Screening for Eating Disorders in Adolescents and Adults: An Evidence Review for the U.S. Preventive Services Task Force

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. HHSA-290-2015-00011-I, Task Order No. 15

Prepared by:
RTI International–University of North Carolina at Chapel Hill Evidence-based Practice Center
Research Triangle Park, NC 27709

Investigators:
Cynthia Feltner, MD, MPH
Christine Peat, PhD
Shivani Reddy, MD, MS
Sean Riley, MA, MSc
Nancy Berkman, PhD
Jennifer Cook Middleton, PhD
Casey Balio, PhD
Manny Coker-Schwimmer, MPH
Daniel E. Jonas, MD, MPH

AHRQ Publication No. 21-05284-EF-1
March 2022
This report is based on research conducted by the RTI International–University of North Carolina at Chapel Hill Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSA-290-2015-00011-I, Task Order No. 15). The findings and conclusions in this document are those of the authors, who are responsible for its contents, and do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help healthcare decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information (i.e., in the context of available resources and circumstances presented by individual patients).

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: Iris Mabry-Hernandez, MD, MPH, AHRQ Medical Officer; current and former member of the U.S. Preventive Services Task Force; expert reviewers: Bryn Austin, ScD, Harvard Medical School; Denise Wilfley, PhD, Washington University School of Medicine; Devan Kansagara, MD, Oregon Health and Science University; Susan Kornstein, MD, Virginia Commonwealth University; federal partners from the Centers for Disease Control and Prevention; Christiane Voisin, MSLS, research librarian; Sharon Barrell, MA, editor; Loraine Monroe, publications specialist; Carol Woodell, EPC Program Manager; Caroline Rains, MPH, research assistant; Kayla Giger, research assistant; Ina Wallace, PhD, senior research analyst.

Suggested Citation

Structured Abstract

**Purpose:** To systematically review the evidence on (1) benefits and harms of screening for eating disorders in adults and adolescents, (2) accuracy of screening tools, and (3) benefits and harms of interventions for eating disorders that were screen detected or not previously treated for populations and settings relevant to primary care in the United States.

**Data Sources:** PubMed/MEDLINE, the Cochrane Library, PsycINFO, and trial registries through December 18, 2020; reference lists of retrieved articles; outside experts; and reviewers, with surveillance of the literature through January 1, 2022.

**Study Selection:** English-language controlled trials for eating disorder screening or evaluation of interventions for screen-detected or previously untreated eating disorders and studies of screening test accuracy.

**Data Extraction:** One investigator extracted data and a second checked accuracy. Two reviewers independently rated quality for all included studies using predefined criteria.

**Data Synthesis:** No studies directly assessed the benefits and harms of screening. Seventeen studies evaluated the accuracy of screening questionnaires for identifying eating disorders. For detecting any eating disorder among adults, the SCOFF (cut point ≥2) had a pooled sensitivity of 84 percent (95% confidence interval [CI], 74% to 90%) and a pooled specificity of 80 percent (95% CI, 65% to 89%) (10 studies; 3,684 participants). At a higher cut point (≥3), the pooled sensitivity was lower (69% [95% CI, 56% to 80%]) and specificity was higher (90% [95% CI, 69% to 98%]) (7 studies; 2,749 participants). In one study enrolling adolescents, the SCOFF (cut point ≥2) had a sensitivity of 73 percent and specificity of 78 percent. Two studies assessed the Eating Disorders Screen for Primary Care among adults using a cut point of 2 or greater; sensitivity was similar (97% and 100%), and specificity varied (40% and 71%). All other screening questionnaires were evaluated using only one study each. Only one other study evaluated a screening questionnaire among adolescents (11 to 18 years); the Adolescent Binge Eating Questionnaire (designed to detect binge eating) had a sensitivity of 100 percent and specificity of 27 percent in a population recruited from a pediatric obesity clinic.

Forty studies assessed interventions for populations with recently detected or previously untreated eating disorders. 17 assessed pharmacotherapy, 22 assessed psychological intervention, and two assessed both. None enrolled a population with screen-detected eating disorders. Four trials of lisdexamfetamine for binge-eating disorder (BED) (900 participants) measured change in eating disorder symptom severity using the Yale–Brown Obsessive Compulsive Scale modified for binge eating (YBOCS-BE) and found larger reductions in changes from baseline scores associated with lisdexamfetamine 50 to 60 mg/day than placebo (pooled mean difference, -5.75 [95% CI, -8.32 to -3.17]). Two trials compared topiramate with placebo for BED and both found significantly larger reductions in YBOCS-BE scores from baseline among the topiramate group than the placebo group, from -6.40 (p<0.001) to -2.55 (p=0.004). Five trials assessed various selective serotonin reuptake inhibitors among persons with BED (not selected based on comorbid depression), only two reported on change in eating disorder symptoms and results were
imprecise. Selective serotonin reuptake inhibitors were associated with a larger reduction in depression symptom scores than placebo over 6 to 16 weeks (pooled standardized mean difference [SMD], -0.6 [95% CI, -0.90 to -0.33]) (5 studies; 208 participants). Three trials assessed fluoxetine for populations with bulimia nervosa; two found benefit favoring fluoxetine for eating disorder symptom severity and depression symptoms. Twenty-four trials assessed a psychological intervention. Guided self-help for BED improved eating disorder symptom severity more than inactive control (pooled SMD, -0.96 [95% CI, -1.26 to -0.67]) (5 studies; 391 participants); pooled estimates for unguided self-help (6 studies; 368 participants) also favored the intervention, but the difference between groups was not statistically significant (SMD, -0.18 [95% CI, -0.38 to 0.03]). Similarly, self-help interventions for BED also reduced depression symptoms more than inactive control, including both guided self-help (pooled SMD, -0.73 [95% CI, -1.04 to -0.43]; 4 studies; 324 participants) and unguided self-help (pooled SMD, -0.37 [95% CI, -0.68 to -0.05]; 3 studies; 156 participants). Group therapy (7 trials; 253 participants) for BED and bulimia nervosa was associated with larger reductions in depression scores from baseline than inactive control (pooled SMD, -0.48 [95% CI, -0.69 to -0.27]). Few trials of self-help or group therapy reported on other outcomes. Four studies assessed different forms of individual therapy and measured heterogeneous outcomes. Nine trials of pharmacotherapy (2,006 participants) reported on adverse effects over a relatively short duration (6 to 16 weeks); few reported on more than one medication. Lisdexamfetamine was associated with higher rates of dry mouth, headache, and insomnia than placebo, and topiramate was associated with significant higher rates of paresthesia, taste perversion, and difficulty with concentration or confusion than placebo.

**Limitations:** Included studies of screening test accuracy primarily enrolled populations of adult women. Few reported on the accuracy of screening tests among men, adolescents or other specific populations. Aside from the SCOFF, most were assessed by only one study each, and some enrolled populations from specialty settings (e.g., obesity clinics) that may have a higher prevalence of binge eating. No included treatment studies enrolled populations who were screen detected; most were recruited via advertisements or enrolled referred populations and reported on outcomes over a relatively short duration (6 to 16 weeks). One treatment study was limited to adolescents, and all others enrolled adults; the majority focused on BED and bulimia nervosa. No eligible treatment studies focused on populations with anorexia nervosa.

**Conclusions:** No studies directly assessed the benefits and harms of screening. Screening questionnaires available for use in primary care have adequate accuracy for detecting eating disorders among adults; the most commonly studied screening questionnaire is the SCOFF. The accuracy of screening questionnaires for detecting eating disorders among adolescents is unclear. No treatment studies were found that enrolled participants who were screen-detected in primary care. Guided self-help interventions are effective for reducing eating disorder symptom severity and depressive symptoms among referred populations. Lisdexamfetamine and topiramate are effective in reducing eating disorder symptom severity in populations with BED but are also associated with adverse effects.
# Table of Contents

**Chapter 1. Introduction** .............................................................................................................. 1  
Scope and Purpose .......................................................................................................................... 1  
Condition Definition ..................................................................................................................... 1  
Prevalence ................................................................................................................................... 2  
  Burden and Natural History ........................................................................................................ 3  
Etiology and Risk ............................................................................................................................ 4  
Rationale for Screening and Screening Strategies ........................................................................... 5  
Interventions/Treatment for Eating Disorders ............................................................................... 5  
Recommendations of Other Organizations and Current Clinical Practice ................................. 6  

**Chapter 2. Methods** .................................................................................................................. 7  
Key Questions and Analytic Framework ....................................................................................... 7  
Data Sources and Searches .......................................................................................................... 7  
Study Selection ............................................................................................................................ 8  
Quality Assessment and Data Abstraction .................................................................................... 9  
Data Synthesis and Analysis ....................................................................................................... 9  
Expert Review and Public Comment ............................................................................................ 10  
USPSTF Involvement ................................................................................................................... 11  

**Chapter 3. Results** .................................................................................................................... 12  
Literature Search ......................................................................................................................... 12  
Results by Key Question ............................................................................................................. 12  
  KQ 1. Does Screening for Eating Disorders in Adolescents and Adults Improve Health Outcomes, Including for Specific Subgroups of Interest? .......................................................... 12  
  KQ 2. What Is the Accuracy of Primary Care–Relevant Screening Tests for Eating Disorders in Adolescents and Adults, Including for Specific Subgroups of Interest? .................. 12  
  KQ 3. What Are the Harms of Screening for Eating Disorders in Adolescents and Adults, Including for Specific Subgroups of Interest? ................................................................. 17  
  KQ 4. How Effective Are Interventions for Improving Health Outcomes in Screen-Detected or Previously Untreated Adolescents and Adults With Eating Disorders, Including for Specific Subgroups of Interest? ..................................................... 17  
  KQ 5. What Are the Harms of Interventions for Eating Disorders, Including for Specific Subgroups of Interest? ........................................................................................................... 23  

**Chapter 4. Discussion** ............................................................................................................... 25  
Summary of Evidence .................................................................................................................. 25  
Evidence for Benefit and Harms of Screening ........................................................................... 25  
Diagnostic Test Accuracy .......................................................................................................... 25  
Benefits of Interventions for Screen-Detected or Recently Diagnosed Eating Disorders .......... 25  
Limitations .................................................................................................................................... 26  
Future Research Needs ................................................................................................................ 27  
Conclusion .................................................................................................................................... 27  

**References** .................................................................................................................................. 29
Figures
Figure 1. Analytic Framework
Figure 2. Summary of Evidence Search and Selection
Figure 3. Results of Randomized, Controlled Trials of Lisdexamphetamine and Topiramate vs. Placebo for BED (KQ 4)
Figure 4. Results of Randomized, Controlled Trials of SSRI vs. Placebo for Eating Disorders (KQ 4)
Figure 5. Results of Randomized, Controlled Trials of Self-Help Interventions for Eating Disorders (KQ 4)
Figure 6. Results of Randomized, Controlled Trials of Group and Individual Therapy for Eating Disorders (KQ 4)

Tables
Table 1. Characteristics of Included Studies for KQ 2
Table 2. Summary of Accuracy for Included Screening Tests (KQ 2)
Table 3. Characteristics of Randomized, Placebo-Controlled Trials of Pharmacotherapy for Eating Disorders (KQ 4)
Table 4. Characteristics of Randomized, Placebo-Controlled Trials of Psychological Interventions for Eating Disorders (KQ 4)
Table 5. Summary of Evidence for Screening in Eating Disorders in Adolescents and Adults

Appendixes
Appendix A. Additional Background
Appendix B. Additional Methods Information
Appendix C. Excluded Studies
Appendix D. Assessments of Study Quality
Appendix E. Additional Evidence Tables
Appendix F. Pooled Analyses of Screening Test Accuracy (KQ 2)
Chapter 1. Introduction

Scope and Purpose

The U.S. Preventive Services Task Force (USPSTF) will use this report to inform a recommendation on screening adolescents and adults for eating disorders in primary care settings to prevent adverse health outcomes. The USPSTF has not previously made a recommendation on this topic.

Condition Definition

Eating disorders refer to a group of psychiatric conditions marked by a disturbance in eating or eating-related behaviors that impairs physical or psychosocial functioning. Current diagnostic criteria are based on the fifth version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5),¹ which divides eating disorders into mutually exclusive diagnoses based on observed symptoms (summarized in Appendix A Table 1). This review focuses on common eating disorders that have the potential to be asymptomatic or undetected in the context of routine primary care, including the following: anorexia nervosa (AN), avoidant/restrictive food intake disorder (ARFID), bulimia nervosa (BN), binge-eating disorder (BED), and other specified eating or feeding disorder (OSFED). DSM-5, released in 2013, included several changes to previous eating disorder diagnostic criteria. Most notably, BED was categorized as a separate diagnosis instead of being included as an eating disorder not otherwise specified (EDNOS), and the frequency and duration criteria for BN and BED were lowered compared with criteria in the previous version of the DSM. Additionally, amenorrhea was removed from the diagnostic criteria for AN.

As with the previous version of the DSM, current eating disorder diagnoses remain mutually exclusive (individuals are not given more than one diagnosis at a time);¹ however, there are similarities between diagnostic criteria, and many individuals experience diagnostic crossover during their lifetime. Individuals may also experience varying levels of eating disorder symptom burden with threshold diagnoses indicating that all diagnostic criteria have been met and subthreshold diagnoses indicating that some but not all criteria are met. Based on strict DSM-5 criteria, these individuals may be diagnosed with OSFED, which includes those with symptoms that cause significant distress or impair psychosocial functioning but do not meet the full criteria for a specific eating disorder. However, in the published research literature, these individuals are often categorized as having a “subthreshold” diagnosis for a specific eating disorder rather than OSFED. Definitions for subthreshold eating disorder diagnoses vary across research studies or clinical settings; however, common definitions include endorsement of key behaviors (e.g., binge-eating episodes) that fall short of the required frequency and duration thresholds. This review includes populations diagnosed with full or subthreshold eating disorder diagnoses to be inclusive of populations likely to be detected by routine screening in primary care.
Prevalence

Details of study cohorts and results of nationally representative surveys of eating disorders’ prevalence in the United States are summarized in Appendix A Table 2.

Age

Among adults (age 18 years or older), the most recent estimates come from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC-III) (n=36,309) fielded between 2012 and 2013 using DSM-5 criteria. Lifetime prevalence for AN, BN, and BED in women were 1.42 percent, 0.46 percent, and 1.25 percent, respectively. Estimates were lower among men (AN, 0.12%; BN, 0.08%; BED, 0.42%). Other prevalence estimates for adults come from smaller surveys conducted between 2001 and 2003 that categorized eating disorders using DSM-IV criteria: the National Comorbidity Replication Survey (NCS-R) (n=2,980) and the Collaborative Epidemiologic Psychological Surveys (CEPS) (n=12,337). Compared with these older surveys, the estimated prevalence of AN among women was higher in the NESARC-III (1.42% vs. 0.90% in the NCS-R) but lower for BN and BED (BN 0.08% vs. 1.50% in the NCS-R; BED 0.42% vs. 3.50% in the NCS-R). The lower prevalences of BN and BED were unexpected given the lower number of binge episodes required for diagnosis by DSM-5 compared with DSM-IV, although NESARC was a larger cohort compared with the other studies.

Across age categories, younger adults are at higher risk of having a lifetime-prevalent eating disorder than older populations. In a national sample, the odds of lifetime risk of AN, BN, and BED for individuals ages 18 to 29 years compared with those older than 60 years was 2.0 (95% confidence interval [CI], 0.7 to 5.3), 16.8 (95% CI, 3.0 to 95.6), and 4.9 (95% CI, 2.1 to 11.5), respectively. Similar correlations of younger age and higher risk of BN and BED were observed in an international study of 14 countries. The odds of BN for participants ages 18 to 29 years compared with those age 60 years was 21.4 (95% CI, 11.5 to 39.6) and for BED 14.6 (95% CI, 9.1 to 23.6).

Among adolescents (12 to 17 years), the NCS-A (adolescent supplement to the NCS-R) (n=10,123) estimated prevalence of eating disorders based on surveys conducted in 2001 and 2004 using DSM-IV criteria. Lifetime prevalences of AN, BN, and BED were 0.3 percent, 1.3 percent, and 2.3 percent, respectively, for females and 0.3 percent, 0.5 percent, and 1.3 percent for males.

Race/Ethnicity, Sexual Identity, and Gender Identity

Prevalence estimates of eating disorders also varied by race/ethnicity, sexual identity, and gender identity. Data from large, national samples indicated that White individuals are approximately twice as likely to receive a lifetime diagnosis of AN compared with Black individuals and 5 times as likely compared with Hispanic individuals. However, contrary to previous conceptualizations of eating disorders, there is a demonstrated burden of eating disorders among racial and ethnic minorities. Data from the NESARC indicate that the odds of lifetime diagnosis
of BN and BED among Black and Hispanic individuals were not significantly different from White counterparts, and the CEPS study found BN was more prevalent among Latinos, Asians, and African Americans than non-Latino Whites.\(^5\) Evidence also suggests that transgender adolescents and young adults have higher rates of eating disorders than same-age peers.\(^6\) In fact, a 2015 survey of U.S. university students (N=289,024, mean age 20 years) found that transgender students had higher rates of self-reported eating disorder diagnoses than cisgender heterosexual women (15.82% vs. 1.85%) and higher rates of past-month vomiting or laxative use (15.01% vs. 3.71%).\(^7\)

**Burden and Natural History**

Eating disorders are associated with significant short-term and long-term adverse health outcomes, including physical, psychological, and social problems.

**Physical Complications**

Eating disorders can lead to physical complications affecting all organs and systems. Specific complications vary by diagnosis and frequency of certain behaviors. For example, purging behaviors (e.g., self-induced vomiting, laxative abuse, and diuretic abuse) are associated with morbidity affecting the teeth, esophagus, gastrointestinal system, kidneys, skin, cardiovascular system, and musculoskeletal system.\(^8\) BN is particularly associated with consequences related to purging, including cardiovascular issues (e.g., arrhythmias, cardiac failure), electrolyte disturbances, pancreatitis, gastric erosions or perforations, dental erosion, and renal injury.\(^8\) BED, if untreated, can contribute to obesity (30% to 45%) and related metabolic disorders.\(^9\) AN is associated with physical complications directly attributed to weight loss and malnutrition, such as low bone density and increased fracture prevalence.\(^10\) The degree of weight loss and chronicity of illness increase the risk of complications.\(^11\)

**Mortality**

Evidence also suggests that individuals with eating disorders have higher mortality rates than the general population, particularly those with AN.\(^12\) One meta-analysis (36 studies) estimated weighted annual mortality for AN, BN, and BED as 5.10 deaths (95% CI, 3.99 to 6.14) per 1,000 person-years, 1.74 deaths (95% CI, 1.09 to 2.44) per 1,000 person-years, and 3.31 deaths (95% CI, 1.48 to 5.75) per 1,000 person-years, respectively. The overall pooled standardized mortality rates (ratio of the crude mortality rate to the expected mortality rate) for AN, BN, and BED were 5.86, 1.93, and 1.92, respectively.\(^13\)

**Psychological and Social Complications**

Eating disorders are commonly comorbid with other psychiatric conditions including mood, anxiety, and substance abuse disorders.\(^14\) Across eating disorder diagnoses, depression is the most common comorbidity followed by alcohol use disorder.\(^14\) Eating disorders have also been associated with disturbances in cognitive and emotional functioning;\(^15\) however, there is debate
about whether deficits are etiologic (and lead to eating disorders) or whether deficits occur because of consequences of the disorder. In one case-control study (N=148), cognitive impairment was more frequent in patients with long-term eating disorders (>10 years) compared with healthy controls, but those with short-term eating disorders (<2 years) had similar performance as controls on neuropsychological testing. In terms of social function, results from one nationally representative sample using the National Institute on Alcohol Abuse and Alcoholism’s Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 to measure function found high rates of any social impairment in persons across all eating disorders, with significantly higher impairment among those with BN (61.4%) and BED (53.7%) than those with AN (30.7%). Similarly, participants with BN (49.5%) and BED (52.5%) reported greater interference with daily activities than participants with AN (23.5%). Other evidence suggests disparities in social functioning across ethnic groups. In one nationally representative survey assessing functional impairment using the World Health Organization’s Disability Assessment Schedule instrument, African Americans with AN, BN, or BED reported significantly greater levels of impairment with respect to days out of role due to mental disorder, cognition, mobility, and role functioning compared with the non-Latino White reference group, with similar findings for men and women.

Etiology and Risk

Much about the natural history and pathogenesis of eating disorders is not well understood. Underlying causes of eating disorders have been categorized into predisposing (background vulnerabilities), precipitating (environmental context at onset), and perpetuating factors (secondary aspects of the illness that cause it to be valued and maintained).

Predisposing risk factors for eating disorders can be biological, psychological, or socioenvironmental. Twin studies and gene studies suggest genetic heritability may contribute to the risk of developing AN or BN. Identified psychological and socioenvironmental factors include childhood adversity or trauma, personality traits such as rigidity and attention to detail, perfectionism, and a high ability to delay reward. Recent evidence suggests an association between food insecurity and higher rates of eating disorder pathology. Additionally, both psychological and societal pressures may serve to maintain an eating disorder including idealization of a thin physique, social isolation, and overcontrol of weight and shape.

Research also suggests that some groups are at disproportionate risk for eating disorders. As summarized in the prevalence section, women have a higher rate of lifetime eating disorders than men, and younger adults (18-29 years) have a higher rate of prevalent eating disorders compared with those older than 50 years. Risk of eating disorder diagnosis also varies by race/ethnicity: White populations have higher rates of prevalent AN and BN, while BED is more prevalent among racial minorities. Similarly, sexual and gender minorities have demonstrated a higher risk for eating disorders than their heterosexual and cisgendered peers. Athletes are another group at higher risk of having an eating disorder. In a study of elite athletes compared with age-matched controls, 13.5 percent of elite athletes met diagnostic criteria for threshold or subthreshold eating disorders compared with controls from the general population (4.6%, p<0.001). In terms of natural history, morbidity and mortality associated with untreated eating
disorders are summarized above. For those diagnosed with one eating disorder, evidence suggests that diagnostic crossover is relatively common over time. For example, in one cohort study of women diagnosed with AN or BN (n=216) 34 percent with an initial diagnosis of AN later developed BN, and 14 percent of those originally diagnosed with BN developed AN over 7 years of followup. 

**Rationale for Screening and Screening Strategies**

Routine screening for eating disorders in populations without signs or symptoms could detect eating disorders early or identify disorders not otherwise known, lead to earlier treatment, and reduce future morbidity and mortality. Assessment of weight, height, and body mass index (BMI) is considered the standard of care in primary care settings, and changes in growth or weight may lead to detection of some eating disorders. Those without obvious physical symptoms may go unrecognized, or symptoms may be attributed to other conditions. In addition, individuals experiencing eating disorder symptoms may not seek care for various reasons. A 2017 systematic review on perceived barriers of help-seeking for eating disorders (k=13 studies) found the following to be the most commonly reported barriers: stigma and shame, denial of and failure to perceive the severity of illness, practical barriers (e.g., cost of treatment), low motivation to change, negative attitudes toward seeking help, lack of encouragement from others to seek help, and lack of knowledge about help resources. Screening questionnaires are available that could be used in primary care settings, including those designed to detect a range of eating disorders, such as the Eating Disorder Screen for Primary Care (EDS-PC), Screen for Disordered Eating, and the SCOFF, which some experts recommend not considering an acronym since signaling questions are based on specific terminology from each signaling questions (e.g., “Have you recently lost more than One stone in a 3 month period?”). Some screening questionnaires that could be used in primary care are designed to detect eating disorders in which binge eating is the hallmark (e.g., BN, BED).

**Interventions/Treatment for Eating Disorders**

Broadly speaking, the recommended treatments for eating disorders involve an interdisciplinary approach encompassing psychological/behavioral, medical, and nutritional interventions. Given the complicated nature of eating disorder symptomatology, screen-detected or newly diagnosed patients are typically referred to specialists for a diagnostic evaluation and specific treatment recommendations based on symptomatology and symptom severity.

Psychological components of treatment include a range of psychotherapies, including family-based therapy (the Maudsley Method), cognitive behavioral therapy, dialectical behavioral therapy, acceptance and commitment therapy, and interpersonal psychotherapy. Many of these were initially developed for other psychiatric conditions but have since been adapted for use in patients with eating disorders. Specific approaches to psychological interventions may vary based on eating disorder diagnosis and severity. Medical management centers around monitoring physical and medical complications of the eating disorder (e.g., cardiac instability, musculoskeletal injury, endocrine function) and providing appropriate medical intervention.
In terms of pharmacotherapy, there is only one Food and Drug Administration–approved pharmacological monotherapy for the treatment of an eating disorder (lisdexamfetamine for BED). However, other psychotropic medications have been evaluated and are sometimes prescribed to patients to address eating disorder symptoms as well as comorbid psychiatric conditions (e.g., depression, anxiety) but are not always indicated.\textsuperscript{37, 38} Nutritional management is focused on providing nutritional rehabilitation to patients who are malnourished and/or helping patients establish a regular pattern of eating that encourages variety and flexibility.\textsuperscript{37, 38} Although the interdisciplinary components of treatment are typical for eating disorder management, guidance may vary by age, with children/adolescents typically requiring a greater involvement of parents/guardians in the treatment process and adults being able to engage in treatment more independently.\textsuperscript{39, 40}

The treatment components above can be administered throughout different levels of care depending on the severity of the eating disorder symptoms. Levels of care can include inpatient medical hospitalization, inpatient psychiatric hospitalization, residential care, partial hospitalization (or day treatment), intensive outpatient therapy, and outpatient therapy. Determination of the appropriate level of care is typically overseen by an eating disorder specialist(s) in conjunction with primary care providers who can provide input on the overall medical status/stability of the patient. Several established guidelines provide guidance on the level of care, and each takes into account medical stability (including weight status), severity and/or frequency of risky eating disorder behaviors (e.g., self-induced vomiting, laxative/diuretic misuse), and overall psychiatric stability (e.g., risk for suicide, chronicity of the disorder).\textsuperscript{34–36} Availability of specialty services and/or geographic limitations are acknowledged as crucial in making level-of-care determinations.

**Recommendations of Other Organizations and Current Clinical Practice**

**Appendix A Table 3** summarizes recommendations from other organizations relevant to screening for eating disorders in clinical settings. Many guidelines mention screening in the context of monitoring for potential signs and symptoms of eating disorders. For example, the American Congress of Obstetricians and Gynecologists’ recommendations highlight gynecologic concerns/symptoms associated with eating disorders and recommend providers be comfortable recognizing and screening “at-risk” patients.\textsuperscript{41} The American Academy of Pediatrics recommends screening for eating disorders by monitoring and assessing risk factors and symptoms at annual and sports physicals, including monitoring for changes in height, weight and BMI longitudinally.\textsuperscript{42} Similarly, guidelines from specialty groups promote awareness of eating disorder symptoms and screening of certain groups who may be at risk (e.g., adolescent and young female athletes). Screening rates for eating disorders in primary care are not clear; we found no recent estimates in the literature related to current clinical practice.
Chapter 2. Methods

Key Questions and Analytic Framework

The scope and key questions (KQs) were developed by the Evidence-based Practice Center (EPC) investigators, USPSTF members, and Agency for Healthcare Research and Quality (AHRQ) Medical Officers. The analytic framework and KQs that guided the review are shown in Figure 1. Five KQs were developed for this review:

1. Does screening for eating disorders in adolescents and adults improve health outcomes, including for specific subgroups of interest?
2. What is the accuracy of primary care–relevant screening tools for eating disorders in adolescents and adults, including for specific subgroups of interest?
3. What are the harms of screening for eating disorders in adolescents and adults, including for specific subgroups of interest?
4. How effective are interventions for improving health outcomes in screen-detected or previously untreated adolescents and adults with eating disorders, including for specific subgroups of interest?
5. What are the harms of interventions for eating disorders, including for specific subgroups of interest?

Data Sources and Searches

We searched PubMed/MEDLINE, the Cochrane Library, and PsycINFO for English-language articles published through June 23, 2020. Medical Subject Headings were used as search terms when available and keywords when appropriate, focusing on terms to describe relevant populations, tests, interventions, outcomes, and study designs. Complete search terms and limits are detailed in Appendix B1. This search was updated during the peer review process by applying the same search strategies, limited from the date of the original searches through December 18, 2020. Targeted searches for unpublished literature were conducted by searching ClinicalTrials.gov. To supplement electronic searches, reference lists of pertinent articles were reviewed. Studies suggested by peer reviewers or public comment respondents were reviewed and, if appropriate, were incorporated into the final review. Since December 18, 2020, ongoing surveillance was conducted through article alerts and targeted searches of journals to identify major studies published in the interim that might affect the conclusions or understanding of the evidence and the related USPSTF recommendation. The last surveillance was conducted on January 1, 2022 and no additional studies meeting eligibility criteria were identified. All literature search results were managed using EndNote™ version 9.2 (Thomson Reuters, New York, NY).
Study Selection

Inclusion and exclusion criteria for populations, interventions, comparators, outcomes, settings, and study designs were developed with input from the USPSTF (Appendix B2). For all KQs, English-language studies of adolescents and adults age 10 years or older conducted in settings generalizable to primary care, including school-based health centers and other community settings, and in countries categorized as “very high” on the United Nations Human Development Index were included.33 Studies limited to populations undergoing evaluation for bariatric surgery or who were identified based on physical signs or symptoms associated with eating disorders were also excluded, including studies limited to populations who are underweight. These include treatment studies (KQ 4) limited to populations with AN who are underweight. Measurement of weight and BMI are part of routine primary care, and assessment of eating disorders would be part of the diagnostic evaluation for those identified as being underweight. The scope of this review is focused on populations with eating disorders who are unlikely to be detected based on physical signs or symptoms in the context of routine care.

For KQs 1 and 3 (direct evidence of benefits and harms of screening), controlled clinical trials enrolling adults with asymptomatic or undetected eating disorders comparing screening with no screening were eligible. Cohort studies with a concurrent control group for KQ 3 (harms of screening) were also eligible. For KQ 2 (accuracy of eating disorder screening tests), cohort or cross-sectional studies of asymptomatic or unselected adolescents or adults comparing one or more screening tests with an acceptable reference standard. Eligible reference standards included a structured or semi-structured diagnostic interview with a mental health clinician or diagnostic questionnaire (e.g., Eating Disorder Examination-Questionnaire). For KQs 1 through 3, eligible screening tests included those used, or feasible for use in primary care settings (e.g., the SCOFF or the EDS-PC. Studies evaluating tests not feasible for routine screening in primary care settings (e.g., longer, time-intensive questionnaires such as the 26-item Eating Attitudes Test) or investigating serologic screening (e.g., using biomarkers) were excluded.

For KQs on benefits (KQ 4) and harms (KQ 5) of treatment, controlled clinical trials of adolescents and adults with screen-detected eating disorders (from primary care or other healthcare settings) were included. Studies enrolling populations from specialty settings who were referred or found the study through advertisements and were diagnosed with an eating disorder but had not been previously treated for eating disorders were also included. As noted above, studies limited to populations identified based on physical signs or symptoms of eating disorders were excluded. Cohort studies with a concurrent control group for KQ 5 (harms of treatment) were also eligible. Eligible studies evaluated psychological interventions (e.g., cognitive behavioral therapy or other forms of therapy) delivered in a group, individual or family-based sessions, including variations of self-help interventions, or pharmacotherapy with Food and Drug Administration–approved medications. Interventions that combine psychological interventions with pharmacotherapy, or those that include other components such as education or nutritional counseling (in addition to therapy or pharmacotherapy) were also eligible. Eligible studies had to compare treatment with an inactive control group (i.e., no treatment, wait-list, minimal intervention [e.g., brief education about eating disorders], or placebo in controlled pharmacotherapy studies). Head-to-head comparisons of two active treatments were excluded. Public awareness campaigns without specific interventions linked to screening, complementary
and alternative therapies, those considered to be adjunctive therapy (e.g., acupuncture, herbal supplements, massage, light therapy), and interventions designed for primary prevention of eating disorders only were ineligible.

Eligible outcomes for KQs on the benefit of screening or treatment (KQs 1 and 4) include measures of health outcomes such as eating disorder remission or symptom reduction based on validated questionnaires (e.g., Eating Disorders Examination-Questionnaire) or diagnostic interviews, health-related quality of life (QOL) or function, depression, and others. Intermediate outcomes were excluded, such as mean change in BMI and frequency of specific behaviors (e.g., change in frequency of binge-eating episodes). For KQ 2 (accuracy of eating disorder screening tests), we included studies reporting on test accuracy, including sensitivity and specificity. Eligible outcomes of KQ 3 (harms of screening) include increased anxiety and labeling and stigma associated with screening, and for KQ 5 (harms of treatment), we included any harms attributed to interventions, such as adverse effects associated with medications.

Two investigators independently reviewed titles and abstracts; those marked for potential inclusion by either reviewer were retrieved for evaluation of the full text. The full texts were then independently reviewed by two investigators to determine final inclusion or exclusion. Disagreements were resolved by discussion and consensus. Covidence systematic review software (Covidence, Melbourne, Australia) was used to assign and track literature review decisions.

Quality Assessment and Data Abstraction

Two reviewers independently assessed each study’s methodological quality. Disagreements in study quality ratings were resolved through discussion or with an independent assessment from a third senior investigator. For randomized, controlled trials (RCTs), the most recent versions of the Cochrane Risk of Bias Tool (RoB 2.0) available for parallel and crossover trials was used. It assessed the following risk-of-bias domains: bias arising from selection or randomization, bias due to missing outcome data, bias due to departures from intended interventions, bias from measurement of outcomes, and bias from selective reporting of results. For studies of diagnostic test accuracy, the QUADAS-2 instrument was used. Our risk-of-bias assessments using these instruments were translated into an overall study quality rating of good, fair, or poor using predefined criteria developed by the USPSTF and adapted for this topic (Appendix B3). Only studies rated as having good or fair quality were included.

For each included study, one investigator extracted pertinent information about the methods, populations, interventions, comparators, outcomes, timing, settings, and study designs. A second team member reviewed all data extractions for completeness and accuracy.

Data Synthesis and Analysis

Findings for each KQ were summarized in tabular and narrative format. The overall strength of the evidence for each KQ was assessed as high, moderate, low, or insufficient based on the
overall quality of the studies, consistency of results between studies, precision of findings, risk of reporting bias, and limitations of the body of evidence, using methods developed for the USPSTF (and the EPC program). Additionally, the applicability of the findings to U.S. primary care populations and settings was assessed. Discrepancies were resolved through consensus discussion.

To determine whether meta-analyses were appropriate, the clinical heterogeneity and methodological heterogeneity of the studies were assessed following established guidance. The populations, tests, treatments, comparators, outcomes, and study designs were assessed qualitatively, looking for similarities and differences. For KQ 2 pooled sensitivities and specificities for screening tests were calculated using a hierarchical summary receiver operating characteristic curve analysis when at least four similar studies were available. For KQ 4, we ran random-effects restricted maximum likelihood models on continuous measures of eating disorder and depression symptom severity (analyzing standardized mean difference or unstandardized mean difference in change between groups) when at least three similar studies were available. For psychological interventions, we pooled studies separately for guided self-help, unguided self-help, and group interventions. When studies reported more than one continuous outcome for eating disorder symptom severity, we preferentially selected the outcome most commonly reported by similar studies. Comprehensive Meta-Analysis version 3.4 (Biostat Inc) and Stata version 16 were used to conduct all quantitative analyses.

**Expert Review and Public Comment**

A draft research plan for this topic was posted on the USPSTF website for public comment from June 25, 2020, to July 28, 2020. In response to comments, the following changes were made: (1) added two specific populations based on a rationale for how these factors may affect outcomes of screening and treatment: “sexual orientation” (in addition to “gender identity”) and “recruitment setting”; (2) clarified which specific populations are of interest in this review by removing the “part b” from each KQ and adding “including for specific populations of interest” at the end of each KQs (3); clarified that interventions designed for primary prevention only are not eligible (KQs 4 and 5); and (4) clarified that eligible interventions could have multiple components. The final version of the research plan was posted on the USPSTF website on September 24, 2020. A draft report was reviewed by four content experts, one representative of Federal partners, USPSTF members, and AHRQ Medical Officers and was revised based on comments received. In response to these comments, we provided additional information regarding the rationale for screening, clarified the scope of the review and limitations of included studies, and expanded the future research needs section. The draft report was posted for public comment from October 19, 2020 to November 16, 2020. Minor revisions were made based on comments received to clarify the rationale for population eligibility criteria and eligibility criteria for studies of treatment. All references suggested by expert or public reviewers were evaluated for inclusion/exclusion.
USPSTF and AHRQ Involvement

The authors worked with USPSTF liaisons at key points throughout the review process to develop and refine the analytic framework and key questions and to resolve issues around scope for the final evidence synthesis.

AHRQ staff provided oversight for the project, coordinated systematic review, reviewed the draft report, and assisted in an external review of the draft evidence synthesis.
Chapter 3. Results

Literature Search

This review identified 15,037 unique records and assessed 1,451 full-text articles for eligibility (Figure 2). The review excluded 1,392 studies for various reasons, detailed in Appendix C, and included 57 unique studies (described in 59 publications). Of the included studies, 17 studies (described in 18 publications) evaluated the accuracy of one or more screening questionnaires for eating disorders (KQ 2). Forty RCTs (described in 41 publications) addressed the benefits (KQ 4) of interventions compared with no treatment for adults and adolescents with screen-detected eating disorders, and nine studies (described in 8 publications) assessed harms of interventions (KQ 5). No studies evaluating the direct benefits or harms of screening (KQs 1 and 3) were found. Details of quality assessments of included studies and studies excluded because of poor quality are in Appendix D Tables 1–4.

Results by Key Question

KQ 1. Does Screening for Eating Disorders in Adolescents and Adults Improve Health Outcomes, Including for Specific Subgroups of Interest?

We identified no eligible studies for this KQ.

KQ 2. What Is the Accuracy of Primary Care–Relevant Screening Tests for Eating Disorders in Adolescents and Adults, Including for Specific Subgroups of Interest?

