent change in binge eating was reported with GSH: -68% (SD ± 46) vs. -55% (SD ± 44) (MD -13 %, p=0.05) Binge Eating, Objective – Baseline->12 wk - 20.25->5.1/mo vs. 18.25->10.4/mo - 16.65->4.2 d/mo vs. 13.65->7.6 d/mo Binge Eating, Subjective - Baseline->12 wk - 16.05->7.75/mo vs. 10.7->7.1/mo - 9.85->5 d/mo vs. 8.6->4.8 d/mo BMI - Baseline->12 wk: 35.39->35.72 kg/m² vs. 36.15->36.12 kg/m²

				Age: 41.5 yr (SD ± 9.42) Gender, Female: 40 (100%) Race - Caucasian: 38 (95%) - Black or African American: 1 (2.5%) - Asian: 1 (2.5%)	Overall Attrition: 33% (13/40)
Peterson et al. (2020)	Design: RCT Setting: Multi-center Country: United States Funding: NR	Randomized N=112 CBT-GSH 17 wk (N= 56) Integrative Cognitive-Affective Therapy 17 wk (N= 56) Follow-up: Baseline – 89 wk	Inclusion: 18-65 years of age; BED Exclusion: BMI < 21 kg/m^2; history of psychotic symptoms or bipolar disorder; substance use disorder; medically or psychiatrically unstable (e.g., acute suicidality); purging behavior more than once per mo; current BN; medical condition impacting eating or weight (e.g., a thyroid condition); history of gastric bypass surgery; pregnant or lactating; receiving weight loss or eating disorder treatment; taking any medication impacting eating or weight (e.g., stimulants); psychotropic medication changes	BED: 112 (100%) BMI: 36.5 kg/m² (SD ± 8.9) vs. 33.7 kg/m² (SD ± 8.4) Age: 39.6 yr (SD ± 13.4) vs. 39.7 yr (SD ± 13.5) Gender - Female: 48 (85.7%) vs. 44 (78.6%) - Male: 8 (14.3%) vs. 12 (21.4%) Race Caucasian: 51 (91.1%) vs. 51 (91.1%)	Both groups showed significant reductions in objective binge-eating episodes during treatment, with modest increases at follow-up: Binge Eating, Objective – Baseline->End of Treatment->Follow-Up: 17.39 (SE ± 1.64)->1.75 (SE ± 0.486)->3.78 (SE ± 0.930) vs. 14.11 (SE ± 1.17)->1.20 (SE ± 0.383)->2.33 (SE ± 0.552) Binge-eating abstinence rates at end of treatment were 42.9% vs. 57.1% and at follow-up were 42.9% vs. 46.4%. Attrition: 27% (15/56) vs. 9% (5/56)
Wilson et al. (2010)	Design: RCT Setting: Outpatient: University clinics Country: United States Funding: Government	Randomized N=205 CBT-GSH 6 mo (N= 66) BWLT 6 mo (N= 64) IPT 6 mo (N= 75) Treatment Setting, Rutgers University subgroup (N= 31 vs. 32 vs. 37)	Inclusion: Aged 18 years and older; BMI 27-45 kg/m^2; overweight or obese; BED Exclusion: Current psychosis; bipolar disorder; current suicidal state; alcohol or drug dependence within the past 6 months; current participation in a weight-control program; taking medication that would affect weight	BED: 205 (100%) Overweight or Obesity: 205 (100%) BMI 27 kg/m²-45 kg/m²: 205 (100%) Age >= 18 yr: 205 (100%) Age: 50.3 yr (SD ± 13.6) vs. 46.2 yr (SD ± 10.9) vs. 48.7 yr (SD ± 11.2) Gender - Female: 54 (82%) vs. 57 (89%) vs. 64 (85%)	BMI- Baseline: 36.2 kg/m² (SD ± 4.3) vs. 36.8 kg/m² (SD ± 5.5) vs. 36.3 kg/m² (SD ± 5.1) BMI - 6 mo: 36.1 kg/m² (SD ± 4.4) vs. 35.4 kg/m² (SD ± 5.7) vs. 35.9 kg/m² (SD ± 5.3) - CBT-GSH vs. BWLT: SMD 0.741 - IPT vs. BWLT: SMD 0.48 - CBT-GSH vs. IPT 6 mo: SMD 0.15 BMI - 30 mo: 35.7 kg/m² (SD ± 5) vs. 36.3 kg/m² (SD ± 6.2) vs. 36.1 kg/m² (SD ± 5.5) - BWLT vs. CBT-GSH: SMD 0.52

		Treatment Setting, Washington University in St. Louis subgroup (N= 35 vs. 32 vs. 38) Follow-up: Baseline – 30 mo		- Male: 12 (18%) vs. 7 (11%) vs. 11 (15%) Race - Caucasian: 54 (82%) vs. 56 (88%) vs. 58 (77%) - Black or African American: 7 (11%) vs. 7 (11%) vs. 13 (17%) - Native American/Alaska Native: 0 (0%) vs. 0 (0%) vs. 1 (1%) Ethnicity, Hispanic/Latino: 5 (8%) vs. 1 (2%) vs. 3 (4%)	- BWLT vs. IPT: SMD 0.29 - IPT vs. CBT-GSH: SMD 0.2 Weight – Baseline->6 mo: 100.3->100 kg vs. 103.5->99.8 kg vs. 100.4->99.1 kg Weight, Decrease >= 5 % - Baseline – 6 mo: 10 (15%) vs. 26 (41%) vs. 11 (15%) - BWLT vs. CBT-GSH: OR 3.9 - BWLT vs. IPT: OR 3.9 Disease Response, Remission - 30 mo: 41 (62.1%) vs. 28 (43.9%) vs. 51 (67.9%) - CBT-GSH vs. BWLT: OR 2.3 - IPT vs. BWLT: OR 2.6 - IPT vs. CBT-GSH: OR 1.2 Attrition: 30% (20/66) vs. 28% (18/64) vs. 7% (5/75)6 mo
Wyssen et al. (2021)	Design: RCT Setting: Outpatient: University of Fribourg Country: Switzerland Funding: Non-profit	Randomized N=63 Internet-Based CBT-GSH NR (N=24) WLC 4 wk > CBT-GSH NR (N=39) - Standard WLC 4 wk > CBT-GSH NR (N= 19) Positive expectation induction WLC 4 wk > CBT-GSH NR (N=20)	Inclusion: BED; 18-70 years of age Exclusion: Pregnancy; another serious psychological or medical condition warranting priority treatment	BED: 63 (100%) Age: 36.5 yr (SD ± 10.52) vs. 38.5 yr (SD ± 11.62) vs. 36.8 yr (SD ± 9.52) Gender - Female: 55 (87%) - Male: 8 (13%) Race: NR	WBQ Weekly Binges – Baseline: 2.79 (SD ± 2.23) vs. 4.11 (SD ± 2.85) vs. 3.30 (SD ± 1.95) WBQ Weekly Binges – End of Treatment: 1.75 (SD ± 1.34) vs. 2.07 (SD ± 2.89) vs. 1.14 (SD ± 1.46) WBQ Weekly Binges – 6 mo Follow-Up: 0.92 (SD ± 0.95) vs. 1.62 (SD ± 3.15) vs. 0.69 (SD ± 0.85) BMI - Baseline: 28.71 kg/m² (SD ± 9.17) vs. 31.90 kg/m² (SD ± 8.31) vs. 32.58 kg/m² (SD ± 8.08) BMI - End of Treatment: 29.13 kg/m² (SD ± 9.05) vs. 31.49 kg/m² (SD ± 8.84) vs. 31.93 kg/m² (SD ± 8.26) (Total N=46) BMI – 6 mo Follow-Up: 28.32 kg/m² (SD ± 6.65) vs. 29.74 kg/m² (SD ± 8.17) vs. 31.17 kg/m² (SD ± 7.94)

					Attrition at follow-up: 46% (11/24) vs. 26% (5/19) vs. 30% (7/20)

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; BWL=behavioral weight loss; BWLT=behavioral weight loss treatment; CBT=cognitive-behavioral therapy; CBT-GSH=cognitive-behavioral therapy guided self-help; CBT-SH=cognitive-behavioral therapy self-help; CI=confidence interval; d=day; GSH=guided self-help; IPT=interpersonal psychotherapy; ITT=intention-to-treat; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; SE=standard error; SH=self-help; SMD=standardized mean difference; WBQ=Weekly Binges Questionnaire; wk=week; WLC=wait-list control; yr=year