Summary

Seventeen studies (reported in 18 articles) evaluated the accuracy of screening tests for detecting eating disorders. Most enrolled only females or were predominantly female (>60% girls or women); two enrolled a population with a majority of males, one study set in the Veterans Health Administration (89% male), and one in a primary and secondary school setting (51% male). Included studies assessed a range of screening questionnaires that varied in length from a single question to as many as 16 questions. Screening questionnaires varied in terms of whether they screened for any eating disorder (e.g., transdiagnostic) or if they screened for specific eating disorder diagnoses (e.g., BED only). Thirteen studies evaluated transdiagnostic questionnaires, and most (k=11 assessed the SCOFF). Four evaluated questionnaires designed to detect eating disorders characterized by binge eating only (e.g., BN or BED). Ten studies (4,348 participants) assessed the accuracy of the SCOFF in adult women or women and men. With a cut point of at least 2, the pooled sensitivity of the SCOFF was 84 percent (95% CI, 74% to 91%), and the pooled specificity was 80 percent.
(95% CI, 65% to 89%) (10 studies; 3,684 participants). Using a cut point of 3 or greater for the SCOFF, the pooled sensitivity was 69 percent (95% CI, 56% to 80%), and pooled specificity was 90 percent (95% CI, 69% to 98%) (7 studies; 3,424 participants). A single study evaluated the SCOFF in a sample of adolescent girls and boys and found a sensitivity of 73 percent and a specificity of 78 percent. The only other transdiagnostic screener evaluated by more than one study was the EDS-PC; in two studies, sensitivity of the EDS-PC ranged from 97 to 100 percent, and specificity ranged from 40 to 71 percent using a cut point of at least 2. All other studies of screening test accuracy were evaluated by only one study each.

Detailed Evidence

Ten good- and 7 fair-quality studies (reported in 18 articles) assessed the accuracy of 10 screening questionnaires for identifying any eating disorder or specific eating disorders (Table 1).

Eleven studies (reported in 12 articles) evaluated the SCOFF; most assessed accuracy among adults, and one study evaluated an adapted version for use in a pediatric population. The accuracy of eight other screening questionnaires was evaluated by one study each. Screening questionnaires varied in terms of the range of eating disorders they detected. Most were evaluated to detect any eating disorder (e.g., transdiagnostic), and four were designed to detect a disorder in which binge eating is the hallmark (e.g., BN, BED).

The screening questionnaires were compared with two categories of reference standards, either a diagnostic clinical interview or a longer self-reported diagnostic questionnaire. Nine studies used a semi-structured interview as the reference standard based on Diagnostic and Statistical Manual (fourth edition) (DSM-IV) or Diagnostic and Statistical Manual fifth edition (DSM-5) diagnostic criteria, either diagnostic clinical interview, or the Eating Disorder Examination (EDE). Seven studies used the following diagnostic questionnaires to diagnose eating disorders: Questionnaire on Weight and Eating Patterns-Revised (QEWP-R), Eating Attitudes Test-26, Questionnaire for Eating Disorder Diagnoses (QEDD), EDE-Questionnaire (EDE-Q), and Eating Disorders Inventory-2.

Most studies (k=6) were set in the United States. The remainder were in the United Kingdom, Italy, Taiwan, Malaysia, and other European countries. Studies enrolled participants from community settings or outpatient clinical settings. Community recruitment of adults most often occurred in a university setting, while one study recruited young adults identified via a population database through the mail, and another recruited students from primary and secondary schools. Outpatient primary care clinics were the most common clinical setting from which patients were recruited (k=5). Of these, two recruited patients from a Veterans Affairs (VA) primary care clinic, and one enrolled from primary care and a university campus. Four studies recruited from specialty or referral settings, including one from an outpatient psychiatry clinic and three from diet or obesity clinics.

Detailed Evidence

Ten good- and 7 fair-quality studies (reported in 18 articles) assessed the accuracy of 10 screening questionnaires for identifying any eating disorder or specific eating disorders (Table 1).

Eleven studies (reported in 12 articles) evaluated the SCOFF; most assessed accuracy among adults, and one study evaluated an adapted version for use in a pediatric population. The accuracy of eight other screening questionnaires was evaluated by one study each. Screening questionnaires varied in terms of the range of eating disorders they detected. Most were evaluated to detect any eating disorder (e.g., transdiagnostic), and four were designed to detect a disorder in which binge eating is the hallmark (e.g., BN, BED).

The screening questionnaires were compared with two categories of reference standards, either a diagnostic clinical interview or a longer self-reported diagnostic questionnaire. Nine studies used a semi-structured interview as the reference standard based on Diagnostic and Statistical Manual (fourth edition) (DSM-IV) or Diagnostic and Statistical Manual fifth edition (DSM-5) diagnostic criteria, either diagnostic clinical interview, or the Eating Disorder Examination (EDE). Seven studies used the following diagnostic questionnaires to diagnose eating disorders: Questionnaire on Weight and Eating Patterns-Revised (QEWP-R), Eating Attitudes Test-26, Questionnaire for Eating Disorder Diagnoses (QEDD), EDE-Questionnaire (EDE-Q), and Eating Disorders Inventory-2.

Most studies (k=6) were set in the United States. The remainder were in the United Kingdom, Italy, Taiwan, Malaysia, and other European countries. Studies enrolled participants from community settings or outpatient clinical settings. Community recruitment of adults most often occurred in a university setting, while one study recruited young adults identified via a population database through the mail, and another recruited students from primary and secondary schools. Outpatient primary care clinics were the most common clinical setting from which patients were recruited (k=5). Of these, two recruited patients from a Veterans Affairs (VA) primary care clinic, and one enrolled from primary care and a university campus. Four studies recruited from specialty or referral settings, including one from an outpatient psychiatry clinic and three from diet or obesity clinics.
Most studies reported on the prevalence of any eating disorder, which ranged from 2 to 46 percent; four studies examined BN or BED only, where prevalence ranged from 8 to 22 percent. The three studies reporting an eating disorder prevalence of greater than 20 percent were conducted in a pediatric obesity clinic, an outpatient diet clinic, and a university health center.

Most study populations were either entirely female or predominantly female, enrolling over 60 percent girls or women. Males outnumbered females in two studies: one study set in the Veterans Health Administration (89% male) and one in a primary and secondary school setting (51% male). One study did not report the sex of participants.

Sample sizes analyzed varied from 51 to 1,541 participants, with a median of 341. Across the 15 studies that reported on the age of enrolled participants (mean, median, or range), age ranged from 11 to 62 years. Two studies focused on study populations under 18 years of age, recruiting from pediatric or school settings. Three studies enrolled an older sample of participants, mean ages ranging from 44 to 62 years; two of these studies were set in the Veterans Health Administration and one recruited participants from an outpatient specialty clinic for metabolic diseases. The remaining 10 studies included participants in which the mean age, median age, or age range primary fell between 20 and 30 years. Among the nine studies reporting BMI (mg/kg^2), four studies reported a mean BMI of less than 25 (underweight or normal), two studies reported mean BMI ranging from 25 to 30 (overweight), two studies reported BMI over 30 (obese), and one study reported BMI range of 16 to 44.

Of the 17 studies, eight reported on race or ethnicity; the percentage of participants who were non-White ranged from 12 to 100. For the six studies set in the United States, non-White participants ranged from 12 to 52 percent. Only two studies reported mental health comorbidity: mood disorder (44%) and dysthymia (29%) were the most common psychiatric disorders reported, respectively.

We rated 10 studies as good quality and the remainder as fair quality (Appendix D Table 1). In the studies rated as fair quality, common limitations included risk of bias in patient selection (e.g., accuracy metrics based on a subset of participants who underwent screening) and the flow and timing domain.

Screening test accuracy results are organized by screener below. Detailed results by screening questionnaire are also shown in Table 2.

**Studies Evaluating the SCOFF**

Ten studies assessed the accuracy of the SCOFF screening questionnaire among adults (Table 1). Of these, two enrolled populations from primary care settings, four enrolled from university-based settings, one recruited veterans, and two recruited from specialty settings including an outpatient diet clinic and an outpatient psychiatric clinic. One study used a nationally representative sample of households in Finland. The SCOFF is a 5-question self-report screening questionnaire assessing the following hallmark criteria for eating disorders: self-induced vomiting, recent weight loss, binge eating, and intrusive thoughts about food or...
body weight or shape. Scores on the SCOFF range from 0 to 5 with each positive response assigned one point. Using the original cut point of at least 2 in studies enrolling adult men and women or women only, the pooled sensitivity based on 10 studies (4,348 participants) was 84 percent (95% CI, 74% to 90%), and the pooled specificity was 80 percent (95% CI, 65% to 89%) (Appendix F Figure 1). Seven studies (3,424 participants) also assessed the accuracy using a cut point of 3. The pooled sensitivity was lower at 69 percent (95% CI, 56% to 80%), but the specificity improved to 90 percent (95% CI, 69% to 98%) (Appendix F Figure 2).

A single study assessed the accuracy of the SCOFF among a sample of 954 female and male adolescents (mean age=13.6, standard deviation=1.31) enrolled from primary and secondary schools. The optimal cut point was at least 2, which yielded a sensitivity of 73 percent (95% CI, 63% to 83%) and a specificity of 78 percent (95% CI, 75% to 80%).

**Specific populations.** Two included studies also evaluated the SCOFF among various populations of adults including sex, age, and BMI; one recruited a sample of 1,541 women and men from an outpatient psychiatric clinic, and another recruited 147 women from a primary care setting. The one study reporting on sex differences found that because it used a cut point of at least 2 the SCOFF demonstrated greater sensitivity in women (95% vs. 86%), but greater specificity in men (74% vs. 66%). With respect to age, younger age (less than 30 years) was associated with better SCOFF performance in women, but the opposite pattern was found among men. Across both studies, the SCOFF performed better at a lower BMI (<27.5 kg/m²) for both women and men (Table 2).

**Studies Evaluating Other Eating Disorder Screeners**

Eight other screening questionnaires were assessed across eight included studies; of these, one (EDS-PC) was assessed in two studies, and all others were evaluated by one study each (Table 1). The screening questionnaires varied in terms of whether they were designed to screen for any eating disorder (i.e., transdiagnostic) or binge eating only. Results are organized by diagnostic category below and in Table 2, highlighting results for commonly used cut points or cut points considered optimal. Many studies reported on the accuracy of screening tests at various cut points, and some reported on the accuracy of different eating disorder reference standard definitions.

Four separate transdiagnostic screening tests were evaluated across four included studies among adults enrolled from various settings including primary care, colleges or universities, and the VA. Two studies evaluated the EDS-PC: one recruited females and males from primary care (77% female) and the other recruited all female veterans. The EDS-PC is a 5-item self-report measure developed specifically for use in the primary care setting. Scores range from 0 to 5, and a cut point of at least 2 is considered a positive screen. Across both studies (627 participants), sensitivity ranged from 97 to 100 percent, and specificity ranged from 40 to 71 percent.

The three remaining transdiagnostic screening tests were assessed in single studies that recruited adult females only. The Stanford-Washington University Eating Disorder Screen (SWED) and the Eating Disturbance Scale (EDS) were both evaluated in college or university settings.
populations. The SWED is an 11-item self-report measure derived from existing questionnaires assessing eating disorder behaviors (e.g., dietary restriction, self-induced vomiting) and weight concerns. Among 549 college-aged women, the SWED demonstrated a sensitivity of 80 percent and a specificity of 82 percent in detecting any eating disorder with a cut point of at least 59 on the Weight Concerns Scale (WCS).52 The EDS is a 5-item self-report questionnaire that assesses eating-disordered thoughts and behaviors (e.g., feeling guilty after eating, engaging in strict diets). Each item is scored on a Likert-type scale from 1 to 7, with seven being the most pathological response. Using a cut point of at least 16, the EDS had a sensitivity of 90 percent and a specificity of 88 percent among a smaller sample of 51 women enrolled in teaching or nursing colleges.56

The Screen for Disordered Eating (SDE) is a 5-item self-report questionnaire derived from existing measures with face valid items that map onto eating disorder diagnoses including AN, BN, and BED.30 The SDE was evaluated in a single study of 402 female veterans and demonstrated a sensitivity of 91 percent (95% CI, 80% to 96%) and a specificity of 58 percent (52% to 63%) using a cut point of at least 2.30

Studies Evaluating Binge-Eating Disorder Screeners

Four studies screened for eating disorders in which binge eating is the hallmark (e.g., BN, BED).59-62 Three screening questionnaires were evaluated in adult populations: the Binge Eating Scale (BES),61 the VA Binge Eating Screener (VA-BES),62 and the Eating Disorder Module of the Patient Health Questionnaire (PHQ-ED).60 The BES is a 16-item self-report questionnaire assessing thoughts and feelings associated with binge-eating episodes. Scores range from 0 to 46 with higher scores indicating a more pathological response. In a sample of 334 adults with obesity recruited from an outpatient metabolic clinic, a cut point score of at least 17 on the BES yielded a sensitivity of 85 percent and a specificity of 75 percent.61 The VA-BES is a single item derived from a more general assessment used in the VA’s weight management program—the VA MOVE! program.62 The binge-eating item (referred to as the VA-BES in this included study) asks participants how often they eat large amounts of food while feeling out of control (i.e., a binge-eating episode) with Likert-type responses ranging from “never” to “five or more times per week.” In a sample of 116 veterans recruited from the MOVE! program (12% female), a cut point of at least 2 on the VA-BES yielded a sensitivity of 89 percent and a specificity of 83 percent.62 The PHQ-ED is a 6-item self-report questionnaire that assesses engagement in binge-eating episodes or inappropriate compensatory behaviors (e.g., self-induced vomiting, abuse of laxatives).60 Two additional items assessing frequency and duration are required if either binge-eating episodes or inappropriate compensatory behaviors are endorsed. In a sample of 348 adults (82% female) recruited from a health maintenance organization, a positive screen on the PHQ-ED demonstrated a sensitivity of 100 percent and a specificity of 28 percent.60

One study assessed a screening for binge eating in adolescents recruited from a pediatric obesity clinic (94 participants, age range 11 to 18 years) using the Adolescent Binge-Eating Disorder Questionnaire (ADO-BED).59 The ADO-BED is a 10-item self-report questionnaire that assesses several diagnostic criteria for BED. The first two questions are interrelated and assess the presence of binge-eating episodes (“Do you sometimes have a strong craving to eat although you are not really hungry or have recently eaten?” and “In this situation, do you sometimes find
yourself starting to eat and then being unable to stop?”). A positive response to both questions yielded a sensitivity of 100 percent, but a specificity of only 27 percent for detecting binge eating.59

KQ 3. What Are the Harms of Screening for Eating Disorders in Adolescents and Adults, Including for Specific Subgroups of Interest?

We identified no eligible studies for this KQ.

KQ 4. How Effective Are Interventions for Improving Health Outcomes in Screen-Detected or Previously Untreated Adolescents and Adults With Eating Disorders, Including for Specific Subgroups of Interest?

Summary

Forty included RCTs assessed heterogeneous interventions, none enrolled populations who were screen-detected. One was limited to adolescents; all others enrolled adults. Most enrolled a majority of women and limited to populations with BED or BN only. Eighteen assessed the benefit of pharmacotherapy. Four trials of lisdexamfetamine for BED (900 participants) measured change in eating disorder symptom severity using the Yale–Brown Obsessive Compulsive Scale modified for binge eating (YBOCS-BE) and found larger reductions in change from baseline scores associated with lisdexamfetamine than placebo (pooled mean difference, -5.75 [95% CI, -8.32 to -3.17]). Two trials compared topiramate with placebo for BED (465 participants) and both found significantly larger reductions in YBOCS-BE scores from baseline among those receiving topiramate than placebo, from -6.40 (p<0.001) to -2.55 (p=0.004). Five trials assessed various selective serotonin reuptake inhibitors (SSRIs) among persons with BED; two reported on change in eating disorder symptoms and results were imprecise. SSRIs were associated with a larger reduction in depression symptom scores than placebo over 6-16 weeks (pooled standardized mean difference [SMD], -0.6 [95% CI, -0.90 to -0.33]) (5 studies; 208 participants). Three trials assessed fluoxetine for populations with BN; two found benefit favoring fluoxetine for eating disorder symptom severity and depression. Twenty-four trials assessed a psychological intervention. Guided self-help improved eating disorder symptom severity more than control (pooled SMD for difference in change from baseline score, -0.96 [95% CI, -1.26 to -0.67]) (5 studies; 391 participants); pooled estimates for unguided self-help (6 studies, 368 participants) also favored the intervention, but the difference between groups was not statistically significant (SMD, -0.18 [95% CI, -0.38 to 0.03]). Similarly, self-help interventions also reduced depression symptoms more than control, including both guided self-help (pooled SMD, -0.73 [95% CI, -1.04 to -0.43]; 4 studies, 324 participants) and unguided self-help (pooled SMD, -0.37 [95% CI, -0.68 to -0.05]; 3 studies, 156 participants). Few trials of self-help measured other outcomes. Group therapy (7 studies, 253 participants) was associated with larger reductions in depression scores from baseline than inactive control (pooled SMD, -0.48 [95% CI, -0.69 to -0.27]). Few studies of group therapy measured other outcomes. Four studies assessed different forms of individual therapy and measured heterogeneous outcomes. Common limitations included risk of bias due to missing outcome data, and few trials
assessed outcomes over a duration longer than 16 weeks (most reported outcomes over 6 to 12 weeks).

**Detailed Evidence**

We identified 40 RCTs comparing an intervention for eating disorders with a control: 18 (19 publications) evaluated pharmacotherapy, and 24 evaluated therapy; two RCTs included both pharmacotherapy and therapy interventions.

**Pharmacotherapy**

Eighteen included trials evaluated the benefit of pharmacotherapy versus placebo among populations with an eating disorder (**Table 3**); most (k=14) enrolled populations with BED defined by DSM-IV or DSM-5 criteria. Four trials enrolled populations with BN defined by DSM-III criteria, and two of these were limited to populations with BN and recurrent binge-eating behavior. No included studies enrolled populations who were screen detected in primary care (or via routine screening in other healthcare settings). Most trials described recruiting participants via media or newspaper advertisements or a combination of both referrals and trial advertisements. Four trials mentioned that advertisements were for studies addressing binge eating and obesity. Two trials mentioned recruiting participants from referrals or from specialty clinical settings only, and one trial invited potentially eligible participants identified by chart reviews of patients previously seen at a university-based outpatient clinic for “weight problems.” Most studies were set in the United States (k=13); three trials (described in 2 publications) enrolled participants from both the United States and other country settings (Canada or various European countries) and one trial each was set in Switzerland and Finland.

All studies enrolled adults, and only one included both adolescents and adults (15 years or older) with BN. Mean ages of enrolled populations ranged from 25 to 44 years. Five trials were limited to women only, and all others enrolled a majority of women (78% to 96%). Across 14 studies describing race or ethnicity, all enrolled a minority of non-White participants ranging from 3 to 27 percent. One study of duloxetine was limited to adults with BED and current depressive disorder. Seventeen trials reported on mean BMI (kg/m²), and most enrolled a population with a mean BMI greater than 33; four studies (all focused on populations with BN) enrolled populations with a lower mean BMI (range 25 to 27). Eleven studies described the proportion of participants with various psychiatric comorbidity using heterogenous definitions (**Table 3**). The most commonly reported comorbid condition was depression (k=8), with rates of lifetime depressive disorder or any mood disorder ranging from 32 to 77 percent, and rates of current depression ranging from 23 to 52 percent. Ten studies focused on BED limited to populations who were overweight or obese (using various criteria), the mean BMI in these studies ranged from 34.9 to 44.

Studies assessed various pharmacotherapy. Three medications were evaluated by two or more trials: fluoxetine (k=4), lisdexamfetamine (k=4), and topiramate (k=2). Medications evaluated by one study include fluvoxetine, sertraline, escitalopram, duloxetine, bupropion, desipramine, and imipramine. The dose of medications evaluated is
shown in Table 2. All trials compared the active medication with placebo. One trial of imipramine included a co-intervention of dietary counseling and psychological support delivered to both groups.78

All included trials were parallel RCTs. Sample sizes ranged from 20 to 404 participants, and most evaluated outcomes over a duration of 6 to 12 weeks; three assessed outcomes over a slightly longer duration (16 weeks).77,83,84 Detailed results of studies are summarized in Appendix E Table 2. Two were rated good quality and the remainder were rated fair. Common limitations included risk of bias due to missing outcome data.

**Binge-Eating Disorder**

*Lisdexamfetamine*

Four trials (described in 3 publications) compared lisdexamfetamine with placebo among adults with BED.66,68,70 One (n=260) randomized participants to three separate doses of medication (30, 50, or 70 mg/day), and three randomized participants to flexible dosing of lisdexamfetamine (20 to 70 mg/day, mean 57 to 60 mg/day). All measured BED symptom severity using the YBOCS-BE; for doses ranging from 50 to 60 mg/day, pooled mean difference in change from baseline score over 11 to 12 weeks (4 studies, 900 participants) was -5.75 (95% CI, -8.32 to -3.17) (Figure 3). This difference falls within the range considered a minimum clinically important change on the YBOCS-BE (-4 to -17).85 Other eligible outcomes were reported by only one or two studies each (Appendix E Table 2). One trial found significantly higher reduction in mean BES scores associated with lisdexamfetamine (difference between groups, -5.4; p=0.002 for 50 mg/day).70 but no significant difference on measures of depression, anxiety or QOL using the 12-Item Short Form Survey (SF-12). Two trials (reported in 1 publication) found no significant differences between groups in functional impairment measured by the Sheehan Disability Scale.68

*Topiramate*

Two trials of topiramate (465 participants) measured reduction in eating disorder symptom severity using the YBOCS-BE over 14 to 15 weeks (mean or median dose of 212 to 300 mg/day),77,81 and both found significant improvement favoring topiramate (Figure 3). The difference between groups in mean change from baseline score was within the range considered a minimum clinically important change (-4 to -17) in one trial (-6.40)77 but not the other (-2.55).81 Neither found a significant difference between groups on depression scores (measured by the Montgomery-Åsberg Depression Rating Scale [MADRS]77 and Hamilton Depression Rating Scale [HAM-D]81) and one found no significant differences between groups for anxiety measured by Hamilton Anxiety Rating Scale (HAM-A) scores.70

*Selective Serotonin Reuptake Inhibitor*

Five trials assessed an SSRI for improving BED: two assessed fluoxetine,71,83 and one study each assessed fluvoxetine,75 sertraline,73 and escitalopram.79 None selected participants based on presence of comorbid depression. Four reported on the proportion with lifetime major depressive
Three trials also reported on the prevalence of current major depressive disorder with rates ranging from 18 to 25 percent. Two measured eating disorder symptom severity and both found reduction in symptom scores favoring SSRIs (Figure 4), although results were imprecise; other trials of SSRIs reported on intermediate outcomes associated with binge eating only. The trial of fluoxetine found no significant difference between groups on EDE-Q scores (SMD -0.29; 95% CI, -0.83 to 0.24) and the trial of escitalopram found significant benefit for YBOCS-BE score reduction (SMD -0.69; 95% CI, -1.30 to -0.08) that did not fall within the range considered a minimum clinically important change (-4 to -17). All five reported on change in depression symptoms over 6-16 weeks (Figure 4); SSRIs were associated with a larger reduction in depression symptom scores than placebo over 6-16 weeks (pooled SMD, -0.6 [95% CI, -0.90 to -0.33]) (5 studies; 208 participants).

Other Medications

One trial each evaluated duloxetine (n=40), bupropion (n=61), and imipramine (n=31) for populations with BED (Appendix E Table 2). No significant difference was found between groups for any health outcome, including measures of eating disorder symptom severity associated with bupropion (EDE-Q) and duloxetine (YBOCS-BE), and measures of depression associated with imipramine (HAM-D and Self-Rating Depression Scale) and bupropion (BDI).

Bulimia Nervosa

Four trials enrolled populations with BN, including two that limited to populations with BN and recurrent binge-eating behavior. Three assessed fluoxetine and one assessed desipramine. One trial of fluoxetine also included an arm comparing fluoxetine plus a self-help intervention with placebo (in addition to fluoxetine alone).

The three trials of fluoxetine assessed a dose of 60 mg/day (Appendix E Table 2); two reported outcomes at 8 weeks and one at 16 weeks. All reported on a measure of eating disorder symptom severity. two found benefit associated with fluoxetine over 8 weeks on the Eating Attitudes Test (EAT); differences between groups were significant in one trial (difference in mean change from baseline: -4.5; p=0.001), but not the second. One found no significant difference between groups over 16 weeks on the Eating Disorder Inventory (EDI) but only reported a p-value (p>0.05). All three measured depression using the HAM-D. Only one found a significant difference favoring fluoxetine (difference in mean change from baseline: -2; p<0.001 at 8 weeks). In the other two trials, one found a similar reduction in HAM-D scores favoring fluoxetine that was not statistically significant (-2.10; p=0.12), and the other reported that changes were not significant but did not provide numerical results for mean score changes. The one trial assessing a combined intervention (fluoxetine plus self-help vs. placebo) found no significant differences between groups on EDI and HAM-D score changes from baseline.

In the trial assessing desipramine, there was no significant difference between groups on measures of BN symptom severity (EAT), depression (HAM-D, BDI), or anxiety (State-Trait Anxiety Inventory Scores) (Appendix E Table 2).
Psychological Interventions

Twenty-four trials assessed the benefit of a psychological intervention for eating disorders compared with an inactive control (Table 4)\textsuperscript{83, 84, 86-107} Most trials enrolled populations with binge eating, either BED defined as those with BN and recurrent binge-eating behavior; one trial enrolled those with BN without mention of binge eating,\textsuperscript{106} and three enrolled women with any eating disorder based on DSM-5.\textsuperscript{86, 92, 98} No included studies enrolled populations who were screen detected in primary care. Most trials recruited participants via media or newspaper advertisements or a combination of both referrals and trial advertisements. Nine trials were set in the United States; others were set in a range of country settings, including the United Kingdom (k=3),\textsuperscript{88, 92, 100} Canada (k=4),\textsuperscript{89, 93, 97, 104} Sweden, (k=2)\textsuperscript{91, 96} Germany (k=3),\textsuperscript{87, 99, 106} and one each in Switzerland,\textsuperscript{91} and Australia.\textsuperscript{86}

One study was limited to adolescents (12 to 18 years; mean age of 15),\textsuperscript{107} and all others enrolled adults; one study enrolled both adults and adolescents (as young as 14 years).\textsuperscript{98} Among studies enrolling adults, mean ages of enrolled populations ranged from 22 to 46 years. Included studies either limited enrollment to women or girls only (k=11) or enrolled a majority of women (72% to 99%). In 17 studies describing race or ethnicity of enrolled populations, one limited to Latinas only,\textsuperscript{105} two enrolled a population that was 54 to 55 percent non-White (from the United States),\textsuperscript{90, 95} and the others enrolled a population with a majority of White participants. Four studies limited to populations who were overweight or obese only (using various criteria).\textsuperscript{83, 90, 95, 96} Most trials reported on mean BMI (kg/m\textsuperscript{2}) (Table 5), and the majority enrolled populations with a BMI in the overweight or obese range; seven enrolled populations with a BMI less than 25.\textsuperscript{86, 88, 92, 94, 97, 100, 106}

All included trials were parallel RCTs. Sample sizes ranged from 17 to 154 participants and most evaluated outcomes over a duration of 8 to 16 weeks. Studies focused on a variety of psychological interventions (Table 4). Most evaluated a form of self-help (k=14) based on CBT, DBT, or other strategies designed to help participants cope with eating disorder symptoms.\textsuperscript{84, 88-91, 93, 95, 97, 98, 100, 104, 105, 108} Seven trials evaluated a type of group therapy for eating disorders,\textsuperscript{86, 96, 99, 101-103, 106} and four evaluated a form of individual CBT.\textsuperscript{83, 92, 94, 107} Appendix E Table 1 provides additional detail related to the intervention components and intensity. Results are summarized below by intervention type.

Self-Help Interventions

Of the 13 trials evaluating a self-help intervention, seven assessed a form of “guided” self-help,\textsuperscript{87, 88, 91, 93, 100, 104, 105} and seven assessed an “unguided” self-help intervention.\textsuperscript{84, 89, 90, 93, 95, 97, 98} One trial compared both guided and unguided self-help interventions with a control.\textsuperscript{92} Guided interventions included ongoing support and guidance; for example, three studies included six to eight brief (25- to 30-minute) individually guided sessions or phone calls for support;\textsuperscript{93, 104, 105} two provided regular email contact with coaches, (1 to 2 per week)\textsuperscript{88, 91} and one provided individual feedback from therapists on 11 completed assignments.\textsuperscript{87} Unguided interventions involved providing the intervention materials with instructions (e.g., providing a written manual or access to online modules with activities). Most (k=9) self-help interventions were based on CBT, two were based on DBT.\textsuperscript{93, 104} One trial each assessed other types of self-help, including an
online form of The Body Project intervention, and one trial assessed two forms of self-help (compassion-focused therapy and behaviorally based self-help).

Most trials of self-help reported on changes in eating disorder severity and depression using various outcome measures. Figure 5 shows results for self-help interventions organized by intervention type (guided or unguided) and outcome. For reduction in eating disorder symptom severity (measured by the EDE or EDE-Q), guided self-help was associated with a larger reduction in eating disorder severity scores than control over 12-24 weeks [pooled SMD, -0.96 [95% CI, -1.26 to -0.67] [5 studies, 391 participants]). Results from studies assessing unguided self-help (6 studies, 368 participants) were consistent in favoring self-help, but pooled results were not statistically significant (SMD, -0.18 [95% CI, -0.38 to 0.03]) (Figure 5). For measures of depression, pooled results demonstrated larger reductions in mean scores than controls for both guided self-help (SMD, -0.73 [95% CI, -1.04 to -0.43]; 4 studies; 324 participants) and unguided self-help (SMD, -0.37 [95% CI, -0.68 to -0.05]; 3 studies; 156 participants).

Few trials of self-help measured other outcomes. One trial of DBT guided self-help (n=60) measured the Eating Disorder Quality of Life Scale (EDQLS) over 13 weeks and found significantly higher improvement in EDQLS scores from baseline among the self-help group than controls (18.37 vs. 0.14; difference in change from baseline: 18.23; p<0.05). A second trial of DBT self-help (assessing both guided and unguided interventions) measured general health-related QOL using the Short Form 6D (SF6D) and found no significant differences between either form of self-help and controls. Finally, one trial of guided CBT-self-help measured QOL using the World Health Organization Quality of Life scale at 12 weeks and only reported on subdomain scores; results were mixed, with significant improvement on some domains but not others. Two studies reported on anxiety (one guided and one unguided) using different outcome measures and time points; one (n=56) found a statistically significant benefit in favor of guided CBT at 12 weeks on the HADS-Anxiety scale (SMD, -0.91 [95% CI, -1.38 to -0.44]) and one trial (n=76) of unguided CBT found no significant differences between groups at 8 weeks on the BAI (SMD, -0.50 [95% CI, -1.03 to -0.03]).

**Group and Individual Interventions**

Outcomes from trials assessing group or individual psychological interventions are shown in Figure 6. Seven trials assessed a group-based psychological intervention, including four CBT-based group interventions, and one trial each of group-based behavioral activation, behaviorally oriented therapy, impulsivity-focused therapy and interpersonal therapy. One trial compared two forms of therapy (CBT based and interpersonal therapy based) with a control group. Most enrolled women with BED or BN and recurrent binge eating; one enrolled women with any eating disorder based on DSM-5. Most offered 8 to 10 weekly group sessions (90 minutes each); one trial offered sessions twice weekly for 8 weeks, then 8 weekly sessions (total of 16). Most studies of group interventions reported on depression using various measures (Figure 6); group therapy was more effective than control for reducing depression symptoms over 8 to 16 weeks (pooled SMD, -0.48 [95% CI, -0.69 to -0.27]) (7 studies; 253 participants). Few studies of group therapy measured other outcomes. Three measured eating disorder symptom severity using the EDE-Q; one study of group-based CBT found a statistically significant benefit compared with controls (SMD, -1.01 [95% CI, -1.71 to -0.31]), and two
trials found no significant differences between groups (Figure 6).\textsuperscript{96, 99} Two trials of group therapy assessed changes in anxiety scores over 8 to 10 weeks, and neither found a significant difference between groups (Figure 6).\textsuperscript{96, 99}

Four trials assessed an individual psychological intervention. One trial of CBT (8 sessions) was limited to adolescents and found no significant difference between CBT and control groups at 12 or 24 weeks on measures of depression (BDI) and psychosocial functioning (Youth Social Adjustment Scale).\textsuperscript{107} Of the three trials enrolling adults, two focused on individual CBT (16 to 21 sessions),\textsuperscript{83, 92} and one evaluated appetite-focused DBT (AF-DBT) delivered over 12 weekly sessions.\textsuperscript{94} One trial evaluated two forms of individual CBT (n=154): one focused on eating disorder pathology only, and a second “broad form” also addressed other problems common with eating disorders\textsuperscript{92} and found no significant differences between groups in the change from baseline EDE-Q scores (Figure 6). The second trial of individual CBT (n=108) found a significant improvement in depression scores (measured by the BDI) among the intervention group compared with controls (SMD, -0.60 [95% CI, -1.14 to -0.06]).\textsuperscript{83} Finally, the trial of AF-DBT (n=32) found improvement in EDE-Q scores (SMD, -1.18 [95% CI, -1.94 to -0.43]) and depression measured by the BDI-II (SMD, -0.92 [95% CI, -1.65 to -0.19]) over 6 weeks (Figure 6).\textsuperscript{94}

**KQ 5. What Are the Harms of Interventions for Eating Disorders, Including for Specific Subgroups of Interest?**

No included studies of psychological interventions reported on harms. Nine trials of pharmacotherapy reported various harms associated with four medications, including lisdexamfetamine (k=4),\textsuperscript{66, 68, 70} topiramate (k=2),\textsuperscript{77, 81} fluoxetine (k=2),\textsuperscript{71, 76} and escitalopram (k=1).\textsuperscript{79} Characteristics are described in KQ 4 and Table 3, and outcomes are shown in Appendix E Table 3.

In one trial of lisdexamfetamine (n=259) over 11 weeks,\textsuperscript{70} one participant died during the study, and postmortem toxicology analysis found that methamphetamine/amphetamine levels were consistent with a methamphetamine overdose (death was not attributed to the study drug). Across all three arms of lisdexamfetamine and placebo, rates of specific adverse effects were low (Appendix E Table 3); combined incidence of any treatment-emergent adverse events was higher (85%) for the treatment group than the placebo group (59%).\textsuperscript{70} Similarly, a second trial (n=50) of lisdexamfetamine reported few rates of specific adverse events over 12 weeks;\textsuperscript{66} those that were significantly higher in the treatment arm versus control included insomnia, jitteriness, and dry mouth. Finally, in one study reporting on two separate trials of lisdexamfetamine (626 total participants), more than 50 percent in each treatment group reported treatment-emergent adverse events; more treatment-emergent adverse events were related to lisdexamfetamine than with placebo, those reported by more than 10 percent treated with lisdexamfetamine included dry mouth, headache, and insomnia (vs. no treatment-emergent adverse events reported by more than 10% of the placebo groups).\textsuperscript{68}

Two trials of topiramate reported rates of specific adverse events over 14-16 weeks, and both found significantly higher rates of paresthesia and taste perversion\textsuperscript{77, 81} among groups receiving
topiramate versus placebo; one also found significantly higher rates of difficulty concentrating\textsuperscript{77} and another found significantly higher rates of confusion.\textsuperscript{81}

In three trials of SSRIs, one assessing fluoxetine found rates of several adverse effects significantly higher in the treatment arm than in the placebo arm (Appendix E Table 3), such as insomnia, nausea, and tremor; however, authors noted that there was no significant difference between groups in the proportion that discontinued because of adverse effects. The second trial of fluoxetine reported no significant differences between groups for any adverse effects over 6 weeks.\textsuperscript{71} The trial of escitalopram (n=44) reported adverse effects experienced by 5 percent or more of enrolled participants (Appendix E Table 3) and reported that no differences between groups were statistically significant.\textsuperscript{79}
Chapter 4. Discussion

Summary of Evidence

Table 5 provides a summary of the main findings in this evidence review organized by KQ along with a description of consistency, precision, quality, limitations, strength of evidence, and applicability.

Evidence for Benefit and Harms of Screening

We did not find direct evidence on the benefits and harms of screening. Potential harms include false-positive screening results that lead to unnecessary referrals (and associated time and economic burden), treatment, or labeling. Based on our pooled analyses of the SCOFF for detecting any eating disorder among adults 10 studies (4,348 participants), the expected rate of false-positives tests would be 20 percent (Table 2). Other harms of screening are likely to be minimal because screening is noninvasive.

Diagnostic Test Accuracy

Screening tools are available for clinical practice that may reasonably identify adults with eating disorders. Of 18 included studies reporting on the accuracy of screening questionnaires, most reported on the SCOFF (k=10) and assessed the accuracy to identify adults with any eating disorder. Most others were assessed by only one study each, limiting our ability to make stronger conclusions about the accuracy of screening tests.

Some studies of screening test accuracy were limited by unclear applicability to populations presenting for routine primary care. For example, the one study assessing the accuracy of a screening questionnaire among adolescents enrolled participants from a pediatric obesity clinic, which is likely to have a higher prevalence of binge eating. Overall, the estimates of SCOFF screening test accuracy were derived from populations with a current prevalence of eating disorders of approximately 4 to 46 percent based on the reference standard.

Benefits of Interventions for Screen-Detected or Recently Diagnosed Eating Disorders

Forty RCTs evaluated benefits of interventions to improve eating disorder symptoms, most limited to adult women with BED and BN. Included interventions were heterogenous. Both lisdexamfetamine and topiramate were effective in reducing eating disorder severity for populations with BED measured by the YBOCS-BE but were also associated with various adverse effects. Studies reported outcomes over a relatively short duration (8 to 12 weeks). Few studies of SSRIs reported an eligible health outcome specific to eating disorder symptoms,
however, results of 5 studies enrolling populations with BED found consistent improvement in depression symptoms associated with various SSRIs based on mean score changes. Although studies did not enroll participants based on depression status, rates of lifetime depression among participants with BED ranged from 37 to 77 percent in four studies that reported on mental health comorbidity. Population cohort studies in the United States indicate high rates of mental health comorbidity among populations with BED. In the 2001–2003 National Comorbidity Replication cohort study (n=9,282), 32 percent of U.S. adults with BED reported a history of major depression and 65 percent reported a history of any anxiety disorder. Results for other medications were mixed; studies included heterogeneous populations and measured different outcomes. Twenty-three trials assessed a psychological intervention. Guided self-help and unguided self-help BED improved eating disorder symptom severity and depressive symptoms; results for guided self-help were generally more precise and larger in magnitude than pooled estimates for unguided self-help. Group therapy for BED (7 trials; 253 participants) was also more effective for improving depression than inactive control. Only three trials of individual therapy met our inclusion criteria. These trials differed in type of therapy and measured various outcomes, limiting our ability to make stronger conclusions about the benefit of individual therapy compared with inactive control for populations meeting our inclusion criteria (i.e., not previously treated, no obvious signs or symptoms of eating disorders). No studies of psychological interventions reported on potential harms of interventions, including whether some participants experienced increased anxiety or distress as a result of the intervention; however, harms associated with these interventions are likely to be minimal. Despite benefits of these interventions, generalizability to populations detected by routine screening in primary care is limited.

Limitations

The limitations of the included studies are discussed above in Results and Summary of Evidence. Here we focus on limitations of this review. We excluded studies limited to persons with signs and symptoms of eating disorders, including populations who are underweight (defined by BMI or other criteria). We also excluded head-to-head comparisons of different interventions because the scope was designed to provide evidence on benefits of treatments compared with no treatment rather than assess the comparative effectiveness of interventions. For these reasons, no included studies focused on populations with AN met our eligibility criteria. Second, for studies related to benefits of screening and interventions for screen-detected populations, we limited the review to study designs that included a control group and those that reported on health outcomes. Intermediate outcomes, including mean changes in the frequency of binge-eating episodes over relatively short durations may not indicate that people identified by routine screening have better long-term health outcomes than those who are identified and referred for treatment in the context of routine primary care. Finally, we excluded studies assessing primary prevention strategies to reduce eating disorders (e.g., among groups considered “at risk” but who do not meet threshold or subthreshold eating disorder condition definitions), many of which are targeted to school or university settings. Our aim was to limit the review to interventions for individuals with eating disorders (or subthreshold conditions) that are appropriate to deliver in primary care settings or refer to from primary care.
Future Research Needs

Trials directly assessing the benefit of screening compared with no screening that focus on health outcomes and enroll asymptomatic or unselected populations from general primary care are needed, as are studies on potential harms of screening such as labeling, harms from false-positive results, burden, and inconvenience. Such studies could also address the acceptability of screening and provide insight on the prevalence of various eating disorders among primary care populations, in addition to clarifying whether routine screening (followed by appropriate referral and treatment) leads to improved health outcomes compared with identification and treatment in the context of usual care. Included trials of eating disorder treatment enrolled treatment-seeking populations, primarily via advertisements. Trials of treatment focused on both adolescent and adult populations that are applicable to U.S. populations would inform future recommendations on the benefit of screening, for example, populations recruited from primary care using brief screening questionnaires. Most included studies of treatment enrolled participants via advertisements and focused on specific eating disorders (primarily BN and BED); the applicability of results to populations that are not seeking care for eating disorder symptoms or who may have a new onset or less severe eating disorder is uncertain.

Accuracy studies enrolling asymptomatic or unselected populations from primary care settings that use consistent definitions of and reference standards to define eating disorder conditions would improve certainty about the accuracy of primary care–relevant screening tests. There is a need for studies assessing screening accuracy among adolescents, given that adolescence is a known time of risk for eating disorder onset. Few included studies enrolled adolescents, and the extent to which screening tools developed for adults are appropriate for adolescents is not clear. Similarly, studies of screening test accuracy that enroll a more diverse population with respect to race, ethnicity, gender, and sexual identity would help assess whether findings are broadly representative of the U.S. population.

The evidence from the current report highlights several important research needs. First, studies directly focusing on health outcomes among screen-detected populations in a general primary care population are needed to better understand the potential benefits and harms of screening. Such studies could also address the potential feasibility and acceptability of screening and provide a better sense of the prevalence of eating disorders in primary care. Some evidence indicates that prognosis is improved when individuals receive diagnosis and treatment early, so future primary care screening studies could confirm whether routine screening in the context of primary care improves outcomes among those who are not seeking treatment and have no obvious signs or symptoms of eating disorders.

Conclusion

No studies directly assessed the benefits and harms of screening. Screening questionnaires available for use in primary care have adequate accuracy for detecting eating disorders among adults; the most commonly studied screening questionnaire is the SCOFF. The accuracy of screening questionnaires for detecting eating disorders among adolescents is unclear. No treatment studies were found that enrolled participants who were screen detected in primary care.
Guided self-help interventions are effective for reducing eating disorder symptom severity and depressive symptoms among referred populations. Lisdexamfetamine and topiramate are effective in reducing eating disorder symptom severity in populations with BED but are also associated with adverse effects.
References


49. StataCorp. Stata statistical software: release 16.0. College Station, TX: StataCorp LP; 2019.


Figure 1. Analytic Framework

Screening for Eating Disorders

Adolescents and Adults

1. Screening
2. Early Detection of Eating Disorders
3. Harms of Screening
4. Intervention
5. Harms of Treatment

Health Outcomes Morbidity and Mortality
Figure 2. Summary of Evidence Search and Selection

Number of records identified through database searching
- PubMed: 8,927
- Cochrane: 2,320
- PsycINFO: 5,453

Number of additional records identified through other sources
- Public Comment: 1
- ClinicalTrials.gov: 262
- Handsearch: 0

Number of records screened after duplicates removed: 15,037

Number of title and abstract records excluded: 13,586

Number of full-text articles excluded, with reasons:
- Wrong language/non-English: 4
- Wrong condition: 18
- Wrong population: 231
- Wrong screening test: 215
- Wrong or no comparator: 261
- Wrong or no outcome: 279
- Wrong intervention: 76
- Wrong setting: 30
- Wrong study design: 174
- Intermediate outcome only: 14
- Wrong country: 13
- Not original research: 21
- Abstract only: 48

Number of full-text articles assessed for eligibility: 1,451

Number of studies (articles) included in qualitative synthesis of systematic review: 57 (59)

Key Question 1: 0
Key Question 2: 17 (10)
Key Question 3: 0
Key Question 4: 40 (41)
Key Question 5: 9 (8)
Figure 3. Results of Randomized, Controlled Trials of Lisdexamphetamine and Topiramate vs. Placebo for BED (KQ 4)

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcome</th>
<th>Weeks</th>
<th>Mean Dose</th>
<th>Med N</th>
<th>Placebo N</th>
<th>Raw Mean Diff. with 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lisdexamphetamine (ED Symptom Severity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guerdjikova, 2016</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>60mg/day</td>
<td>25</td>
<td>25</td>
<td>-2.80 [-7.28, 1.68]</td>
</tr>
<tr>
<td>McElroy, 2016a</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>57mg/day</td>
<td>190</td>
<td>184</td>
<td>-7.40 [-8.92, -5.88]</td>
</tr>
<tr>
<td>McElroy, 2016b</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>56mg/day</td>
<td>174</td>
<td>176</td>
<td>-7.94 [-9.51, -6.37]</td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>YBOCS-BE</td>
<td>11</td>
<td>50mg/day</td>
<td>64</td>
<td>62</td>
<td>-3.25 [-5.81, -0.69]</td>
</tr>
<tr>
<td><strong>Topiramate (ED Symptom Severity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McElroy, 2007</td>
<td>YBOCS-BE</td>
<td>16</td>
<td>300mg/day</td>
<td>202</td>
<td>202</td>
<td>-6.40 [-8.14, -4.66]</td>
</tr>
<tr>
<td>McElroy, 2003</td>
<td>YBOCS-BE</td>
<td>14</td>
<td>212mg/day</td>
<td>30</td>
<td>31</td>
<td>-2.55 [-4.29, -0.81]</td>
</tr>
<tr>
<td><strong>Topiramate (Depression, Anxiety)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McElroy, 2003</td>
<td>HAM-D</td>
<td>14</td>
<td>212mg/day</td>
<td>30</td>
<td>31</td>
<td>-0.55 [-1.57, 0.46]</td>
</tr>
<tr>
<td>McElroy, 2007</td>
<td>MADRS</td>
<td>16</td>
<td>300mg/day</td>
<td>202</td>
<td>202</td>
<td>-0.50 [-1.79, 0.79]</td>
</tr>
<tr>
<td>McElroy, 2007</td>
<td>HAM-A</td>
<td>16</td>
<td>300mg/day</td>
<td>202</td>
<td>202</td>
<td>-0.60 [-1.52, 0.32]</td>
</tr>
</tbody>
</table>

**Abbreviations:** BED=binge-eating disorder; CI=confidence interval; ED=eating disorder; HA=Hamilton Anxiety Rating Scale; HAM-D=Hamilton Depression Rating Scale—Depression; KQ=key question; MADRS=Montgomery-Åsberg Depression Rating Scale; N=number; YBOCS-BE=Yale–Brown Obsessive Compulsive Scale modified for binge eating.
Figure 4. Results of Randomized, Controlled Trials of SSRI vs. Placebo for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcome</th>
<th>Weeks</th>
<th>SSRI</th>
<th>Dose</th>
<th>Med N</th>
<th>Placebo N</th>
<th>Std. Mean Diff. with 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guardjikova, 2008</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>Escitalopram</td>
<td>27mg/day</td>
<td>21</td>
<td>23</td>
<td>-0.69 [-1.30, -0.08]</td>
</tr>
<tr>
<td>Grillo, 2005</td>
<td>EDE-Q</td>
<td>16</td>
<td>Fluoxetine</td>
<td>60mg/day</td>
<td>27</td>
<td>27</td>
<td>-0.29 [-0.83, 0.24]</td>
</tr>
</tbody>
</table>

**Depression Symptom Severity**

<table>
<thead>
<tr>
<th>Study</th>
<th>Abbreviation(s)</th>
<th>Week(s)</th>
<th>SSRI</th>
<th>Dose</th>
<th>Med N</th>
<th>Placebo N</th>
<th>Std. Mean Diff. with 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guardjikova, 2008</td>
<td>HAM-D</td>
<td>12</td>
<td>Escitalopram</td>
<td>27mg/day</td>
<td>21</td>
<td>23</td>
<td>-0.34 [-0.94, 0.25]</td>
</tr>
<tr>
<td>Arnold, 2002</td>
<td>HAM-D</td>
<td>6</td>
<td>Fluoxetine</td>
<td>71mg/day</td>
<td>30</td>
<td>30</td>
<td>-1.03 [-1.57, -0.49]</td>
</tr>
<tr>
<td>McElroy, 2009</td>
<td>HAM-D</td>
<td>6</td>
<td>Sertraline</td>
<td>50-200mg/day</td>
<td>18</td>
<td>16</td>
<td>-0.46 [-1.14, 0.23]</td>
</tr>
<tr>
<td>Pearlstein, 2003</td>
<td>BDI</td>
<td>12</td>
<td>Fluvoxamine</td>
<td>239mg/day</td>
<td>7</td>
<td>9</td>
<td>-0.49 [-1.50, 0.51]</td>
</tr>
<tr>
<td>Grillo, 2005</td>
<td>BDI</td>
<td>16</td>
<td>Fluoxetine</td>
<td>60mg/day</td>
<td>27</td>
<td>27</td>
<td>-0.56 [-1.10, -0.02]</td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.01$, $Q = 5.92$, $H^2 = 1.06$
Test of $b = 9$; $Q(4) = 3.41, p = 0.49$

**Anxiety Symptom Severity**

<table>
<thead>
<tr>
<th>Study</th>
<th>Abbreviation(s)</th>
<th>Week(s)</th>
<th>SSRI</th>
<th>Dose</th>
<th>Med N</th>
<th>Placebo N</th>
<th>Std. Mean Diff. with 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearlstein, 2003</td>
<td>HAM-A</td>
<td>12</td>
<td>Fluvoxamine</td>
<td>239mg/day</td>
<td>8</td>
<td>8</td>
<td>-0.59 [-1.59, 0.41]</td>
</tr>
</tbody>
</table>

Abbreviations: BDI=Beck Depression Inventory; CI=confidence interval; ED=eating disorder; EDE-Q=Eating Disorder Examination Questionnaire; HAM-A=Hamilton Anxiety Rating Scale; HAM-D=Hamilton Depression Rating Scale--Depression; KQ=key question; N=number; SSRI=selective serotonin reuptake inhibitor; YBOCS-BE=Yale-Brown Obsessive Compulsive Scale modified for binge eating.
### Figure 5. Results of Randomized, Controlled Trials of Self-Help Interventions for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Measure</th>
<th>Weeks</th>
<th>Std. Mean Diff. with 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Guided Self-Help (ED Symptom Severity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carter, 2020</td>
<td>DBT</td>
<td>EDE-Q</td>
<td>24</td>
<td>-0.40 [-0.98, 0.18]</td>
</tr>
<tr>
<td>Wagner, 2016</td>
<td>CBT</td>
<td>EDE-Q</td>
<td>16</td>
<td>-1.18 [-1.54, -0.82]</td>
</tr>
<tr>
<td>Ljotsson, 2007</td>
<td>CBT</td>
<td>EDE-Q</td>
<td>12</td>
<td>-1.13 [-1.64, -0.62]</td>
</tr>
<tr>
<td>Masson, 2013</td>
<td>DBT</td>
<td>EDE-Q</td>
<td>13</td>
<td>-0.70 [-1.22, -0.18]</td>
</tr>
<tr>
<td>Sánchez-Ortiz, 2011</td>
<td>CBT</td>
<td>EDE-Q</td>
<td>12</td>
<td>-1.24 [-1.73, -0.75]</td>
</tr>
<tr>
<td><strong>Unguided Self-Help (ED Symptom Severity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carter, 2020</td>
<td>DBT</td>
<td>EDE-Q</td>
<td>24</td>
<td>-0.42 [-0.99, 0.16]</td>
</tr>
<tr>
<td>Grilo, 2013</td>
<td>CBT</td>
<td>EDE-Q</td>
<td>16</td>
<td>-0.24 [-0.81, 0.33]</td>
</tr>
<tr>
<td>Grilo, 2014</td>
<td>CBT</td>
<td>EDE-Q</td>
<td>26</td>
<td>-0.11 [-0.65, 0.44]</td>
</tr>
<tr>
<td>Schmitz, 2008</td>
<td>CBT</td>
<td>EDE-Q</td>
<td>12</td>
<td>-0.08 [-0.48, 0.31]</td>
</tr>
<tr>
<td>Kelly, 2014</td>
<td>Comp. focused</td>
<td>EDE-Q</td>
<td>8</td>
<td>-0.23 [-0.87, 0.41]</td>
</tr>
<tr>
<td>Green, 2018</td>
<td>Body Project</td>
<td>EDE-Q</td>
<td>8</td>
<td>-0.15 [-0.58, 0.29]</td>
</tr>
<tr>
<td><strong>Guided Self-Help (Depression)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cachalia, 2019</td>
<td>CBT</td>
<td>BDI-II</td>
<td>12</td>
<td>-0.39 [-1.02, 0.24]</td>
</tr>
<tr>
<td>Wagner, 2016</td>
<td>CBT</td>
<td>BDI</td>
<td>16</td>
<td>-0.53 [-0.87, -0.19]</td>
</tr>
<tr>
<td>Ljotsson, 2007</td>
<td>CBT</td>
<td>MADRS</td>
<td>12</td>
<td>-0.93 [-1.43, -0.43]</td>
</tr>
<tr>
<td>Sánchez-Ortiz, 2011</td>
<td>CBT</td>
<td>HADS-Dep</td>
<td>12</td>
<td>-1.09 [-1.57, -0.60]</td>
</tr>
<tr>
<td><strong>Unguided Self-Help (Depression)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carter, 2003</td>
<td>CBT</td>
<td>BDI</td>
<td>8</td>
<td>-0.53 [-1.06, 0.01]</td>
</tr>
<tr>
<td>Grilo, 2013</td>
<td>CBT</td>
<td>BDI</td>
<td>16</td>
<td>-0.41 [-0.98, 0.16]</td>
</tr>
<tr>
<td>Grilo, 2014</td>
<td>CBT</td>
<td>BDI</td>
<td>26</td>
<td>-0.16 [-0.70, 0.38]</td>
</tr>
</tbody>
</table>

**Abbreviations:** BDI=Beck Depression Inventory; BDI-II=Beck Depression Inventory-II; CBT=cognitive behavioral therapy; CI=confidence interval; DBT=dialectical behavioral therapy; ED=eating disorder; EDE=Eating Disorder Examination; EDE-Q=Eating Disorder Examination Questionnaire; HADS-Dep=Hospital Anxiety and Depression Score—Depression score; KQ=key question; MADRS=Montgomery-Åsberg Depression Rating Scale.
Figure 6. Results of Randomized, Controlled Trials of Group and Individual Therapy for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Measure</th>
<th>Weeks</th>
<th>Std. Mean Diff. with 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group (ED Symptom Severity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schag, 2019</td>
<td>Impul.</td>
<td>EDE-Q</td>
<td>12</td>
<td>-0.30 [-0.74, 0.14]</td>
</tr>
<tr>
<td>Wade, 2017</td>
<td>CBT</td>
<td>EDE-Q</td>
<td>8</td>
<td>-1.01 [-1.71, -0.31]</td>
</tr>
<tr>
<td>Alfonsen, 2015</td>
<td>Beh. Act.</td>
<td>EDE-Q</td>
<td>10</td>
<td>-0.10 [-0.50, 0.30]</td>
</tr>
<tr>
<td><strong>Group (Depression)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laassle, 1987</td>
<td>Behav.</td>
<td>BDI</td>
<td>16</td>
<td>-0.25 [-1.21, 0.71]</td>
</tr>
<tr>
<td>Schag, 2019</td>
<td>Impul.</td>
<td>BDI-II</td>
<td>12</td>
<td>-0.38 [-0.82, 0.06]</td>
</tr>
<tr>
<td>Schlup, 2009</td>
<td>CBT</td>
<td>BDI</td>
<td>8</td>
<td>-0.64 [-1.31, 0.03]</td>
</tr>
<tr>
<td>Telch, 1990</td>
<td>CBT</td>
<td>BDI</td>
<td>10</td>
<td>-0.48 [-1.08, 0.12]</td>
</tr>
<tr>
<td>Willey, 1993</td>
<td>IPT</td>
<td>BDI</td>
<td>16</td>
<td>-0.81 [-1.48, -0.15]</td>
</tr>
<tr>
<td>Willey, 1993</td>
<td>CBT</td>
<td>BDI</td>
<td>16</td>
<td>-0.26 [-0.90, 0.37]</td>
</tr>
<tr>
<td>Alfonsen, 2015</td>
<td>Beh. Act.</td>
<td>HADS-Dep</td>
<td>10</td>
<td>-0.51 [-0.92, -0.10]</td>
</tr>
<tr>
<td>Heterogeneity: $I^2 = 0.00$, $Q^2 = 0.00$, $H^2 = 1.00$</td>
<td></td>
<td></td>
<td></td>
<td>-0.48 [-0.69, -0.27]</td>
</tr>
<tr>
<td><strong>Group (Anxiety)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schlup, 2009</td>
<td>CBT</td>
<td>BAI</td>
<td>8</td>
<td>-0.34 [-0.99, 0.32]</td>
</tr>
<tr>
<td>Alfonsen, 2015</td>
<td>Beh. Act.</td>
<td>HADS-Anx</td>
<td>10</td>
<td>-0.03 [-0.43, 0.38]</td>
</tr>
<tr>
<td><strong>Individual (ED Symptom Severity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairburn, 2009</td>
<td>CBT-Eb</td>
<td>EDE-Q</td>
<td>8</td>
<td>-0.19 [-0.59, 0.20]</td>
</tr>
<tr>
<td>Fairburn, 2009</td>
<td>CBT-Ef</td>
<td>EDE-Q</td>
<td>8</td>
<td>-0.15 [-0.53, 0.24]</td>
</tr>
<tr>
<td>Hill, 2011</td>
<td>DBT-AF</td>
<td>EDE-Q</td>
<td>6</td>
<td>-1.18 [-1.94, -0.43]</td>
</tr>
<tr>
<td><strong>Individual (Depression)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grilo, 2005</td>
<td>CBT</td>
<td>BDI</td>
<td>16</td>
<td>-0.60 [-1.14, -0.06]</td>
</tr>
<tr>
<td>Hill, 2011</td>
<td>DBT-AF</td>
<td>BDI-II</td>
<td>6</td>
<td>-0.92 [-1.65, -0.19]</td>
</tr>
</tbody>
</table>

**Abbreviations:** BAI=Beck Anxiety Inventory; BDI=Beck Depression Inventory; BDI-II=Beck Depression Inventory-II; CBT=cognitive behavioral therapy; CBT-Ef=focused form of enhanced cognitive behavioral therapy; CBT-Eb=broad form of enhanced cognitive behavioral therapy; CI=confidence interval; DBT=dialectical behavioral therapy; DBT-AF=appetite focused dialectical behavioral therapy; ED=eating disorder; EDE=Eating Disorder Examination; EDE-Q=Eating Disorder Examination Questionnaire; HADS-Anx=Hospital Anxiety and Depression Score—Anxiety score; HADS-Dep=Hospital Anxiety and Depression Score—Depression score; IPT=interpersonal therapy; KQ=key question; MADRS=Montgomery-Asberg Depression Rating Scale.
Table 1. Characteristics of Included Studies for KQ 2

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Screener</th>
<th>Reference Standard: ED Diagnoses Assessed</th>
<th>Type of Study</th>
<th>Recruitment Setting and Country</th>
<th>N</th>
<th>Population</th>
<th>Prevalence of ED (%)</th>
<th>Age, Mean (SD)</th>
<th>Sex (% F)</th>
<th>% Non-White</th>
<th>Mean BMI (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lui, 2015\textsuperscript{?} Good</td>
<td>SCOFF</td>
<td>SCID (DSM-IV): AN, BN, BED, EDNOS</td>
<td>Cohort</td>
<td>Outpatient psychiatric clinics Taiwan</td>
<td>1,541</td>
<td>Adults (18-45 y) recruited at their first outpatient psychiatric visit</td>
<td>Any ED: 16</td>
<td>31 (7.9)</td>
<td>61</td>
<td>NR\textsuperscript{†}</td>
<td>22.2 (5.4)</td>
</tr>
<tr>
<td>Graham, 2019\textsuperscript{?} Good</td>
<td>SWED</td>
<td>EDE (DSM-5): AN, BN, BED</td>
<td>Cohort</td>
<td>University campuses United States</td>
<td>549</td>
<td>College-age women (18-25 y) responding to recruitment ads and flyers for an ED prevention trial</td>
<td>Any ED: 19\textsuperscript{†,‡}</td>
<td>21 (1.97)</td>
<td>100</td>
<td>44\textsuperscript{†}</td>
<td>24.5 (5.02)</td>
</tr>
<tr>
<td>Maugen, 2018\textsuperscript{?} Good</td>
<td>EDS-PC, SCOFF, SDE</td>
<td>EDE-Q (DSM-5): AN, BN, BED</td>
<td>Cross-sectional</td>
<td>VHA medical center United States</td>
<td>402</td>
<td>Female veterans (18-70 y) responding to mailed questionnaires</td>
<td>Any ED: 16\textsuperscript{†}</td>
<td>49 (NR)</td>
<td>100</td>
<td>52\textsuperscript{†}</td>
<td>NR</td>
</tr>
<tr>
<td>Chamay-Weber, 2017\textsuperscript{?} Good</td>
<td>ADO-BED</td>
<td>SCID (DSM-IV): BED</td>
<td>Cross-sectional</td>
<td>Outpatient pediatric obesity center Switzerland</td>
<td>94</td>
<td>Adolescents (12-18 y) recruited at their outpatient pediatric visit</td>
<td>BED, sub: 28 BED, full: 22 BED, overall: 50</td>
<td>Median 14 (Range 11-18)</td>
<td>60</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Dorflinger, 2017\textsuperscript{?} Good</td>
<td>VA-BES</td>
<td>QEWP-R: BED</td>
<td>Cohort</td>
<td>VHA medical center United States</td>
<td>116</td>
<td>Veterans recruited at primary care–based weight management group</td>
<td>BED: 8</td>
<td>62 (8.73)</td>
<td>11</td>
<td>26</td>
<td>37.9 (7.35)</td>
</tr>
<tr>
<td>Luck, 2002; Hill, 2010\textsuperscript{?} Good</td>
<td>SCOFF</td>
<td>Clinical interview (DSM-IV): AN, BN, EDNOS</td>
<td>Cross-sectional</td>
<td>Primary care practices United Kingdom</td>
<td>341</td>
<td>Women (18-50 y) attending primary care practices</td>
<td>Any ED: 4\textsuperscript{†}</td>
<td>NR</td>
<td>100</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Siervo, 2005\textsuperscript{?} Fair</td>
<td>SCOFF</td>
<td>“Clinical diagnosis” (DSM-IV): BN, BED</td>
<td>Cross-sectional</td>
<td>Outpatient diet clinics Italy</td>
<td>162</td>
<td>Women (16-35 y) recruited at an outpatient dietetic clinic</td>
<td>Any ED: 46\textsuperscript{†}</td>
<td>24\textsuperscript{†}</td>
<td>100</td>
<td>NR</td>
<td>29.6\textsuperscript{†}</td>
</tr>
<tr>
<td>Parker, 2005\textsuperscript{?} Good</td>
<td>SCOFF</td>
<td>EDE-Q (DSM-IV): AN, BN, EDNOS</td>
<td>Cross-sectional</td>
<td>University Health Center United States</td>
<td>297</td>
<td>Adults (20-51 y) recruited at their campus health visit</td>
<td>Any ED: 20</td>
<td>&lt;23 y: 10 23-26 y: 66 &gt;26 y: 23</td>
<td>72</td>
<td>33</td>
<td>(Range: 16-44)</td>
</tr>
</tbody>
</table>
Table 1. Characteristics of Included Studies for KQ 2

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Quality</th>
<th>Screener</th>
<th>Reference Standard: ED Diagnoses Assessed</th>
<th>Type of Study</th>
<th>Recruitment Setting and Country</th>
<th>N</th>
<th>Population</th>
<th>Prevalence of ED (%)</th>
<th>Age, Mean (SD)</th>
<th>Sex (% F)</th>
<th>% Non-White</th>
<th>Mean BMI (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ricca, 2000</td>
<td>Fair</td>
<td>BES</td>
<td>SCID (DSM-IV): BN, BED</td>
<td>Cross-sectional</td>
<td>Outpatient clinic for metabolic diseases Italy</td>
<td>344</td>
<td>Patients recruited at an outpatient clinic for metabolic diseases including obesity</td>
<td>BED: 8</td>
<td>43.5 (13.6)</td>
<td>83</td>
<td>NR</td>
<td>35.8 (6.1)</td>
</tr>
<tr>
<td>Rosenvinge, 2001</td>
<td>Fair</td>
<td>EDS-5</td>
<td>SCID (DSM-III-R): Any ED</td>
<td>Cohort University campuses Norway</td>
<td>51</td>
<td>College-age women (20-42 y) recruited at their teaching and nursing colleges</td>
<td>CED: 20¹</td>
<td>25.2 (5.33)</td>
<td>100</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Wan Wahida, 2017</td>
<td>Good</td>
<td>SCOFF</td>
<td>EAT-26</td>
<td>Cross-sectional</td>
<td>University Malaysia</td>
<td>292</td>
<td>Undergraduate students (18-22 y) who understood English</td>
<td>Any ED: 11</td>
<td>20 (0.5)</td>
<td>65</td>
<td>Malay: 44 Chinese: 42 Indian: 14</td>
<td>NR</td>
</tr>
<tr>
<td>Mond, 2008</td>
<td>Fair</td>
<td>SCOFF</td>
<td>EDE</td>
<td>Cohort Primary care practices United States</td>
<td>147</td>
<td>Adult women (18-40 y) recruited at their primary care visit</td>
<td>Any ED: 17</td>
<td>28 (6.50)</td>
<td>100</td>
<td>12</td>
<td>28.10 (7.20)</td>
<td></td>
</tr>
<tr>
<td>Cotton, 2003</td>
<td>Good</td>
<td>SCOFF, ESP</td>
<td>QEDD</td>
<td>Cohort University campuses and primary care United Kingdom</td>
<td>225</td>
<td>Students (18-65 y) recruited from posters and lecture announcements and adults (18-65 y) recruited at a primary care visit</td>
<td>Any ED: 12</td>
<td>29</td>
<td>77</td>
<td>NR</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Garcia, 2010</td>
<td>Good</td>
<td>SCOFF</td>
<td>MINI (DSM-IV-TR): Any ED, AN, BN</td>
<td>Cross-sectional</td>
<td>University clinic France</td>
<td>400</td>
<td>Female undergraduate students (18-35 y)</td>
<td>Any ED: 9</td>
<td>21 (2.5)</td>
<td>100</td>
<td>14¹</td>
<td>21.98 (3.5)</td>
</tr>
<tr>
<td>Muro-Sans, 2008</td>
<td>Good</td>
<td>SCOFF</td>
<td>EDI-2</td>
<td>Cross-sectional</td>
<td>Primary and secondary schools Spain</td>
<td>954</td>
<td>Adolescents (10-17 y) recruited from schools</td>
<td>Any ED: 8¹</td>
<td>14 (1.31)</td>
<td>49</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Lähteenmäki, 2009</td>
<td>Fair</td>
<td>SCOFF</td>
<td>SCID (DSM-IV): AN, BN, EDNOS</td>
<td>Cohort Households Finland</td>
<td>541</td>
<td>Young adults recruited via mail</td>
<td>Current AN, BN, EDNOS: 1¹</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Characteristics of Included Studies for KQ 2

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Screener</th>
<th>Reference Standard: ED Diagnoses Assessed</th>
<th>Type of Study</th>
<th>Recruitment Setting and Country</th>
<th>N</th>
<th>Population</th>
<th>Prevalence of ED (%)</th>
<th>Age, Mean (SD)</th>
<th>Sex (% F)</th>
<th>% Non-White</th>
<th>Mean BMI (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Striegel-Moore, 2010 Fair</td>
<td>PHQ-ED EDE (DSM-IV): BN, BED</td>
<td>Cross-sectional</td>
<td>Health maintenance organization United States</td>
<td>348</td>
<td>Adults (18-35 y) selected from the EHR of a HMO via letter</td>
<td>Life-time AN, BN, EDNOS: 4†</td>
<td>28 (5.38)</td>
<td>82</td>
<td>13</td>
<td>NR</td>
<td></td>
</tr>
</tbody>
</table>

*“Full” refers to meeting the full diagnostic criteria for a given eating disorder; “sub” refers to a subthreshold condition definition.
† Conducted in Taiwan and required to understand Mandarin.
‡ Also conducted analysis including subthreshold BN, BED, and purging disorder (in addition to threshold AN, BN, and BED).
§ BED defined as an average of one or more objective binge episodes per week without compensatory or purging behaviors.
†† Enrolled all screen-positive participants and a random sample of those who screened negative.
¶ Computed by data abstractors.

**Abbreviations:** ADO-BED=Adolescent Binge-Eating Disorder Questionnaire; AN=anorexia nervosa; BED=binge eating disorder; BES=Binge Eating Scale; BMI=body mass index (kg/m²); BN=bulimia nervosa; CED=clinical eating disorder; DSM=Statistical Manual of Mental Disorders; EAT-26=Eating Attitudes Test; ED=end eating disorder; EDE=Eating Disorder Examination; EDE-Q=Eating Disorder Examination Questionnaire; EDNOS=Eating disorder not otherwise specified; EDS-5=Eating Disturbance Scale-5; EDS-PC=Eating Disorder Screen for Primary Care; EHR=electronic health record; ESP=Eating Disorder Screen for Primary Care; HMO=health maintenance organization; KQ=key question; MINI=Mini International Neuropsychiatric Interview; NR=not reported; PHQ-ED=Eating disorder module of the Patient Health Questionnaire; QEDD=Questionnaire for Eating Disorder Diagnoses; QEWPR=Questionnaire of Eating and Weight Patterns-Revised; SCID=Structured Clinical Interview for DSM Disorders; SCOFF=not an acronym; SD=standard deviation; SDE=Screen for Disordered Eating; SWED=Stanford-Washington University Eating Disorder screen; VA-BES=Veterans Affairs Binge Eating Screener; VHA=Veterans Health Administration.
Table 2. Summary of Accuracy for Included Screening Tests (KQ 2)

<table>
<thead>
<tr>
<th>Screener (Cut Point)</th>
<th>ED Diagnosis</th>
<th>N Studies (Participants)</th>
<th>Sensitivity (%) (95% CI)</th>
<th>Specificity (%) (95% CI)</th>
<th>PLR (95% CI)</th>
<th>NLR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOFF (≥3)</td>
<td>Any</td>
<td>7 (2,749)</td>
<td>Pooled: 69 (56 to 80)</td>
<td>Pooled: 90 (69 to 98)</td>
<td>Pooled: 7.3 (2.2 to 24.0)</td>
<td>Pooled: 0.34 (0.25 to 0.46)</td>
</tr>
<tr>
<td>SCOFF (≥2)</td>
<td>Any</td>
<td>10 (3,684)</td>
<td>Pooled: 84 (74 to 90)</td>
<td>Pooled: 80 (65 to 89)</td>
<td>Pooled: 4.1 (2.3 to 7.3)</td>
<td>Pooled: 0.2 (0.12 to 0.33)</td>
</tr>
<tr>
<td>SWED (&gt;59)</td>
<td>Any</td>
<td>1 (549)</td>
<td>80 (NR)</td>
<td>82 (NR)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>EDS-PC (≥2)</td>
<td>Any</td>
<td>2 (627)</td>
<td>97 (88 to 100)</td>
<td>40 (35–46)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>SDE (≥2)</td>
<td>Any</td>
<td>1 (402)</td>
<td>91 (80 to 96)</td>
<td>58 (80 to 96)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>EDS-5 (≥16)</td>
<td>Any</td>
<td>1 (51)</td>
<td>90 (NR)</td>
<td>88 (NR)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>PHQ-ED (NA)</td>
<td>BN, BED</td>
<td>1 (348)</td>
<td>100 (NR)</td>
<td>30 (NR)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>ADO-BED (NA)</td>
<td>BED</td>
<td>1 (94 adolescents)</td>
<td>100 (NR)</td>
<td>27 (NR)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>VA-BES (≥1)</td>
<td>BED</td>
<td>1 (162)</td>
<td>89 (NR)</td>
<td>65 (NR)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>BES (≥17)</td>
<td>BED</td>
<td>1 (344)</td>
<td>85 (NR)</td>
<td>75 (NR)</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

* Value calculated based on individual cell frequencies differs from reported specificity value reported in study (91.7% vs. 27.7%, respectively).