Dialectical Behavior Therapy

Carter et al. (2020)	Design: RCT Setting: NR Country: NR Funding: Academic	Randomized N=71 DBT-GSH 12 wk (N= 24) DBT-USH 12 wk (N= 24) Self-Esteem Unguided Self-Help 12 wk (N= 23) Follow-up: Baseline – 24 wk	Inclusion: 19-65 years of age; BED; BMI >=18.5 kg/m^2 Exclusion: Current treatment for binge eating; major medical illness (e.g., diabetes); current pregnancy; exceeding a cut-off of 5 on the Drug Abuse Screening Test or 16 on the Alcohol Use Disorders Identification Test	BED: 74 (100%) BMI: 36.49 kg/m² (SD ± 7.05) vs. 36.11 kg/m² (SD ± 10.44) vs. 39.50 kg/m² (SD ± 10.60) Age: 40.21 yr (SD ± 11.46) vs. 40.88 yr (SD ± 12.57) vs. 41.04 yr (SD ± 10.73) Gender - Female: 22 (91.7%) vs. 22 (91.7%) vs. 22 (95.7%) - Male: 2 (8.3%) vs. 2 (8.3%) vs. 1 (4.3%) Race - Caucasian: 22 (92%) vs. 22 (92%) vs. 22 (96%) Other: 2 (8%) vs. 2 (8%) vs. 1 (4%)	Both groups showed a significant decrease in binge frequency from baseline to post-treatment (p<0.0005) and the decrease was maintained at follow-up in all three conditions (p=0.657). There was a statistically significant difference in the proportion of individuals who were binge eating at or above diagnostic levels at posttreatment (i.e., 4 or more episodes in the past 28 days) between the DBT-GSH and self-esteem unguided self-help conditions (25% vs. 57%, p=0.028). Binge Eating, Days – Baseline->End of Treatment->Follow-Up: 13.5->3.1->4.8 vs. 11.8->3.2->2.7 vs. 12.8->5.0->6.3 Binge Eating, Episodes – Baseline->End of Treatment->Follow-Up: 16.2->3.4-> 3.8 vs. 13.6->3.6-> 2.8 vs. 21.7->4.9-> 5.2 Overall Attrition at End of Treatment: 35% (25/71)
Dastan et al. (2020)	Design: RCT Setting: Outpatient	Randomized N=40 DBT 20 wk (N= 20)	Inclusion: Female; 18-50 years of age; BED; BMI >=30 kg/m^2 Exclusion: Other eating disorders; current pregnancy; substance	BED: 40 (100%) Age 18-50 yr: 40 (100%) Gender, Female: 40 (100%)	DBT group showed lower BMI than the control group at post-treatment (p<0.001): BMI – Baseline->End of Treatment:

	Country: Iran Funding: NR	Control 20 wk (N= 20)	users; severe physical illnesses; receiving any psychotherapy or dietary regimens	Race: NR	34.20->29.10 kg/m² vs. 35.25->34.96 kg/m²
Klein et al. (2013)	Design: RCT Setting: NR Country: United States Funding: Academic	Randomized N=36 Group-Based DBT 16 wk (N= 22) Diary Card Self-Monitoring + Individual Sessions 16 wk (N= 14) Among completers, BED and Obesity subgroup (N= 4 vs. 6) BED and not Obesity subgroup (N= 0 vs. 4) BN subgroup (N= 4 vs. 2) Purging subgroup (N= 3 vs. 2)	Inclusion: BN or BED Exclusion: Borderline personality disorder; BMI < 18.5 kg/m^2	BED or BN: 36 (100%) - BED: 10 (45.4%) vs. 9 (64.3%) - BED, Subclinical: 5 (22.7%) vs. 1 (7.1%) - BN: 6 (27.2%) vs. 4 (28.6%) - BN, Subclinical: 1 (4.5%) vs. 0 (0%) Eating, Binge Eating, Duration: 16.02 yr (SD ± 16.18) vs. 14.65 yr (SD ± 13.95) Age: 33.05 yr (SD ± 13.73) (N= 20) - 36.67 yr (SD ± 14.95) vs. 32.14 yr (SD ± 11.34) Gender, Female: 36 (100%) Race, Non-Caucasian: 2 (9%) vs. 5 (36%)	Eating, Binge Eating - Baseline: 1.96/wk (SD ± 1.26, N= 8) vs. 3.54/wk (SD ± 2.25, N= 12) - 16 wk: 0/wk (N= 8) vs. 1.64/wk (SD ± 1.62, N= 12) (RR 0.6849, p=0.311) Purging – Baseline->16 wk - Purging subgroup: 3.83->1.67/wk vs. 1->0/wk BED and BN, None - 16 wk: 4 (50%, N= 8) vs. 3 (27%, N= 12) BED or BN, Subclinical - 16 wk: 4 (50%, N= 8) vs. 3 (27%, N= 12) Attrition: 64% (14/22) vs. 14% (2/14)
Safer et al. (2010)	Design: RCT Setting: NR Country: United States Funding: NR	Randomized N=101 DBT Adapted for Binge Eating 21 wk (N= 50) Active Comparison Group Therapy 21 wk (N= 51) Follow-up: Baseline – 73 wk	Inclusion: Aged 18 and older; BED Exclusion: BMI less than 17.5 kg/m^2; concurrent psychotherapy treatment; unstable dosage of psychotropic medications over the 3 months prior to initial assessment; regular use of other compensatory behaviors over the past 6 months; psychosis; current alcohol or drug dependence; current alcohol or drug abuse; severe depression with recent suicidality; current use of weight-altering medications; current use of phentermine; severe medical	BED: 101 (100%) BMI: 36.38 kg/m² (SD ± 8.62) Age >= 18 yr: 101 (100%) Age: 52.2 yr (SD ± 10.6) - 51.9 yr (SD ± 11.6) vs. 52.35 yr (SD ± 9.52) Gender - Female: 43 (86%) vs. 43 (84%) - Male: 7 (14%) vs. 8 (16%) Race - Caucasian: 40 (80%) vs. 37 (73%)	There was significantly lower dropout for DBT (p<0.001) Eating, Binge Eating – Baseline->21 wk->73 wk: 15.9->1.48->2.76 d/mo vs. 15.9->4.62->3.14 d/mo Binge Eating, Abstinence – 21 wk: 32 (64%) vs. 17 (33.3%) Weight – Baseline->21 wk->73 wk: 216.91->212.61->213.23 lbs vs. 224.03->221.87->221.61 lbs

			condition affecting weight; severe medical condition affecting appetite; Insulin dependent diabetes; cancer requiring active chemotherapy; imminently planning gastric bypass surgery; undergoing gastric bypass surgery; current pregnancy; current breast feeding; regular use of purging over the past 6 months	- Asian: 2 (4%) vs. 3 (6%) - Black or African American: 0 (0%) vs. 3 (6%) Ethnicity - Hispanic/Latino : 8 (16%) vs.5 (10%) Ethnicity or Race, Unknown: 0 (0%) vs. 3 (6%)	BMI – Baseline->21 wk->73 wk: 35.84->35.13->35.29 kg/m² vs. 36.9->36.65->36.45 kg/m² BDI – Baseline->21 wk: 17.94->9.1 units (SD ± 9.21) vs. 15.27->10.84 units (SD ± 6.86) Attrition: 4% (2/50) vs. 33% (17/51)
Telch et al. (2001)	Design: RCT Setting: NR Country: United States Funding: Government	Randomized N=44 DBT 20 wk (N= 22) WLC 20 wk (N= 22) Follow-up: Baseline – 46 wk	Inclusion: Female; BED; 18-65 years of age Exclusion: Current involvement in psychotherapy; current involvement in weight loss treatment; use of psychotropic medications; current substance abuse; current substance dependence; current suicidality; current psychosis; pregnancy	BED: 44 (100%) Eating, Binge Eating, Duration: 29.2 yr (SD ± 11.7) BN, Lifetime: 3 (6%) BDI: 12.8 units (SD ± 7.4, N= 18) vs. 13.8 units (SD ± 9.1, N= 16) BMI: 36.4 kg/m² (SD ± 6.6) Age 18 yr-65 yr: 44 (100%) Age: 50 yr (SD ± 9.1) Gender, Female: 44 (100%) Race, Caucasian: 41 (94%)	Significantly greater improvement on BES was reported with DBT at 20 wk: WLC vs. DBT – SMD 1.16 (p=0.001) BES – Baseline->20 wk: 28.8->15.7 (N= 18) vs. 31.8->28.2 units (N= 16) Binge Eating, Abstinence – 20 wk: 16 (89%, N= 18) vs. 2 (12.5%, N= 16) Weight – Baseline->20 wk: 214.7->209.2 lbs (N= 18) vs. 223.4->223.8 lbs (N= 16) WLC vs. DBT at 20 wk: SMD 0.33 (p=0.13) Attrition: 18% (4/22) vs. 27% (6/22)

Abbreviations: BDI=Beck Depression Inventory; BED=binge-eating disorder; BES=Binge Eating Scale; BMI=body mass index; BN=bulimia nervosa; DBT=dialectical behavior therapy; DBT-GSH=dialectical behavior therapy guided self-help; DBT-USH=dialectical behavior therapy unguided self-help; NR=not reported; RCT=randomized controlled trial; RR=risk ratio; SD=standard deviation; SMD=standardized mean difference; wk=week; WLC=wait-list control; yr=year

Other Psychotherapies

Alfonsson et al. (2015)	Design: RCT Setting: Obesity Clinic	Randomized N=100 Behavioral Activation 10 wk (N=50)	Inclusion: Obese; BMI > 30 kg/m^2; BED Exclusion: Severe mental illness; schizophrenia; suicidal ideation; untreated bipolar disorder;	BED: 100 (100%) Obesity: 100 (100%) BMI > 30 kg/m²: 100 (100%)	There was significantly more dropout with behavioral activation (p=0.01)