**Abbreviations:** ADO-BED=Adolescent Binge-Eating Disorder Questionnaire; BED=binge-eating disorder; BES=Binge Eating Scale; BN=bulimia nervosa; CI=confidence interval; ED=eating disorder; EDS-5=Eating Disturbance Scale-5; EDS-PC=Eating Disorder Screen for Primary Care; KQ=key question; NA=not available; NLR=negative likelihood ratio; PHQ-ED=Eating disorder module of the Patient Health Questionnaire; PLR=positive likelihood ratio; SCOFF=not an acronym; SDE=Screen for Disordered Eating; SWED=Stanford-Washington University Eating Disorder screen; VA-BES=Veterans Affairs Binge Eating Screener; vs.=versus.
Table 3. Characteristics of Randomized, Placebo-Controlled Trials of Pharmacotherapy for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Medication</th>
<th>Dose</th>
<th>Duration (Weeks)</th>
<th>Recruitment</th>
<th>Population</th>
<th>Country Setting</th>
<th>N</th>
<th>Mean Age, y (SD)</th>
<th>% Female</th>
<th>% Non-White</th>
<th>Mean BMI (SD)</th>
<th>% With Comorbid Psych. Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnold, 2002*</td>
<td>Fluoxetine</td>
<td>20-80 mg/day</td>
<td>6</td>
<td>Advertisements for a binge-eating trial</td>
<td>Adults (18-60 y) with BED (DSM-IV), ≥3 binge-eating episodes/week for ≥6 m, &gt;85% ideal body weight</td>
<td>United States</td>
<td>60</td>
<td>41 (93)</td>
<td>12</td>
<td>38.2</td>
<td>Current MDD: 25 Lifetime MDD: 65</td>
<td></td>
</tr>
<tr>
<td>Fluoxetine Bulimia Nervosa Collaborative Study Group, 1992</td>
<td>Fluoxetine</td>
<td>20-60 mg/day</td>
<td>8</td>
<td>Advertisements and referrals</td>
<td>Adult women (≥18 y) with BN (DSM-III-R), ≥3 binge-eating episodes/week for ≥6 m, no other BN treatment initiated in previous month</td>
<td>United States, Canada</td>
<td>387</td>
<td>27 (100)</td>
<td>3</td>
<td>22.6</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Grilo, 2005*</td>
<td>Fluoxetine</td>
<td>60 mg/day, and fluoxetine 60 mg/day + CBT</td>
<td>16</td>
<td>Referrals to a university outpatient treatment center</td>
<td>Adults (18-60 y) with BED (DSM-IV) classified as overweight/obese (100-200% ideal weight for height)*</td>
<td>United States</td>
<td>108</td>
<td>44 (8.6)</td>
<td>78</td>
<td>11</td>
<td>36.3 (7.9)</td>
<td>Lifetime MDD: 50%. Lifetime anxiety d/o: 37%</td>
</tr>
<tr>
<td>Guerdjikova, 2012*</td>
<td>Duloxetine</td>
<td>60-90 mg/day† (mean: 77 mg/day)</td>
<td>12</td>
<td>Newspaper advertisements for a study on medications for binge eating and depression</td>
<td>Adults (18-65 y) with BED and current depressive disorder (DSM-IV-TR) for ≥4 weeks</td>
<td>United States</td>
<td>40</td>
<td>40 (12.0)</td>
<td>88</td>
<td>17</td>
<td>40.6 (7.4)</td>
<td>MDD, recurrent: 63 MDD, single episode: 23 Dysthymic disorder: 13 Lifetime anxiety disorder: 30 Lifetime substance use disorder: 13</td>
</tr>
<tr>
<td>Guerdjikova, 2008*</td>
<td>Escitalopram</td>
<td>10-30 mg/day (mean: 27 mg/day)</td>
<td>12</td>
<td>Advertisements for trial of binge eating and obesity</td>
<td>Adults (18-60 y) with BED (DSM-IV) and BMI ≥30</td>
<td>United States</td>
<td>44</td>
<td>39</td>
<td>96</td>
<td>25</td>
<td>40.2</td>
<td>Current MDD: 23 Lifetime disorder: MDD: 77 Alcohol use: 11 Anxiety: 23</td>
</tr>
<tr>
<td>Guerdjikova, 2016*</td>
<td>Lisdexam-fetamine</td>
<td>20-70 mg/day (mean: 60 mg/day)</td>
<td>12</td>
<td>Newspaper advertisements for a study on binge eating</td>
<td>Adults (18-55 y) with BED (DSM-IV) and ≥3 binge-eating episodes/week for ≥2 weeks before enrollment</td>
<td>United States</td>
<td>50</td>
<td>38 (8.9)</td>
<td>92</td>
<td>22</td>
<td>39.8 (9.3)</td>
<td>Lifetime disorder: Depressive: 32 Anxiety: 4 Substance: 2 Probable childhood ADHD: 6</td>
</tr>
</tbody>
</table>
Table 3. Characteristics of Randomized, Placebo-Controlled Trials of Pharmacotherapy for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year Quality Rating</th>
<th>Medication Dose</th>
<th>Duration (Weeks)</th>
<th>Recruitment</th>
<th>Population</th>
<th>Country Setting</th>
<th>N</th>
<th>Mean Age, y (SD)</th>
<th>% Female</th>
<th>% Non-White</th>
<th>Mean BMI (SD)</th>
<th>% With Comorbid Psych. Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaneva, 1995 Fair</td>
<td>Fluoxetine 60 mg/day</td>
<td>8</td>
<td>Referrals to an outpatient university psychiatry clinic</td>
<td>Adults and adolescent females (&gt;15 y) with BN (DSM-III-R) and BMI ≥ 16</td>
<td>Finland</td>
<td>50</td>
<td>25 (range: 16-55)</td>
<td>100</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Laederach-Hofmann, 1998 Fair</td>
<td>Imipramine 75 mg/day (+ co-intervention)</td>
<td>8</td>
<td>Chart reviews of patients seen at a university outpatient clinic for &quot;weight problems&quot;</td>
<td>Adults (20-60 y) with BED (DSM-IV) and BMI &gt;27.5, seen for weight problems within the previous 2 y</td>
<td>Switzerland</td>
<td>31</td>
<td>38</td>
<td>87</td>
<td>NR</td>
<td>39.5</td>
<td>NR</td>
</tr>
<tr>
<td>McElroy, 2000 Fair</td>
<td>Sertraline 50-200 mg/day</td>
<td>6</td>
<td>NR</td>
<td>Adults (18-60 y) with BED (DSM-IV), ≤3 binge-eating episodes/week for ≥6 m, and 85% of ideal body weight</td>
<td>United States</td>
<td>34</td>
<td>42 (NR)</td>
<td>94</td>
<td>NR</td>
<td>36.1</td>
<td>Lifetime MDD: 53 Current MDD: 18</td>
</tr>
<tr>
<td>McElroy, 2003 Fair</td>
<td>Topiramate 25-600 mg/day; median: 212 mg/day (range: 50-600)</td>
<td>14</td>
<td>Radio advertisement for a study of binge eating associated with obesity</td>
<td>Adults (18-60 y) with BED (DSM-IV-TR), BMI ≥30, and marked distress related to binge-eating (YBOCS-BE score ≥15)</td>
<td>United States</td>
<td>61</td>
<td>41</td>
<td>87</td>
<td>NR</td>
<td>43</td>
<td>Lifetime disorder: Depressive: 54 Bipolar: 10</td>
</tr>
<tr>
<td>McElroy, 2007 Fair</td>
<td>Topiramate 25-400 mg/day (mean: 300 mg/day)</td>
<td>16</td>
<td>Radio and newspaper advertisements for study of medication for binge eating and obesity, and clinical referrals</td>
<td>Adults (18-65 y) with BED (DSM-IV-TR), ≥3 binge days/week 2 weeks before randomization, and BMI ≥30 and ≤ 50</td>
<td>United States</td>
<td>404</td>
<td>44</td>
<td>84</td>
<td>22</td>
<td>38.5</td>
<td>NR</td>
</tr>
<tr>
<td>McElroy, 2015; McElroy, 2016</td>
<td>Lisdexamfetamine 30, 50, or 70 mg/day</td>
<td>11</td>
<td>Clinical research centers, university-affiliated clinics, and psychiatric practices (methods of recruitment NR)</td>
<td>Adults (18-55 y) with BED (DSM-IV-TR), BMI ≥25 and ≤45, and ≥3 binge-eating days/week during a 2-4 week screening period</td>
<td>United States</td>
<td>260</td>
<td>39 (10.2)</td>
<td>82</td>
<td>22</td>
<td>34.9 (5.3)</td>
<td>NA</td>
</tr>
<tr>
<td>McElroy, 2016a, Sheehan, 2018</td>
<td>Lisdexamfetamine 30-70 mg/day (mean: 57 mg/day)</td>
<td>12</td>
<td>Investigators’ databases, local/central advertisements</td>
<td>Adults (18-55 y) with BED (DSM-IV-TR) and ≥3 binge-eating days/week over prior 2 weeks, CGI-S score ≥4, and BMI ≥18 and ≤45</td>
<td>United States, Sweden, Spain, Germany</td>
<td>379</td>
<td>38</td>
<td>87</td>
<td>22</td>
<td>33.4</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 3. Characteristics of Randomized, Placebo-Controlled Trials of Pharmacotherapy for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Quality Rating</th>
<th>Medication Dose</th>
<th>Duration (Weeks)</th>
<th>Recruitment</th>
<th>Population</th>
<th>Country Setting</th>
<th>N</th>
<th>Mean Age, y (SD)</th>
<th>% Female</th>
<th>% Non-White</th>
<th>Mean BMI (SD)</th>
<th>% With Comorbid Psych. Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>McElroy, 2016b&lt;sup&gt;a&lt;/sup&gt;, Sheehan, 2018&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Fair</td>
<td>Lisdexam-fetamine 30-70 mg/day (mean: 58 mg/day)</td>
<td>12</td>
<td>Investigators' databases, local/central advertisements</td>
<td>Adults (18-55 y) with BED (DSM-IV-TR) and ≥3 binge-eating days/week for ≥2 consecutive weeks before baseline, CGI-S score ≥4, and BMI ≥18 and ≤45</td>
<td>United States, Germany</td>
<td>366</td>
<td>38 (15)</td>
<td>85</td>
<td>27</td>
<td>33.5 (3.9)</td>
<td>NA &lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mitchell, 2001&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Fair</td>
<td>Fluoxetine‡ 60 mg/day</td>
<td>16</td>
<td>Referrals to an outpatient university ED program, local newspaper advertisements</td>
<td>Adult women (≥18 y) with BN (DSM-III-R) and binge eating coupled with self-induced vomiting ≥3 times weekly over previous 6m, within 85% of ideal body weight, not currently receiving other ED treatment</td>
<td>United States</td>
<td>91</td>
<td>27 (7.1)</td>
<td>100</td>
<td>3</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Pearlstein, 2003&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Fair</td>
<td>Fluvoxamine up to 300 mg/day (mean 239 mg/day)</td>
<td>12</td>
<td>Media advertisement and referral from health professionals</td>
<td>Adults (age range NR) with BED (DSM-IV)</td>
<td>United States</td>
<td>20</td>
<td>41 (20)</td>
<td>85</td>
<td>10</td>
<td>41.2</td>
<td>NR</td>
</tr>
<tr>
<td>Walsh, 1991&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Fair</td>
<td>Desipramine 200-300 mg/day</td>
<td>8</td>
<td>Advertisements in local media and referrals</td>
<td>Adult women (18-45 y) with BN (DSM-III-R) between 85-120% of ideal weight&lt;sup&gt;f&lt;/sup&gt;</td>
<td>United States</td>
<td>78</td>
<td>25 (NR)</td>
<td>100</td>
<td>NR</td>
<td>22.2</td>
<td>Currently depressed: 52</td>
</tr>
<tr>
<td>White, 2013&lt;sup&gt;f&lt;/sup&gt;</td>
<td>Good</td>
<td>Bupropion 300 mg/day</td>
<td>8</td>
<td>NR</td>
<td>Adult women (18-65 y) with BED (DSM-IV-TR) and BMI 25-50</td>
<td>United States</td>
<td>61</td>
<td>44 (12.5)</td>
<td>100</td>
<td>16</td>
<td>35.8 (6.8)</td>
<td>Lifetime disorder: Mood: 53 Anxiety: 38 Substance: 25</td>
</tr>
</tbody>
</table>

<sup>a</sup> Based on the 1959 Metropolitan Life Insurance Company Tables.
<sup>b</sup> Medication started at 30 mg/day for the first 7 days. At the beginning of the second week, medication was increased, as tolerated, to 60 mg/day and this dose was kept constant for 2 weeks. In the absence of remission of binge eating or depressive symptoms and intolerable side effects, the dose could be further increased as follows: 90 mg/day at the beginning of the fourth treatment week and 120 mg/day at the beginning of the sixth treatment week.
<sup>c</sup> Both treatment and placebo groups received 30 min of individual diet counseling by a dietician on a biweekly basis in addition to behavioral-oriented psychological support. Individual evaluation of problems (weight, relations in the family, professional concern, personal difficulties) took place after each diet counseling session (for approximately 15–35 min). Participants were also invited to a behavioral-oriented group therapy that took place monthly guided by an assistant dietitian and supervised by a physician.
<sup>d</sup> Participants with the following conditions were excluded: current BN, AN, ADHD, or another psychiatric disorder; a lifetime history of bipolar disorder or psychosis or other conditions that may confound efficacy and safety assessments; a total Montgomery-Åsberg Depression Rating Scale score of at least 18 at screening or baseline visits.
Exclusion criteria included current AN or BN, comorbid current psychiatric disorders either controlled with prohibited medications or uncontrolled and associated with significant symptoms or any condition/symptom that may confound clinical assessment; psychotherapy or weight loss support (including peer support) for BED ≤3 months before screening; use of psychostimulants for fasting or dieting for BED ≤6 months before screening; Montgomery–Åsberg Depression Rating Scale total score ≥18 at screening; being considered a suicide risk by the investigator, having previously made a suicide attempt, or currently demonstrating active suicidal ideation; lifetime histories of psychosis, mania, hypomania, dementia, or ADHD.

Fluoxetine + self-help, self-help manual + placebo.

**Abbreviations:** ADHD=attention-deficit hyperactivity disorder; AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index (kg/m²); BN=bulimia nervosa; CBT=cognitive behavioral therapy; CGI-S= Clinical Global Impressions scale-Severity; DSM=Diagnostic and Statistical Manual of Mental Disorders; ED=eating disorder; MDD= major depressive disorder; n=sample size randomized; NA=not applicable; NR=not reported; Psych.=psychiatric; SD=standard deviation; YBOCS-BE=Yale–Brown Obsessive Compulsive Scale modified for binge eating.
Table 4. Characteristics of Randomized, Placebo-Controlled Trials of Psychological Interventions for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Intervention Type No. of Sessions</th>
<th>Duration (Weeks)</th>
<th>Control</th>
<th>Country</th>
<th>Population</th>
<th>N</th>
<th>Mean Age, Years (SD)</th>
<th>Sex (% Female)</th>
<th>Race/Ethnicity (% Non-White)</th>
<th>Mean BMI (kg/m²) (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfonsson, 2015&lt;sup&gt;56&lt;/sup&gt; Fair</td>
<td>Group psychotherapy (Behavioral Activation) 10 weekly 90-minute group sessions</td>
<td>10</td>
<td>Wait-list</td>
<td>Sweden</td>
<td>Adults with BED (DSM-5) and obesity (BMI &gt;30) presenting to an initial assessment at an outpatient obesity clinic</td>
<td>96</td>
<td>44 (10.74)</td>
<td>94</td>
<td>NR</td>
<td>41.17 (5.32)</td>
</tr>
<tr>
<td>Cachelin, 2019&lt;sup&gt;55&lt;/sup&gt; Fair</td>
<td>Culturally adapted CBT-based guided self-help 8 25-min guided sessions</td>
<td>12</td>
<td>Wait-list</td>
<td>United States</td>
<td>Adult Latinas (18-55 y) who met criteria for BED and had a BMI ≥18 mg/kg, responding to trial advertisement</td>
<td>40</td>
<td>27</td>
<td>100</td>
<td>100 (Latinas)</td>
<td>29.4</td>
</tr>
<tr>
<td>Carter, 2003&lt;sup&gt;57&lt;/sup&gt; Fair</td>
<td>Unguided CBT-based self-help (Overcoming Binge Eating manual) 8 Wait-list and unguided* non-specific self-help control</td>
<td>8</td>
<td>Self-esteem† unguided self-help</td>
<td>Canada</td>
<td>Women (≥17 y) with BN (DSM-IV, but inclusive of those with one binge-eating episode and compensatory behaviors per week vs. 2) and BMI ≥18, recruited from a wait-list of patients referred for outpatient ED treatment</td>
<td>85</td>
<td>27 (8)</td>
<td>100</td>
<td>17</td>
<td>23 (5)</td>
</tr>
<tr>
<td>Carter, 2020&lt;sup&gt;51&lt;/sup&gt; Fair</td>
<td>DBT self-help, guided and unguided (The DBT Solution for Emotional Eating manual) 6 30-min sessions in guided group</td>
<td>12</td>
<td>Self-esteem† unguided self-help</td>
<td>Canada</td>
<td>Adults (19-65 y) with BED (DSM-5) and BMI of ≥18.5, recruited from the community or health centers via advertisements</td>
<td>71</td>
<td>41 (11.46)</td>
<td>93</td>
<td>8</td>
<td>37.3 (9.4)</td>
</tr>
<tr>
<td>DeBar, 2013&lt;sup&gt;107&lt;/sup&gt; Fair</td>
<td>Individual CBT (adolescent specific) 8 in-person sessions, option of supplemental sessions to</td>
<td>24</td>
<td>Wait-list</td>
<td>United States</td>
<td>Female adolescents (12-18 y) enrolled in an HMO with BN or BED (DSM-IV) and at least one binge-eating episode over the previous 3 months</td>
<td>25</td>
<td>15 (1.9)</td>
<td>100</td>
<td>29</td>
<td>26.6 (5.7)</td>
</tr>
<tr>
<td>Author, Year Quality</td>
<td>Intervention Type No. of Sessions</td>
<td>Duration (Weeks)</td>
<td>Control</td>
<td>Country</td>
<td>Population</td>
<td>N</td>
<td>Mean Age, Years (SD)</td>
<td>Sex (% Female)</td>
<td>Race/Ethnicity (% Non-White)</td>
<td>Mean BMI (kg/m²)(SD)</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------------------------</td>
<td>------------------</td>
<td>---------</td>
<td>---------</td>
<td>------------</td>
<td>----</td>
<td>---------------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Fairburn, 2009*² Fair</td>
<td>Two forms of individual CBT: focused form targeting ED pathology only or broad form also addressing other problems common with ED</td>
<td>20</td>
<td>Wait-list</td>
<td>United Kingdom</td>
<td>Adults (18-65 y) with any ED (DSM-IV) requiring treatment (judged by referring provider and ED specialist) and BMI &gt;17.5</td>
<td>154</td>
<td>26 (7.0)</td>
<td>96</td>
<td>10</td>
<td>Lowest adult BMI: 18.7 (2.9) Highest adult BMI: 26.4 (4.8)</td>
</tr>
<tr>
<td>Green, 2018*² Fair</td>
<td>Unguided online dissonance-based intervention (online version of the Body Project) 8 online modules and activities</td>
<td>NR</td>
<td>Wait-list</td>
<td>United States</td>
<td>Adults and adolescents (14-52 y) any ED (DSM-5) or subclinical disorder, recruited via online advertisements and local flyers</td>
<td>82</td>
<td>26 (6.09)</td>
<td>100</td>
<td>10</td>
<td>NR</td>
</tr>
<tr>
<td>Grilo, 2005*³ Fair</td>
<td>Individual CBT (+ placebo), or CBT + fluoxetine 60 mg/day 16 1-hr individual weekly sessions</td>
<td>16</td>
<td>Placebo</td>
<td>United States</td>
<td>Adults (18-60 y) with BED (DSM-IV) and overweight/obese (between 100-200% ideal weight for height based on the 1959 Metropolitan Life Insurance Company Tables)</td>
<td>108</td>
<td>44 (8.6)</td>
<td>78</td>
<td>11</td>
<td>36.3 (7.9)</td>
</tr>
<tr>
<td>Grilo, 2013*⁰ Good</td>
<td>Self-help CBT via a structured manual, initiated by PCP One introductory session with PCP based on trial script⁴</td>
<td>16</td>
<td>Usual primary care</td>
<td>United States</td>
<td>Adults with BED (DSM-IV-TR)⁶ and obesity (BMI ≥30) recruited from primary care settings (via advertisements or referral)</td>
<td>48</td>
<td>46 (11.0)</td>
<td>79</td>
<td>54</td>
<td>37.62 (4.79)</td>
</tr>
</tbody>
</table>
Table 4. Characteristics of Randomized, Placebo-Controlled Trials of Psychological Interventions for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Quality</th>
<th>Intervention Type</th>
<th>No. of Sessions</th>
<th>Duration (Weeks)</th>
<th>Control</th>
<th>Country</th>
<th>Population</th>
<th>N</th>
<th>Mean Age, Years (SD)</th>
<th>Sex (% Female)</th>
<th>Race/Ethnicity (% Non-White)</th>
<th>Mean BMI (kg/m^2)(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grilo, 2014</td>
<td>Fair</td>
<td>Self-help CBT (+ placebo)§ (Overcoming Binge Eating manual) PCP training and to assist with assigning the program</td>
<td>12</td>
<td>Placebo#</td>
<td>United States</td>
<td>Adults (18-65 y) with BED (DSM-5 except for duration of 6 vs.3 months) and obesity (BMI≥30), recruited from advertisements and referrals in primary care</td>
<td>53</td>
<td>44</td>
<td>72</td>
<td>55</td>
<td>37.9</td>
<td></td>
</tr>
<tr>
<td>Hill, 2011</td>
<td>Fair</td>
<td>Individual DBT-appetite focused (DBT-AF) 12 weekly individual sessions</td>
<td>12</td>
<td>Wait-list</td>
<td>United States</td>
<td>Adult women (≥ 18 y) with BN (DSM-IV) or subthreshold** BN (at least 1 binge eating and 1 vomit episode per week over the previous 12 weeks); excluded those with BED or AN and those in concurrent therapy for BN</td>
<td>32</td>
<td>22</td>
<td>100</td>
<td>6</td>
<td>22.6 (NR)</td>
<td></td>
</tr>
<tr>
<td>Kelly, 2014*</td>
<td>Fair</td>
<td>Two forms of self-help: compassion-focused therapy, and behaviorally based self-help Both groups; 1 in-lab self-help session</td>
<td>3</td>
<td>Wait-list</td>
<td>Canada</td>
<td>Adults (&gt;18 y) with BED (DSM-5) not currently receiving therapy, recruited via advertisements</td>
<td>41</td>
<td>45 (15)</td>
<td>83</td>
<td>24</td>
<td>33 (1.05)</td>
<td></td>
</tr>
<tr>
<td>Laessle, 1987</td>
<td>Fair</td>
<td>Behaviorally oriented group treatment Twice weekly for 8 weeks, then once weekly 8 weeks</td>
<td>16</td>
<td>Wait-list</td>
<td>Germany</td>
<td>Adult women with bulimia (DSM-III) seeking treatment at an outpatient psychiatry clinic</td>
<td>17</td>
<td>23</td>
<td>100</td>
<td>NR</td>
<td>21.3</td>
<td></td>
</tr>
<tr>
<td>Ljotsson, 2007</td>
<td>Fair</td>
<td>Internet-assisted CBT-self-help (Overcoming Binge Eating)</td>
<td>12</td>
<td>Wait-list</td>
<td>Sweden</td>
<td>Adults (≥ 18 y) with full or subthreshold BN or BED (criteria NR),††BMI ≥18 not receiving current treatment, responding to study</td>
<td>73</td>
<td>34</td>
<td>94</td>
<td>NR</td>
<td>28.7</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Characteristics of Randomized, Placebo-Controlled Trials of Psychological Interventions for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Intervention Type</th>
<th>Duration (Weeks)</th>
<th>Control</th>
<th>Country</th>
<th>Population</th>
<th>N</th>
<th>Mean Age, Years (SD)</th>
<th>Sex (% Female)</th>
<th>Race Ethnicity (% Non-White)</th>
<th>Mean BMI (kg/m²)(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masson, 2013</td>
<td>DBT guided self-help</td>
<td>13</td>
<td>Wait-list</td>
<td>Canada</td>
<td>Adults (≥ 18 y) with BED (DSM-IV, also including those with binge-eating frequency once weekly for 6 months vs. twice weekly), responding to trial advertisements</td>
<td>60</td>
<td>43 (10.5)</td>
<td>88</td>
<td>8</td>
<td>37.97 (NR)</td>
</tr>
<tr>
<td>Mitchell, 2001</td>
<td>Unguided self-help; provision of manual (manual with elements of CBT, behavioral strategies, and meal planning)</td>
<td>16</td>
<td>Placebo</td>
<td>United States</td>
<td>Adult women (≥18 y) with BN (DSM-III-R) and binge eating coupled with self-induced vomiting 3 times weekly over previous 6 months, within 85% of ideal body weight, not currently receiving treatment</td>
<td>91</td>
<td>27 (7.1)</td>
<td>100</td>
<td>3</td>
<td>NR</td>
</tr>
<tr>
<td>Sánchez-Ortiz, 2011</td>
<td>Internet-based CBT (Overcoming Bulimia Online)</td>
<td>12</td>
<td>Wait-list</td>
<td>United Kingdom</td>
<td>Students (college age) with BN or EDNOS** (DSM-IV, but no min number of binge or purge episodes) and BMI &gt;18.5 kg/m², recruited from higher education institutions</td>
<td>76</td>
<td>24 (5.9)</td>
<td>99</td>
<td>NR</td>
<td>22.0 (2.8)</td>
</tr>
<tr>
<td>Schag, 2019</td>
<td>Impulsivity-focused group intervention</td>
<td>8</td>
<td>Wait-list</td>
<td>Germany</td>
<td>Adults with BED (DSM-5) recruited by email, flyers, press releases, and from an outpatient psychiatry department of a university hospital</td>
<td>80</td>
<td>40</td>
<td>84</td>
<td>NR</td>
<td>36.9</td>
</tr>
<tr>
<td>Schlup, 2009</td>
<td>Group CBT</td>
<td>8</td>
<td>Wait-list</td>
<td>Switzerland</td>
<td>Adults (ages 18-70 y) with BED (DSM-IV-TR), responding to study advertisements</td>
<td>36</td>
<td>44 (10.3)</td>
<td>100</td>
<td>NR</td>
<td>33.4 (7.6)</td>
</tr>
<tr>
<td>Schmidt, 2008</td>
<td>CD-ROM-based unguided CBT</td>
<td>12</td>
<td>Wait-list</td>
<td>United Kingdom</td>
<td>Adults with BN and EDNOS (DSM-IV), referred</td>
<td>97</td>
<td>27 (7.6)</td>
<td>97</td>
<td>27</td>
<td>23.6 (5.2)</td>
</tr>
</tbody>
</table>
**Table 4. Characteristics of Randomized, Placebo-Controlled Trials of Psychological Interventions for Eating Disorders (KQ 4)**

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Intervention Type No. of Sessions</th>
<th>Duration (Weeks)</th>
<th>Control</th>
<th>Country</th>
<th>Population</th>
<th>N</th>
<th>Mean Age, Years (SD)</th>
<th>Sex (% Female)</th>
<th>Race/Ethnicity (% Non-White)</th>
<th>Mean BMI (kg/m²)(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fair</td>
<td>8 modules</td>
<td></td>
<td></td>
<td></td>
<td>by general practitioners to an ED outpatient center</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telch, 1990†</td>
<td>Group CBT 10 weekly 90 min sessions</td>
<td>10</td>
<td>Wait-list</td>
<td>United States</td>
<td>Adult women (18-65 y) with compulsive binge eating but no purging (DSM-II-R BN criteria except for purging criterion), responding to trial advertisements</td>
<td>44</td>
<td>43 (8.4)</td>
<td>100</td>
<td>9</td>
<td>32.6 (5.1)</td>
</tr>
<tr>
<td>Wade, 2017³</td>
<td>Group CBT 18 2h group sessions, one 1-hr individual session</td>
<td>8</td>
<td>Wait-list</td>
<td>Australia</td>
<td>Adult women (18-36 y) with any ED (DSM-5), referred by a clinician or responding to trial advertisements</td>
<td>40</td>
<td>24 (5)</td>
<td>100</td>
<td>5</td>
<td>22.12 (0.54)</td>
</tr>
<tr>
<td>Wagner, 2016³</td>
<td>Web-based CBT 11 assignments, individual feedback from therapists</td>
<td>16</td>
<td>Wait-list</td>
<td>Germany</td>
<td>Adults (18-65 y) with BED (DSM-IV) and no current AN or BN, responding to trial advertisements</td>
<td>139</td>
<td>35 (9.9)</td>
<td>96</td>
<td>NR</td>
<td>32.4 (7.4)</td>
</tr>
<tr>
<td>Wilfley, 1993†</td>
<td>Two forms of group therapy: Group-CBT and group interpersonal therapy 15 90-min sessions</td>
<td>16</td>
<td>Wait-list</td>
<td>United States</td>
<td>Adult women (18-65 y) with BN and binge eating (DSM-III BN criteria, and those meeting all BN criteria except for purging), responding to trial advertisements</td>
<td>56</td>
<td>44 (8.3)</td>
<td>100</td>
<td>14</td>
<td>32.8 (5.2)</td>
</tr>
</tbody>
</table>

* Provision of Self-Assertion for Women manual of similar length/difficulty as intervention but does not address BN.
† Control group received a self-help book on cognitive behavioral techniques for improving self-esteem that does not address binge eating.
‡ Definition used for subclinical eating disorder included endorsement of high levels of body dissatisfaction and one or more of the following behaviors at subclinical levels for weight-control purposes on the Questionnaire for Eating Disorder Diagnoses: binging, laxative use, diuretic use, 24-h fasting, appetite control pill use, strict dieting, or maladaptive exercise (exercising despite injury or exercise which interfered with other important activities).
§ Full or subthreshold criteria; subthreshold criteria defined as binge eating greater than once weekly in frequency and with a duration of at least 6 months (42% randomized).
¶ PCP introduced study and did informed consent; gave patients manual and indicated treatment would be 4 months and they would receive monthly assessments during treatment period.
‖ In addition to groups randomized to self-help + placebo and placebo only, this study included two additional arms involving a medication that is no longer FDA approved (Sibutramine) and therefore not eligible. The sample size and characteristics here refer to only those randomized to self-help + placebo and placebo.
¶ incorporation of structures and processes that are not eligible. The sample size and characteristics here refer to only those randomized to self-help + placebo and placebo. Overall, 35 percent had a subclinical diagnosis; of those with a full clinical diagnosis (65%), specific diagnoses included AN (n=1), BN (n=29), BED (n=13), and OSFED (n=10).
Table 4. Characteristics of Randomized, Placebo-Controlled Trials of Psychological Interventions for Eating Disorders (KQ 4)

** All but six met full DSM-IV criteria for BN.
†† DSM (or other condition definition) used were not reported. Subthreshold BN was defined as at least twice-monthly episodes of binge eating and compensatory behaviors during the last 3 months. Subthreshold BED required at least 2 days with objective bulimic episodes per month during the past 6 months, with binge-eating episodes rated as “markedly stressful.” Overall, 48 percent had BN and 52 percent had BED.
§§ Of those randomized, 51 percent met criteria for BN and 49 percent met criteria for EDNOS.
ǂǂ Of those randomized, 62 percent had BN and 38 percent had EDNOS.
‖‖ Eight (20%) had AN, 23 (58%) had BN, two (5%) had BED, and seven (18%) had OSFED.

Abbreviations: AN=anorexia nervosa; BED=binge eating disorder; BMI=body mass index (kg/m²); BN=bulimia nervosa; CBT=cognitive behavioral therapy; DBT=dialectical behavioral therapy; DBT-AF=dialectical behavioral therapy-appetite focused; DSM=Diagnostic and Statistical Manual of Mental Disorders; ED=eating disorder; EDNOS=Eating disorder not otherwise specified; FDA=Food and Drug Administration; HMO=health maintenance organization; KQ=key question; N=number of participants; No.=number; NR=not reported; OSFED=Other Specified Feeding and Eating Disorder; PCP=primary care physician; SD=standard deviation.
Table 5. Summary of Evidence for Screening in Eating Disorders in Adolescents and Adults

<table>
<thead>
<tr>
<th>Key Question and Topic</th>
<th>No. of Studies; No. of Participants (n)</th>
<th>Summary of Findings</th>
<th>Consistency and Precision</th>
<th>Study Quality</th>
<th>Limitations (Including Reporting Bias)</th>
<th>Overall Strength of Evidence</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ 1. Benefits of screening</td>
<td>0 (0)</td>
<td>No eligible studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Insufficient</td>
<td>NA</td>
</tr>
<tr>
<td>KQ 2. Accuracy of screening tests for detecting eating disorders</td>
<td>SCOFF (≥2) 10 (3,684)</td>
<td>Pooled: Sn: 84 (74 to 90) Sp: 80 (65 to 89)</td>
<td>Consistent and precise for Sn; Inconsistent and imprecise for Sp*</td>
<td>7 Good; 3 Fair</td>
<td>Potential bias related to participant selection. Reference standards varied across studies</td>
<td>Moderate for adequate accuracy</td>
<td>Studies enrolled adults and either limited to women or enrolled a majority of women. Several studies enrolled from specialty clinics or college campuses</td>
</tr>
<tr>
<td></td>
<td>SCOFF (≥3) 7 (2,749)</td>
<td>Pooled: Sn: 69 (56 to 80) Sp: 90 (69 to 98)</td>
<td>Inconsistent and imprecise for both Sn and Sp†</td>
<td>4 Good; 3 Fair</td>
<td>Potential bias related to participant selection. Reference standards varied across studies</td>
<td>Low for adequate accuracy</td>
<td>All studies enrolled adults and either limited to women or enrolled a majority of women. Several studies enrolled from specialty clinics or college campuses</td>
</tr>
<tr>
<td></td>
<td>EDS-PC (≥ 2) 2 (627)</td>
<td>Sn: 97 (88-100) 100 (90-100) Sp: 40 (35 to 46) 71 (64-77)</td>
<td>Consistent and precise for Sn; inconsistent and imprecise for Sp</td>
<td>2 Good</td>
<td>Studies used different reference standards and enrolled diverse populations</td>
<td>Insufficient</td>
<td>One study recruited females and males from primary care and college campuses in the U.K. (77% females), and the other recruited female U.S. veterans.</td>
</tr>
</tbody>
</table>

KQ 3: Harms of screening | 0 (0) | No eligible studies | NA | NA | NA | Insufficient | NA |
<table>
<thead>
<tr>
<th>Key Question and Topic</th>
<th>No. of Studies; No. of Participants (n)</th>
<th>Summary of Findings</th>
<th>Consistency and Precision</th>
<th>Study Quality</th>
<th>Limitations (Including Reporting Bias)</th>
<th>Overall Strength of Evidence</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ 4. Benefits pharmacotherapy for screen-detected or previously untreated ED</td>
<td>LDX (BED) 4 (900)</td>
<td>Pooled mean difference for reduction in YBOCS-BE scores larger in LDX group vs. placebo: -5.75 (-9.32 to -3.17) Other outcomes assessed by one trial each (depression, anxiety, QOL, and function)</td>
<td>YBOCS-BE: consistent, precise Other health outcomes: unknown consistency and imprecise</td>
<td>4 Fair</td>
<td>Outcomes assessed over relatively short duration (11-12 weeks)</td>
<td>Moderate for benefit in ED symptom severity; insufficient for other health outcomes</td>
<td>Studies enrolled adults with BED and obesity recruited via study advertisements.</td>
</tr>
<tr>
<td>Topiramate (BED) 2 (465)</td>
<td>Larger reduction in YBOCS-BE in topiramate groups vs. placebo; difference between groups in score change -6.40 (p&lt;0.001) and -2.55 (p=0.004). Other outcomes assessed by one trial each (depression, anxiety).</td>
<td>YBOCS-BE; consistent, imprecise* Other outcomes: unknown consistency, imprecise</td>
<td>2 Fair</td>
<td>Outcomes assessed over a relatively short duration (14-16 weeks)</td>
<td>Low for benefit in ED symptom severity; insufficient for other outcomes</td>
<td>Studies enrolled adults with BED and obesity recruited via study advertisements.</td>
<td></td>
</tr>
<tr>
<td>Key Question and Topic</td>
<td>No. of Studies; No. of Participants (n)</td>
<td>Summary of Findings</td>
<td>Consistency and Precision</td>
<td>Study Quality</td>
<td>Limitations (Including Reporting Bias)</td>
<td>Overall Strength of Evidence</td>
<td>Applicability</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------------------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>---------------</td>
<td>----------------------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>SSRIs (BED) 5 (208)</td>
<td>Two reported on ED symptom severity (SMD): fluoxetine (EDE-Q): -0.69 (-1.30 to -0.08), and escitalopram (YBOCS-BE): -0.29 (-0.83 to -0.24). Larger reduction in depression symptoms among SSRI groups vs. placebo (5 trials): pooled SMD -0.61 (-0.90 to -0.33)</td>
<td>ED symptom severity: unknown consistency, imprecise Depression: consistent, imprecise</td>
<td>5 Fair</td>
<td>Studies assessed different SSRIs and reported outcomes over 6-16 weeks. Study eligibility criteria varied in terms of body weight and duration/frequency of binge-eating episodes.</td>
<td>Insufficient for ED symptom severity Low for benefit in depression symptom severity</td>
<td>Studies enrolled adults with BED, most recruited via advertisements. Two limited to populations that were obese, and one limited to those with concurrent depression.</td>
<td></td>
</tr>
<tr>
<td>Fluoxetine (BN) 3 (528)</td>
<td>Two found larger reduction in EAT scores among fluoxetine group vs. placebo, difference was statistically significant in one trial. Two found larger reductions in HAM-D scores among fluoxetine vs. placebo, difference was statistically significant in one trial.</td>
<td>ED symptom severity: consistent; imprecise Depression symptom severity: consistent; imprecise</td>
<td>3 Fair</td>
<td>Studies reported outcomes at different durations (8 and 16 weeks).</td>
<td>Low for benefit (ED and depression symptom severity)</td>
<td>All enrolled populations with BN recruited via advertisements; one limited to those with BN and recurrent binge-eating.</td>
<td></td>
</tr>
<tr>
<td>Key Question and Topic</td>
<td>No. of Studies; No. of Participants (n)</td>
<td>Summary of Findings</td>
<td>Consistency and Precision</td>
<td>Study Quality</td>
<td>Limitations (Including Reporting Bias)</td>
<td>Overall Strength of Evidence</td>
<td>Applicability</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------------------</td>
<td>---------------------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>----------------------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>KQ 4. Benefits therapy interventions for screen-detected or previously untreated ED</td>
<td>Guided self-help 7 (431)</td>
<td>Guided self-help reduced ED symptom severity more than control (k=5; 391): pooled SMD -0.96 (-1.26 to -0.67) Guided self-help reduced depression symptoms more than control (k=4, 324): pooled SMD, -0.73 (-1.04 to -0.43)</td>
<td>ED symptom severity: consistent, precise Depression symptom severity: consistent, precise</td>
<td>7 Fair</td>
<td>Frequency and mode of delivering guidance varied (e.g., emails, individual sessions); studies assessed ED and depression symptoms using different measures over a relatively short duration (8-16 weeks).</td>
<td>Moderate for benefit (ED and depression symptom severity)</td>
<td>All enrolled adults with BED recruited primarily via advertisements; several limited to populations that were obese</td>
</tr>
<tr>
<td></td>
<td>Unguided self-help 7 (421)</td>
<td>Pooled results (k=6; 368) favored self-help for reduction in ED symptom severity but difference was not statistically significant: SMD, -0.18 (-0.38 to 0.03) Unguided self-help reduced depression symptoms more than control (k=3,156 participants): SMD 0.37 (-0.68 to -0.05)</td>
<td>ED symptom severity: consistent, imprecise Depression symptom severity: consistent, imprecise</td>
<td>7 Fair</td>
<td>Studies assessed ED and depression symptoms using various measures over a relatively short duration (8-16 weeks). Content and underlying theory of some interventions varied.</td>
<td>Low for benefit (ED and depression symptom severity)</td>
<td>All enrolled adults with BED recruited primarily via advertisements; several limited to populations that were obese</td>
</tr>
<tr>
<td>Key Question and Topic</td>
<td>No. of Studies; No. of Participants (n)</td>
<td>Summary of Findings</td>
<td>Consistency and Precision</td>
<td>Study Quality</td>
<td>Limitations (Including Reporting Bias)</td>
<td>Overall Strength of Evidence</td>
<td>Applicability</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------------------</td>
<td>---------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Group interventions</td>
<td>7 (253)</td>
<td>Group therapy reduced depression symptoms more than control (k=7; 253): pooled SMD, -0.48 (-0.69 to -0.27). Three measured ED symptom severity using the EDE-Q; one found a statistically significant benefit vs. control (SMD -1.01) and two found no significant differences between groups (SMD -0.10 and -0.30)</td>
<td>ED symptom severity: inconsistent, imprecise Depression symptom severity: consistent, precise</td>
<td>7 Fair</td>
<td>Type of group therapy differed across studies (e.g., CBT-based, interpersonal therapy). Outcomes were measured over a relatively short duration (8-16 weeks). Number, length, and frequency of sessions varied.</td>
<td>Moderate for benefit in depression symptom severity; insufficient for ED symptom severity</td>
<td>All enrolled adults with BED recruited primarily via advertisements; several limited to populations that were obese</td>
</tr>
</tbody>
</table>
### Table 5. Summary of Evidence for Screening in Eating Disorders in Adolescents and Adults

<table>
<thead>
<tr>
<th>Key Question and Topic</th>
<th>No. of Studies; No. of Participants (n)</th>
<th>Summary of Findings</th>
<th>Consistency and Precision</th>
<th>Study Quality</th>
<th>Limitations (Including Reporting Bias)</th>
<th>Overall Strength of Evidence</th>
<th>Applicability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual interventions 4 (319)</td>
<td>One trial assessing two forms of individual CBT found no significant differences between groups on EDE-Q scores; one trial of AF-DBT found significant improvement in EDE-Q scores (SMD, -1.18; -1.94 to -0.43). One trial of CBT found improvement on BDI scores (SMD, -0.60, -1.14 to -0.06) and a trial of DBT-AF found significant improvement on BDI-II scores (SMD, -0.92, -1.65 to -0.19). One trial limited to adolescents found no significant improvement in depression (BDI scores) or psychosocial function (SCARED scores).</td>
<td>Unknown consistency; imprecise</td>
<td>Fair</td>
<td>Trials addressed different types of individual therapy (e.g., CBT, DBT) and reported on different outcomes over a relatively short duration (6 to 16 weeks)</td>
<td>Insufficient</td>
<td>All enrolled adults with BED (or BED and BN) referred or recruited via trial advertisements.</td>
<td></td>
</tr>
<tr>
<td>Key Question and Topic</td>
<td>No. of Studies; No. of Participants (n)</td>
<td>Summary of Findings</td>
<td>Consistency and Precision</td>
<td>Study Quality</td>
<td>Limitations (Including Reporting Bias)</td>
<td>Overall Strength of Evidence</td>
<td>Applicability</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>---------------</td>
<td>----------------------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>KQ 4. Harms of pharmacotherapy for screen-detected or previously untreated ED</td>
<td>9 (2,006)</td>
<td>LDX (k=4) is associated with higher rates of dry mouth, headache, and insomnia vs. placebo. Topiramate (k=2) is associated with significant higher rates of paresthesia, taste, and difficulty with concentration/confusion vs. placebo. Trials of other medications were assessed by only one study each and results were imprecise.</td>
<td>LDX: consistent; imprecise</td>
<td>Fair</td>
<td>Some trials did not prespecify adverse events or describe how they were ascertained; trials assessed adverse events over a relatively short duration.</td>
<td>Topiramate and LDX: Moderate for increased rates of various adverse effects Other medications: insufficient</td>
<td>All studies enrolled adults with BED and obesity, recruited via referrals or study advertisements. Most studies of LDX limited to populations without ADHD, substance abuse or other psychiatric comorbidity</td>
</tr>
</tbody>
</table>

---

* Based on Appendix F Figure 1, the 95 percent prediction region indicates the results are mostly consistent for sensitivity and somewhat inconsistent for specificity; based on the 95 percent confidence region, estimates are precise for sensitivity and somewhat imprecise for specificity.
† Based on Appendix F Figure 2, the 95 percent prediction region indicates results are inconsistent; based on the 95 percent confidence region, estimates are imprecise.
‡ Although results were in same direction of effect (favoring SSRI), only two studies assessed change in ED symptom reduction. Each assessed a different medication using different measures of ED symptom burden (YBOCS-BE vs. EDE-Q) and reported outcomes at slightly different durations (12 vs. 16 weeks), limiting ability to assess consistency for this outcome.
§ One additional trial (n=42) assessed Fluoxetine 60mg/day for BN and reported no significant difference between groups for ED symptom severity (EDI) and depression (HAM-D), p>0.05

**Abbreviations:** ADHD=attention deficit hyperactivity disorder; AF-DBT=appetite focused-dialectical behavior therapy; BDI=Beck Depression Inventory; BED= binge eating disorder; BN=bulimia nervosa; CBT=cognitive behavioral therapy; DBT=dialectical behavior therapy; DBT-AF=dialectical behavior therapy, appetite focused; EAT=Eating Attitudes Test; ED=eating disorders; EDE-Q=Eating Disorder Examination Questionnaire; EDI=Eating Disorder Inventory; EDS-PC=Eating Disorder Screen for Primary Care; k=number of studies; HAM-D=Hamilton Depression Rating Scale—Depression; KQ=key question; LDX=lisdexamfetamine; MCID=Minimal clinically important change; n=number of participants; NA=not applicable; No.=number; QOL=quality of life; SCARED=Screen for Child Anxiety Related-Emotional Disorders; SMD=standardized mean difference; Sn=sensitivity; Sp=specificity; SSRI=selective serotonin reuptake inhibitor; U.K.=United Kingdom; US=United States; YBOCS-BE=Yale–Brown Obsessive Compulsive Scale modified for binge eating; vs.=versus.
### Detailed Summary of Diagnostic Criteria

Table 1 summarizes key diagnostic criteria of eating disorders likely to be undetected and thus relevant to screening in primary care based on the 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5). Two diagnoses in this diagnostic category of feeding and eating disorders (pica and rumination disorder) are not included below because their symptomatology is often readily identifiable and unlikely to benefit from early detection through routine screening.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Key Diagnostic Criteria</th>
</tr>
</thead>
</table>
| **AN** subtypes: Restricting type | • Restriction of energy intake relative to requirements, leading to a significantly low body weight in the context of age, sex, developmental trajectory, and physical health.  
  • Intense fear of gaining weight or of becoming fat, or persistent behavior that interferes with weight gain, even though at a significantly low weight.  
  • Disturbance in the way in which one’s body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or persistent lack of recognition of the seriousness of the current low body weight. |
| Binge-eating/purging type | |
| **BN** | • Recurrent episodes of binge eating that are characterized by both of the following:  
  o Eating, in a discrete period of time (e.g., within any 2-hour period), an amount of food that is definitely larger than what most individuals would eat in a similar period of time under similar circumstances  
  o A sense of lack of control overeating during the episode (e.g., a feeling that one cannot stop eating or control what or how much one is eating)  
  • Recurrent inappropriate compensatory behaviors to prevent weight gain (e.g., self-induced vomiting; misuse of laxative, diuretics, or other medications; fasting; or excessive exercise)  
  • Self-evaluation is unduly influenced by body shape and weight |
| **BED** | • Recurrent episodes of binge eating (as defined above)  
  • Binge eating is associated with:  
    o Eating more rapidly than usual  
    o Eating until feeling uncomfortably full  
    o Eating large amounts of food when not feeling physically hungry  
    o Eating alone because of feeling embarrassed by how much one is eating  
    o Feeling disgusted with oneself, depressed, or very guilty afterward  
  • Marked distress about the binge-eating episodes |
| Other specified feeding and eating disorder | • An eating or feeding disturbance that causes clinically significant distress or impairment but does not meet the full criteria for any of the disorders in this diagnostic class. This might include (but is not limited to):  
  o Atypical AN  
  o BN or BED of low frequency or limited duration  
  o Purging disorder  
  o Night eating syndrome |
| Avoidant/restrictive food intake disorder | • An eating or feeding disturbance (e.g., apparent lack of interest in eating or food, avoidance based on sensory characteristics of food, concern about aversive consequences of eating) associated with one or more of the following:  
  o Significant weight loss or failure to achieve expected weight gain or faltering growth in children  
  o Significant nutritional deficiency  
  o Dependent on enteral feeding or oral nutritional supplements  
  o Marked interference with psychosocial functioning |

**Abbreviations:** AN=anorexia nervosa; BED=binge-eating disorder; BN=bulimia nervosa; DSM-5= *Diagnostic and Statistical Manual of Mental Disorders* (5th edition).
## Detailed Summary of Eating Disorders’ Prevalence

Details of study cohorts and results of nationally representative surveys of eating disorders’ prevalence in the United States are summarized in Table 2.

<table>
<thead>
<tr>
<th>Study Cohort</th>
<th>Cohort Description</th>
<th>DSM</th>
<th>Study Population</th>
<th>Lifetime Prevalence % (SE)</th>
<th>Information on Survey Questions</th>
</tr>
</thead>
</table>
| National Comorbidity Replication Survey (NCS-R)² (n=2,980) | Survey of the U.S. household population (2001 and 2003) based on a multistage clustered-area probability design. English-speaking adults ages 18 years or older. Subsample assessed for eating disorders. | DSM-IV  | In the weighted sample³
Ages 18-34 years: 31.5%
Ages 35-49 years: 30.9%
Female: 5.30%
White: 72.3%
Black: 12.4%
Hispanic: 11.1% | Women
AN: 0.9% (0.3)
BN: 1.5% (0.3)
BED: 3.5% (0.5)
Men
AN: 0.3% (0.1)
BN: 0.5% (0.3)
BED: 2.0% (0.5) | NCS-R diagnoses were based on Version 3.0 of the WHO Composite International Diagnostic Interview (CIDI), a structured lay-administered diagnostic interview that generates diagnoses according to both ICD-10 and DSM-IV criteria.³⁴ |
| National Comorbidity Survey Replication Adolescent Supplement (NCS-A)³ (n=10,123) | Nationally representative survey (2001 and 2004) of U.S. adolescents ages 13 to 17 years, extension of NCS-R.⁵ | DSM-IV  | In the weighted sample³
Age 13: 20.7%
Age 17: 18.8%
Female: 48.8%
White: 65.5%
Black: 15.1%
Hispanic: 14.4% | Women
AN: 0.3% (0.1)
BN: 1.3% (0.3)
BED: 2.3% (0.4)
Men
AN: 0.3% (0.1)
BN: 0.5% (0.2)
BED: 1.3% (0.4) | Although the version of the CIDI used in the NCS-A is very similar to the NCS-R version, a number of important modifications were made for the NCS-A to make sure the instrument was relevant to the special experiences and language of youth.⁶ |
| Collaborative Psychiatric Epidemiology Surveys⁷ (n=12,337) | Comprising three nationally representative U.S. surveys: NCS-R, National Survey of American Life (NSAL), and National Latino and Asian American Study (NLAAS); NSC-R described above; NSAL is a survey of U.S. adults (2001 and 2003) self-identifying as African American or of Caribbean descent; NLAAS is a nationally representative survey (2001 and 2003) of Latinos and Asian Americans. | DSM-IV  | In the weighted sample³
Ages 18-24: 14-16%
Ages 25-34: 19%
Ages 35-44: 20-22%
Ages 45-54: 19-20%
Female: 58.7%
White: 58-59%
Black: 16-18%
Hispanic: 16-20% | Women
AN: 0.69%
BN: 1.68%
BED: 2.97%
Men
AN: 0.19%
BN: 0.55%
BED: 1.59% | All three of these surveys used similar versions of the WHO CIDI. |
Appendix A Table 2. Lifetime Prevalence Estimates of Eating Disorders in the United States

<table>
<thead>
<tr>
<th>Study Cohort</th>
<th>Cohort Description</th>
<th>DSM</th>
<th>Study Population</th>
<th>Lifetime Prevalence % (SE)</th>
<th>Information on Survey Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Epidemiologic Survey on</td>
<td>Nationally representative survey of non-institutionalized U.S. civilians, age 18</td>
<td>DSM-V</td>
<td>In the weighted sample²</td>
<td>Women</td>
<td>NESARC-III used the National Institute on Alcohol Abuse and Associated Disabilities Interview Schedule-5 (AUDADIS-5) (14) to assess DSM-5-defined psychiatric disorders and their criteria, including AN, BN, and BED.†</td>
</tr>
<tr>
<td>Alcohol and Related Conditions</td>
<td>years or older, based on multistage, probabilistic sampling (2012 and 2013).</td>
<td></td>
<td>Ages 18-24: 13.0%</td>
<td>AN: 1.42% (0.12)</td>
<td></td>
</tr>
<tr>
<td>(NESARC-III)</td>
<td></td>
<td></td>
<td>Ages 25-34: 17.4%</td>
<td>BN: 0.46% (0.06)</td>
<td></td>
</tr>
<tr>
<td>(n=36,309)</td>
<td></td>
<td></td>
<td>Ages 35-44: 17.1%</td>
<td>BED: 1.25% (0.10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ages 45-54: 18.5%</td>
<td>Men</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Female: 51.9%</td>
<td>AN: 0.12% (0.04)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>White: 66.2%</td>
<td>BN: 0.08% (0.03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Black: 11.6%</td>
<td>BED: 0.42% (0.06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hispanic: 14.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

² Most of the CIDI questions closely paralleled DSM-IV criteria; the exception was BED, where DSM-IV required at least 6 months of regular eating binges compared with CIDI (3 months of symptoms).
† For an AN diagnosis, respondents were required to meet the following criteria: 1) had a self-reported lowest BMI of 18.5; 2) tried not to gain weight or restricted food intake despite low weight; 3) were afraid of gaining weight or “getting fat” despite low weight; and 4) reported at least one of the following while their BMI was lowest: a) thought they “looked fat,” b) thought their weight or shape was one of the most important things about them, c) did not think they might have been unhealthy, d) did not believe others who thought their weight was unhealthy, or e) were constantly weighing themselves or measuring body parts. For BN and BED diagnoses, respondents were required to report recurrent binge eating, which was defined by three criteria: 1) had ever eaten an unusually large amount of food within a 2-hour period, not including during the holidays; 2) had ever eaten unusually large amounts of food, on average, at least once weekly for at least 3 months; and 3) while eating an unusually large amount of food, had felt unable to stop eating or control how much or what they were eating. For a BN diagnosis, in addition to meeting criteria for recurrent binge eating, respondents were required to report whether during any of those times that they were binge eating they 1) tried to keep from gaining weight by vomiting; using enemas, laxatives, diuretics, or other medicines; fasting; or exercising excessively; 2) engaged in the weight compensatory behaviors at least once weekly for at least 3 months; or 3) thought their weight or shape was one of the most important things about them. For a BED diagnosis, in addition to meeting criteria for recurrent binge eating, respondents were required to report 1) eating an unusually large amount of food that made them very upset and 2) at least three of the following five features during the times they ate unusually large amounts of food: a) eating much more quickly than usual; b) eating until uncomfortably full; c) eating despite not being hungry; d) eating alone because they were embarrassed by how much they were eating; or e) feeling disgusted, depressed, or very guilty about the overeating.

Abbreviations: AN=anorexia nervosa; AUDADIS-5=Associated Disabilities Interview Schedule-5; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CIDI=Composite International Diagnostic Interview; DSM-IV=Statistical Manual of Mental Disorders, Version 4; ICD-10=International Classification of Diseases, 10th Revision; NCS-A=National Comorbidity Survey Replication Adolescent Supplement; NCS-R=National Comorbidity Replication Survey; NESARC-III=National Epidemiologic Survey on Alcohol and Related Conditions; NLAAS=National Latino and Asian American Study; NSAL=National Survey of American Life; SE=standard error; U.S.=United States; WHO=World Health Organization.
### Detailed Summary of Recommendations From Organizations

Table 3 summarizes recommendations from other organizations relevant to screening for eating disorders in clinical settings.

<table>
<thead>
<tr>
<th>Organization, Year</th>
<th>Screening Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP, 2020&lt;sup&gt;10&lt;/sup&gt;</td>
<td>Pediatrics should screen for eating disorders by monitoring and assessing risk factors and symptoms at annual and sports physicals. Pediatrics should monitor and identify changes in height, weight, BMI, and vital signs longitudinally. If findings indicate that an eating disorder may be present, the pediatrician should conduct thorough medical and psychological assessments to identify if an eating disorder diagnosis is appropriate. After diagnosis, the pediatrician may continue to monitor the patient, help set weight goals, refer the patient to eating disorder specialists ideally with expertise among this age group, and continue to care for the patient as part of a multidisciplinary team.</td>
</tr>
<tr>
<td>ACOG, 2018&lt;sup&gt;11&lt;/sup&gt;</td>
<td>ACOG recommends that practitioners be able to identify signs of disordered eating and screen at risk patients, especially considering the presence of many gynecological symptoms, including irregular menstrual cycles, amenorrhea, pelvic pain, atrophic vaginitis, and breast atrophy.</td>
</tr>
<tr>
<td>APA, 2012&lt;sup&gt;12&lt;/sup&gt;</td>
<td>The APA recommends that practitioners working with young athletes pay special attention to disordered eating. Assessment of weight, body image, amenorrhea, and nutrition can help screen and identify athletes suffering from or at risk for eating disorders.</td>
</tr>
<tr>
<td>NICE, 2017&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Patients who present symptoms of eating disorders should be assessed and treated as soon as possible. Guidelines highlight that reliance on screening tools alone, such as SCOFF, is not sufficient for diagnosis.</td>
</tr>
<tr>
<td>AED, 2016&lt;sup&gt;14&lt;/sup&gt;</td>
<td>All high-risk patients should be monitored for symptoms of eating disorders, which may present in patients of any age, race, gender, or size. Screening with validated tools, such as SCOFF, can help identify patients who may need treatment or referral to specialty care.</td>
</tr>
<tr>
<td>SAHM, 2015&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Guidelines state that medical providers should be able to recognize and diagnose eating disorders in adolescents and young adults and highlight the importance of medical providers in monitoring for medical complications in the context of multidisciplinary care for those with eating disorders.</td>
</tr>
<tr>
<td>AACAP, 2015&lt;sup&gt;16&lt;/sup&gt;</td>
<td>AACAP recommends mental health practitioners screen all preteen and adolescent patients for eating disorders through height and weight assessments and screener questions about eating patterns and body image. Concern about these results should lead to referral for further evaluation. For older patients, the following screening instruments are recommended: The Eating Disorder Examination Questionnaire, Eating Disorder Inventory, and Eating Attitudes Test. For younger children, the following screening instruments are recommended: The Kids' Eating Disorder Survey, the Children's Eating Disorder Questionnaire, the Child-Eating Attitudes Test, and the Eating Disorder Inventory for Children.</td>
</tr>
</tbody>
</table>

**Abbreviations:** AACAP=American Academy of Child and Adolescent Psychiatry; AAP=American Academy of Pediatrics; ACOG=American Congress of Obstetricians and Gynecologists; AED=Academy for Eating Disorders; APA=American Psychological Association; EDI=Eating Disorder Inventory; EDI-C=Eating Disorder Inventory-Children; NICE=National Institute for Health and Care Excellence; SAHM=Society for Adolescent Health and Medicine.
### Appendix B2. Eligibility Criteria

**PubMed, 6/18/2020**

**Total Unduplicated Yield = 8,570**

**Screening Benefits (KQ 1) and Harms (KQ 3) Searches**

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>#1 AND #2</td>
<td>22,089</td>
</tr>
<tr>
<td>#5</td>
<td>#1 AND #4</td>
<td>2,683</td>
</tr>
<tr>
<td>#6</td>
<td>#3 OR #5</td>
<td>22,285</td>
</tr>
<tr>
<td>#7</td>
<td>#3 OR #5 Filters: English</td>
<td>20,911</td>
</tr>
<tr>
<td>#8</td>
<td>#3 OR #5 Filters: English, Child: 6-12 years</td>
<td>2,783</td>
</tr>
<tr>
<td>#9</td>
<td>#3 OR #5 Filters: English, Adolescent: 13-18 years, Child: 6-12 years</td>
<td>8,454</td>
</tr>
<tr>
<td>#10</td>
<td>#3 OR #5 Filters: English, Adult: 19+ years, Child: 6-12 years, Adolescent: 13-18 years</td>
<td>14,932</td>
</tr>
<tr>
<td>#11</td>
<td>#7 AND (publisher[wb] OR inprocess[wb] OR pubstatusaheadofprint)</td>
<td>1,239</td>
</tr>
<tr>
<td>#13</td>
<td>#10 OR #12</td>
<td>16,103</td>
</tr>
<tr>
<td>#15</td>
<td>#13 NOT #14</td>
<td>14,483</td>
</tr>
<tr>
<td>#17</td>
<td>#15 AND #16</td>
<td>1,415</td>
</tr>
<tr>
<td>#19</td>
<td>#15 AND #18</td>
<td>3,083</td>
</tr>
<tr>
<td>#20</td>
<td>#19 NOT #17</td>
<td>2,743</td>
</tr>
</tbody>
</table>
## Appendix B2. Eligibility Criteria

### Screening Test Accuracy (KQ 2)

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>#1 AND #2</td>
<td>22,089</td>
</tr>
<tr>
<td>#5</td>
<td>#1 AND #4</td>
<td>2,683</td>
</tr>
<tr>
<td>#6</td>
<td>#3 OR #5</td>
<td>22,285</td>
</tr>
<tr>
<td>#7</td>
<td>#3 OR #5 Filters: English</td>
<td>20,911</td>
</tr>
<tr>
<td>#8</td>
<td>#3 OR #5 Filters: English. Child: 6-12 years</td>
<td>2,783</td>
</tr>
<tr>
<td>#9</td>
<td>#3 OR #5 Filters: English. Adolescent: 13-18 years, Child: 6-12 years</td>
<td>8,454</td>
</tr>
<tr>
<td>#10</td>
<td>#3 OR #5 Filters: English. Adult: 19+ years, Child: 6-12 years, Adolescent: 13-18 years</td>
<td>14,932</td>
</tr>
<tr>
<td>#11</td>
<td>#7 AND (publisher*[sb] OR inprocess*[sb] OR pubstatusaheadofprint)</td>
<td>1,239</td>
</tr>
<tr>
<td>#13</td>
<td>#10 OR #12</td>
<td>16,103</td>
</tr>
<tr>
<td>#15</td>
<td>#13 NOT #14</td>
<td>15,938</td>
</tr>
<tr>
<td>#17</td>
<td>#15 AND #16</td>
<td>2,085</td>
</tr>
</tbody>
</table>
### Intervention Benefits (KQ 4) and Harms (KQ 5)

<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#3</td>
<td>#1 AND #2</td>
<td>4,869</td>
</tr>
<tr>
<td>#5</td>
<td>#1 AND #4</td>
<td>42,286</td>
</tr>
<tr>
<td>#6</td>
<td>&quot;Antidepressive Agents, Second-Generation*[Mesh] OR (&quot;Antipsychotic Agents*[Mesh] AND atypical) OR (antidepressant*[tiab] AND &quot;second-generation*[tiab] OR (antidepressive AND &quot;second-generation*[tiab]) OR &quot;atypical antipsychotic*[tiab] OR lisdexamfetamine OR topiramate</td>
<td>24,534</td>
</tr>
<tr>
<td>#7</td>
<td>#1 AND #6</td>
<td>514</td>
</tr>
<tr>
<td>#8</td>
<td>#3 OR #5 OR #7</td>
<td>42,658</td>
</tr>
<tr>
<td>#9</td>
<td>#3 OR #5 Filters: English</td>
<td>37,423</td>
</tr>
<tr>
<td>#10</td>
<td>#3 OR #5 Filters: English, Child: 6-12 years</td>
<td>4,000</td>
</tr>
<tr>
<td>#11</td>
<td>#3 OR #5 Filters: English, Adolescent: 13-18 years, Child: 6-12 years</td>
<td>9,856</td>
</tr>
<tr>
<td>#12</td>
<td>#3 OR #5 Filters: English, Adult: 19+ years, Child: 6-12 years, Adolescent: 13-18 years</td>
<td>20,296</td>
</tr>
<tr>
<td>#13</td>
<td>#9 AND (publisher*[sb] OR inprocess*[sb] OR pubstatusaheadofprint</td>
<td>1.713</td>
</tr>
<tr>
<td>#15</td>
<td>#12 OR #14</td>
<td>21,935</td>
</tr>
<tr>
<td>#17</td>
<td>#15 NOT #16</td>
<td>16,165</td>
</tr>
<tr>
<td>#19</td>
<td>#17 AND #18</td>
<td>2,327</td>
</tr>
<tr>
<td>#20</td>
<td>&quot;Observational Study*[Publication Type] OR &quot;Prospective Studies*[Mesh] OR &quot;Cohort Studies*[Mesh] OR &quot;observational study*[tiab] OR &quot;observational studies*[tiab] OR prospective*[tiab] OR cohort*[tiab]</td>
<td>2,547,404</td>
</tr>
<tr>
<td>#21</td>
<td>#17 AND #20</td>
<td>3,959</td>
</tr>
<tr>
<td>#22</td>
<td>#21 NOT #19</td>
<td>3,395</td>
</tr>
</tbody>
</table>
### Cochrane Review, 6/20/2020

**KQ 2 Search for Diagnostic Accuracy, 6/20/2020**

All results are trials = 269; 179 imported

<table>
<thead>
<tr>
<th>Search Number</th>
<th>String</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>[mh &quot;Anorexia&quot;] OR [mh &quot;Anorexia Nervosa&quot;] OR [mh &quot;Avoidant Restrictive Food Intake Disorder&quot;] OR [mh &quot;Binge-Eating Disorder&quot;] OR [mh Bulimia] OR [mh &quot;Feeding and Eating Disorders&quot;] OR anorexi*:ti,ab OR &quot;avoidant restrictive food intake disorder&quot;:ti,ab,kw OR ARFID:ti,ab OR &quot;binge-eating&quot;:ti,ab OR bulimia:ti,ab OR bulimic:ti,ab OR (eating:ti,ab AND disorder*:ti,ab) OR (feed*:ti,ab OR &quot;night eating&quot;:ti,ab OR &quot;purging disorder&quot;:ti,ab OR EDNOS:ti,ab OR OSFED:ti,ab)</td>
<td>8581</td>
</tr>
<tr>
<td>#2</td>
<td>[mh &quot;Surveys and Questionnaires&quot;] OR [mh &quot;Psychiatric Status Rating Scales&quot;] OR identifying:ti,ab OR instrument:ti,ab OR instruments:ti,ab OR measure:ti,ab OR measures:ti,ab OR questionnaire:ti,ab OR questionnaires:ti,ab OR inventory:ti,ab OR inventories:ti,ab OR scale:ti,ab OR scales:ti,ab OR screen:ti,ab OR screening:ti,ab</td>
<td>433635</td>
</tr>
<tr>
<td>#3</td>
<td>#1 AND #2</td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td>&quot;Adolescent Binge Eating Scale&quot;:ti,ab OR &quot;Eating Attitudes Test&quot;:ti,ab OR &quot;EAT-26&quot;:ti,ab OR &quot;Eating Disorder Inventory&quot;:ti,ab OR &quot;Eating Disorder Screen for Primary Care&quot;:ti,ab OR &quot;EDS-PC&quot;:ti,ab OR &quot;Primary Care Evaluation of Mental Disorders Patient Health Questionnaire&quot;:ti,ab OR SCOFF:ti,ab OR &quot;Sick, Control, One, Fat and Food&quot;:ti,ab OR &quot;Dutch Eating Behavior Questionnaire&quot;:ti,ab OR DEBQ:ti,ab OR &quot;Eating Disorder Examination&quot;:ti,ab OR &quot;Minnesota Eating Behavior Survey&quot;:ti,ab OR &quot;Patient Health Questionnaire&quot;:ti,ab OR PHQ:ti,ab OR PRIMEMD:ti,ab OR &quot;Screening for Disordered Eating&quot;:ti,ab</td>
<td>2944</td>
</tr>
<tr>
<td>#5</td>
<td>#1 AND #4</td>
<td></td>
</tr>
<tr>
<td>#6</td>
<td>#3 OR #5</td>
<td>4015</td>
</tr>
<tr>
<td>#7</td>
<td>[mh child] OR [mh adolescent] OR [mh adult]</td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>#6 AND #7</td>
<td>633</td>
</tr>
<tr>
<td>#9</td>
<td>(#6 AND (adolescent:ti,ab OR adolescents:ti,ab OR adult:ti,ab OR adults:ti,ab OR elderly:ti,ab OR teen:ti,ab OR teens:ti,ab OR teenage:ti,ab OR teenaged:ti,ab)) OR (#6 NOT (newborn:ti,ab OR newborns:ti,ab OR infant:ti,ab OR infants:ti,ab))</td>
<td>3869</td>
</tr>
<tr>
<td>#10</td>
<td>#8 OR #9</td>
<td>3870</td>
</tr>
<tr>
<td>#11</td>
<td>address:pt OR &quot;autobiography&quot;:pt OR &quot;bibliography&quot;:pt OR &quot;biography&quot;:pt OR &quot;case control&quot; OR &quot;case report&quot; OR &quot;case reports&quot; OR &quot;case series&quot; OR &quot;comment&quot;:pt OR &quot;comment on&quot; OR congress:pt OR &quot;cross-sectional&quot; OR &quot;dictionary&quot;:pt OR &quot;directory&quot;:pt OR &quot;editorial&quot;:pt OR &quot;estschrift&quot;:pt OR &quot;historical article&quot;:pt OR &quot;interview&quot;:pt OR lecture:pt OR &quot;legal case&quot;:pt OR &quot;legislation&quot;:pt OR letter:pt OR &quot;news&quot;:pt OR &quot;newspaper article&quot;:pt OR &quot;patient education handout&quot;:pt OR &quot;periodical index&quot;:pt OR &quot;retrospective cohort&quot;:pt OR (mh &quot;Humans&quot;) NOT (mh &quot;Animals&quot;) OR rats OR cow OR cows OR chicken OR chickens OR horse OR horses OR mice OR mouse OR bovine OR sheep OR ovine OR murine OR murinae</td>
<td>65740</td>
</tr>
<tr>
<td>#12</td>
<td>#10 NOT #11</td>
<td>3714</td>
</tr>
<tr>
<td>#13</td>
<td>[mh &quot;Sensitivity and Specificity&quot;] OR [mh &quot;Predictive Value of Tests&quot;] OR [mh &quot;ROC Curve&quot;] OR [mh &quot;Reproducibility of Results&quot;] OR [mh &quot;False Negative Reactions&quot;] OR [mh &quot;False Positive Reactions&quot;] OR [mh &quot;predictive value&quot;:ti,ab,kw OR sensitivity:ti,ab,kw OR specificity:ti,ab,kw OR accuracy:ti,ab,kw OR ROC:ti,ab,kw OR ROC:ti,ab,kw OR reproducibility:ti,ab,kw OR &quot;false positive&quot;:ti,ab,kw OR &quot;false negative&quot;:ti,ab,kw OR &quot;likelihood ratio&quot;:ti,ab,kw</td>
<td>92405</td>
</tr>
<tr>
<td>#14</td>
<td>#12 AND #13</td>
<td></td>
</tr>
<tr>
<td>#15</td>
<td>[mh &quot;Anorexia&quot;] OR [mh &quot;Anorexia Nervosa&quot;] OR [mh &quot;Avoidant Restrictive Food Intake Disorder&quot;] OR [mh &quot;Binge-Eating Disorder&quot;] OR [mh Bulimia] OR [mh &quot;Feeding and Eating Disorders&quot;] OR anorexi*:ti,ab OR &quot;avoidant restrictive food intake disorder&quot;:ti,ab,kw OR ARFID:ti,ab OR &quot;binge-eating&quot;:ti,ab OR bulimia:ti,ab OR bulimic:ti,ab OR (eating:ti,ab AND disorder*:ti,ab) OR (feed*:ti,ab OR &quot;AND disorder&quot;:ti,ab) OR (food*:ti,ab AND neophobia*:ti,ab) OR &quot;night eating&quot;:ti,ab OR &quot;purging disorder&quot;:ti,ab OR EDNOS:ti,ab OR OSFED:ti,ab</td>
<td>8581</td>
</tr>
</tbody>
</table>
## Cochrane Library, Screening, 6/20/2020

RCTs = 2,993; 1,735 imported

Observational (all are controlled trials): 34; 13 imported

<table>
<thead>
<tr>
<th>Search Number</th>
<th>String</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>[mh &quot;Anorexia&quot;] OR [mh &quot;Anorexia Nervosa&quot;] OR [mh &quot;Avoidant Restrictive Food Intake Disorder&quot;] OR [mh &quot;Binge-Eating Disorder&quot;] OR [mh Bulimia] OR [mh &quot;Feeding and Eating Disorders&quot;] OR anorexi*:ti,ab OR &quot;avoidant restrictive food intake disorder&quot;:ti,ab,kw OR ARFID:ti,ab OR &quot;binge-eating&quot;:ti,ab OR bulimia:ti,ab OR (eating:ti,ab AND disorder*:ti,ab) OR (feed*:ti,ab AND disorder*:ti,ab) OR (food*:ti,ab AND neophobia*:ti,ab) OR &quot;night eating&quot;:ti,ab OR &quot;purging disorder&quot;:ti,ab OR EDNOS:ti,ab OR OSFED:ti,ab</td>
<td>8581</td>
</tr>
<tr>
<td>#2</td>
<td>[mh &quot;Surveys and Questionnaires&quot;] OR [mh &quot;Psychiatric Status Rating Scales&quot;] OR identifying:ti,ab OR identification:ti,ab OR instrument:ti,ab OR instruments:ti,ab OR measure:ti,ab OR measures:ti,ab OR questionnaire:ti,ab OR questionnaires:ti,ab OR inventory:ti,ab OR inventories:ti,ab OR scale:ti,ab OR scales:ti,ab OR screen:ti,ab OR screening:ti,ab</td>
<td>433635</td>
</tr>
<tr>
<td>#3</td>
<td>#1 AND #2</td>
<td>3978</td>
</tr>
<tr>
<td>#4</td>
<td>&quot;Adolescent Binge Eating Scale&quot;:ti,ab OR &quot;Eating Attitudes Test&quot;:ti,ab OR &quot;EAT-26&quot;:ti,ab OR &quot;Eating Disorder Inventory&quot;:ti,ab OR &quot;Eating Disorder Screen for Primary Care&quot;:ti,ab OR &quot;EDS-PC&quot;:ti,ab OR &quot;Primary Care Evaluation of Mental Disorders Patient Health Questionnaire&quot;:ti,ab OR SCOFF:ti,ab OR &quot;Sick, Control, One, Fat and Food&quot;:ti,ab OR &quot;Dutch Eating Behavior Questionnaire&quot;:ti,ab OR DEBQ:ti,ab OR &quot;Eating Disorder Examination&quot;:ti,ab OR &quot;Minnesota Eating Behavior Survey&quot;:ti,ab OR &quot;Patient Health Questionnaire&quot;:ti,ab OR PHQ:ti,ab OR PRIMEMD:ti,ab OR &quot;Screening for Disordered Eating&quot;:ti,ab</td>
<td>2944</td>
</tr>
<tr>
<td>#5</td>
<td>#1 AND #4</td>
<td>523</td>
</tr>
<tr>
<td>#6</td>
<td>#3 OR #5</td>
<td>4015</td>
</tr>
<tr>
<td>#7</td>
<td>[mh child] OR [mh adolescent] OR [mh adult]</td>
<td>120404</td>
</tr>
<tr>
<td>#8</td>
<td>#6 AND #7</td>
<td>633</td>
</tr>
<tr>
<td>#9</td>
<td>(#6 AND (adolescent:ti,ab OR adolescents:ti,ab OR adult:ti,ab OR adults:ti,ab OR elderly:ti,ab OR teen:ti,ab OR teens:ti,ab OR teenage:ti,ab OR teenaged:ti,ab)) OR (#6 NOT (newborn:ti,ab OR newborns:ti,ab OR infant:ti,ab OR infants:ti,ab))</td>
<td>3869</td>
</tr>
<tr>
<td>#10</td>
<td>#8 OR #9</td>
<td>3870</td>
</tr>
<tr>
<td>#11</td>
<td>address:pt OR &quot;autobiography&quot;:pt OR &quot;bibliography&quot;:pt OR &quot;biography&quot;:pt OR &quot;case control&quot; OR &quot;case report&quot; OR &quot;case reports&quot; OR &quot;case series&quot; OR &quot;comment&quot;:pt OR &quot;comment on&quot; OR congress:pt OR &quot;cross-sectional&quot; OR &quot;dictionary&quot;:pt OR &quot;directory&quot;:pt OR &quot;editorial&quot;:pt OR &quot;festschrift&quot;:pt OR &quot;historical article&quot;:pt OR &quot;interview&quot;:pt OR &quot;lecture&quot;:pt OR &quot;legal case&quot;:pt OR &quot;legislation&quot;:pt OR letter:pt OR &quot;news&quot;:pt OR &quot;newspaper article&quot;:pt OR &quot;patient education handout&quot;:pt OR &quot;periodical index&quot;:pt OR &quot;retrospective cohort&quot; OR [(mh &quot;Animals&quot;) NOT (mh &quot;Humans&quot;)] OR rats OR cow OR cows OR chicken OR chickens OR horse OR horses OR mice OR mouse OR bovine OR sheep OR ovine OR murine OR murinae</td>
<td>65740</td>
</tr>
<tr>
<td>#12</td>
<td>#10 NOT #11</td>
<td>3714</td>
</tr>
<tr>
<td>#13</td>
<td>#12 AND (&quot;randomized controlled trial&quot;:pt OR &quot;randomized controlled trial as topic&quot;:pt OR &quot;single-blind method&quot;:pt OR &quot;double-blind method&quot;:pt OR &quot;random allocation&quot;:pt OR (random* AND control* AND trial*) OR &quot;single-blind&quot; OR &quot;double-blind&quot; OR (random* AND allocat*))</td>
<td>2993</td>
</tr>
<tr>
<td>#14</td>
<td>#12 AND [(mh &quot;Cohort Studies&quot;) OR [mh &quot;Follow-Up Studies&quot;] OR &quot;prospective cohort&quot; OR [mh &quot;prospective studies&quot;] OR (prospective* AND cohort AND (study OR studies)) OR (cohort AND (study OR studies)) OR (&quot;follow-up&quot; AND (study OR studies))]</td>
<td>372</td>
</tr>
<tr>
<td>#15</td>
<td>#14 NOT #13</td>
<td>34</td>
</tr>
</tbody>
</table>
## Cochrane Library, Interventions, 6/23/2020