	Country: Sweden Funding: NR	WLC 10 wk (N=50) Follow-up: Baseline – 36 wk ITT (N=96) - 50 vs. 46	ongoing alcohol abuse; ongoing drug abuse	BMI: 41.17 kg/m² (SD ± 5.32, N=96) - 41.26 kg/m² (SD ± 5.02) vs. 41.07 kg/m² (SD ± 5.66, N=46) Age: 44.34 yr (SD ± 10.74, N=96) - 45.5 yr (SD ± 10.71) vs. 44.17 yr (SD ± 10.9, N=46) Gender - Female: 46 (92%) vs. 44 (95.7%, N=46) - Male: 4 (8%) vs. 2 (4.3%, N=46) Race: NR	Disease Response, Recovery - Baseline – 10 wk: 7 (21%, N=34) vs. NR (N=38) EDE-I, BED, Diagnosis, None - Baseline – 10 wk: 10 (29%, N=34) vs. 10 (26%, N=38) Attrition: 32% (16/50) vs. 26% (12/50)
Allen et al. (1999)	Design: RCT Setting: Education System: University of North Carolina; University of Colorado Country: United States Funding: NR	Randomized N=29 Appetite Awareness Training 8 wk (N=15) WLC 8 wk (N=14) Current Analysis (N=20) - 11 vs. 9	Inclusion: Female; BED Exclusion: Below 90% of IBW; over 160% of IBW; history of anorexia; currently in treatment for eating difficulties; significantly underweight; purging; BN	BED: 29 (100%) %IBW: 122.82% (SD ± 22.86, N=11) vs. 116.5% (SD ± 21.98, N=9) Age: NR Gender, Female: 29 (100%) Race: NR	Significant improvement in BDI score was reported with appetite awareness training at 8 wk: BDI - Baseline: 17.09 units (SD ± 5.05, N=11) vs. 16.22 units (SD ± 8.54, N=9) - 8 wk: 6.91 units (SD ± 3.21, N=11) vs. 12.33 units (SD ± 5.15, N=9) (MD -5.42 units, p<0.03) Binge Eating - Baseline: 4.86/wk (SD ± 2, N=11) vs. 3.91/wk (SD ± 2.28, N=9) Binge Eating, Change - Baseline – 8 wk: -4.14/wk (SD ± 1.46, N=11) vs. 1.04/wk (SD ± 1.75, N=9) Attrition: 27% (4/15) vs. 36% (5/14)
Brennan et al. (2020)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=72 Yoga 8 wk (N=36) WLC 8 wk (N=36)	Inclusion: Age above 18 years; BN or BED; no or limited yoga experience Exclusion: Suicidal ideation, psychosis, or substance abuse, or a pre-existing diagnosis of BPD	BN or BED: 53 (100%) - BN: 40 (75.5%) - BED: 13 (24.5%) BMI - Overweight: 8 (15.1%) - Obese: 17 (32.1%)	Compared to WLC, yoga decreased binge eating frequency: Binge Eating Episodes: 11.46->5.11/28d vs. 12.92->12.11/28d

		Current Analysis (Completers; N=53): 26 vs. 27		Age > 18 yr: 53 (100%) Gender, Female: 53 (100%) Race - Caucasian: 38 (72%) - Asian: 9 (17%) Other: 6 (11%)	Binge Eating Days: 11.63->4.58 d/28d vs. 11.70->10.60 d/28d Attrition: 28% (10/36) vs. 25% (9/36)
Compare et al. (2013a, 2013b)	Design: Non-Randomized Controlled Trial; Post-hoc Analysis Setting: Outpatient: Outpatient department for the treatment of eating disorders Country: Italy Funding: NR	Total N=189 EFT 5 mo (N=63) EFT+ DC 3 mo > DC 5 mo (N=63) Follow-up: Baseline – 11 mo Current Analysis (N=189) - 63 vs. 63 vs. 63	Inclusion: 35-60 years of age; BED; BMI of 30 kg/m^2 or greater Exclusion: Concurrent treatment for eating disorder; concurrent treatment for weight; concurrent treatment for psychiatric illness; severe psychosis requiring treatment; severe bipolar disorder requiring treatment	BED, Severe: 189 (100%) Binge Eating: 17.5/mo (SD ± 3.33) vs. 17.5/mo (SD ± 1.33) vs. 16.9/mo (SD ± 3.85) BMI >= 30 kg/m²: 189 (100%) BMI: 33 kg/m² (SD ± 1.6) vs. 33.6 kg/m² (SD ± 2.6) vs. 32.3 kg/m² (SD ± 1.3) Age 35 yr-60 yr: 189 (100%) Age: 50.8 yr (SD ± 6) vs. 51.1 yr (SD ± 4.1) vs. 50.4 yr (SD ± 4.7) Gender - Female: 42 (66.7%) vs. 26 (41.3%) vs. 26 (41.3%) - Male: 21 (33.3%) vs. 37 (58.7%) vs. 37 (58.7%) Race: NR	Significantly greater decreases in BES were reported with EFT and EFT + DC compared with DC alone. BES – Baseline: 34.2 units (SD ± 4.2) vs. 33.8 units (SD ± 4.6) vs. 32.9 units (SD ± 3.9) BES - 5 mo: 17 units (SD ± 2.9, N=55) vs. 15.1 units (SD ± 1.9) vs. 32.3 units (SD ± 1.6, N=46) - EFT vs. DC: MD -15.3 units (p<0.016) - EFT + DC > EFT vs. DC: MD -17.2 units (p<0.016) - EFT + DC > EFT vs. EFT: MD -1.9 units (p<0.016) BES <= 16 units - 11 mo: 29 (46%) vs. 45 (71%) vs. 0 (0%) Disease Response, Remission - Baseline – 11 mo: 27 (42.9%) vs. 39 (61.9%) vs. 13 (20.6%) - EFT vs. DC: OR 2.93 (95% CI 1.33 – 6.47) - DC + EFT vs. DC: OR 6.66 (95% CI 2.97 – 14.94) - DC + EFT vs. EFT: OR 2.27 (95% CI 1.1 – 4.67)

					Weight – Baseline: 99.9 kg vs. 97.9 kg vs. 99.1 kg Weight, % Change - Baseline – 11 mo: -11.4% vs. -13.2% vs. -5.4% Attrition: 13% (8/63) vs. 0% (0/63) vs. 27% (17/63)
Glisenti et al. (2021)	Design: RCT Setting: NR Country: Australia Funding: NR	Randomized N=21 EFT 12 wk (N=11) WLC 12 wk (N=10) Follow-up: Baseline – 24 wk Current Analysis (N=20) 10 vs. 10	Inclusion: BED; 18-65 years of age Exclusion: Current psychosis, intellectual disability, high suicide risk, drug or alcohol abuse; concurrent treatment for obesity; pregnancy; the presence of AN or BN.	BED: 21 (100%) Age: 44.5 yr (SD ± 11.9) Gender - Female: 17 (81%) - Male: 4 (19%) Race: NR	EFT intervention group showed significantly greater reductions in objective binge episodes (p=0.017) and binge episode days (p=0.01): Binge Eating Episodes, Objective – Baseline->End of Treatment: 6.9->2.9 vs. 5.1-> 5.1 Binge Eating Days, Objective – Baseline->End of Treatment: 4.2->1.5 vs. 4.4-> 4.7 Attrition: 9% (1/11) vs. 0% (0/10)
Lewer et al. (2017)	Design: RCT Setting: Single Center: Mental Health Research and Treatment Center of the Ruhr-Universität Bochum Country: Germany Funding: Non-profit and government	Randomized N=36 Cognitive-Behavioral Exposure Based Body Image Therapy 10 wk (N= 15) WLC 10 wk (N= 21) Current Analysis (N=34) - 15 vs. 19	Inclusion: BED; female; 18-60 years of age; overweight; BMI > 25 kg/m^2 Exclusion: Suffered from a personality disorder; displayed suicidal tendencies; showed deliberate self-harm behavior; pregnant; personality disorders; current psychotherapy; intake of psychotropic drugs	BED: 36 (100%) Overweight: 36 (100%) BMI > 25 kg/m²: 36 (100%) Age 18 yr-60 yr: 36 (100%) Gender, Female: 36 (100%) Race: NR	BMI – Baseline: 31.98 kg/m² (SD ± 4.7) vs. 36.8 kg/m² (SD ± 5.08) BMI, Change - Baseline – 10 wk: -0.09 kg/m² (SD ± 3.67) vs. 0.58 kg/m² (SD ± 3.95) Attrition: 0% (0/15) vs. 10% (2/21)
Tasca et al. (2019)	Design: RCT Setting: NR Country: NR	Randomized N=85 CBT-USH 10 wk > Group Psychodynamic Interpersonal	Inclusion: BED Exclusion: Pregnancy; other psychotherapies/weight loss programs; comorbid bipolar,	BED: 85 (100%) BMI: 34.8 kg/m² (SD ± 7.25) vs. 37.5 kg/m² (SD ± 9.31)	CBT-USH resulted in a significant reduction in binge eating frequency and in eating disorder psychopathology. Binge Eating Episode