RCTs = 2,134; 316 imported

Observational Studies (all are also trials): 564; 11 imported

<table>
<thead>
<tr>
<th>Search Number</th>
<th>String</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>[mh “Anorexia”] OR [mh “Anorexia Nervosa”] OR [mh “Avoidant Restrictive Food Intake Disorder”] OR [mh “Binge-Eating Disorder”] OR [mh Bulimia] OR [mh “Feeding and Eating Disorders”] OR anorex[i]:ti,ab OR “avoidant restrictive food intake disorder”:ti,ab,kw OR ARFID:ti,ab OR “binge-eating”:ti,ab OR bulimia:ti,ab OR bulimic:ti,ab OR (eating:ti,ab AND disorder*:ti,ab) OR (feed*:ti,ab AND disorder*:ti,ab) OR (food*:ti,ab AND neophobia*:ti,ab) OR “night eating”:ti,ab OR “purging disorder”:ti,ab OR EDNOS:ti,ab OR OSFED:ti,ab</td>
<td>8581</td>
</tr>
<tr>
<td>#2</td>
<td>[mh “Cognitive Behavioral Therapy”] OR [mh “Dialectical Behavior Therapy”] OR [mh “Distance Counseling”] OR [mh “Emotion-Focused Therapy”] OR [mh “Family Therapy”] OR [mh “Internet-Based Intervention”] OR [mh “Mediation”] OR [mh “Mindfulness”] OR [mh “Physical Therapy Modalities”] OR CBT:ti,ab;ab OR “cognitive behavior therapy”:ti,ab OR “cognitive behavioral therapy”:ti,ab OR “cognitive behavioral therapy”:ti,ab OR “cognitive analytic therapy”:ti,ab OR “cognitive orientation therapy”:ti,ab OR “Dialectical Behavior Therapy”:ti,ab OR e-therapy:ti,ab OR (emotion*:ti,ab AND therap*:ti,ab) OR “exposure and response prevention therapy”:ti,ab OR “Family Therapy”:ti,ab OR “group therapy”:ti,ab OR “internet-based intervention”<em>:ti,ab OR meditation:ti,ab OR “Maudsley Method”:ti,ab OR mindfulness:ti,ab OR (nutrition</em>:ti,ab AND counsel*:ti,ab) OR “physical therapy”:ti,ab OR psychotherap*:ti,ab</td>
<td>63662</td>
</tr>
<tr>
<td>#3</td>
<td>#1 AND #2</td>
<td>1666</td>
</tr>
<tr>
<td>#4</td>
<td>[mh /DT] OR [mh “Early Medical Intervention”] OR [mh Therapeutics] OR [mh /TH] OR treatment*:ti,ab OR intervention*:ti,ab OR pharmacotherap*:ti,ab</td>
<td>1070790</td>
</tr>
<tr>
<td>#5</td>
<td>#1 AND #4</td>
<td>7147</td>
</tr>
<tr>
<td>#6</td>
<td>[mh “Antidepressive Agents, Second-Generation”] OR ([mh “Antipsychotic Agents”] AND atypical) OR (antidepressant*:ti,ab AND “second-generation”:ti,ab) OR (antidepressive” AND “second-generation”:ti,ab) OR “atypical antipsychotic”*:ti,ab OR lisdexamfetamine OR topiramate</td>
<td>4560</td>
</tr>
<tr>
<td>#7</td>
<td>#1 AND #6</td>
<td>175</td>
</tr>
<tr>
<td>#8</td>
<td>#3 OR #5 OR #7</td>
<td>7248</td>
</tr>
<tr>
<td>#9</td>
<td>(#8 AND (adolescent:ti,ab OR adolescents:ti,ab OR adult:ti,ab OR adults:ti,ab OR elderly:ti,ab OR teen:ti,ab OR teens:ti,ab OR teenage:ti,ab OR teenaged:ti,ab)) OR (#8 NOT (newborn:ti,ab OR newborns:ti,ab OR infant:ti,ab OR infants:ti,ab))</td>
<td>7013</td>
</tr>
<tr>
<td>#10</td>
<td>#8 AND ((mh child) OR [mh adolescent] OR [mh adult])</td>
<td>983</td>
</tr>
<tr>
<td>#11</td>
<td>#9 OR #10</td>
<td>7017</td>
</tr>
<tr>
<td>#12</td>
<td>Address:pt OR “autobiography”:pt OR “bibliography”:pt OR “biography”:pt OR “case control” OR “case report” OR “case reports” OR “case series” OR “comment”:pt OR “comment on” OR congress:pt OR “cross-sectional” OR “dictionary”:pt OR “directory”:pt OR “editorial”:pt OR “festschrift”:pt OR “historical article”:pt OR “interview”:pt OR “lecture”:pt OR “legal case”:pt OR “legislation”:pt OR “letter”:pt OR “news”:pt OR “newspaper article”:pt OR “patient education handout”:pt OR “periodical index”:pt OR “retrospective cohort” OR ([mh “Animals”] NOT [mh “Humans”]) OR rats OR cows OR cows OR chicken OR chickens OR horse OR horses OR mice OR mouse OR bovine OR sheep OR ovine OR murine OR murinae</td>
<td>65742</td>
</tr>
<tr>
<td>#13</td>
<td>#11 NOT #12</td>
<td>6698</td>
</tr>
<tr>
<td>#14</td>
<td>“randomized controlled trial”:pt OR “randomized controlled trial as topic”:pt OR “single-blind method”:pt OR “double-blind method”:pt OR “random allocation”:pt</td>
<td>486307</td>
</tr>
<tr>
<td>#15</td>
<td>#13 AND #14</td>
<td>2134</td>
</tr>
<tr>
<td>#16</td>
<td>[mh “Cohort Studies”] OR [mh “Follow-Up Studies”] OR “prospective cohort” OR [mh “prospective studies”] OR (prospective* AND cohort AND (study OR studies))</td>
<td>156405</td>
</tr>
<tr>
<td>#17</td>
<td>#13 AND #16</td>
<td>564</td>
</tr>
</tbody>
</table>
Appendix B2. Eligibility Criteria

PsycInfo, all searches done 6/22/2020

KQ 2 Search for Diagnostic Accuracy, 6/22/2020

Results: 3,168; 2,109 imported

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>DE “Anorexia Nervosa” OR “Avoidant Restrictive Food Intake Disorder” OR DE “Binge Eating” OR DE “Binge Eating Disorder” OR DE “Bulimia” OR DE “Feeding Disorders” OR anorexia” OR “avoidant restrictive food intake disorder” OR ARFID OR “binge-eating” OR “binge eating” OR bulimia OR bulimic OR (eating AND disorder*) OR (feed* AND disorder*) OR (food* AND neophobia*) OR “night eating” OR “purging disorder” OR EDNOS OR OSFED</td>
<td>Search modes - Boolean/Phrase</td>
<td>73,015</td>
</tr>
<tr>
<td>S2</td>
<td>DE “Surveys” OR DE “Consumer Surveys” OR DE “Mail Surveys” OR DE “Online Surveys” OR DE “Questionnaires” OR DE “Telephone Surveys” OR DE “Psychiatric Evaluation” OR DE “Screening Tests” OR DE “Psychological Screening Inventory” OR questionnaire OR questionnaires OR identifying OR identification OR instrument OR instruments OR inventory OR inventories OR measure OR measures OR scale OR scales OR screen OR screening</td>
<td>Search modes - Boolean/Phrase</td>
<td>1,526,313</td>
</tr>
<tr>
<td>S3</td>
<td>S1 AND S2</td>
<td>Search modes - Boolean/Phrase</td>
<td>33,494</td>
</tr>
<tr>
<td>S4</td>
<td>“Adolescent Binge Eating Scale” OR “Eating Attitudes Test” OR “EAT-26” OR “Eating Disorder Inventory” OR “Eating Disorder Screen for Primary Care” OR “EDS-PC” OR “Primary Care Evaluation of Mental Disorders” OR “Patient Health Questionnaire” OR SCOFF OR “Sick, Control, One, Fat and Food” OR “Dutch Eating Behavior Questionnaire” OR DEBQ OR “Eating Disorder Examination” OR “Minnesota Eating Behavior Survey” OR “Patient Health Questionnaire” OR PHQ OR PRIMEMD OR “Screening for Disordered Eating”</td>
<td>Search modes - Boolean/Phrase</td>
<td>11,646</td>
</tr>
<tr>
<td>S5</td>
<td>S1 AND S4</td>
<td>Search modes - Boolean/Phrase</td>
<td>6,909</td>
</tr>
<tr>
<td>S6</td>
<td>S3 OR S5</td>
<td>Search modes - Boolean/Phrase</td>
<td>33,783</td>
</tr>
<tr>
<td>S7</td>
<td>S6</td>
<td>Limiters - English; Boolean/Phrase</td>
<td>32,049</td>
</tr>
<tr>
<td>S8</td>
<td>S7</td>
<td>Limiters - Age Groups: School Age (6-12 yrs), Adolescence (13-17 yrs), Adulthood (18 yrs &amp; older)</td>
<td>26,131</td>
</tr>
<tr>
<td>S9</td>
<td>(S7 AND (adolescent OR adolescents OR adult OR adults OR elderly OR teen OR teens OR teenage OR teenaged)) OR (S7 NOT (newborn OR newborns OR infant OR infants))</td>
<td>Search modes - Boolean/Phrase</td>
<td>30,274</td>
</tr>
<tr>
<td>S10</td>
<td>S8 OR S9</td>
<td>Expanders - Apply equivalent subjects</td>
<td>31,728</td>
</tr>
<tr>
<td>S11</td>
<td>PZ Abstract Collection OR PZ Bibliography OR PZ Column/Opinion OR PZ Comment/Reply OR PZ Editorial OR PZ Encyclopedia Entry OR PZ Letter OR PZ Obituary OR PZ Poetry OR rats OR cow OR cows OR chicken OR chickens OR horse OR horses OR mice OR mouse OR bovine OR sheep OR ovine OR murine OR murinae</td>
<td>Expanders - Apply equivalent subjects</td>
<td>465,772</td>
</tr>
<tr>
<td>S12</td>
<td>S10 NOT S11</td>
<td>Expanders - Apply equivalent subjects</td>
<td>30,783</td>
</tr>
</tbody>
</table>
## Appendix B2. Eligibility Criteria

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S13</td>
<td>S12 NOT (PO Animal NOT PO Human)</td>
<td>Expanders - Apply equivalent subjects</td>
<td>30,705</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Search modes - Boolean/Phrase</td>
<td></td>
</tr>
<tr>
<td>S14</td>
<td>“predictive value” OR sensitivity OR specificity OR accuracy OR ROC OR reproducibility OR “false positive” OR “false negative” OR “likelihood ratio”</td>
<td>Expanders - Apply equivalent subjects</td>
<td>271,247</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Search modes - Boolean/Phrase</td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>S13 AND S14</td>
<td>Expanders - Apply equivalent subjects</td>
<td>3,168</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Search modes - Boolean/Phrase</td>
<td></td>
</tr>
</tbody>
</table>

**PsycInfo, Screening, 6/22/2020**

RCTs = 1,519; **1,122** imported

Observational Studies = 2,304; **1,551** imported

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>DE “Anorexia Nervosa” OR “Avoidant Restrictive Food Intake Disorder” OR DE “Binge Eating” OR DE “Binge Eating Disorder” OR DE “Bulimia” OR DE “Feeding Disorders” OR anorexia OR “avoidant restrictive food intake disorder” OR ARFID OR “binge-eating” OR “binge eating” OR bulimia OR bulimic OR (eating AND disorder*) OR (feed* AND disorder*) OR (food* AND neophobia*) OR “night eating” OR “purging disorder” OR EDNOS OR OSFED</td>
<td>Search modes - Boolean/Phrase</td>
<td>73,015</td>
</tr>
<tr>
<td>S2</td>
<td>DE “Surveys” OR DE “Consumer Surveys” OR DE “Mail Surveys” OR DE “Online Surveys” OR DE “Questionnaires” OR DE “Telephone Surveys” OR DE “Psychiatric Evaluation” OR DE “Screening Tests” OR DE “Psychological Screening Inventory” OR questionnaire OR questionnaires OR identifying OR identification OR instrument OR instruments OR inventory OR inventories OR measure OR measures OR scale OR scales OR screen OR screening</td>
<td>Search modes - Boolean/Phrase</td>
<td>1,526,313</td>
</tr>
<tr>
<td>S3</td>
<td>S1 AND S2</td>
<td>Search modes - Boolean/Phrase</td>
<td>33,494</td>
</tr>
<tr>
<td>S4</td>
<td>“Adolescent Binge Eating Scale” OR “Eating Attitudes Test” OR “EAT-26” OR “Eating Disorder Inventory” OR “Eating Disorder Screen for Primary Care” OR “EDS-PC” OR “Primary Care Evaluation of Mental Disorders Patient Health Questionnaire” OR SCOFF OR “Sick, Control, One, Fat and Food” OR “Dutch Eating Behavior Questionnaire” OR DEBQ OR “Eating Disorder Examination” OR “Minnesota Eating Behavior Survey” OR “Patient Health Questionnaire” OR PHQ OR PRIME MD OR “Screening for Disordered Eating”</td>
<td>Search modes - Boolean/Phrase</td>
<td>11,646</td>
</tr>
<tr>
<td>S5</td>
<td>S1 AND S4</td>
<td>Search modes - Boolean/Phrase</td>
<td>6,909</td>
</tr>
<tr>
<td>S6</td>
<td>S3 OR S5</td>
<td>Search modes - Boolean/Phrase</td>
<td>33,783</td>
</tr>
<tr>
<td>S7</td>
<td>S6</td>
<td>Limiters - English; Boolean/Phrase</td>
<td>32,049</td>
</tr>
<tr>
<td>S8</td>
<td>S7</td>
<td>Limiters - Age Groups: School Age (6-12 yrs), Adolescence (13-17 yrs), Adulthood (18 yrs &amp; older)</td>
<td>Search modes - Boolean/Phrase</td>
</tr>
</tbody>
</table>
Appendix B2. Eligibility Criteria

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S9</td>
<td>(S7 AND (adolescent OR adolescents OR adult OR adults OR elderly OR teen OR teens OR teenage OR teenaged)) OR (S7 NOT (newborn OR newborns OR infant OR infants))</td>
<td>Search modes - Boolean/Phrase</td>
<td>31,460</td>
</tr>
<tr>
<td>S10</td>
<td>S8 OR S9</td>
<td>Search modes - Boolean/Phrase</td>
<td>31,728</td>
</tr>
<tr>
<td>S11</td>
<td>PZ Abstract Collection OR PZ Bibliography OR PZ Column/Opinion OR PZ Comment/Reply OR PZ Editorial OR PZ Encyclopedia Entry OR PZ Letter OR PZ Obituary OR PZ Poetry OR rats OR cow OR cows OR chicken OR chickens OR horse OR horses OR mice OR mouse OR bovine OR sheep OR ovine OR murine OR murinae</td>
<td>Search modes - Boolean/Phrase</td>
<td>465,772</td>
</tr>
<tr>
<td>S12</td>
<td>S10 NOT S11</td>
<td>Search modes - Boolean/Phrase</td>
<td>31,310</td>
</tr>
<tr>
<td>S13</td>
<td>S12 NOT (PO Animal NOT PO Human)</td>
<td>Expanders - Apply equivalent subjects</td>
<td>30,705</td>
</tr>
<tr>
<td>S14</td>
<td>DE &quot;Randomized Controlled Trials&quot; OR &quot;Single-Blind&quot; OR &quot;Double-Blind&quot; OR &quot;Random Allocation&quot; OR ((randomized OR randomised) AND controlled AND trial)</td>
<td>Search modes - Boolean/Phrase</td>
<td>78,358</td>
</tr>
<tr>
<td>S15</td>
<td>S13 AND S14</td>
<td>Search modes - Boolean/Phrase</td>
<td>1,519</td>
</tr>
<tr>
<td>S16</td>
<td>DE &quot;Observation Methods&quot; OR DE &quot;Direct Observation&quot; OR DE &quot;Prospective Studies&quot; OR DE &quot;Cohort Analysis&quot; OR &quot;observational study&quot; OR &quot;observational studies&quot; OR prospective* OR cohort*</td>
<td>Search modes - Boolean/Phrase</td>
<td>162,174</td>
</tr>
<tr>
<td>S17</td>
<td>S13 AND S16</td>
<td>Search modes - Boolean/Phrase</td>
<td>2,402</td>
</tr>
<tr>
<td>S18</td>
<td>S17 NOT S15</td>
<td>Search modes - Boolean/Phrase</td>
<td>2,304</td>
</tr>
</tbody>
</table>
### PsyInfo, Interventions, 6/22/2020

Line S17 - RCTs = 1,834; **470** imported

Line S20 - Observational Studies = 1,508; **398** imported

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>DE “Anorexia Nervosa” OR “Avoidant Restrictive Food Intake Disorder” OR DE “Binge Eating” OR DE “Binge Eating Disorder” OR DE “Bulimia” OR DE “Feeding Disorders” OR anorexi* OR “avoidant restrictive food intake disorder” OR ARFID OR “binge-eating” OR “binge eating” OR bulimia OR bulimic OR (eating AND disorder*) OR (feed* AND disorder*) OR (food* AND neophobia*) OR “night eating” OR “purging disorder” OR EDNOS OR OSFED</td>
<td>Search modes - Boolean/Phrase</td>
<td>73,015</td>
</tr>
<tr>
<td>S2</td>
<td>DE “Cognitive Behavior Therapy” OR DE “Cognitive Restructuring” OR DE “Dialectical Behavior Therapy” OR DE “Emotion Focused Therapy” OR DE “Family Therapy” OR DE “Group Psychotherapy” OR DE “Encounter Group Therapy” OR DE “Therapeutic Community” OR DE “Meditation” OR DE “Mindfulness” OR DE “Mindfulness-Based Interventions” OR DE “Online Therapy” OR DE “Physical Therapy” OR DE “Psychotherapeutic Counseling” OR DE “Psychotherapy” OR CBT OR “cognitive behavior therapy” OR “cognitive behavioral therapy” OR “cognitive analytic therapy” OR “cognitive orientation therapy” OR “Dialectical Behavior Therapy” OR “distance counseling” OR “e-therapy” OR (emotion* AND therap*) OR “exposure and response prevention therapy” OR “Family Therapy” OR “group therapy” OR “internet-based intervention” OR meditation OR “Maudsley Method” OR mindfulness OR (nutrition* AND counsel*) OR “physical therapy” OR psychotherap*</td>
<td>Search modes - Boolean/Phrase</td>
<td>392,450</td>
</tr>
<tr>
<td>S3</td>
<td>S1 AND S2</td>
<td>Search modes - Boolean/Phrase</td>
<td>12,948</td>
</tr>
<tr>
<td>S4</td>
<td>DE “Drug Therapy” OR DE “Early Intervention” OR DE “Intervention” OR DE “Prescribing (Drugs)” OR “drug therapy” OR therapeutics OR therapy OR treatment* OR intervention* OR pharmacotherap*</td>
<td>Search modes - Boolean/Phrase</td>
<td>1,383,885</td>
</tr>
<tr>
<td>S5</td>
<td>S1 AND S4</td>
<td>Search modes - Boolean/Phrase</td>
<td>38,158</td>
</tr>
<tr>
<td>S7</td>
<td>S1 AND S6</td>
<td>Search modes - Boolean/Phrase</td>
<td>693</td>
</tr>
<tr>
<td>S8</td>
<td>S1S3 OR S5 OR S7</td>
<td>Search modes - Boolean/Phrase</td>
<td>38,215</td>
</tr>
<tr>
<td>S9</td>
<td>S8</td>
<td>Limiters - English</td>
<td>35,903</td>
</tr>
<tr>
<td>S10</td>
<td>S9</td>
<td>Limiters - Age Groups: School Age (6-12 yrs), Adolescence (13-17 yrs), Adulthood (18 yrs &amp; older) Search modes - Boolean/Phrase</td>
<td>20,945</td>
</tr>
</tbody>
</table>
Appendix B2. Eligibility Criteria

<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Limiters/Expanders</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S11</td>
<td>(S9 AND (adolescent OR adolescents OR adult OR adults OR elderly OR teen OR teens OR teenage OR teenaged)) OR (S9 NOT (newborn OR newborns OR infant OR infants))</td>
<td>Search modes - Boolean/Phrase</td>
<td>35,161</td>
</tr>
<tr>
<td>S12</td>
<td>S10 OR S11</td>
<td>Search modes - Boolean/Phrase</td>
<td>35,403</td>
</tr>
<tr>
<td>S13</td>
<td>PZ Abstract Collection OR PZ Bibliography OR PZ Column/Opinion OR AB “case control” OR TI “case control” OR DE “Case Report” OR AB “case series” OR TI “case series” OR PZ Comment/Reply OR PZ Editorial OR PZ Encyclopedia Entry OR PZ Letter OR MR NONCLINICAL CASE STUDY OR PZ Obituary OR PZ Poetry OR rats OR cows OR chicken OR chickens OR horse OR horses OR mice OR mouse OR bovine OR sheep OR ovine OR murine OR murinae</td>
<td>Search modes - Boolean/Phrase</td>
<td>535,313</td>
</tr>
<tr>
<td>S14</td>
<td>S12 NOT S13</td>
<td>Search modes - Boolean/Phrase</td>
<td>31,624</td>
</tr>
<tr>
<td>S15</td>
<td>S14 NOT (PO Animal NOT PO Human)</td>
<td>Search modes - Boolean/Phrase</td>
<td>31,467</td>
</tr>
<tr>
<td>S16</td>
<td>DE “Randomized Controlled Trials” OR “Single-Blind Method” OR “Double-Blind Method” OR “Random Allocation” OR (randomized OR randomised) AND controlled AND trial</td>
<td>Search modes - Boolean/Phrase</td>
<td>70,104</td>
</tr>
<tr>
<td>S17</td>
<td>S15 AND S16</td>
<td>Search modes - Boolean/Phrase</td>
<td>1,834</td>
</tr>
<tr>
<td>S18</td>
<td>DE “Observation Methods” OR DE “Direct Observation” OR DE “Prospective Studies” OR DE “Cohort Analysis” OR observational study OR “observational studies” OR prospective* OR cohort*</td>
<td>Search modes - Boolean/Phrase</td>
<td>162,174</td>
</tr>
<tr>
<td>S19</td>
<td>S15 AND S18</td>
<td>Search modes - Boolean/Phrase</td>
<td>1,621</td>
</tr>
<tr>
<td>S20</td>
<td>S19 NOT S17</td>
<td>Search modes - Boolean/Phrase</td>
<td>1,508</td>
</tr>
</tbody>
</table>

Gray Literature Searches, 6/23/2020

ClinicalTrials.gov Searches, 6/23/2020

Screening and KQ 2 Diagnostic Accuracy Search, 6/23/2020

214 results; 214 imported to EndNote

Condition box:

“Avoidant Restrictive Food Intake Disorder” OR Bulimia OR anorexi* OR ARFID OR “binge eating” OR “binge-eating” OR bulimia OR bulimic OR (eating AND disorder*) OR (feed* AND disorder*) OR (food* AND neophobia*) OR “night eating” OR “purging disorder” OR EDNOS OR OSFED

Other box:

(identifying OR identification OR instrument OR instruments OR measure OR measures OR questionnaire OR questionnaires OR inventory OR inventories OR scale OR scales OR screen OR screening OR survey OR surveys OR “Eating Attitudes Test” OR “EAT-26” OR “EDS-PC” OR “Primary Care Evaluation of Mental Disorders Patient Health Questionnaire” OR SCOFF
Appendix B2. Eligibility Criteria

OR “Sick, Control, One, Fat and Food” OR “Dutch Eating Behavior Questionnaire” OR DEBQ OR “Minnesota Eating Behavior Survey” OR PHQ OR PRIMEMD)

214 Studies found for: (identifying OR identification OR instrument OR instruments OR measure OR measures OR questionnaire OR questionnaires OR inventory OR inventories OR scale OR scales OR screen OR screening OR survey OR surveys OR EXPAND[Concept] “Eating Attitudes Test” OR EXPAND[Concept] “EAT-26” OR EXPAND[Concept] “EDS-PC” OR EXPAND[Concept] “Primary Care Evaluation of Mental Disorders Patient Health Questionnaire” OR SCOFF OR EXPAND[Concept] “Sick, Control, One, Fat and Food” OR EXPAND[Concept] “Dutch Eating Behavior Questionnaire” OR DEBQ OR EXPAND[Concept] “Minnesota Eating Behavior Survey” OR PHQ OR PRIMEMD ) | EXPAND[Concept] “Avoidant Restrictive Food Intake Disorder” OR Bulimia OR anorexi* OR ARFID OR “binge eating” OR EXPAND[Concept] “binge-eating” OR bulimia OR bulimic OR eating AND disorder* OR feed* AND disorder* OR food* AND neophobia* OR EXPAND[Concept] “night eating” OR EXPAND[Concept] “purging disorder” OR EDNOS OR OSFED

Treatment (Interventions) Search, 6/23/2020

110 results; 30 imported to EndNote

Conditions box:

(“Avoidant Restrictive Food Intake Disorder” OR Bulimia OR anorexi* OR ARFID OR “binge eating” OR “binge-eating” OR bulimia OR bulimic OR (eating AND disorder*) OR (feed* AND disorder*) OR (food* AND neophobia*) OR “night eating” OR “purging disorder” OR EDNOS OR OSFED)

Interventions box:

(“Distance Counseling” OR CBT OR “cognitive behavior therapy” OR “cognitive behavioral therapy” OR “cognitive analytic therapy” OR “cognitive orientation therapy” OR “Dialectical Behavior Therapy” OR “e-therapy” OR (emotion* AND therap*) OR “exposure and response prevention therapy” OR “Family Therapy” OR “group therapy” OR “internet-based intervention*” OR meditation OR “Maudsley Method” OR mindfulness OR (nutrition* AND counsel*) OR “physical therapy” OR psychotherap* OR “Therapeutics” OR “therapy” OR treatment* OR intervention* OR pharmacotherap*)

110 Studies found for: EXPAND[Concept] “Avoidant Restrictive Food Intake Disorder” OR Bulimia OR anorexi* OR ARFID OR “binge eating” OR EXPAND[Concept] “binge-eating” OR bulimia OR bulimic OR eating AND disorder* OR feed* AND disorder* OR food* AND neophobia* OR EXPAND[Concept] “night eating” OR EXPAND[Concept] “purging disorder” OR EDNOS OR OSFED | “Distance Counseling” OR CBT OR “cognitive behavior therapy” OR “cognitive analytic therapy” OR “cognitive orientation therapy” OR “Dialectical Behavior Therapy” OR “e-therapy” OR emotion* AND therap* OR “exposure and response prevention therapy” OR “Family Therapy” OR “group therapy” OR “internet-based intervention*” OR meditation OR “Maudsley Method” OR mindfulness OR
Appendix B2. Eligibility Criteria

nutrition* AND counsel* OR “physical therapy” OR psychotherap* OR “Therapeutics” OR “therapy” OR treatment* OR intervention* OR pharmacotherap*

Treatment (Interventions)

Condition box:

(“Avoidant Restrictive Food Intake Disorder” OR Bulimia OR anorexi* OR ARFID OR “binge eating” OR “binge-eating” OR bulimia OR bulimic OR (eating AND disorder*) OR (feed* AND disorder*) OR (food* AND neophobia*) OR “night eating” OR “purging disorder” OR EDNOS OR OSFED)

Intervention box:

(“Distance Counseling” OR CBT OR “cognitive behavior therapy” OR “cognitive behavioral therapy” OR “cognitive analytic therapy” OR “cognitive orientation therapy” OR “Dialectical Behavior Therapy” OR “e-therapy” OR (emotion* AND therap*) OR “exposure and response prevention therapy” OR “Family Therapy” OR “group therapy” OR “internet-based intervention*” OR meditation OR “Maudsley Method” OR mindfulness OR (nutrition* AND counsel*) OR “physical therapy” OR psychotherap* OR “Therapeutics” OR “therapy” OR treatment* OR intervention* OR pharmacotherap*)
## Appendix B2. Eligibility Criteria

<table>
<thead>
<tr>
<th>Condition definition</th>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td>All KQs: Anorexia nervosa, bulimia nervosa, binge eating disorder, other specified feeding and eating disorder, and avoidant restrictive food intake disorder, based on DSM-5 criteria or other valid diagnostic criteria (e.g., DSM-IV) KQs 4, 5: Studies enrolling populations with subthreshold conditions (e.g., meeting most but not all diagnostic criteria for the disorders above, as defined by study authors) are also eligible</td>
<td>Other DSM-5 categories of eating disorders (e.g., pica, rumination disorder) or potentially unhealthy eating behaviors or syndromes not recognized by DSM-5 (e.g., orthorexia)</td>
<td></td>
</tr>
<tr>
<td>Populations KQs 1–3: Unselected or explicitly asymptomatic adolescents and adults (age ≥10 years) without signs or symptoms of an eating disorder, including populations selected for increased risk of an eating disorder (e.g., based on age, sex, race/ethnicity, gender identity, or mental health comorbidity) and populations selected based on high BMI KQs 2, 4, 5: Studies enrolling adolescents and adults (age ≥10 years) who screen positive for eating disorders in a primary care setting or are identified through other population-based screening; studies enrolling populations from specialty settings who have not been previously treated for eating disorders are also eligible</td>
<td>Studies limited to participants undergoing evaluation for bariatric surgery; studies limited to individuals who are underweight (BMI &lt;18.5 kg/m² for adults or BMI &lt;5th percentile on growth charts for age and sex for adolescents) or with other physical signs or symptoms of an eating disorder</td>
<td></td>
</tr>
<tr>
<td>Screening KQs 1–3: Screening questionnaires designed to detect eating disorders or risk of eating disorders that are feasible for use for screening in primary care (i.e., brief, easy to interpret)</td>
<td>KQs 1–3: Serologic screening (e.g., using leptin or other biomarkers)</td>
<td></td>
</tr>
<tr>
<td>Interventions KQs 4, 5: Individual, group, or family therapy (cognitive behavioral therapy or other forms of psychotherapy); pharmacotherapy with FDA-approved medications</td>
<td>KQs 4, 5: Public awareness campaigns without specific interventions linked to screening, complementary and alternative therapies, or those considered to be adjunctive therapy (e.g., acupuncture, herbal supplements, massage, light therapy)</td>
<td></td>
</tr>
<tr>
<td>Comparisons KQs 1, 3: Screened vs. nonscreened groups KQ 2: Comparison with acceptable reference standard (structured or semistructured diagnostic interview or a nonbrief [&gt;5 minutes] unstructured interview with mental health clinician) KQs 4, 5: No treatment, attention control, wait-list control, or minimal intervention (e.g., brief education about eating disorders); placebo-controlled studies of pharmacotherapy</td>
<td>KQs 4, 5: Head-to-head comparisons of two active interventions</td>
<td></td>
</tr>
<tr>
<td>Outcomes KQs 1, 4: Eating disorder remission or symptom reduction, general health-related quality of life or function, eating disorder–related quality of life or function, depression, anxiety, suicide, and mortality KQ 2: Sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and diagnostic odds ratios KQ 3: Anxiety, labeling, stigma, psychosocial harms, and false-positive results KQ 5: Any harms that result as an effect of interventions such as increased depression, increased anxiety, worsening of the eating disorder, or adverse effects from medications</td>
<td>KQs 1, 4: Screening or referral rates, attitudes about screening; intermediate outcomes (e.g., weight change, frequency of menses, frequency of specific behaviors [e.g., change in frequency of binge eating episodes]) KQ 2: Theory or survey development and validation without correlation to eating disorder outcomes, studies that focus only on particular risk factors, or assessment of provider or participant attitudes toward the instrument</td>
<td></td>
</tr>
<tr>
<td>Study designs All KQs: RCTs KQ 2: Cross-sectional and cohort studies of screening test accuracy are also eligible KQs 3, 5: Cohort studies with a concurrent control group are also eligible</td>
<td>All other study designs, including case-series and case-control studies, systematic reviews, and others</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix B2. Eligibility Criteria

<table>
<thead>
<tr>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical setting</strong></td>
<td>Settings or institutions not applicable to primary care (e.g., school classrooms, bariatric surgery centers) or inpatient/residential settings</td>
</tr>
<tr>
<td>All KQs: Primary care clinics or other settings applicable to primary care, including school-based health centers and other community settings</td>
<td></td>
</tr>
<tr>
<td>KQs 2, 4, 5: Settings referable from primary care are also eligible</td>
<td></td>
</tr>
<tr>
<td><strong>Country setting</strong></td>
<td></td>
</tr>
<tr>
<td>Research conducted in the United States or in populations similar to U.S. populations with services and interventions applicable to U.S. practice (countries categorized as &quot;very high&quot; on the Human Development Index [as defined by the United Nations Development Programme])</td>
<td>Research not relevant to the United States in countries categorized as less than &quot;very high&quot; on the Human Development Index</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td></td>
</tr>
<tr>
<td>Full text published in English</td>
<td>Non-English</td>
</tr>
<tr>
<td><strong>Quality</strong></td>
<td></td>
</tr>
<tr>
<td>Studies rated good or fair quality</td>
<td>Studies rated poor quality</td>
</tr>
</tbody>
</table>

**“Very high” on Human Development Index:** Andorra, Argentina, Australia, Austria, Bahamas, Bahrain, Barbados, Belarus, Belgium, Brunei Darussalam, Bulgaria, Canada, Chile, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong China (SAR), Hungary, Iceland, Ireland, Israel, Italy, Japan, Kazakhstan, Korea (Republic of), Kuwait, Latvia, Liechtenstein, Lithuania, Luxembourg, Malaysia, Malta, Montenegro, Netherlands, New Zealand, Norway, Oman, Poland, Portugal, Qatar, Romania, Russian Federation, Saudi Arabia, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, United States, Uruguay.

**Abbreviations:** BMI=body mass index; DSM-5=Diagnostic and Statistical Manual of Mental Disorders (5th edition); DSM-IV=Diagnostic and Statistical Manual of Mental Disorders (4th edition); FDA=U.S. Food and Drug Administration; KQ=key question; RCT=randomized, controlled trial.
Randomized, Controlled Trials and Cohort Studies

Criteria

- Initial assembly of comparable groups
- Randomized, controlled trials (RCTs)—adequate randomization, including concealment and whether potential confounders were distributed equally among groups; cohort studies—consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, crossovers, adherence, and contamination)
- Important differential loss to followup or overall high loss to followup
- Measurements that are equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: Adjustment for potential confounders for cohort studies or intention-to-treat analysis for RCTs; for cluster RCTs, correction for correlation coefficient
Appendix B3. U.S. Preventive Services Task Force Quality Rating Criteria

Definition of Ratings Based on Above Criteria

**Good:** Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (followup ≥80%); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes are considered; and appropriate attention is given to confounders in analysis. In addition, intention-to-treat analysis is used for RCTs.

**Fair:** Studies will be graded “fair” if any or all of the following problems occur, without the important limitations noted in the “poor” category below: Generally comparable groups are assembled initially, but some question remains on whether some (although not major) differences occurred in followup; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention-to-treat analysis is lacking for RCTs.

**Poor:** Studies will be graded “poor” if any of the following major limitations exist: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. Intention-to-treat analysis is lacking for RCTs.


### Diagnostic Accuracy Studies

**Criteria:**

- Screening test relevant, available for primary care, and adequately described
- Credible reference standard, performed regardless of test results
- Reference standard interpreted independently of screening test
- Indeterminate results handled in a reasonable manner
- Spectrum of patients included in study
- Sample size
- Reliable screening test
Definition of Ratings Based on Above Criteria:

**Good:** Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; assesses reliability of test; has few or handles indeterminate results in a reasonable manner; includes large number (greater than 100) of broad-spectrum patients with and without disease.

**Fair:** Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; has moderate sample size (50 to 100 subjects) and a “medium” spectrum of patients.

**Poor:** Has a fatal flaw, such as: uses inappropriate reference standard; improperly administers screening test; biased ascertainment of reference standard; has very small sample size or very narrow selected spectrum of patients.

Appendix C. Excluded Studies

X1: Non-English
X2: Ineligible condition
X3: Ineligible population
X4: Ineligible screening
X5: Ineligible intervention
X6: Ineligible comparison
X7: Ineligible outcome
X8: Ineligible clinical setting
X9: Ineligible study design
X10: Intermediate outcome only
X11: Ineligible country
X12: Not original research
X13: Abstract only
X14: Poor quality rating


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


53. Aoun A, Azzam J, Jabbour FE, et al. Validation of the Arabic version of the SCOFF questionnaire for the


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


124. Bodell LP, Keel PK. Current treatment for anorexia nervosa:
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


De Young KP, Anderson DA. An interactive, graphical tool for retrospectively assessing symptom
Appendix C. Excluded Studies


Appendix C. Excluded Studies


293. DiVasta AD, Feldman HA, Rubin CT, et al. The ability of low-magnitude mechanical signals to normalize bone turnover in adolescents hospitalized for anorexia nervosa. *Osteoporos Int.* 2017
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


341. Evans C, Dolan B. Body shape questionnaire: derivation of shortened 'alternative forms'. *Int J
Appendix C. Excluded Studies


Appendix C. Excluded Studies


375. Fitzsimmons-Craft EE, Bardone-Cone AM. One-year temporal stability and predictive and incremental validity of the body,
Appendix C. Excluded Studies


Appendix C. Excluded Studies


410. Garfinkel PE, Kline SA, Stancer HC. Treatment of anorexia nervosa using...


422. Ghaderi A, Scott B. Pure and guided self-help for full and sub-threshold
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies

efficacy and social support during treatment of binge eating disorder. *Int J Eat Disord.* 1999


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


524. Herman BK, Deal LS, Di Benedetti DB, et al. The use and value of the 7-
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


569. Innamorati M, Imperatori C, Balsamo M, et al. Food Cravings Questionnaire-Trait (FCQ-T) discriminates between obese and overweight patients with and without...
Appendix C. Excluded Studies


Appendix C. Excluded Studies

592. Joiner TE, Jr., Heatherton TF, Keel PK. Ten-year stability and predictive
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


673. Lee SW, Stewart SM, Striegel-Moore RH, et al. Validation of the eating disorder diagnostic scale for use with Hong Kong adolescents. *Int
Appendix C. Excluded Studies


684. Lewis DM, Cachelin FM. Body image, body dissatisfaction, and
Appendix C. Excluded Studies


695. Lobera IJ, Santed MA, Shafran R, et al. Psychometric properties of the Spanish version of the Thought-Shape Fusion Questionnaire. The
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies

Appendix C. Excluded Studies


Appendix C. Excluded Studies


798. Miller KK, Grieco KA, Klibanski A. Testosterone administration in women with anorexia nervosa. *J Clin
Appendix C. Excluded Studies

Exclusion Code: X3.


Appendix C. Excluded Studies

treatment of anorexia nervosa?


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


912. Pettersen G, Rosenvinge JH, Bakland M, et al. Patients' and


922. Pliatskidou S, Samakouri M, Kalamara E, et al. Validity of the Greek Eating Disorder Examination Questionnaire 6.0 (EDE-Q-6.0) among Greek adolescents. Psychiatriki. 2015 Jul-
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


952. Rausch Herscovici C. Lunch session, weight gain and their interaction with the psychopathology of anorexia nervosa in adolescents.