	Funding: NR	Psychotherapy 16 wk (N=39) CBT-USH 10 wk > No Treatment 16 wk (N=46)	psychotic, or substance use disorders	Age: 44.97 yr (SD ± 12.70) vs. 42.98 yr (SD ± 12.80) Gender, Female: 33 (84.6%) vs. 40 (87%) Race, Caucasian: 37 (94.9%) vs. 41 (89.1%)	- Pre CBT-USH: 13.30 (SD ± 6.87, N=135) - Pre-Randomization: 6.13 (N=38) vs. 5.84 (N=43) - End of Treatment: 6.09 (N=31) vs. 5.90 (N=31) - 6 mo: 5.50 (N=28) vs. 6.28 (N=28) Compared with the control condition, receiving group psychodynamic interpersonal psychotherapy resulted in 1.04 greater odds of changing from non-abstinent to abstinent status at 6 mo post-treatment: Pre-Treatment->End of Treatment->6 mo: 3 (7.9% N=38)->3 (9.4%, N=32)->7 (25.0% N=28) vs. 10 (23.3% N=43)->10 (32.3% N=31)->6 (21.4% N=28) Attrition: 33% (13/39) vs. 24% (11/46)
Reeves et al. (2001)	Design: RCT Setting: NR Country: NR Funding: NR	Randomized N=98 Behavioral Self-Management 6 mo (N= 59) WLC 6 mo (N= 39)	Inclusion: 25-50 years of age; 30-90 lbs overweight; purging in the past 6 mo; BES score >20 Exclusion: Prescribed medications; using tobacco products; consuming more than 2 alcoholic beverages per day; receiving treatment for psychological problems or medical problems	Weight: 197 lbs (SD ± 21, N= 46) vs. 191 lbs (SD ± 23, N= 36) Age 25 yr-50 yr: 98 (100%) Gender, Female: 39 (100%) Race - Caucasian: 42 (91%, N= 46) vs. 34 (94%, N= 36) - Mexican American: 3 (7%, N= 46) vs. 0 (0%, N= 36) - Black or African American: 1 (2%, N= 46) vs. 2 (6%, N= 36)	Behavioral self-management was associated with significantly less binge eating d/wk at 6 mo:(MD -0.7 d/wk, p=0.03). Binge Eating – Baseline->6 mo: 3.2->1 d/wk (N= 46) vs. 2.8->1.7 d/wk (N= 36) Binge Eating, Change - Baseline – 6 mo: -2.2 d/wk (N= 46) vs. -1.1 d/wk (N= 36) Weight - 6 mo: 195 lbs (SD ± 22, N= 46) vs. 191 lbs (SD ± 25, N= 36) (MD 4 lbs, p=0.47) Attrition: 22% (13/59) vs. 8% (3/39)
Wadden et al. (2011); Chao et	Design: Prospective Cohort Study; Follow-up	Total N= 208	Inclusion: At least 18 years of age; BMI >=40 kg/m^2 or >=35 kg/m^2 in	BMI >= 40 kg/m² or BMI >= 35 kg/m² and Comorbidities: 85 (100%)	Significantly greater weight reduction was reported with bariatric surgery.

al. (2016)	Setting: Single Center: Hospital of the University of Pennsylvania Country: United States Funding: Government; product donation by industry	Those with BED N=119 Bariatric Surgery 0 wk (N= 62) Lifestyle Modification 52 wk (N= 57) Current Analysis (N= 85) - 36 vs. 49 Surgery, Bariatric, Laparoscopic subgroup (N=14) Surgery, Bariatric, Bypass subgroup (N=19) Follow-up: Baseline – 24 mo	the presence of a comorbid condition Exclusion: Pregnancy; lactation; use of medications known to affect body weight; steroids; weight loss ≥5% of initial weight in the prior 6 months; use of anorectic agents in the prior 6 months; BDI score >28; type 1 diabetes	Eating, Binge Eating: 36 vs. 49 Diabetes Mellitus, Type 2: 11 (30.6%, N= 36) vs. 5 (10.2%, N= 49) BMI: 48.9 kg/m² (SD ± 6.6, N= 36) vs. 44.3 kg/m² (SD ± 4.9, N= 49) Age >= 18 yr: 144 (100%) Age: 47 yr (SD ± 9.6, N= 36) vs. 43.8 yr (SD ± 9.8, N= 49) Gender - Female: 26 (72.2%, N= 36) vs. 39 (79.6%, N= 49) - Male: 10 (27.8%, N= 36) vs. 10 (20.4%, N= 49) Race - Caucasian: 27 (75%, N= 36) vs. 18 (36.7%, N= 49) - Black or African American: 8 (22.2%, N= 36) vs. 26 (53.1%, N= 49) - Other: 0 (0%, N= 36) vs. 2 (4.1%, N= 49) Ethnicity, Hispanic/Latino: 1 (2.8%, N= 36) vs. 3 (6.1%, N= 49)	Weight - Baseline: 138.7 kg (SD ± 24, N= 36) vs. 125.8 kg (SD ± 20.3, N= 49) Weight, Decrease - Baseline – 12 mo - >= 5 %: 27 (75%, N= 36) vs. 20 (40.8%, N= 49) (p=0.004) - >= 10 %: 24 (66.7%, N= 36) vs. 13 (26.5%, N= 49) (p<0.001) - >= 20 %: 16 (44.4%, N= 36) vs. 6 (12.2%, N= 49) (p<0.001) Weight, % Change - Baseline – 2 mo: -9.93% (N= 36) vs. -4.82% (N= 49) (MD -5.11 % (p<0.001) - Baseline – 12 mo: -22.1% (N= 36) vs. -10.3% (N= 49) (MD -11.8 % (p<0.001) - Baseline – 24 mo: -18.6% vs. -5.6% (MD -13 % (p<0.001) - Bariatric, Bypass subgroup: -24.2% - Bariatric, Laparoscopic subgroup: -9.1% Disease Response, Remission - 6 mo: 34 (94.4%, N= 36) vs. 43 (87.8%, N= 49) - 12 mo: 33 (91.7%, N= 36) vs. 42 (85.7%, N= 49) Attrition: 41% (21/51) vs. 18% (9/51)

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BES=Binge Eating Scale; BMI=body mass index; BN=bulimia nervosa; CBT-USH=cognitive-behavioral therapy unguided self-help; CI=confidence interval; d=day; DC=dietary counseling; EDE=Eating Disorder Examination; EFT=Emotionally Focused Therapy; IBW=ideal body weight; ITT=intention-to-treat; MD=mean difference; mo=month; NR=not reported; OR=odds ratio; RCT=randomized controlled trial; SD=standard deviation; wk=week; WLC=wait-list control; yr=year

Pharmacotherapies

Anticonvulsants

Topiramate

Claudino et al. (2007)	Design: RCT Setting: 4 university centers Country: Brazil Funding: Industry	Randomized N=73 Placebo + CBT 21 wk (N=36) Topiramate 200-300mg (up-titrate) + CBT 21 wk (N=37)	Inclusion: Obese; 18-60 years of age; BMI >= 30 kg/m^2; BED; Score of > 17 on the BES Exclusion: Clinically significant schizophrenia, major affective disorders, or alcohol or drug abuse; unstable schizophrenia, major affective disorders, or alcohol or drug abuse; high potential suicide risk; concurrent use of antipsychotics, cyproheptadine, antiepileptics, systemic steroids, or antiobesity agents; psychotherapy for weight loss within 3 months	BED: 73 (100%) BES > 17 units: 73 (100%) Obesity: 73 (100%) BMI >= 30 kg/m²: 73 (100%) BMI: 37.4 kg/m² (SD ± 3.5) vs. 37.4 kg/m² (SD ± 4.9) Weight: 98.4 kg (SD ± 10.9) vs. 96.6 kg (SD ± 16.7) Age 18 yr-60 yr: 73 (100%) Age: 35.4 yr (SD ± 10.7) vs. 41.1 yr (SD ± 9.9) Gender - Female: 34 (94.4%) vs. 36 (97.3%) - Male: 2 (5.6%) vs. 1 (2.7%) Race, Caucasian:19 (52.8%) vs. 23 (62.1%)	Amount and rate of weight reduction was greater with topiramate: -0.9 kg with CBT vs. -6.8 kg with topiramate; 11.5% vs. 36.7% lost more than 10% of body weight (p=0.05). More patients with topiramate achieved remission (61.1% vs. 83.8% (p=0.03)) though reductions in binge frequency did not differ. - Baseline: 3.8/wk (SD ± 1.5) vs. 4.7/wk (SD ± 3.3) - Baseline: 3.4 d/wk (SD ± 1.3) vs. 4.2 d/wk (SD ± 3.4) - % Change - Baseline – 21 wk: -92.9% (SD ± 17.7, N=24) vs. -99.5% (SD ± 2.6, N=29) (MD 6.6 %, p=0.08) BDI - Baseline: 15.9 units (SD ± 9.4) vs. 16.8 units (SD ± 8.3) BDI, Change - Baseline – 21 wk: -6.7 units (SD ± 11.26) vs. -5.9 units (SD ± 10.48) (MD -0.66 units) Study withdrawal rates did not differ significantly but topiramate had more paresthesia and dysgeusia and placebo had more insomnia. Attrition: 28% (10/36) vs. 19% (7/37)
McElroy et al. (2003)	Design: RCT Setting: Outpatient: University of	Randomized N=61	Inclusion: 18-60 years of age; BED; obese; BMI >= 30 kg/m^2; score >=15 on the Yale-Brown Obsessive Compulsive Scale	BED: 61 (100%) Obesity: 61 (100%)	Topiramate was associated with significant reduction in binge eating. Eating, Binge Eating - Baseline:

	Cincinnati Medical Center Country: United States Funding: Industry	Topiramate 25-600 mg (up-titrate) 14 wk (N= 30) Placebo 14 wk (N= 31) Follow-up: Baseline – 16 wk	Exclusion: Substance use disorder within the past 6 months; unstable bipolar disorder within the past 3 months; treatment with psychoactive medications within 2 weeks; clinically significant suicidality; current psychiatric disorder that could interfere with diagnostic assessment, treatment, or study adherence; stimulants; antidepressants; carbonic anhydrase inhibitors; treatment with any medication that might adversely interact with topiramate; treatment with any medication that might obscure the action of topiramate; treatment with an experimental drug within 30 days of random assignment; treatment with an experimental device within 30 days of random assignment	BMI >= 30 kg/m²: 61 (100%) Age 18 yr-60 yr: 61 (100%) Age: 40.9 yr (SD ± 8.2) vs. 40.7 yr (SD ± 9.1) Gender - Female: 53 (86.9%) - Male: 8 (13.1%) Race: NR	- 5.3/wk (SD ± 2.8) vs. 6.3/wk (SD ± 2.8) - 4.3 d/wk (SD ± 1.8) vs. 4.8 d/wk (SD ± 1.8) Binge Eating, % Change - Baseline – 14 wk - episodes/wk: -94% vs. -46% (MD -48 %, p<0.02) - d/wk: -93% vs. -46% (MD -47 %, p<0.02) Weight - Baseline: 120.4 kg (SD ± 18.8) vs. 123.4 kg (SD ± 24.4) Weight, Change - Baseline – 14 wk: -5.9 kg vs. -1.2 kg (MD -3.2 kg) Disease Response, Remission - Baseline – 14 wk: 18 (64%, N= 28) vs. 9 (30%, N= 30) Topiramate was significantly more associated with the following adverse events: - Confusion: 5 (17%) vs. 0 (0%) (p<0.05) - Dysgeusia: 6 (20%) vs. 0 (0%) (p<0.05) - Paresthesia: 21 (70%) vs. 3 (10%) (p<0.05) Study Withdrawal - Baseline – 14 wk - Adverse Events: 6 (20%) vs. 3 (9.68%) - Lack of Efficacy: 1 (3.33%) vs. 2 (6.45%) Attrition: 47% (14/30) vs. 39% (12/31)
McElroy et al. (2007b)	Design: RCT Setting: Multi-center, outpatient: Private	Randomized N=407	Inclusion: Moderate to severe BED; >=3 binge days/wk during	BED, Moderate to Severe: 407 (100%) Binge Eating >= 3 d/wk: 407 (100%)	Topiramate was associated with decreased binge eating. Binge Eating - Baseline:

	practice and university settings Country: United States Funding: Industry	Topiramate 25-400 mg 16 wk (25 mg induction) (N=204) Placebo 16 wk (N=203) ITT (N=401) - 199 vs. 202	screening; 18-65 years of age; obesity; BMI ≥30 and ≤50 kg/m^2 Exclusion: Current major organic psychiatric disease; lifetime history of major organic psychiatric disease; current psychotic disorder; lifetime history of psychotic disorder; current bipolar disorder; lifetime history of bipolar disorder; clinically significant depression; MADRS score >24 at the screening or baseline visits; substance use disorder within 3 months of start of medication; enrollment in a formal psychotherapy program ≤6 months before screening; enrollment in a CBT program ≤6 months before screening; enrollment in an interpersonal therapy program ≤6 months before screening; enrollment in a behavioral therapy program ≤6 months before screening; history of factitious disorder; history of a personality disorder	BMI >= 30 kg/m²-<= 50 kg/m²: 407 (100%) Obesity: 407 (100%) BED, Symptoms, Duration: 18.6 yr (SD ± 14.3, N=195) vs. 20.6 yr (SD ± 14.5, N=199) Weight: 106 kg (SD ± 18.5, N=202) vs. 107 kg (SD ± 18.3, N=202) Age 18 yr-65 yr: 407 (100%) Age: 44 yr (SD ± 11.5, N=202) vs. 45 yr (SD ± 11.6, N=202) Gender - Female: 170 (84.2%, N= 202) vs. 170 (84.2%, N= 202) - Male: 32 (15.8%, N= 202) vs. 32 (15.8%, N= 202) Race - Caucasian: 153 (75.7%, N= 202) vs. 164 (81.2%, N= 202) - Black or African American: 38 (18.8%, N= 202) vs. 27 (13.4%, N= 202) - Other: 11 (5.4%, N= 202) vs. 11 (5.4%, N= 202)	- 4.6 d/wk (SD ± 1.3, N=195) vs. 4.6 d/wk (SD ± 1.3, N=199) - 6.6/wk (SD ± 4.6, N=195) vs. 6.3/wk (SD ± 3.6, N=199) (Total N=394) Binge Eating, Change - Baseline – 16 wk - -3.5 d/wk (SD ± 1.9, N=195) vs. -2.5 d/wk (SD ± 2.1, N=199) (MD -1 d/wk, p<0.001) - -5/wk (SD ± 4.3, N=195) vs. -3.4/wk (SD ± 3.8, N=199) (MD -1.6/wk, p<0.001) Topiramate was significantly more associated with the following adverse events: - Paresthesia: 113 (55.9%, N=202) vs. 25 (12.4%, N=202) (p<0.001) - Dysgeusia: 28 (13.9%, N=202) vs. 2 (1%, N=202) (p<0.001) - Concentration or Attention, Difficulty: 26 (12.9%, N=202) vs. 5 (2.5%, N=202) (p<0.001) - Memory Impairment: 25 (12.4%, N=202) vs. 12 (5.9%, N=202) (p=0.037) - Infection, Upper Respiratory Tract: 37 (18.3%, N=202) vs. 20 (9.9%, N=202) (p=0.022) Treatment Discontinuation - Baseline – 16 wk - Adverse Events: 29 (15%, N=199) vs. 16 (8%, N=202) - Lack of Efficacy: 1 (0.5%, N=199) vs. 3 (1.49%, N=202) Attrition: 29% (57/199) vs. 30% (61/202)
Safer et al. (2020)	Design: RCT Setting: Outpatient	Randomized N=22 Phentermine/Topiramate ER 3.75 mg/23 mg-15	Inclusion: BED or BN; 18-60 years of age; obesity; BMI ≥ 21 kg/m^2;	BED or BN: 22 (100%) - BED: 18 (81.8%) - BN: 4 (18.2%)	Phentermine/Topiramate ER showed significantly greater reductions in binge

	Country: United States Funding: Government; product donation by industry	mg/92 mg 12 wk > Washout 2 wk > Placebo 12 wk (N=12) Placebo 12 wk > Washout 2 wk > Phentermine/Topiramate ER 3.75 mg/23 mg-15 mg/92 mg 12 wk (N=10) Follow-up: Baseline – 34 wk	refractory to prior psychotherapy or pharmacologic treatment Exclusion: bipolar disorder or schizophrenia; use of a mood stabilizer or antipsychotic medication; history of AN; prescription weight loss medication or over the counter weight-reducing agent; psychological weight-loss intervention; psychostimulant use; change in thyroid, psychiatric, or hypertensive medications; use of a potassium-wasting diuretic, carbonic anhydrase inhibitor, insulin, or insulin secretagogue; abnormal baseline labs; substance abuse or dependence; stimulant misuse; suicidal ideation; nephrolithiasis; pregnancy; cardiovascular disease	BMI: 31.1 kg/m² (SD ± 6.2) Obese: 12 (54.5%) Weight: 86 kg (SD ± 19.8) Age 18 yr-60 yr: 22 (100%) Age: 42.9 yr (SD ± 10.1) Gender - Female: 21 (95.5%) - Male: 1 (0.5%) Race - Caucasian: 12 (54.5%) - Black or African American: 3 (13.6%) - Other: 7 (31.8%) Ethnicity - Hispanic: 5 (22.7%) - Non-Hispanic: 17 (77.3%)	day frequency and significantly higher abstinence rates compared to placebo. Binge Eating, Objective – Baseline (mean): 16.2 d/28 days (SD ± 7.8) Binge Eating, Objective – 12 wk 4.2 d/28 days vs. 14.5 d/28 days (p<0.0001) Binge Eating, Subjective – Baseline (mean): 6.3 d/28 days (SD ± 9.0) Binge Eating, Subjective – 12 wk 3.7 d/28 days vs. 6.8 d/28 days Binge Eating Episodes, Objective – Baseline (mean): 23.5/28 days (SD ± 15.4) Binge Eating Episodes, Objective – 12 wk: 6.6/28 days vs. 20.1/28 days (p=0.0002) Binge Eating, Abstinence – 12 wk: 14 (63.6%) vs. 2 (9.1%) (p<0.0001) The average weight loss for all participants while on phentermine/topiramate ER was 6.4%. When patients crossed over to placebo, they regained weight on average about 1.5%. Adverse Event (while on phentermine/topiramate ER) - Dry mouth 11 (52.4%) - Insomnia 6 (28.6%) - Paresthesia/tingling 6 (28.6%) - Dysgeusia 5 (23.8%) Attrition: 17% (2/12) vs. 20% (2/10)