Appendix C. Excluded Studies


Appendix C. Excluded Studies


981. Robinson AL, Strahan E, Girz L, et al. ‘I know I can help you’: parental self-efficacy predicts adolescent outcomes in family-based therapy for
Appendix C. Excluded Studies


Appendix C. Excluded Studies


1004. Rothschild L, Stein D. Treatment monitoring: changes in affective distress and dependency following symptom alleviation of eating
Appendix C. Excluded Studies


Appendix C. Excluded Studies


1039. Sarto HM, Barcelo-Soler A, Herrera-Mercadal P, et al. Efficacy of a mindful-eating programme to reduce emotional eating in patients suffering...
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


117. Stanford SC, Lemberg R. A clinical comparison of men and women on the Eating Disorder Inventory-3
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


1150. Stice E, Trost A, Chase A. Healthy weight control and dissonance-based eating disorder prevention programs:
Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies

Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


Appendix C. Excluded Studies


<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Trial Name</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agras, 1989</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>No analyses to correct for possible bias related to attrition. Greater number of dropouts in active treatment group vs. waitlist. No description of allocation sequence concealment.</td>
</tr>
<tr>
<td>Alfonsson, 2015</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Differential missing data between treatment and outcome groups. Lack of information about allocation concealment. Some differences in baseline measures.</td>
</tr>
<tr>
<td>Alger, 1991</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Differential early termination across treatment groups. Baseline characteristics only presented for study completes (n=55), rather than entire initial sample (n=69).</td>
</tr>
<tr>
<td>Arnold, 2002</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High amount of missing data in both groups; however, different reasons for each group. Potential for missingness to be dependent on true value of outcome.</td>
</tr>
<tr>
<td>Bachar, 1999</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Randomization process not described. No table of baseline characteristics or description in text by groups randomized. For both AN/BN (randomized separately), overall attrition was 25%. Missingness may depend on true outcome and affect results. Those who withdrew were excluded, no additional analyses to assess missing data.</td>
</tr>
<tr>
<td>Barlow, 1988</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>No information about randomization process or balance at baseline. Approximately 50% population had data missing with different reasons for dropping out.</td>
</tr>
<tr>
<td>Cachelin, 2019</td>
<td>High</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>No reporting of randomization methods and suboptimal analyses to adjust for missing outcome data.</td>
</tr>
<tr>
<td>First Author, Year</td>
<td>Randomization Process Domain</td>
<td>Deviations From Intended Interventions Domain</td>
<td>Missing Outcome Data Domain</td>
<td>Outcome Measurement Domain</td>
<td>Selection of Reported Result Domain</td>
<td>Overall ROB</td>
<td>Comments</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------</td>
<td>---------------------------------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>------------------------------------</td>
<td>------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Carrard, 2011&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Overall, there were few baseline differences, except for EDI-2 body dissatisfaction. Risk of bias due to missing data. Assessors not blinded to treatment allocation. Participants and carers aware of assigned intervention as expected. Baseline assessors were not blind to treatment allocation.</td>
<td></td>
</tr>
<tr>
<td>Carter, 1998&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Lack of further description of those who dropped out by group or baseline characteristics; dropout was higher among intervention group, although measured differently. Unclear if there was a prespecified analysis plan.</td>
<td></td>
</tr>
<tr>
<td>Carter, 2003&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No information on whether there was a prespecified analysis plan.</td>
<td></td>
</tr>
<tr>
<td>Carter, 2019&lt;sup&gt;28&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>No information on the randomization process. Treatment status not blinded. Groups randomized to control rated treatment “less suitable” than those in active arm. Attrition ranged from 35-37% for posttreatment and 3-month followup. Noncompleters differed on some baseline scores compared with completers.</td>
<td></td>
</tr>
<tr>
<td>DeBar, 2011&lt;sup&gt;30&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>No information on concealment. Baseline difference by race-ethnicity. Participants were not prohibited from using treatment resources offered by the HMO throughout the study, and usual care involved advising participants at trial assignment of treatment options within the HMO. Approximately 13% had missing data. Unclear if there was a prespecified data analysis plan.</td>
<td></td>
</tr>
<tr>
<td>DeBar, 2013&lt;sup&gt;30&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No information on randomization and concealment. High amount of missing data (15%) on top of an already small sample not addressed in analyses. Unclear if there was a prespecified data analysis plan.</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duarte, 2017</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Some concerns</td>
<td>High</td>
<td>No information on randomization and concealment. Dropout was high between those who were randomized and those who actually completed posttreatment assessment. No information on whether assessors of other scales had knowledge of intervention groups.</td>
</tr>
<tr>
<td>Fairburn, 2009</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Lack of information about patient blinding, but carers were aware of assignments. Limited information on missing data. No information provided on analysis plan.</td>
</tr>
<tr>
<td>Fitzsimmons-Craft, 2020</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High loss to follow-up, and different by group assignment. Those that didn't see a benefit may have been more likely to drop out from intervention group in particular. May expect differences in drop out for those in the control based on access to other resources.</td>
</tr>
<tr>
<td>Golay, 2005</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Authors do not report on specific adverse effects or whether participants suspected they were receiving the active medication. Methods note ITT analysis, but authors excluded 18% of randomized participants who discontinued the trial prematurely and did not provide data at the final outcome assessment (24 weeks).</td>
</tr>
<tr>
<td>Goldstein, 1995</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition of 43%, and significant differential attrition (8% fluoxetine vs. 26% placebo) suggests a high risk of attrition bias.</td>
</tr>
<tr>
<td>Grant, 2019</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Risk off bias due to lack of blinding. Unclear if a prespecified plan; only 45% of overall sample had efficacy data at baseline and week 12.</td>
</tr>
<tr>
<td>Green, 2016</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Lack of information about the randomization process. Large amount of missing data with suboptimal handling with mean imputation.</td>
</tr>
<tr>
<td>First Author, Year</td>
<td>Randomization Process Domain</td>
<td>Deviations From Intended Interventions Domain</td>
<td>Missing Outcome Data Domain</td>
<td>Outcome Measurement Domain</td>
<td>Selection of Reported Result Domain</td>
<td>Overall ROB</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------</td>
<td>----------------------------</td>
<td>-------------------------------------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>Green, 2017(^{16})</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Lack of information about the randomization process. No reporting of the extent of missing data and suboptimal handling of missing data that was present with mean imputation. Lack of blinding of outcome assessment.</td>
</tr>
<tr>
<td>Green, 2018(^{16})</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>No information on randomization or allocation concealment. No table or description of whether randomized groups were similar at baseline or whether sample analyzed differed in these factors (authors note that a higher proportion randomized to control group chose not to enroll in the trial vs. intervention group). Missing data were addressed by single imputation, and additional analyses were conducted using multiple imputation. Authors noted a difference in results—findings from multiple imputations led to no statistically significant interactions between groups.</td>
</tr>
<tr>
<td>Grilo, 2005(^{40})</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Potential risk of bias due to lack of blinding, selection of reported results.</td>
</tr>
<tr>
<td>Grilo, 2013(^{41})</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Randomization not described. Most baseline characteristics are similar (age, ethnicity, comorbidity); however, participants in the placebo group had a lower rate of college education (5%) than the treatment group (14%). Overall attrition was 15%, slightly higher in placebo group than treatment arm. Analyses addressed missing data using LOCF.</td>
</tr>
<tr>
<td>Grilo, 2014(^{42})</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td></td>
</tr>
<tr>
<td>Guerdjikova, 2008(^{43})</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Most enrolled participants had at least one baseline assessment (43/44 randomized); however, 21% withdrew early, including 25% in the treatment and 17% in placebo groups. Analyses included time trends for treatment response, as well as LOCF. No other analyses were conducted to assess bias.</td>
</tr>
</tbody>
</table>
### Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guerdjikova, 200944</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Of those randomized (n=51), most (n=49) had at least one postrandomization outcome measure; however, overall attrition was 35% and was higher in the lamotrigine group vs. placebo (44% vs. 29%). ITT analysis relied on LOCF; no other analyses performed to assess bias. Those who did not complete the study may have differed in terms of ED severity.</td>
</tr>
<tr>
<td>Guerdjikova, 201245</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High overall attrition (33%), and unclear how successfully statistical methods accounted for missing data (longitudinal analysis, LOCF modified ITT). Still, low differential attrition of only 5%, and reasons for discontinuation were mostly or entirely unrelated to the true value of the study’s outcomes.</td>
</tr>
<tr>
<td>Guerdjikova, 201646</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Hedges, 200347</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition (41%), and although attrition was not significantly different between groups (12%, with 35% topiramate vs. 47% placebo), there were significantly more placebo patients than topiramate patients who dropped out because of “patient choice” (p=0.028). Unclear if this was related to severity of BN, but if so, unlikely that the study’s ITT analysis accounted for resulting attrition bias.</td>
</tr>
</tbody>
</table>
### Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hill, 2011^{48}</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Authors conducted ITT analysis that included all participants randomized; missing data were imputed based on an estimate of mean values from the sample. This may not address bias if those with missing data had worse outcomes; however, attrition was relatively low (13%) and not differential. Not clear if results were analyzed based on prespecified plan that was finalized before unblinded data were available.</td>
</tr>
<tr>
<td>Hoopes, 2003^{49}</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition of 42%. Only ITT analyses were used to account for potential attrition bias. Relatively low differential attrition overall (10%), but a significantly greater N of patients discontinued placebo than topiramate because of “patient choice” (7 vs. 1, respectively). Unclear if allocation concealment was used.</td>
</tr>
<tr>
<td>Horne, 1988^{50}</td>
<td>High</td>
<td>Some concerns</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Baseline differences based on baseline frequency of binge eating; results focused on completers, but reanalysis of data showed a greater impact of intervention. Measures were self-reported. ITT was not used. High attrition. No evidence that there was bias based on missing data. No information on why participants dropped out. Unclear if there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>Hudson, 1998^{51}</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Potential bias from overall attrition (21%) and significant differential attrition with more fluvoxamine patients than placebo patients dropping out for adverse medical events (12% vs. 0%, respectively, ( p=0.03 )) or any reason (31% vs. 12%, respectively, ( p=0.04 )). Unclear that the study’s modified ITT and random regression analyses were successful at managing potential attrition bias.</td>
</tr>
</tbody>
</table>

KQ4
<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hughes, 1986\textsuperscript{52}</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Unclear information about missing data. Deviations from the intended intervention.</td>
</tr>
<tr>
<td>Jacobi, 2012\textsuperscript{53}</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Average adherence across all intervention components was 66.2% and varied substantially by component. Study focused on completer's analysis only. Lack of reporting about allocation concealment or whether there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>Kanerva, 1995\textsuperscript{54}</td>
<td>Some concerns</td>
<td>High</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No information on randomization and allocation sequence. Few baseline characteristics reported. Placebo group had higher total EDI score (80.5 vs. 69.4) than fluoxetine group at baseline. No analyses to address missing data; however, overall attrition was relatively low (8%) and did not differ between groups.</td>
</tr>
<tr>
<td>Kelly, 2014\textsuperscript{55}</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Unclear whether assessors are blinded. Unclear if there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>Laederach-Hofmann, 1999\textsuperscript{56}</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Participants randomized to placebo vs. imipramine were slightly older (mean age 36 vs. 41 years) and had a significantly higher BMI and body weight (authors attributed this to two participants in the placebo group with a very high BMI/weight). Data appear to be available for nearly all participants, except for two who discontinued treatment (6.8%, one from each group).</td>
</tr>
<tr>
<td>Laessle, 1987\textsuperscript{57}</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Limited information about analyses or dropout rates. Small sample so high likelihood of missingness being dependent on true value of outcome. Limited selection of reported results.</td>
</tr>
<tr>
<td>Ljotsson, 2007\textsuperscript{58}</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>All eligible outcomes were self-reported or self-conveyed to an assessor, and all patients knew what their assigned condition was.</td>
</tr>
</tbody>
</table>
### Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linardon, 2020[60]</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High dropout rates, differences in follow-up by group and baseline characteristics, and changes to the analysis make this high. Also of note, have those with subthreshold but not OSFED. Also have 44% of the total sample with prior ED treatment and 16% with current ED treatment, not sure if these overlap or not, but close to the 50% with prior treatment. May be excluded for several reasons</td>
</tr>
<tr>
<td>Masson, 2013[61]</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>No information on concealment. High attrition, but reasons provided. No information on whether there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>McCann, 1990[61]</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>No ITT analysis. Overall attrition was 23% and was higher among the intervention vs. control group (33% vs. 13%). Completers analysis only. No analyses to address missing data.</td>
</tr>
<tr>
<td>McElroy, 2000[62]</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Potential bias from overall attrition (24%).</td>
</tr>
<tr>
<td>McElroy, 2003[63]</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>The small sample size makes the study especially susceptible to attrition bias, even from a moderate amount of attrition (18% overall). Possible that the placebo group’s 10% higher rate of withdrawal due to worsening depression may have affected the statistical significance of outcomes, especially given that the placebo group had a significantly worse baseline CGI-S score than the citalopram group. Also unclear why completers analysis results were only reported for binge response, but not any other outcome.</td>
</tr>
<tr>
<td>McElroy, 2003[64]</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Despite overall attrition of 43%, authors reported adverse events among all those randomized.</td>
</tr>
<tr>
<td>First Author, Year</td>
<td>Randomization Process Domain</td>
<td>Deviations From Intended Interventions Domain</td>
<td>Missing Outcome Data Domain</td>
<td>Outcome Measurement Domain</td>
<td>Selection of Reported Result Domain</td>
<td>Overall ROB</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------</td>
<td>---------------------------</td>
<td>-------------------------------------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>McElroy, 2006 $^6^5$</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition (50%) and differential attrition (20%) with more zonisamide patients withdrawing than placebo patients (60% vs. 40%, respectively). Authors acknowledged their analyses may not have accounted for data missing not at random, which means BED severity or other outcomes measured may have contributed to problematic attrition bias.</td>
</tr>
<tr>
<td>McElroy, 2007 $^6^6$</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>No description of randomization. Of those randomized, all but one had at least one postrandomization outcome measure and were included in analyses; however, 37% withdrew early (30% in treatment group and 45% in placebo group). No other analyses to assess bias or assumption that those who withdrew had worse outcomes.</td>
</tr>
<tr>
<td>McElroy, 2007 $^6^7$</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Overall, 30% of participants randomized did not complete the study (no differential attrition). Authors described a &quot;repeated measures&quot; random regression model to account for missing data, which implies imputation of missing data values as well as a two-way analysis that implies LOCF. However, assumptions made in models about missing data values are not clear and do not provide clear evidence that results were not biased by missing data.</td>
</tr>
</tbody>
</table>
### Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year, Trial Name</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>McElroy, 2011</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall (40%) and differential (34%) attrition with significantly more placebo than acamprosate patients dropping out (55% vs. 21%, respectively). As a result, the study’s analyses relied heavily on assumptions about missing data underlying the study’s random regression and LOCF ITT analyses. More placebo patients than acamprosate patients withdrew due to lack of efficacy (4/11 vs. 0/5) or were lost to followup for unknown reasons (5/11 vs. 1/5). Likely that the placebo group’s higher attrition was related to the severity of BED and that the study’s analyses did not account for the resulting bias.</td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No assessment of whether baseline characteristics differed. While it does mention those who dropped out for adverse events, details of this are not included in this study but rather in the protocol paper.</td>
</tr>
<tr>
<td>McElroy, 2016</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No statistical assessments of differences at baseline. Unclear if there was prespecified statistical analysis or not.</td>
</tr>
<tr>
<td>McElroy, 2016</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Lack of sensitivity analyses for secondary outcomes of interest</td>
</tr>
<tr>
<td>Mitchell, 1990</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Randomization process altered periodically based on attrition. No information on baseline differences. High attrition and analyses focused on completers, no ITT or analyses to address missing data. No information on reasons for missingness. Unclear if there was a prespecified analysis plan.</td>
</tr>
</tbody>
</table>

<p>| Keel, 2002                   | High                           | High                                        | High                        | Some concerns                | Some concerns                     | High        |          |</p>
<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitchell, 2001⁷⁶</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No information about allocation concealment or whether patient, provider, or outcome assessor blinding of treatment assignment was used, but patients were probably blinded to their assigned drug treatments. Unclear if assessors of clinician-measured outcomes were blinded to treatment assignment. Relatively low risk of bias from overall attrition (9%), despite lack of detail about differential attrition or reasons for dropout. Appears that only a completers analysis was used to analyze EDI scores, HAM-D-21 scores, laxative abuse, diuretic abuse, and fasting days.</td>
</tr>
<tr>
<td>Fluoxetine Bulimia Nervosa Collaborative Study Group, 1992⁷⁷</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Lack of information about the randomization process. Unclear whether outcome assessors were blinded.</td>
</tr>
<tr>
<td>Pearlstein, 2003⁷⁸</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No ITT analysis: of 25 who met criteria, authors noted that 20 completed the protocol. Unclear if this means they did not complete the study (after randomization) or the intake assessment.</td>
</tr>
<tr>
<td>Peterson, 1998⁷⁹</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>No information on randomization or allocation concealment. Participants were assigned to groups, and then four groups were randomized to treatment. Groups differed in terms of baseline frequency of objective and subjective binge-eating episodes. High risk of bias due to attrition; 16% of active treatment groups and 18% of the wait-list group did not complete the protocol. Analysis conducted using last observation but no other analyses to address bias.</td>
</tr>
</tbody>
</table>
## Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peterson, 2011</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High risk of attrition bias affecting the self-help group (40% overall, 21% differential vs. waitlist), the therapist-assisted group (32% overall, 13% differential vs. waitlist), and to a lesser extent, the therapist-led group (12% overall, 7% differential vs. waitlist). Reasons for dropout not provided, which makes it impossible to determine if attrition was related to BED severity or other outcomes of interest. However, there was a notable pattern in attrition rates among the active groups.</td>
</tr>
<tr>
<td>Pope, 1989</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Randomization process not described. No ITT analysis: those who did not complete 4 weeks of treatment were excluded from analysis (10% of those randomized). An additional five participants (12%) withdrew early for various reasons—some related to factors associated with outcomes being measured (no improvement in ED symptoms, hospitalization for mental health comorbidity). Authors used LOCF in analyses for subjects who withdrew after week 4, which may not address bias.</td>
</tr>
<tr>
<td>Robinson, 2008</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition (37%). Differential attrition for any reason not statistically significant between groups (21% for therapy vs. wait-list comparison and 12% for therapy vs. self-directed writing), but no data available on specific reasons for attrition that might have indicated missingness based on ED severity. Similar results from ITT and completers analyses do not rule out the possibility that attrition bias had a major impact on the findings. Lack of allocation concealment because the co-investigator who randomized participants may have also been an outcome assessor aware of treatment assignments.</td>
</tr>
</tbody>
</table>
### Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanchez-Ortiz, 2010</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Schag, 2019</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Overall, 16-22% did not complete study or had missing data across both time points. ITT analysis using imputed values for missing data was similar to per-protocol analyses. No differential attrition.</td>
</tr>
<tr>
<td>Schlup, 2009</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No information on concealment. No information on baseline differences. Assessors were not blinded. Unclear if there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>Schmidt, 2008</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Minor differences in baseline characteristics. Overall, approximately 17% of sample did not have 3-month assessment, but no differential attrition. Authors conducted analyses that assess whether persons with missing data were associated with baseline variables. No imputation of missing data.</td>
</tr>
<tr>
<td>Stice, 2019</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Authors reported using random numbers table to randomize participants; however, the sample (n=12, or 12%) was not randomized, but rather assigned to the wait-list because it was not possible to implement groups due to a holiday. Some minor differences between groups in proportion that were Hispanic and proportion that met criteria for full or subthreshold AN (not statistically significant). Rates of missing data were 9%-27% at posttest; authors described using imputation for MRI-related outcomes. Proportion with missing data for self-reported eligible health outcomes and how this was addressed are not clear.</td>
</tr>
</tbody>
</table>
### Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year, Trial Name</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sundbald, 2015</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>No information on randomization process, concealment, or baseline values between groups. High attrition (21%), but reasons for dropouts provided. No information on whether there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>Telch, 1990</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Lack of information on randomization or concealment process. Baseline measures appear to be similar but no real information. Carers and participants aware, as expected. Measures were self-reported. ITT was not used. Differential attrition between groups but provided reasons. Unclear if there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>Telch, 2001</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition (23%) in this relatively small sample of 44; only those who completed study were analyzed. No detail given about reasons for attrition. Lack of information about allocation concealment.</td>
</tr>
<tr>
<td>Traviss, 2011</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>No information on randomization and allocation sequence. Approximately 30% who completed baseline assessment were lost to followup. Analyses primarily based on last observation carried forward. Participants with worse outcomes or continued disordered eating may have been less likely to continue in the trial.</td>
</tr>
<tr>
<td>Wade, 2017</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Missing information about randomization process and outcome measurement.</td>
<td></td>
</tr>
<tr>
<td>Wagner, 2016</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Deviations from intended interventions. High amount of missing outcome data.</td>
</tr>
<tr>
<td>Walsh, 1985</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Limited information on randomization and allocation process. High attrition. Deviations from the intended intervention.</td>
</tr>
<tr>
<td>Walsh, 1987</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>
Appendix D Table 1. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walsh, 1991&lt;sup&gt;96&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Potential bias from overall attrition (19%), but relatively low differential attrition (7%). Reasons for attrition were generally similar. Only a single patient (1/38, or 2.6%) withdrew from the placebo group because of lack of efficacy. LOCF ITT used to manage potential attrition bias.</td>
</tr>
<tr>
<td>Walsh, 2004&lt;sup&gt;97&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Lack of information on randomization or concealment process. High attrition used LOCF for missing data. Unclear if there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>White, 2013&lt;sup&gt;98&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Wilfley, 1993&lt;sup&gt;99&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No information on randomization and allocation process. Assessors were not blinded. Unclear if there was a prespecified analysis plan.</td>
</tr>
</tbody>
</table>

Abbreviations: AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CGI-S=Clinical Global Impression-Severity; ED=eating disorder; EDI=Eating Disorder Inventory; HAM-D-21=Hamilton Depression Rating Scale-21 Item; HMO=health maintenance organization; ITT=intent-to-treat; KQ=key question; LOCF=last observation carried forward; MRI=magnetic resonance imaging; NA=not available; ROB=risk of bias; vs.=versus.
<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnold, 2002&lt;sup&gt;21&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>High amount of missing data in both groups, but different reasons for each group. High likelihood of missingness being dependent on true value of outcome.</td>
</tr>
<tr>
<td>Goldstein, 1995&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition of 43%, and significant differential attrition (8% fluoxetine vs. 26% placebo) for lack of efficacy suggests a high risk of attrition bias.</td>
</tr>
<tr>
<td>Grant, 2019&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some self-report and assuming assessors were blind but unclear; unclear if a prespecified plan.</td>
</tr>
<tr>
<td>Guerdjikova, 2008&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Most enrolled participants had at least one baseline assessment (43/44 randomized); however, overall attrition was 35% and was higher in the lamotrigine group vs. placebo (44% vs. 29%). ITT analysis relied on LOCF; no other analyses were conducted to assess bias.</td>
</tr>
<tr>
<td>Guerdjikova, 2009&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Of those randomized (n=51), most (n=49) had at least one post randomization outcome measure; however, overall attrition was 35% and was higher in the lamotrigine group vs. placebo (44% vs. 29%). ITT analysis relied on LOCF; no other analyses performed to assess bias. Those who did not complete the study may have differed in terms of ED severity.</td>
</tr>
<tr>
<td>Guerdjikova, 2012&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High overall attrition (33%), and unclear how successfully statistical methods accounted for missing data (longitudinal analysis, LOCF modified ITT). Still, low differential attrition of only 5%, and reasons for discontinuation were mostly or entirely unrelated to the true value of the study’s outcomes.</td>
</tr>
<tr>
<td>Guerdjikova, 2016&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>First Author, Year</td>
<td>Randomization Process Domain</td>
<td>Deviations From Intended Interventions Domain</td>
<td>Missing Outcome Data Domain</td>
<td>Outcome Measurement Domain</td>
<td>Selection of Reported Result Domain</td>
<td>Overall ROB</td>
<td>Comments</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------</td>
<td>---------------------------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td>-----------------------------------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>Hedges, 2003⁷⁷</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Substantial overall attrition (41%), and although differential attrition was significantly different between groups (12%, with 35% topiramate vs. 47% placebo), there were significantly more placebo patients than topiramate patients who dropped out because of “patient choice” (p=0.028). Unclear if this was related to severity of BN, but if so, unlikely that the study’s ITT analysis accounted for resulting attrition bias.</td>
</tr>
<tr>
<td>Hoopes, 2003⁴⁰</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition of 42%. Only ITT analyses were used to account for potential attrition bias. Relatively low differential attrition overall (10%), but a significantly greater N of patients discontinued placebo than topiramate because of “patient choice” (7 vs. 1, respectively). Also unclear if allocation concealment was used.</td>
</tr>
<tr>
<td>Horne, 1988⁴⁰</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Differences based on baseline frequency of binge eating; focused on completers but reanalysis of data showed a greater impact of intervention. Measures were self-reported. ITT was not used. High attrition. No evidence that there was bias based on missing data. No information on why participants dropped out. Unclear if there was a prespecified analysis plan.</td>
</tr>
<tr>
<td>Hudson, 1998⁵¹</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Potential bias from overall attrition (21%) and significant differential attrition with more fluvoxamine patients than placebo patients dropping out for adverse medical events (12% vs. 0%, respectively, p=0.03) or any reason (31% vs. 12%, respectively, p=0.04). Unclear that the study’s modified ITT and random regression analyses were successful at managing potential attrition bias.</td>
</tr>
</tbody>
</table>
## Appendix D Table 2. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 5)

<table>
<thead>
<tr>
<th>First Author, Year Trial Name</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>McElroy, 2003&lt;sup&gt;64&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Despite overall attrition of 43%, authors reported adverse events among all those randomized.</td>
</tr>
<tr>
<td>McElroy, 2006&lt;sup&gt;65&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>High overall attrition (50%) and differential attrition (20%) with more zonisamide patients withdrawing than placebo patients (60% vs. 40%, respectively). Authors acknowledged their analyses may not have accounted for data missing not at random, which means BED severity or other outcomes measured may have contributed to problematic attrition bias.</td>
</tr>
<tr>
<td>McElroy, 2007&lt;sup&gt;67&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Overall, 30% of participants randomized did not complete the study (no differential attrition). Authors described a &quot;repeated measures&quot; random regression model to account for missing data, which implies imputation of missing data values as well as a two-way analysis that implies LOCF. However, assumptions made in models about missing data values are not clear and do not provide clear evidence that results were not biased by missing data.</td>
</tr>
<tr>
<td>McElroy, 2011&lt;sup&gt;68&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Substantial overall (40%) and differential (34%) attrition with significantly more placebo than acamprosate patients dropping out (55% vs. 21%, respectively). As a result, the study’s analyses relied heavily on assumptions about missing data underlying the study’s random regression and LOCF ITT analyses. More placebo patients than acamprosate patients withdrew because of lack of efficacy (4/11 vs. 0/5) or were lost to followup for unknown reasons (5/11 vs. 1/5). Likely that the placebo group’s higher attrition was related to the severity of BED and that the study’s analyses did not account for the resulting bias.</td>
</tr>
</tbody>
</table>
### Appendix D Table 2. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 5)

<table>
<thead>
<tr>
<th>First Author, Year Trial Name</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>McElroy, 2015&lt;sup&gt;69&lt;/sup&gt; McElroy, 2016&lt;sup&gt;70&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No statistical assessment of baseline characteristics. While it does mention those who dropped out for adverse events, details of this are not included in this study, but rather the protocol paper.</td>
</tr>
<tr>
<td>McElroy, 2016&lt;sup&gt;71&lt;/sup&gt; Sheehan, 2017&lt;sup&gt;72&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No statistical assessments of differences at baseline. Unclear if there was prespecified statistical analysis or not.</td>
</tr>
<tr>
<td>Milano, 2004&lt;sup&gt;100&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Some concerns</td>
<td>High</td>
<td>No information on randomization or allocation process. No information on baseline characteristics or if there were differences. No information on if ITT was used. No information on whether participants and carers were aware. No information on missing data. No information on how outcomes were measured. No information on whether assessors were blind. Unclear if there was a prespecified analysis plan</td>
</tr>
<tr>
<td>Milano, 2005&lt;sup&gt;101&lt;/sup&gt;</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>High</td>
<td>Small sample size (n=12). Lack of numerical results of baseline demographic information. Participants and carers aware of assigned intervention as expected. No description in methods related to whether certain adverse effects were prespecified or how they were assessed.</td>
</tr>
<tr>
<td>Fluoxetine Bulimia Nervosa Collaborative Study Group, 1992&lt;sup&gt;73&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Lack of information about the randomization process. Unclear whether outcome assessors were blinded.</td>
</tr>
<tr>
<td>Pearlstein, 2003&lt;sup&gt;78&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>No ITT analysis: of 25 who met criteria, authors noted that 20 completed the protocol. Unclear if this means they did not complete the study (after randomization) or the intake assessment.</td>
</tr>
</tbody>
</table>
### Appendix D Table 2. Quality Assessment of Parallel Randomized, Controlled Trials (KQ 5)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pope, 1989&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Some concerns</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td>Randomization process not described. No ITT analysis: those who did not complete 4 weeks of treatment were excluded from analysis (10% of those randomized). An additional five participants (12%) withdrew early for various reasons—some related to factors associated with outcomes being measured (no improvement in ED symptoms, hospitalization for mental health comorbidity). Authors used LOCF in analyses for subjects who withdrew after week 4, which may not address bias.</td>
</tr>
<tr>
<td>Walsh, 1991&lt;sup&gt;16&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Some concerns</td>
<td>Low</td>
<td>Some concerns</td>
<td>Some concerns</td>
<td>Potential bias from overall attrition (19%), but relatively low differential attrition (7%). Reasons for attrition were generally similar. Only a single patient (1/38, or 2.6%) withdrew from the placebo group because of lack of efficacy. LOCF ITT used to manage potential attrition bias.</td>
</tr>
</tbody>
</table>

**Abbreviations:** BED=binge-eating disorder; BN=bulimia nervosa; ED=eating disorder; ITT=intent-to-treat; KQ=key question; LOCF=last observation carried forward; N=number; NA=not available; ROB=risk of bias; vs.=versus.
# Appendix D Table 3. Quality Assessment of Crossover Randomized, Controlled Trials (KQ 4)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Trial Name</th>
<th>Outcome of Interest</th>
<th>Randomization Process Domain</th>
<th>Deviations From Intended Interventions Domain</th>
<th>Missing Outcome Data Domain</th>
<th>Outcome Measurement Domain</th>
<th>Selection of Reported Result Domain</th>
<th>Overall ROB</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corwin, 2012102</td>
<td></td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Study drug is associated with known drowsiness, fatigue, itching, and other adverse effects. There was sufficient time for carry-over effects to have disappeared in terms of blood levels of the active medication, but changes in disordered eating behavior may not have disappeared or returned to baseline.</td>
</tr>
<tr>
<td>Safer, 2019103</td>
<td></td>
<td>Some concerns</td>
<td>High</td>
<td>Some concerns</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lack of information on randomization and allocation concealment.</td>
</tr>
<tr>
<td>Mitchell, 1988104</td>
<td></td>
<td>Some concerns</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Some concerns</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No information on whether analysis plan was prespecified. Small sample size and all withdrawals were from the placebo group. No information on randomization or concealment. No information on whether there were baseline differences.</td>
</tr>
</tbody>
</table>

**Abbreviation:** ROB=risk of bias.
### Appendix D Table 4. Quality Assessment of Screening Test Accuracy Studies (KQ 2)

<table>
<thead>
<tr>
<th>First Author, Year</th>
<th>Index Test</th>
<th>Bias due to patient selection?</th>
<th>Bias due to index test?</th>
<th>Bias due to the reference standard?</th>
<th>Bias due to flow and timing?</th>
<th>Quality Rating</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anstine, 2000&lt;sup&gt;105&lt;/sup&gt;</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Poor</td>
<td>Study assessed accuracy of five questions vs. EAT-26, but analysis grouped participants into high- vs. low-risk groups based on reference standard score and used correlation of index test responses to EAT-26 results to determine threshold for positive response. Only four of five questions were correlated, and accuracy results are only provided for these four questions separately (not for the four or five index questions as a bundle).</td>
<td></td>
</tr>
<tr>
<td>Chamay-Weber, 2017&lt;sup&gt;106&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotton, 2003&lt;sup&gt;107&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dorflinger, 2017&lt;sup&gt;108&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duarte, 2015&lt;sup&gt;109&lt;/sup&gt;</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Poor</td>
<td>The only ones who received the reference standards were those in a small subsample recruited from the general sample, and details about the timing were unclear.</td>
<td></td>
</tr>
<tr>
<td>Franklin, 2019&lt;sup&gt;110&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>Poor</td>
<td>Only participants who had elevated screening results at intake on one or more questionnaires (depression, anxiety, eating, or sleeping) or needed social work assistance were given the reference test.</td>
<td></td>
</tr>
<tr>
<td>Freund, 1999&lt;sup&gt;111&lt;/sup&gt;</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
<td>High</td>
<td>Poor</td>
<td>Only a subset of those who completed the screening test consented to participate in the diagnostic interview. The timing between the screening test and reference standard is not clear (methods indicate it may have been as long as 2 years before the reference standard interview).</td>
<td></td>
</tr>
<tr>
<td>Garcia, 2010&lt;sup&gt;112&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graham, 2019&lt;sup&gt;113&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hill, 2010&lt;sup&gt;114&lt;/sup&gt;</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lähteenmäki, 2009&lt;sup&gt;115&lt;/sup&gt;</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Fair</td>
<td>Study was based on a large population cohort study. Of those completing screening questionnaires, a subset of participants was selected to participate in the accuracy study (see patient selection domain). Of those who completed the screening questionnaire and were invited to participate, only 55% agreed. However, authors noted that mean SCOFF scores were similar among participants and nonparticipants.</td>
<td></td>
</tr>
<tr>
<td>Lui, 2015&lt;sup&gt;116&lt;/sup&gt;</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Screening for Eating Disorders 228 RTI-UNC EPC
### Appendix D Table 4. Quality Assessment of Screening Test Accuracy Studies (KQ 2)