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BES=Binge Eating Scale; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; d=day; ER=extended release; ITT=intention-to-treat; MADRS=Montgomery-Asberg Depression Rating Scale; MD=mean difference; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Lamotrigine

Guerdjikova et al. (2009)	Design: RCT Setting: Single Center: University of Cincinnati Medical Center Country: United States Funding: Industry	Randomized N=51 Lamotrigine 25-400 mg 16 wk (N= 26) Placebo 16 wk (N= 25) ITT (N=49) - 25 vs. 24 Follow-up: Baseline – 17 wk	Inclusion: BED; obese; 18-65 years of age; BMI >=30 kg/m^2 Exclusion: Concurrent AN or BN; concurrent substance abuse or dependence; substance abuse or dependence within 6 months of study entry; lifetime history of a psychotic disorder or dementia; history of any psychiatric disorder; personality disorder that could interfere with diagnostic assessment, treatment, or compliance; displayed a current clinically unstable depressive; displayed a current clinically unstable bipolar disorder; MADRS >24; YMRS >8; Displayed clinically significant suicidality or homicidality; received interpersonal therapy, CBT, or other behavioral therapy for BED within 3 months of study entry; clinically unstable medical illness; required treatment with any drug that might interact adversely with study medication; required treatment with any drug that might obscure the action of the study medication; stimulants; sympathomimetics; antidepressants; Carbonic mood stabilizers; antiobesity agents; had received psychoactive medication within 1 wk of randomization; had received MAOIs within 4 weeks of randomization; had received investigational medications or depot antipsychotics within 3 months of randomization; had been treated with lamotrigine in the past; had less than 2 binge	BED: 51 (100%) Obesity: 51 (100%) BMI >= 30 kg/m²: 51 (100%) BMI: 38.72 kg/m² (SD ± 5.38) vs. 41.52 kg/m² (SD ± 7.24) Age 18 yr-65 yr: 51 (100%) Age: 46.08 yr (SD ± 12.62) vs. 42.88 yr (SD ± 12.74) Gender - Female: 21 (80.77%) vs.18 (72%) - Male: 5 (19.23%) vs. 7 (28%) Race, Caucasian: 21 (80.77%) vs. 20 (80%)	Placebo was associated with higher remission rate: 8 (57%, N= 14) vs.16 (94%, N= 17) (p=0.03) Eating, Binge Eating - Baseline: - 3.92/wk (SD ± 1.47) vs. 28/wk (SD ± 1.31) - 3.81 d/wk (SD ± 1.39) vs. 3.2 d/wk (SD ± 1.26) Eating, Binge Eating, Change - Baseline – 16 wk - Lamotrigine vs. Placebo: MD -0.1/wk, 95% CI -0.24 – 0.04; MD -0.1 d/wk, 95% CI -0.23 – 0.04 Weight - Baseline: 105.93 kg (SD ± 19.08) vs. 120 kg (SD ± 25.39) Weight, Change - Baseline – 16 wk: -1.17 kg (SD ± 2.96, N= 25) vs. -0.15 kg (SD ± 3.61, N= 24) (MD -1.32 kg, 95% CI -3.2 – 0.56) Adverse Events, Serious - Baseline – 16 wk: 0 (0%) vs. 1 (4%) Treatment Discontinuation - Baseline – 16 wk: 11 (42.31%) vs. 7 (28%) Treatment Discontinuation - Baseline – 16 wk - Adverse Events: 3 (11.54%) vs. 1 (4%) - Lack of Efficacy: 3 (11.54%) vs. 1 (4%) Attrition: 46% (12/26) vs. 32% (8/25)

			days in the week before randomization; pregnant; lactating		

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; CI=confidence interval; d=day; MADRS=Montgomery-Asberg Depression Rating Scale; MAOIs=monoamine oxidase inhibitors; MD=mean difference; RCT=randomized controlled trial; SD=standard deviation; wk=week; YMRS=Young Mania Rating Scale; yr=year

Zonisamide

McElroy et al. (2006)	Design: RCT Setting: Outpatient: University of Cincinnati Medical Center Country: United States Funding: Industry	Randomized N=60 Zonisamide 100-600 mg 16 wk (N= 30) Placebo 16 wk (N= 30)	Inclusion: 18-62 years of age; BED; obese; BMI >= 30 kg/m^2; ≥2 days with binge eating episodes (binge days) in the wk before receiving study medication Exclusion: Concurrent AN or BN; substance use disorder within 6 months of study entry; lifetime history of a psychotic disorder, bipolar disorder, dementia, or cognitive disorder; personality disorder; clinically significant suicidality or homicidality; received CBT or interpersonal psychotherapy within 3 months of study entry; received behavioral weight management for BED within 3 months of study entry; history of seizures; received psychoactive medication within 2 weeks of study medication initiation; previously been treated with zonisamide	BED: 60 (100%) Obesity: 60 (100%) Binge Eating >= 2 d, In the Previous 1 wk: 60 (100%) BED, Duration: 19 yr (SD ± 13.8) vs. 17.9 yr (SD ± 12.9) BMI >= 30 kg/m²: 60 (100%) Age 18 yr-62 yr: 60 (100%) Age: 44.8 yr (SD ± 9.3) vs. 43 yr (SD ± 10.7) Gender - Female: 27 (90%) vs. 26 (86.7%) - Male: 3 (10%) vs. 4 (13.3%) Race - Caucasian: 23 (76.6%) vs. 20 (66.7%) - Black or African American: 17 (28.33%)	Significantly shortened time to recovery from binge eating was reported with zonisamide: HR 2.76 (p=0.033) Eating, Binge Eating - Baseline - 4.7/wk (SD ± 1.4) vs. 4.4/wk (SD ± 2) - 3.9 d/wk (SD ± 1.1) vs. 3.9 d/wk (SD ± 1.3) Eating, Binge Eating, Change - Baseline – 16 wk - Zonisamide vs. Placebo: MD 0.002/wk, 95% CI -0.143 – 0.171; MD -0.04 d/wk, 95% CI -0.176 – 0.119 Weight - Baseline: 118 kg (SD ± 30.7) vs. 112.8 kg (SD ± 24.3) Weight, Change - Baseline – 16 wk: -8.97 kg vs. -1.25 kg (MD -3.68 kg, 95% CI -5.91 – -1.45) Adverse Events, Serious, Requiring Hospitalization - Baseline – 17 wk: 1 (3.33%) vs. 1 (3.33%) Study Withdrawal- Baseline – 16 wk - Adverse Events: 8 (26.7%) vs. 4 (13.3%) (p=0.33) - Lack of Efficacy: 1 (3.33%) vs. 0 (0%) Attrition: 60% (18/30) vs. 40% (12/30)

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive-behavioral therapy; CI=confidence interval; d=day; HR=hazard ratio; MD=mean difference; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Stimulants

McElroy et al. (2015a)	Design: RCT Setting: Unclear: Lindner Center of HOPE Country: United States Funding: Industry	Randomized N=60 Armodafinil 150-250 mg 10 wk (N= 30) Placebo 10 wk (N= 30) Current Analysis (N=55) - 27 vs. 28	Inclusion: 18-65 years of age; BED; at least 3 binge-eating days/wk for the 2 weeks before receiving study medication; BMI of at least 25 kg/m^2 Exclusion: Current AN or BN; clinically significant suicidality; substance use disorder within 6 months of study entry; lifetime history of psychosis, mania, hypomania, or dementia; psychotropic medications within 4 weeks before randomization	BED: 60 (100%) Binge Eating >= 3 d/wk, In the Previous 2 wk: 60 (100%) Weight: 110 kg (SD ± 25.2) BMI >= 25 kg/m²: 60 (100%) BMI, Obesity >= 30 kg/m²: 27 (90%) vs. 28 (93%) BMI: 40.1 kg/m² (SD ± 8) Age 18 yr-65 yr: 60 (100%) Age: 41.3 yr (SD ± 12) - 40.8 yr (SD ± 12.7) vs. 41.9 yr (SD ± 11.4) Gender - Female: 28 (93%) vs.23 (77%) - Male: 2 (7%) vs. 7 (23%) Race - Caucasian: 22 (73%) vs. 24 (80%) - Black or African American: 8 (27%) vs. 6 (20%)	Armodafinil associated with significantly more feeling jittery and xerostomia: - Feeling Jittery: 9 (30%) vs. 0 (0%) (p<0.01) - Xerostomia: 7 (23%) vs. 1 (3%) (p=0.05) Eating, Binge Eating - Baseline - 5.9/wk (SD ± 4.8) vs. 5/wk (SD ± 1.6) - 4.5 d/wk (SD ± 1.5) vs. 4.4 d/wk (SD ± 1) Eating, Binge Eating, Change - Baseline – 10 wk - -4.2/wk (SD ± 3.1, N= 27) vs. -2.8/wk (SD ± 1.8, N= 28) (MD -1.3/wk, 95% CI -2.7 – 0) - -3.1 d/wk (SD ± 2.1, N= 27) vs. -2.4 d/wk (SD ± 1.6, N= 28) (MD -0.7 d/wk, 95% CI -1.7 – 0.3) Weight – Baseline: 108.3 kg (SD ± 25) vs. 113.6 kg (SD ± 25.6) Weight, Change - Baseline – 10 wk: -1.6 kg (SD ± 2.4, N= 27) vs. 0 kg (SD ± 3.6, N= 28) (MD -1.6 kg, 95% CI -3.3 – 0) Treatment Discontinuation - Baseline – 10 wk - Lack of Efficacy: 1 (3.33%) vs. 2 (6.67%) - Adverse Events: 2 (6.67%) vs. 2 (6.67%) Attrition: 47% (14/30) vs. 50% (15/30)

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CI=confidence interval; d=day; MD=mean difference; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Other Pharmacotherapies

Brownley et al. (2013)	Design: RCT Setting: Multi-center Country: United States Funding: Non-profit; product donation by industry	Randomized N=24 Chromium Picolinate 600 mcg 6 mo (N=9) Chromium Picolinate 1000 mcg 6 mo (N=8) Placebo 6 mo (N=7) Chromium Picolinate 600/1000 mcg 6 mo (pooled) (N=17) Current Analysis (N=21) - 8 vs. 7 vs. 6 Follow-up: Baseline – 9 mo	Inclusion: BED; overweight Exclusion: BMI < 25; BMI > 45; age <18 years; age >60 years; current chromium use; current use of insulin; current use of other medications to control glucose metabolism; current use of medications known to significantly influence appetite or weight; over-the-counter appetite suppressants that contain phentermine or sibutramine; atypical antipsychotic agents with high weight gain liability; olanzapine; risperidone; prednisone; current psychotropic medication use; current suicidal intent; current homicidal intent; other psychiatric condition that required acute intervention	BED: 24 (100%) Eating, Binge Eating, Duration: 16 yr Overweight: 24 (100%) BMI: 34.2 kg/m² (SD ± 5.4) - 33.5 kg/m² (SD ± 7) vs. 34.9 kg/m² (SD ± 4.1) vs. 34.3 kg/m² (SD ± 5.4) Age: 36.6 yr (SD ± 10.6) - 35.1 yr (SD ± 12.4) vs. 41.4 yr (SD ± 8.5) vs. 37.9 yr (SD ± 10.8) Gender - Female: 7 (78%) vs. 6 (75%) vs. 7 (100%) - Male: 2 (22%) vs. 2 (25%) vs. 0 (0%) Race, Caucasian: 7 (78%) vs. 8 (100%) vs. 6 (86%)	Weight – Baseline: 95.1 kg (SD ± 18.36, N=8) vs. 104.1 kg (SD ± 17.79, N=7) vs. 102.8 kg (SD ± 16.86, N=6) Weight, Change - Baseline – 6 mo: -1.6 kg (SD ± 19.15, N=8) vs. NR (N=7) vs. 5.9 kg (SD ± 12.77, N=6) Adverse Events - Baseline – 6 mo: 0 (0%) vs. 0 (0%) vs. 5 (71.43%) Attrition: 11% (1/9) vs. 13% (1/8) vs. 14% (1/7)
Golay et al. (2005)	Design: RCT Setting: Multi-center Country: Switzerland Funding: Industry	Randomized N=89 Orlistat 120 mg + Hypocaloric Diet 24 wk (N=44) Placebo + Hypocaloric Diet 24 wk (N=45) (All received Hypocaloric Diet) Current Analysis (N=73) - 39 vs. 34	Inclusion: BED; BMI >= 30 kg/m^2; 18- 65 years of age; obese Exclusion: Drug-treated diabetes mellitus; taking antidepressants; taking appetite suppressants; taking tranquilizer; psychological counseling; psychological therapy; taking medications known to alter body weight; history of significant psychological illness; significant psychological illness	BED: 89 (100%) Obesity: 89 (100%) BMI >= 30 kg/m²: 89 (100%) BMI: 35.7 kg/m² (SD ± 3.32) vs. 37.3 kg/m² (SD ± 5.37) Age 18 yr-65 yr: 89 (100%) Age: 41.2 yr (SD ± 6.2) vs. 40.6 yr (SD ± 6.1) Gender - Female: 40 (91%) vs. 41 (91%) - Male: 4 (9%) vs. 4 (9%) Race, Caucasian: 86 (97%)	Orlistat was associated with higher rates of weight loss and reduced fat mass. Weight - Baseline: 96.9 kg (SD ± 15.26) vs. 99.8 kg (SD ± 14.09) Weight, % Change - Baseline – 24 wk: -7.4% vs. -2.3% (MD -5.1 %, p=0.0001) Weight, Fat Mass, Change - Baseline – 24 wk: -5.58 kg vs. -2.79 kg (MD -2.79 kg, p<0.01) Eating, Binge Eating - Baseline: 5.4/wk vs. 6.2/wk - 24 wk: 1/wk (N=39) vs. 1.7/wk (N=34)

					Attrition: 11% (5/44) vs. 29% (13/45)
Grilo et al. (2005b)	Design: RCT Setting: NR Country: United States Funding: Supported by non-profit	Randomized N=50 Placebo + CBT-GSH 12 wk (N=25) Orlistat 120 mg + CBT-GSH 12 wk (N=25) Follow-up: Baseline – 6 mo	Inclusion: 35-60 years of age; BMI of 30 kg/m^2 or greater; BED; obese Exclusion: Concurrent treatment for eating disorder, weight disorder, or psychiatric illness; severe current psychiatric conditions, psychosis, or bipolar disorder requiring treatment	BED: 50 (100%) Obesity: 50 (100%) BMI >= 30 kg/m²: 50 (100%) BMI: 36 kg/m² (SD ± 4.7) - 36.8 kg/m² (SD ± 5.1) vs. 36.2 kg/m² (SD ± 4.7) Age 35 yr-60 yr: 50 (100%) Age: 47 yr (SD ± 7) - 47 yr (SD ± 7) - (N=2) vs. 45.2 yr (SD ± 7.4) Gender - Female: 23 (92%) vs. 21 (84%) - Male: 2 (8%) vs. 4 (16%) Race - Caucasian: 22 (88%) vs. 22 (88%) - Black or African American: 2 (8%) vs. 1 (4%) Ethnicity, Hispanic/Latino: 1 (4%) vs. 2 (8%)	Significantly greater remission rate was reported with orlistat at the end of treatment (36% vs. 64%, p=0.048) but comparable (52%) at 6-mo follow-up Orlistat was also associated with significantly greater percent achieving at least 5% weight loss at both the end of treatment and 6-mo follow-up: 8% vs. 36% (p=0.017); 8% vs. 32% (p=0.034), respectively. Study Withdrawal, Adverse Events - Baseline – 12 wk: 0 (0%) vs. 2 (8%) Attrition: 20% (5/25) vs. 24% (6/25)
Grilo and White (2013)	Design: RCT Setting: Single Center: community mental health center Country: United States Funding: Government and non-profit	Randomized N=79 Of 40 with BED, Orlistat 120 mg + BWLT4 mo (N=20) Placebo + BWL 4 mo (N=20) Follow-up: Baseline – 10 mo	Inclusion: Obese; Latino; 21-65 years of age; BMI of 30 kg/m^2 or greater Exclusion: Serious mental illnesses; psychotic disorders; schizophrenia; current severe bipolar illness; uncontrolled current substance dependence; suicidality; unstable medication regimens; changing medication regimens; current antipsychotic medications; current cardiac disease; current neurologic diseases	Obesity: 79 (100%) Eating, Binge Eating: 20 vs. 20 BMI >= 30 kg/m²: 79 (100%) BMI: 37.57 kg/m² (SD ± 6.62) Age 21 yr-65 yr: 79 (100%) Age: 46.32 yr (SD ± 9.68) - 45.9 yr (SD ± 9) vs. 45.6 yr (SD ± 7.6)	Disease Response, Remission - 4 mo: 12 (60%) vs. 14 (70%) (p=0.51) - 10 mo: 10 (50%) vs. 10 (50%) Weight, % Change - Baseline – 4 mo: -3.9% vs.-2.1% BMI – Baseline: 39 kg/m² (SD ± 7) vs. 37.2 kg/m² (SD ± 5.3) BMI, % Change - Baseline – 10 mo: -3.89% (N=18) vs. -1.47% (N=19)

				Gender of those with BED - Female: 17 (85%) vs. 14 (70%) - Male: 3 (15%) vs.6 (30%) Ethnicity, Hispanic/Latino: 79 (100%)	Attrition: 30% (6/20) vs. 25% (5/20)
McElroy et al. (2011)	Design: RCT Setting: Outpatient: Lindner Center of HOPE Country: United States Funding: Industry	Randomized N=40 Acamprosate 333-999 mg 10 wk (N= 20) Placebo 10 wk (N= 20) Intention-to-Treat (N=39) - 19 vs. 20	Inclusion: 18-65 years of age; BED; weighed >= 85% of the midpoint of IBW for height; had >= 3 binge-eating episodes in the wk before receiving study medication; had >= 2 binge days in the wk before receiving study medication Exclusion: Concurrent AN or BN; substance use disorder within 6 months of study entry; lifetime history of a psychotic disorder; lifetime history of a bipolar disorder, dementia, or other cognitive disorder; personality disorder that could interfere with diagnostic assessment, treatment, or compliance; clinically significant suicidality or homicidality; cognitive-behavioral psychotherapy or interpersonal psychotherapy within 3 months of study entry; behavioral weight management for BED within 3 months of study entry; clinically unstable medical illness; history of seizures; required treatment with any drug that might adversely interact with acamprosate; required treatment with any drug that might obscure the action of the acamprosate; monoamine oxidase inhibitors, tricyclics, lithium, antipsychotics, or fluoxetine within 4 weeks prior to randomization; other psychoactive medication within 1 week of study medication initiation; investigational medications or	BED: 40 (100%) Binge Eating >= 3 episodes, In the Previous 1 wk: 40 (100%) Binge Eating >= 2 d, In the Previous 1 wk: 40 (100%) Weight: 116.5 kg (SD ± 27.3) vs. 107.7 kg (SD ± 23.7) Age 18 yr-65 yr: 40 (100%) Age: 46.2 yr (SD ± 12.2) vs. 45.8 yr (SD ± 9.1) Gender - Female: 16 (80%) vs. 18 (90%) - Male: 4 (20%) vs. 2 (10%) Race - Caucasian: 18 (90%) vs. 17 (85%) - Black or African American: 4 (10%) Ethnicity, Hispanic/Latino: 1 (2.5%)	Acamprosate was significantly more associated with diarrhea: 11 (55%) vs. 5 (25%) (p=0.05) Binge Eating – Baseline->10 wk - 4.5->1.9/wk (N= 19) vs. 4.5->2.8/wk - 4.2->1.8 d/wk (N= 19) vs. 3.8->2.6 d/wk Binge Eating, Change - Baseline – 10 wk - Acamprosate vs. Placebo: MD -0.98/wk, 95% CI -2.13 – 0.18; MD -1.14 d/wk, 95% CI -2.22 – -0.05 BMI – Baseline->10 wk: 39.7->39.7 kg/m² (N= 19) vs. 39.2->39.7 kg/m² Weight – Baseline->10 wk: 116->116.3 kg (N= 19) vs. 107.7->108.9 kg Adverse Events, Serious - Baseline – 10 wk: 0 (0%) vs. 0 (0%) Treatment Discontinuation - Baseline – 10 wk: 6 (30%) vs. 9 (45%) (p=0.51) Attrition: 25% (5/20) vs. 55% (11/20)

			depot antipsychotics within 3 months prior to randomization; treated with acamprosate; pregnant; lactating		

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; BWL=behavioral weight loss; CBT-GSH=cognitive-behavioral therapy guided self-help; CI=confidence interval; d=day; ITT=intention-to-treat; MD=mean difference; mo=month; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; wk=week; yr=year

Night Eating Syndrome Studies

O'Reardon (2006, 2008)	Design: RCT Setting: NR Country: United States Funding: Government and industry	Randomized N=65 Current Analysis (N=34) Sertraline 50-200 mg 8 wk (N=17) Placebo 8 wk (N=17) Weight, Normal subgroup (N=3 vs. 3) Weight, Overweight subgroup (N=14 vs. 14)	Inclusion: >= 18 years old; NES; BMI >= 18 kg/m^2 Exclusion: Severely depressed; lifetime diagnosis of bipolar disorder; psychotic disorder; currently taking psychotropic medications; currently taking hypnotics; current diagnosis of AN or BN; substance abuse or dependence within the preceding 6 months; weight reduction program	NES: 65 (100%) BED: 0 (0%) vs. 3 (17.65%) Night Eating, Duration: 17.6 yr (SD ± 15.5) vs. 15.3 yr (SD ± 12.7) BDI: 14.4 units (SD ± 9.7) vs. 12.1 units (SD ± 9.5) BMI > 18 kg/m²: 65 (100%) BMI: 32.4 kg/m² (SD ± 6.5) vs. 32.9 kg/m² (SD ± 9) Weight, Normal: 3 (17.6%) vs. 3 (17.6%) Weight, Overweight: 14 (82.35%) vs. 14 (82.35%) Age >= 18 yr: 65 (100%) Age: 45.1 yr (SD ± 11) vs. 44.2 yr (SD ± 10.6) Gender - Female: 11 (64.7%) vs. 12 (70.6%) - Male: 6 (35.3%) vs. 5 (29.4%) Race - Caucasian: 12 (70.6%) vs. 15 (88.2%)	Sertraline showed greater reductions in night eating and weight. Night Eating - Baseline: 8.3/wk (SD ± 8.5) vs. 6.4/wk (SD ± 2.6) Night Eating, Change - Baseline – 8 wk: -6.9/wk (SD ± 3.9) vs. -0.4/wk (SD ± 0.6) (MD -6.5/wk, p<0.0125) Night Eating, % Change - Baseline – 8 wk: -81% vs. -14% (MD -67 %, p=0.01) NESS - Baseline: 31.7 units (SD ± 5.6) vs. 30.5 units (SD ± 6.2) - 8 wk: 13.7 units (SD ± 3.3) vs. 25.2 units (SD ± 5.2) (MD -11.5 units, p<0.0125) NESS, Change - Baseline – 8 wk: -18.1 units vs. -5 units (MD 13.1 units, p<0.0001) Weight, Change - Baseline – 8 wk: -2.85 kg (SD ± 1.9) vs. -0.26 kg (SD ± 1.1) (MD -2.59 kg, p<0.0125) - Weight, Overweight subgroup: -2.9 kg (SD ± 3.8) vs. -0.3 kg (SD ± 2.7) (MD -2.6 kg, p=0.009) Disease Response, Remission - 8 wk: 7 (41.18%) vs. 1 (5.88%) Study Withdrawal - Baseline – 8 wk

				- Black or African American: 5 (29.4%) vs. 2 (11.8%)	- Lack of Efficacy: 1 (5.88%) vs. 1 (5.88%) - Adverse Events: 0 (0%) vs. 0 (0%) Attrition: 6% (1/17) vs. 6% (1/17)
Vander Wal et al. (2012)	Design: RCT Setting: 2 academic centers Country: United States Funding: Industry	Randomized N=40 Escitalopram 10-20mg 12 wk (N=20) Placebo 12 wk (N=20) Caucasian subgroup (N=10 vs. 13) Black or African American subgroup (N=9 vs. 7)	Inclusion: NES; 18-70 years of age; BMI 25-50 kg/m^2; minimum score of 25 on the NEQ; overweight or obese Exclusion: Alcohol or drug abuse; AN; BN; BED; major depressive disorder; suicidal ideation; lifetime history of schizophrenia; escitalopram in the past year; psychotropic medications in the past month; nonresponse to SSRI for NES	NES: 40 (100%) NEQ >= 25 units: 40 (100%) Night Eating, Duration: 11.1 yr (SD ± 12.4) vs. 11 yr (SD ± 9.6) Overweight or Obesity: 40 (100%) Weight: 97.8 kg (SD ± 17.5) vs. 95 kg (SD ± 21.1) BMI 25 kg/m²-50 kg/m²: 40 (100%) BMI: 33.3 kg/m² (SD ± 6.4) vs. 32.6 kg/m² (SD ± 7.4) Age 18 yr-70 yr: 40 (100%) Age: 45 yr - 45.2 yr (SD ± 13.7) vs. 44.8 yr (SD ± 12.3) Gender - Female: 11 (55%) vs. 10 (50%) - Male: 9 (45%) vs. 10 (50%) Race - Caucasian: 10 (50%) vs. 13 (65%) - Black or African American: 9 (45%) vs. 7 (35%) Ethnicity, Hispanic/Latino: 1 (5%) vs. 0 (0%)	NEQ - Baseline: 31.8 units (SD ± 4) vs. 34.1 units (SD ± 6.4) NEQ, Change - Baseline – 12 wk: -13 units (SD ± 7.16) vs. -10.6 units (SD ± 9.84) (MD -2.4 units, p=0.124) NEQ, Change - Baseline – 12 wk - Caucasian subgroup: -13.67 units (SD ± 3.16) vs. -6.75 units (SD ± 3.24) (MD -6.92 units, p=0.024) - Black or African American subgroup: -12.44 units (SD ± 2.94) vs. -17.9 units (SD ± 2.38) (MD 5.46 units, p=0.453) Weight, Change - Baseline – 12 wk: -0.43 kg (SD ± 3.13) vs. 1.12 kg (SD ± 2.68) (MD -1.55 kg (p=0.086) Overall Attrition: 0%

Abbreviations: AN=anorexia nervosa; BDI=Beck Depression Inventory; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; MD=mean difference; NES=night eating syndrome; NESS=Night Eating Symptom Scale; NEQ=Night Eating Questionnaire; NR=not reported; RCT=randomized controlled trial; SD=standard deviation; SSRI= serotonin reuptake inhibitor; wk=week; yr=year