<table>
<thead>
<tr>
<th>First Author, Year Index Test</th>
<th>Bias due to patient selection?</th>
<th>Bias due to index test?</th>
<th>Bias due to the reference standard?</th>
<th>Bias due to flow and timing?</th>
<th>Quality Rating</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maugen, 2018 EDS-PC</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Fair</td>
<td>No description of missing data or proportion of participants who did not respond to mailed surveys.</td>
</tr>
<tr>
<td>Maugen, 2018 SCOFF, SDE</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Fair</td>
<td>No description of missing data or proportion of participants who did not respond to mailed surveys.</td>
</tr>
<tr>
<td>Mond, 2008 SCOFF</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Fair</td>
<td>Unclear whether results of index and reference standard interpreted independently. Potential selection bias due to flow/timing: of those completing index test, 33% did not participate in diagnostic interview. Authors stated that participants and nonparticipants did not differ significantly on study variables but did not provide this data or whether participants and nonparticipants differed in terms of index test scores.</td>
</tr>
<tr>
<td>Muro-Sans, 2008 SCOFF-c</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Parker, 2005 SCOFF</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td>Unclear if diagnostic interviews were done before patients completed the BES and whether clinicians were blind to BES results. Limited information across multiple domains makes a good quality rating difficult.</td>
</tr>
<tr>
<td>Ricca, 2000 BES</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Low</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Rosenvinge, 2001 EDS-5</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Fair</td>
<td>Possible selection bias because two groups of 46 students each from a larger sample of 835 students (N=92/835, or 11%) were invited to participate based on having EDS-5 scores in the “upper” or “lower” range. This is different than a random or consecutive selection process. Unclear if any inappropriate exclusions took place.</td>
</tr>
<tr>
<td>Siervo, 2005 SCOFF</td>
<td>Unclear</td>
<td>Low</td>
<td>Unclear</td>
<td>Low</td>
<td>Fair</td>
<td>Some potential for selection bias because nine patients (4.8% of 189 total who underwent “initial screening”) with AN, restrictive EDNOS, and bulimia were excluded because of small numbers in each diagnostic category, even though all included EDs were lumped together to evaluate SCOFF test accuracy. Some risk of bias because unclear if the reference standard—an unstructured DSM-IV clinical interview—led to any misclassification of EDs. Unclear if diagnostic interviews were done before patients completed the SCOFF and whether clinicians were blind to SCOFF results.</td>
</tr>
<tr>
<td>Solmi, 2015 SCOFF</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Poor</td>
<td>Selection of index test sample was nonrandom, attrition was high, and details about the timing were unclear.</td>
</tr>
<tr>
<td>Striegel-Moore, 2010 SCOFF</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Unclear</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Wan Wahida, 2012 SCOFF</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Good</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** ADO-BES=Adolescent Binge Eating Scale; AN=anorexia nervosa; BES=Binge Eating Scale; DSM-IV=Statistical Manual of Mental Disorders, Version 4; EAT-26=Eating Attitudes Test-26; ED=eating disorder; EDNOS=Eating Disorder Not Otherwise Specified; EDS-5=Eating Disturbance Scale-5; EDS-PC=Eating Disturbance Scale for Primary Care; ESP=Eating disorder Scale for Primary Care; KQ=key question; SDE=Screen for Disordered Eating; SWED=Stanford-Washington University Eating Disorder screen; VA-BES=Veterans Administration Binge Eating Scale; vs.=versus.
<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Description of Intervention(s)</th>
<th>Intensity of Intervention (No. and Length of Sessions)</th>
<th>Duration (Weeks)</th>
<th>Recruitment Setting Country</th>
<th>Population</th>
<th>N</th>
<th>Proportion With Comorbidity Psychiatric Disorder (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfonsson, 2015&lt;sup&gt;19&lt;/sup&gt; Fair</td>
<td>Group psychotherapy (Behavioral activation)</td>
<td>10 weekly 90-minute group sessions</td>
<td>10</td>
<td>Obesity clinic providing both behavioral interventions and bariatric surgery Sweden</td>
<td>Adults with BED (DSM-5) and obesity (BMI&gt;30) presenting to an initial assessment at an outpatient obesity clinic</td>
<td>96</td>
<td>NR</td>
</tr>
<tr>
<td>Cachelin, 2019&lt;sup&gt;24&lt;/sup&gt; Fair</td>
<td>Culturally adapted CBT-based guided self-help</td>
<td>Eight guided or supported sessions (each session 25 min in duration); four weekly, then four biweekly sessions</td>
<td>12</td>
<td>Advertisements for a study on overeating or binge eating placed in community health and mental health settings, local organizations (e.g., churches, markets, laundromats), and at an urban university campus</td>
<td>Adult Latinas (18-55 y) who met criteria for BED and had a BMI ≥18 mg/kg, responding to trial advertisement</td>
<td>40</td>
<td>NR</td>
</tr>
<tr>
<td>Carter, 2003&lt;sup&gt;37&lt;/sup&gt; Fair</td>
<td>Unguided CBT-based self-help (based on Overcoming Binge Eating manual)</td>
<td>Provision of manual only; participants encouraged to read the book and follow the advice contained in it over the following 8 weeks</td>
<td>8</td>
<td>Patients recruited from wait-list for treatment at ED clinic in a hospital, they met BN diagnostic criteria and were seeking specialized treatment for the first time Canada</td>
<td>Women (≥17 y) with BN (DSM-IV), but inclusive of those with one binge-eating episode and compensatory behaviors per week vs. two) and BMI ≥18, recruited from a wait-list of patients referred for outpatient ED treatment</td>
<td>85</td>
<td>NR</td>
</tr>
<tr>
<td>Carter, 2020&lt;sup&gt;23&lt;/sup&gt; Fair</td>
<td>DBT self-help, guided and unguided</td>
<td>Both groups provided with a self-help manual, The DBT Solution for Emotional Eating. The guided self-help group also received six 30-min video sessions to provide support (weekly X 2 weeks, biweekly X 3 weeks, and final session at 12 weeks)</td>
<td>12</td>
<td>Advertisements placed in universities, hospitals, and medical clinics as well as on social media, local radio station websites, and newspapers</td>
<td>Adults (19-65 y) with BED (DSM-5) and BMI of ≥18.5, recruited from the community or health centers via advertisements</td>
<td>71</td>
<td>NR</td>
</tr>
</tbody>
</table>
## Appendix E Table 1. Detailed Evidence Table of Psychological Intervention Characteristics (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Quality</th>
<th>Description of Intervention (s)</th>
<th>Intensity of Intervention (No. and Length of Sessions)</th>
<th>Duration (Weeks)</th>
<th>Recruitment Setting</th>
<th>Country</th>
<th>Population</th>
<th>Proportion With Comorbidity Psychiatric Disorder (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeBar, 2013</td>
<td>Fair</td>
<td>Individual CBT (adolescent specific)</td>
<td>Eight individual sessions scheduled at participants’ general medical clinics at the family’s convenience (e.g., after school), phone sessions when in-person attendance was not possible; supplemental sessions offered when appropriate to address mood and interpersonal relationships</td>
<td>24</td>
<td>Adolescents enrolled in an HMO via advertisement and invitations to patients identified as having an incident case of a binge eating–related disorder via monitoring of an EMR</td>
<td>United States</td>
<td>Female adolescents (12-18 y) enrolled in an HMO with BN or BED and at least one binge-eating episode over the previous 3 months (DSM-IV)</td>
<td>25</td>
</tr>
<tr>
<td>Fairburn, 2009</td>
<td>Fair</td>
<td>Two forms of Enhanced (transdiagnostic) CBT: focused form targeting ED pathology only and broad form also addressing other common problems with ED</td>
<td>Both treatment groups received 20 50-min sessions preceded by one 90-min preparatory session and followed by a review session 20 weeks after treatment</td>
<td>20</td>
<td>Referrals to two outpatient ED clinics by family doctors and other clinicians</td>
<td>United Kingdom</td>
<td>Adults (18-65 y) with any ED (DSM-IV) requiring treatment (judged by referring provider and ED specialist) and BMI &gt;17.5</td>
<td>154</td>
</tr>
<tr>
<td>Green, 2018</td>
<td>Fair</td>
<td>Online version of the Body Project (series of exercises aimed at reducing thin-ideal internalization, maladaptive social comparison, and self-objectification via dissonance induction)</td>
<td>Series of eight online modules and 15 activities (individual format)</td>
<td>NR</td>
<td>Recruited via advertisements (newspaper, social media, flyers, local health practitioners, sorority houses) to participate in an ED treatment and prevention study</td>
<td>United States</td>
<td>Adults and adolescents (14-52 y) with a full criteria ED (DSM-5) or subclinical disorder,* recruited via online advertisements and local flyers</td>
<td>82</td>
</tr>
<tr>
<td>Grilo, 2005</td>
<td>Fair</td>
<td>CBT, CBT + fluoxetine or fluoxetine only</td>
<td>CBT: 16 weekly sessions (1 hour each) Fluoxetine: 60 mg/day</td>
<td>16</td>
<td>Referrals to an outpatient university treatment center</td>
<td>United States</td>
<td>Adults (18-60 y) with BED (DSM-IV) and overweight/obese (between 100 and 200% ideal weight for height based on the 1959 Metropolitan Life Insurance Company Tables)</td>
<td>108</td>
</tr>
<tr>
<td>Author, Year Quality</td>
<td>Description of Intervention (s)</td>
<td>Intensity of Intervention (No. and Length of Sessions)</td>
<td>Duration (Weeks)</td>
<td>Recruitment Setting Country</td>
<td>Population</td>
<td>N</td>
<td>Proportion With Comorbid Psychiatric Disorder (%)</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------------</td>
<td>------------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------------------</td>
<td>------------</td>
<td>---</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Grilo, 2013: Good</td>
<td>Self-help CBT via a structured manual, initiated by primary care physicians + usual care</td>
<td>One introductory session with PCP based on trial script; PCP introduced study, obtained informed consent, gave patient manual, and indicated treatment would be 4 months and they would receive monthly assessments during treatment</td>
<td>16</td>
<td>Primary care settings via posters/flyers and clinician referrals to a treatment study for weight loss and binge eating United States</td>
<td>Adults with BED (DSM-IV-TR)^† and obesity (BMI ≥30) recruited from primary care settings (via advertisements or referral)</td>
<td>48</td>
<td>Mood disorder (DSM-IV): 50 Anxiety disorder (DSM-IV): 50 Substance use disorder (DSM-IV): 21</td>
<td></td>
</tr>
<tr>
<td>Grilo, 2014: Fair</td>
<td>Self-help CBT (+placebo)</td>
<td>Provision of a structured self-help manual by primary care physicians (Overcoming Binge Eating); primary care physicians received brief training and a script to assist with assigning the program</td>
<td>12</td>
<td>Advertisements in primary care settings at a university medical center, mailings, and referrals initiated by primary care physicians United States</td>
<td>Adults (18-65 y) with BED (DSM-5 except for duration of 6 vs. 3 months) and obesity (BMI ≥30), recruited from advertisements and referrals in primary care</td>
<td>53</td>
<td>Mood disorders (DSM-IV): 47 Anxiety disorders (DSM-IV): 38 Substance use disorders: 23</td>
<td></td>
</tr>
<tr>
<td>Hill, 2011: Fair</td>
<td>Dialectical behavior therapy, appetite focused (DBT-AF)</td>
<td>Twelve weekly individual sessions (total 15 h) over 12 weeks; first 6 were 90 min, others were 60 min. Comparison with wait-list control at 6 weeks only (after 6 sessions)</td>
<td>12</td>
<td>Recruited via advertisements, set in university outpatient treatment center United States</td>
<td>Adult women (≥18 y) with BN (DSM-IV) or subthreshold BN (at least one binge eating and one vomit episode per week over the previous 12 weeks); excluded those with BED or AN and those in concurrent therapy for BN. All but six met full DSM-IV criteria for BN.</td>
<td>32</td>
<td>NR</td>
<td></td>
</tr>
</tbody>
</table>

Notes: ^†: Data showing proportion of participants with a comorbid psychiatric disorder are not available.
### Appendix E Table 1. Detailed Evidence Table of Psychological Intervention Characteristics (KQ 4)

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Quality</th>
<th>Description of Intervention(s)</th>
<th>Intensity of Intervention (No. and Length of Sessions)</th>
<th>Duration (Weeks)</th>
<th>Recruitment Setting Country</th>
<th>Population</th>
<th>N</th>
<th>Proportion With Comorobidity Psychiatric Disorder (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kelly, 2014&lt;sup&gt;55&lt;/sup&gt; Fair</td>
<td>Fair</td>
<td>Two forms of self-help for BED: CFT-based self-help and behaviorally based self-help differing in approach for managing difficulties, coping with urges to binge (behavioral strategies vs. the cultivation of self-compassion)</td>
<td>Two treatments were identical in intensity, structure, and target (irregular and unbalanced eating). Both groups participated in a single in-lab self-help exercise learning session, then were told to practice exercises daily for 3 weeks</td>
<td>3</td>
<td>Advertisements in hospitals, ED community centers, and online</td>
<td>Canada</td>
<td>41</td>
<td>NR</td>
</tr>
<tr>
<td>Laessle, 1987&lt;sup&gt;57&lt;/sup&gt; Fair</td>
<td>Fair</td>
<td>Behaviorally oriented group treatment</td>
<td>Twice weekly for first 8 weeks, then once weekly for final 8 weeks</td>
<td>16</td>
<td>Outpatient university psychiatry clinic</td>
<td>Germany</td>
<td>17</td>
<td>NR</td>
</tr>
<tr>
<td>Ljotsson, 2007&lt;sup&gt;76&lt;/sup&gt; Fair</td>
<td>Fair</td>
<td>Internet-assisted self-help/CBT</td>
<td>Self-help manual with assignments (Overcoming Binge Eating), e-mail contact/feedback from coach (one to two emails per week), online private discussion forum, completed over 12 weeks</td>
<td>12</td>
<td>Advertisements in newspapers and AN/BN online patient association website</td>
<td>Sweden</td>
<td>73</td>
<td>NR</td>
</tr>
<tr>
<td>Masson, 2013&lt;sup&gt;60&lt;/sup&gt; Fair</td>
<td>Fair</td>
<td>DBT guided self-help</td>
<td>Single in-person 45-min orientation session and 6 biweekly 20-min support phone calls</td>
<td>13</td>
<td>Advertisements in local media</td>
<td>Canada</td>
<td>60</td>
<td>NR</td>
</tr>
<tr>
<td>Mitchell, 2001&lt;sup&gt;76&lt;/sup&gt; Fair</td>
<td>Fair</td>
<td>Fluoxetine, fluoxetine + self-help manual or self-help manual + placebo</td>
<td>Self-help manual with 14 readings and assignments to be completed over an hour each evening (manual included elements of CBT, behavioral strategies, and meal-planning)</td>
<td>16</td>
<td>Referrals to an outpatient university ED program and advertisements in local newspapers</td>
<td>United States</td>
<td>91</td>
<td>NR</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Quality</td>
<td>Description of Intervention (s)</td>
<td>Intensity of Intervention (No. and Length of Sessions)</td>
<td>Duration (Weeks)</td>
<td>Recruitment Setting Country</td>
<td>Population</td>
<td>N</td>
<td>Proportion With Comorobidity Psychiatric Disorder (%)</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>--------------------------------</td>
<td>---------------------------------------------------</td>
<td>-----------------</td>
<td>----------------------------</td>
<td>------------</td>
<td>----</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Sánchez-Ortiz, 2011 ³³</td>
<td>Good</td>
<td>Internet-based CBT (Overcoming Bulimia Online)</td>
<td>Eight self-guided sessions (45 min each) accompanied by workbooks and assignments; emails from therapists every 1-2 weeks during the first 3 months to support and encourage participants to use the intervention</td>
<td>12</td>
<td>Recruitment e-mails to students attending six higher education institutions, advertisements in posters and pamphlets</td>
<td>United Kingdom</td>
<td>Students (college age) with BN or EDNOS (DSM-IV, but no minimum number of binge or purge episodes) and BMI &gt;18.5 kg/m², recruited from higher education institutions. Of those randomized, 51% met criteria for BN and 49% met criteria for EDNOS</td>
<td>76</td>
</tr>
<tr>
<td>Schag, 2019 ²⁵</td>
<td>Fair</td>
<td>IMPULS: cognitive behavioral group intervention focused on impulsive eating</td>
<td>Eight weekly 90-min group sessions</td>
<td>8</td>
<td>University medical center</td>
<td>Germany</td>
<td>Adults with BED (DSM-5) recruited by email, flyers, press releases, and from an outpatient psychiatry department of a university hospital</td>
<td>80</td>
</tr>
<tr>
<td>Schlup, 2009 ²⁵</td>
<td>Fair</td>
<td>Group CBT §</td>
<td>Eight weekly 90-min group sessions</td>
<td>8</td>
<td>Newspaper advertisements and flyers for a study on binge eating and obesity</td>
<td>Switzerland</td>
<td>Adults (ages 18-70 y) with BED (DSM-IV-TR), responding to study advertisements</td>
<td>36</td>
</tr>
<tr>
<td>Schmidt, 2008 ²⁵</td>
<td>Fair</td>
<td>CD-ROM-based CBT (unguided)</td>
<td>Nine modules completed over 8-12 weeks</td>
<td>12</td>
<td>New referrals to an adult ED outpatient clinic</td>
<td>United Kingdom</td>
<td>Adults with BN and EDNOS (DSM-IV), referred by general practitioners to an ED outpatient center. Of those randomized, 62% had BN and 38% had EDNOS.</td>
<td>97</td>
</tr>
<tr>
<td>Telch, 1990 ³³</td>
<td>Fair</td>
<td>Group CBT</td>
<td>10 weekly, 90-min group sessions</td>
<td>10</td>
<td>Newspaper advertisements offering free treatment for compulsive binge-eating</td>
<td>United Kingdom</td>
<td>Adult women (18-65 y) with compulsive binge-eating but no purging (DSM-II-R BN criteria except for purging criterion), responding to study advertisements</td>
<td>44</td>
</tr>
</tbody>
</table>
**Appendix E Table 1. Detailed Evidence Table of Psychological Intervention Characteristics (KQ 4)**

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Description of Intervention(s)</th>
<th>Intensity of Intervention (No. and Length of Sessions)</th>
<th>Duration (Weeks)</th>
<th>Recruitment Setting Country</th>
<th>Population</th>
<th>N</th>
<th>Proportion With Comorbid Psychiatric Disorder (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wade, 2017(^{2})</td>
<td>Enhanced CBT for a group setting</td>
<td>18 2-hour group sessions and one 1-hour individual session in week 8</td>
<td>8</td>
<td>Outpatient university psychology clinic via clinician referrals, advertisements in papers, posters, and via e-mails to undergraduate students</td>
<td>Adult women (18-36 y) who met criteria for any ED (DSM-5), referred by a clinician (ED specialist, psychologist, or general practitioner) or responding to trial advertisements. Eight (20%) had AN, 23 (58%) had BN, two (5%) had BED, and seven (18%) had OSFED.</td>
<td>40</td>
<td>Current depressive disorder: 35, Current anxiety disorder: 23, Current substance misuse disorder: 5</td>
</tr>
<tr>
<td>Wagner, 2016(^{1,3})</td>
<td>Web-based CBT with individual feedback from therapists</td>
<td>11 assignments with Internet, press information, links posted on psychology and ED websites</td>
<td>16</td>
<td>Germany</td>
<td>Adults (18-65 y) with BED (DSM-IV) and no current AN or BN, responding to trial advertisements</td>
<td>139</td>
<td>Depression: 6, Anxiety: 4</td>
</tr>
<tr>
<td>Wilfley, 1993(^{13})</td>
<td>Group CBT and group IPT</td>
<td>15 weekly 90-min sessions</td>
<td>16</td>
<td>United States</td>
<td>Adult women (ages 18-65 y) with BN and binge eating (DSM-III BN criteria, and those meeting all BN criteria except for purging), responding to study advertisements</td>
<td>56</td>
<td>NR: Concurrent DSM-III-R diagnosis of unipolar or bipolar affective disorder, psychosis, drug abuse, alcoholism, or current major depressive episode was a reason for exclusion</td>
</tr>
</tbody>
</table>

\(^{*}\) Definition used for subclinical ED included endorsement of high levels of body dissatisfaction and one or more of the following behaviors at subclinical levels for weight-control purposes on the Questionnaire for Eating Disorder Diagnoses: binging, laxative use, diuretic use, 24-h fasting, appetite control pill use, strict dieting, or maladaptive exercise (exercising despite injury or exercise that interferes with other important activities).

\(^{1}\) Full or subthreshold criteria; subthreshold criteria defined as binge eating greater than once weekly in frequency and with a duration of at least 6 months (42% randomized).

\(^{2}\) DSM (or other condition definition) used not reported. Subthreshold BN was defined as at least twice-monthly episodes of binge eating and compensatory behaviors during the last 3 months. Subthreshold BED required at least 2 days with objective bulimic episodes per month during the past 6 months, with binge-eating episodes rated as “markedly stressful.”

\(^{3}\) The intervention is 8 weeks of CBT followed by five booster sessions across the 12-month period; there is only a control group for the first 8 weeks and uncontrolled outcomes of intervention only are not eligible.

**Abbreviations:** AN=anorexia nervosa; BED=binge-eating disorder; BMI=body mass index; BN=bulimia nervosa; CBT=cognitive behavioral therapy; CFT=compassion-focused therapy; DBT=dialectical behavior therapy; DBT-AF=dialectical behavior therapy, appetite focused; DSM-II-R=Statistical Manual of Mental Disorders, Version 2 Revised; DSM-III=Statistical Manual of Mental Disorders, Version 3; DSM-III-R=Statistical Manual of Mental Disorders, Version 3 Revised; DSM-IV=Statistical Manual of Mental Disorders, Version 4; DSM-IV-TR=Statistical Manual of Mental Disorders, Version 4 Multiaxial System; DSM-5=Statistical Manual of Mental Disorders, Version 5; ED=eating disorder; EDNOS=eating disorder not otherwise specified; IMPULS=cognitive behavioral group intervention focused on impulsive eating; IPT=interpersonal therapy; KQ=key question;
Appendix E Table 1. Detailed Evidence Table of Psychological Intervention Characteristics (KQ 4)

MDD=major depressive disorder; NR=not reported; OCD=obsessive compulsive disorder; OSFED=Other Specified Feeding and Eating Disorders; PCP=primary care physician; PTSD=posttraumatic stress disorder; vs.=versus.
<table>
<thead>
<tr>
<th>T (Diagnosis)</th>
<th>Author, Year</th>
<th>Dose (mg/day)</th>
<th>Outcome Measure</th>
<th>Time Point (Weeks)</th>
<th>TN</th>
<th>T Baseline Mean Score (SD)</th>
<th>T Score Change From Baseline</th>
<th>Placebo N</th>
<th>Placebo Baseline Mean Score</th>
<th>Placebo Score Change From Baseline</th>
<th>Between-Group Difference in Mean Change</th>
<th>Between-Group P Value for Change From Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lisdexamfetamine (BED)</td>
<td>Guerdjikova, 2016</td>
<td>20-70 (mean: 30)</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>25</td>
<td>20.30 (4.5)</td>
<td>-12.10 (7.3)</td>
<td>25</td>
<td>20.90 (2.60)</td>
<td>-9.30 (8.1)</td>
<td>-2.8 (-7.4 to 1.8)</td>
<td>0.23</td>
</tr>
<tr>
<td>McElroy, 2016a</td>
<td>30-70 (mean: 57)</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>192</td>
<td>21.8 (4.9)</td>
<td>-15.68 SE: 0.55</td>
<td>187</td>
<td>21.6 (4.8)</td>
<td>-8.28 SE: 0.55</td>
<td>-7.4 (-8.93 to -9.51)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2016b</td>
<td>30-70 (mean: 58)</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>195</td>
<td>21.2 (4.4)</td>
<td>-15.36 SE: 0.57</td>
<td>195</td>
<td>21.6 (4.8)</td>
<td>-7.42 SE: 0.57</td>
<td>-7.94 (-9.51 to -6.36)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2016a</td>
<td>30-70 (mean: 57)</td>
<td>SDS</td>
<td>12</td>
<td>192</td>
<td>10.5 (7.2)</td>
<td>-7.76 SE: 0.42</td>
<td>187</td>
<td>10.8 (7.5)</td>
<td>-4.96 SE: 0.43</td>
<td>-2.8 (-3.98 to -1.61)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2016b</td>
<td>30-70 (mean: 58)</td>
<td>SDS</td>
<td>12</td>
<td>195</td>
<td>10.9 (7.8)</td>
<td>-8.74 (SE: 0.4)</td>
<td>195</td>
<td>11.3 (7.3)</td>
<td>-5.04 (SE: 0.41)</td>
<td>-3.7 (-4.81 to -2.58)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>50</td>
<td>YBOCS-BE</td>
<td>11</td>
<td>65</td>
<td>19.5 (5.2)</td>
<td>-15.3 SE: 0.83</td>
<td>63</td>
<td>20.9 (4.61)</td>
<td>-12 SE: 0.87</td>
<td>-3.25</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>50</td>
<td>BES</td>
<td>11</td>
<td>65</td>
<td>27.4 (7.2)</td>
<td>-17.6 SE: 1.24</td>
<td>63</td>
<td>27 (8.62)</td>
<td>-12.2 SE: 1.28</td>
<td>-5.4</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>50</td>
<td>MADRS</td>
<td>11</td>
<td>65</td>
<td>3.6 (3.3)</td>
<td>-1.3 SE: 0.33</td>
<td>63</td>
<td>3.4 (3.39)</td>
<td>-1.7 SE: 0.35</td>
<td>0.49</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>50</td>
<td>HAM-A</td>
<td>11</td>
<td>65</td>
<td>2.3 (2.6)</td>
<td>-1.1 SE: 0.29</td>
<td>63</td>
<td>2.5 (3.01)</td>
<td>-1.5 SE: 0.30</td>
<td>0.4</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>50</td>
<td>SF-12 Phy.</td>
<td>11</td>
<td>65</td>
<td>49.16 (9.1)</td>
<td>2.4 SE: 0.74</td>
<td>63</td>
<td>49.54 (7.875)</td>
<td>1.3 SE: 0.78</td>
<td>1.1</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>McElroy, 2015</td>
<td>50</td>
<td>SF-12 Men.</td>
<td>11</td>
<td>65</td>
<td>46.74 (9.6)</td>
<td>5.5 SE: 0.99</td>
<td>63</td>
<td>48.74 (10.24)</td>
<td>4.9 SE: 1.03</td>
<td>0.6</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Fluoxetine (BED)</td>
<td>Grilo, 2005</td>
<td>60</td>
<td>EDE-Q</td>
<td>16</td>
<td>27</td>
<td>3.90 (1.2)</td>
<td>-0.80</td>
<td>27</td>
<td>3.50 (1.50)</td>
<td>-0.90</td>
<td>0.10</td>
<td>NS</td>
</tr>
<tr>
<td>Grilo, 2005</td>
<td>60</td>
<td>BDI</td>
<td>16</td>
<td>27</td>
<td>16.90 (8.4)</td>
<td>-5.10</td>
<td>27</td>
<td>18.70 (9.7)</td>
<td>-7.00</td>
<td>1.90</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Arnold, 2002</td>
<td>20-80</td>
<td>HAM-D</td>
<td>6</td>
<td>30</td>
<td>4.80 (4.3)</td>
<td>-2.20</td>
<td>30</td>
<td>4.20 (2.90)</td>
<td>1.30</td>
<td>-3.01</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Fluoxetine + CBT (BED)</td>
<td>Grilo, 2005</td>
<td>60</td>
<td>EDE-Q</td>
<td>16</td>
<td>26</td>
<td>4.00 (1.1)</td>
<td>-1.80</td>
<td>27</td>
<td>3.50 (1.50)</td>
<td>-0.90</td>
<td>-0.90</td>
<td>0.002</td>
</tr>
</tbody>
</table>
### Appendix E Table 2. Results of Randomized, Placebo-Controlled Trials of Pharmacotherapy for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>T (Diagnosis)</th>
<th>Author, Year</th>
<th>Dose (mg/day)</th>
<th>Outcome Measure</th>
<th>Time Point (Weeks)</th>
<th>T N</th>
<th>T Baseline Mean Score (SD)</th>
<th>T Score Change From Baseline</th>
<th>Placebo N</th>
<th>Placebo Baseline Mean Score</th>
<th>Placebo Score Change From Baseline</th>
<th>Between-Group Difference in Mean Change</th>
<th>Between-Group P Value for Change From Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bupropion (BED)</td>
<td>Grilo, 2005&lt;sup&gt;61&lt;/sup&gt;</td>
<td>60</td>
<td>BDI</td>
<td>16</td>
<td>26</td>
<td>20.20 (12.1)</td>
<td>-11.00</td>
<td>27</td>
<td>18.70 (9.7)</td>
<td>-7.00</td>
<td>-4.00</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Fluvoxamine (BED)</td>
<td>mean: 239</td>
<td>BDI</td>
<td>12</td>
<td>7</td>
<td>0.44 (0.2)</td>
<td>-0.12</td>
<td>9</td>
<td>0.68 (0.57)</td>
<td>-0.31</td>
<td>0.19</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td></td>
<td>Pearlstein, 2003&lt;sup&gt;26&lt;/sup&gt;</td>
<td>mean: 239</td>
<td>HAM-D</td>
<td>12</td>
<td>8</td>
<td>10.78 (9.2)</td>
<td>-1.40</td>
<td>8</td>
<td>14.27 (12.40)</td>
<td>-6.89</td>
<td>5.49</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>Pearlstein, 2003&lt;sup&gt;26&lt;/sup&gt;</td>
<td>mean: 239</td>
<td>HAM-D</td>
<td>12</td>
<td>21</td>
<td>4.6 (3.8)</td>
<td>-2.2</td>
<td>23</td>
<td>5.7 (4.5)</td>
<td>-0.9</td>
<td>-1.3</td>
<td>0.097</td>
</tr>
<tr>
<td>Sertraline (BED)</td>
<td>McElroy, 2000&lt;sup&gt;62&lt;/sup&gt;</td>
<td>50-200</td>
<td>HAM-D</td>
<td>6</td>
<td>18</td>
<td>6.4 (3.9)</td>
<td>NR</td>
<td>16</td>
<td>7.5 (8.4)</td>
<td>NR</td>
<td>1.33 (SE: 1.0)</td>
<td>0.19</td>
</tr>
<tr>
<td>Escitalopram (BED)</td>
<td>Guerdjikova, 2008&lt;sup&gt;63&lt;/sup&gt;</td>
<td>10-30</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>21</td>
<td>19.1 (5.3)</td>
<td>-11.2</td>
<td>23</td>
<td>19 (3.6)</td>
<td>-7.7</td>
<td>-3.5</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>Guerdjikova, 2008&lt;sup&gt;63&lt;/sup&gt;</td>
<td>10-30</td>
<td>HAM-D</td>
<td>12</td>
<td>21</td>
<td>4.6 (3.8)</td>
<td>-2.2</td>
<td>23</td>
<td>5.7 (4.5)</td>
<td>-0.9</td>
<td>-1.3</td>
<td>0.097</td>
</tr>
<tr>
<td>Topiramate (BED)</td>
<td>McElroy, 2003&lt;sup&gt;64&lt;/sup&gt;</td>
<td>25-600</td>
<td>YBOCS-BE</td>
<td>14</td>
<td>30</td>
<td>21.5 (3.9)</td>
<td>NR</td>
<td>31</td>
<td>21.60 (4.6)</td>
<td>NR</td>
<td>-2.55</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>McElroy, 2007&lt;sup&gt;65&lt;/sup&gt;</td>
<td>25-400</td>
<td>YBOCS-BE</td>
<td>16</td>
<td>202</td>
<td>21.1 (4.9)</td>
<td>-14.3 (8.9)</td>
<td>202</td>
<td>16.80 (4.5)</td>
<td>-7.90 (SD: 8.9)</td>
<td>6.40</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>McElroy, 2003&lt;sup&gt;64&lt;/sup&gt;</td>
<td>25-600</td>
<td>HAM-D</td>
<td>14</td>
<td>30</td>
<td>5.9 (5.1)</td>
<td>NR</td>
<td>31</td>
<td>5.80 (4.80)</td>
<td>NR</td>
<td>NR</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>McElroy, 2007&lt;sup&gt;65&lt;/sup&gt;</td>
<td>25-400</td>
<td>MADRS</td>
<td>16</td>
<td>202</td>
<td>5.9 (5.4)</td>
<td>-0.2 (7)</td>
<td>202</td>
<td>6.70 (5.50)</td>
<td>-0.70 (6.2)</td>
<td>0.50</td>
<td>0.893</td>
</tr>
<tr>
<td></td>
<td>McElroy, 2007&lt;sup&gt;65&lt;/sup&gt;</td>
<td>25-400</td>
<td>HAM-A</td>
<td>16</td>
<td>202</td>
<td>5.1 (4.8)</td>
<td>-0.7 (4.9)</td>
<td>202</td>
<td>5.50 (5.10)</td>
<td>-1.30 (4.5)</td>
<td>0.60</td>
<td>0.493</td>
</tr>
<tr>
<td>Bupropion (BED)</td>
<td>White, 2013&lt;sup&gt;66&lt;/sup&gt;</td>
<td>300</td>
<td>EDE-Q</td>
<td>8</td>
<td>31</td>
<td>2.5 (1.1)</td>
<td>-0.70</td>
<td>30</td>
<td>2.70 (0.80)</td>
<td>-0.70</td>
<td>0.00</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>White, 2013&lt;sup&gt;66&lt;/sup&gt;</td>
<td>300</td>
<td>BDI</td>
<td>8</td>
<td>31</td>
<td>13.4 (9.8)</td>
<td>-5.40</td>
<td>30</td>
<td>10.80 (6.10)</td>
<td>-2.10</td>
<td>-3.30</td>
<td>0.84</td>
</tr>
<tr>
<td>Duloxetine (BED)</td>
<td>Guerdjikova, 2012&lt;sup&gt;67&lt;/sup&gt;</td>
<td>60-90</td>
<td>IDS</td>
<td>12</td>
<td>20</td>
<td>35.6 (7.9)</td>
<td>-16.5</td>
<td>20</td>
<td>35.4 (5.4)</td>
<td>-13.8</td>
<td>-2.7</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Appendix E Table 2. Results of Randomized, Placebo-Controlled Trials of Pharmacotherapy for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>T (Diagnosis)</th>
<th>Author, Year</th>
<th>Dose (md/day)</th>
<th>Outcome Measure</th>
<th>Time Point (Weeks)</th>
<th>T N</th>
<th>T Baseline Mean Score (SD)</th>
<th>T Score Change From Baseline</th>
<th>Placebo N</th>
<th>Placebo Baseline Mean Score</th>
<th>Placebo Score Change From Baseline</th>
<th>Between-Group Difference in Mean Change</th>
<th>Between-Group P Value for Change From Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guerdjikova, 2012&lt;sup&gt;54&lt;/sup&gt;</td>
<td>Imipramine (BED)</td>
<td>60-90</td>
<td>HAM-A</td>
<td>12</td>
<td>20</td>
<td>16.9 (9.1)</td>
<td>-7.3</td>
<td>20</td>
<td>16.2 (5.7)</td>
<td>-9</td>
<td>1.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Laederach-Hofmann, 1998&lt;sup&gt;56&lt;/sup&gt;</td>
<td></td>
<td>75</td>
<td>HAM-D</td>
<td>8</td>
<td>15</td>
<td>22.6 (9.8)</td>
<td>-9.6 (7.1)</td>
<td>16</td>
<td>21.30 (12.00)</td>
<td>-3.50 (8.9)</td>
<td>-6.10</td>
<td>NR</td>
</tr>
<tr>
<td>Laederach-Hofmann, 1998&lt;sup&gt;56&lt;/sup&gt;</td>
<td></td>
<td>75</td>
<td>SDS</td>
<td>8</td>
<td>15</td>
<td>35.3 (6.3)</td>
<td>-6.3 (5)</td>
<td>16</td>
<td>35.00 (5.80)</td>
<td>-4.60 (4.3)</td>
<td>-1.70</td>
<td>NR</td>
</tr>
<tr>
<td>Guerdjikova, 2012&lt;sup&gt;54&lt;/sup&gt;</td>
<td>Duloxetine (BED)</td>
<td>60-90</td>
<td>YBOCS-BE</td>
<td>12</td>
<td>20</td>
<td>22.3 (3.5)</td>
<td>-12.9</td>
<td>20</td>
<td>21.6 (2.8)</td>
<td>-11.3</td>
<td>-1.6</td>
<td>0.47</td>
</tr>
<tr>
<td>Walsh, 1991&lt;sup&gt;50&lt;/sup&gt;</td>
<td>Desipramine (BN)</td>
<td>200-300</td>
<td>HAM-D</td>
<td>8</td>
<td>40</td>
<td>8.3 (4.6)</td>
<td>-2.3</td>
<td>38</td>
<td>7.3 (4.6)</td>
<td>-0.8</td>
<td>-1.5</td>
<td>NS</td>
</tr>
<tr>
<td>Walsh, 1991&lt;sup&gt;50&lt;/sup&gt;</td>
<td></td>
<td>200-300</td>
<td>BDI</td>
<td>8</td>
<td>40</td>
<td>10.4 (7.3)</td>
<td>-1.2</td>
<td>38</td>
<td>15 (11.1)</td>
<td>-2</td>
<td>0.8</td>
<td>NS</td>
</tr>
<tr>
<td>Walsh, 1991&lt;sup&gt;50&lt;/sup&gt;</td>
<td></td>
<td>200-300</td>
<td>STAI-State</td>
<td>8</td>
<td>40</td>
<td>48.1 (12.2)</td>
<td>-2.6</td>
<td>38</td>
<td>51.5 (14.3)</td>
<td>-2.2</td>
<td>-0.4</td>
<td>NS</td>
</tr>
<tr>
<td>Walsh, 1991&lt;sup&gt;50&lt;/sup&gt;</td>
<td></td>
<td>200-300</td>
<td>STAI-Trait</td>
<td>8</td>
<td>40</td>
<td>51.9 (10.5)</td>
<td>-5.4</td>
<td>38</td>
<td>54.3 (10.3)</td>
<td>-0.2</td>
<td>-5.2</td>
<td>0.01</td>
</tr>
<tr>
<td>Kaneva, 1995&lt;sup&gt;54&lt;/sup&gt;</td>
<td>Fluoxetine (BN)</td>
<td>60</td>
<td>EAT</td>
<td>8</td>
<td>24</td>
<td>40.3 (15.6)</td>
<td>-10.70</td>
<td>26</td>
<td>42.50 (15.40)</td>
<td>-6.60</td>
<td>-4.10</td>
<td>NR</td>
</tr>
<tr>
<td>Fluoxetine Bulimia Nervosa Collaborative Study Group, 1992&lt;sup&gt;77&lt;/sup&gt;</td>
<td></td>
<td>60</td>
<td>EAT</td>
<td>8</td>
<td>129</td>
<td>31.5 (12.5)</td>
<td>-8.5</td>
<td>129</td>
<td>35.9 (13.3)</td>
<td>-4</td>
<td>-4.5 (median)</td>
<td>0.001</td>
</tr>
<tr>
<td>Kaneva, 1995&lt;sup&gt;54&lt;/sup&gt;</td>
<td></td>
<td>60</td>
<td>EDI</td>
<td>8</td>
<td>24</td>
<td>69.4 (22.5)</td>
<td>-19.40</td>
<td>26</td>
<td>80.50 (25.10)</td>
<td>-18.60</td>
<td>-0.80</td>
<td>NR</td>
</tr>
<tr>
<td>Mitchell, 2001&lt;sup&gt;76&lt;/sup&gt;</td>
<td></td>
<td>60</td>
<td>EDI</td>
<td>16</td>
<td>26</td>
<td>66.8 (16.2)</td>
<td>NR</td>
<td>22</td>
<td>72.11 (14.59)</td>
<td>NR</td>
<td>NR</td>
<td>&gt;0.15</td>
</tr>
<tr>
<td>Kaneva, 1995&lt;sup&gt;54&lt;/sup&gt;</td>
<td></td>
<td>60</td>
<td>BITE</td>
<td>8</td>
<td>24</td>
<td>24.3 (2.3)</td>
<td>-2.00</td>
<td>26</td>
<td>23.90 (3.50)</td>
<td>-1.80</td>
<td>-0.20</td>
<td>NR</td>
</tr>
<tr>
<td>Kaneva, 1995&lt;sup&gt;54&lt;/sup&gt;</td>
<td></td>
<td>60</td>
<td>HDRS-17</td>
<td>8</td>
<td>24</td>
<td>9.3 (4.5)</td>
<td>-3.80</td>
<td>26</td>
<td>9.40 (4.90)</td>
<td>-1.70</td>
<td>-2.10</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Appendix E Table 2. Results of Randomized, Placebo-Controlled Trials of Pharmacotherapy for Eating Disorders (KQ 4)

<table>
<thead>
<tr>
<th>T (Diagnosis)</th>
<th>Author, Year</th>
<th>Dose (mg/day)</th>
<th>Outcome Measure</th>
<th>Time Point (Weeks)</th>
<th>T N</th>
<th>T Baseline Mean Score (SD)</th>
<th>T Score Change From Baseline</th>
<th>Placebo N</th>
<th>Placebo Baseline Mean Score</th>
<th>Placebo Score Change From Baseline</th>
<th>Between-Group Difference in Mean Change</th>
<th>Between-Group P Value for Change From Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine (BN) con't</td>
<td>Fluoxetine Bulimia Nervosa Collaborative Study Group, 1992</td>
<td>60</td>
<td>HDRS</td>
<td>8</td>
<td>129</td>
<td>11.9 (7.3)</td>
<td>-5</td>
<td>129</td>
<td>11.8 (7.7)</td>
<td>-3</td>
<td>-2 (median)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Mitchell, 2001</td>
<td>60</td>
<td>HAM-D</td>
<td>16</td>
<td>26</td>
<td>8.85 (6.8)</td>
<td>NR</td>
<td>22</td>
<td>10.91 (5.89)</td>
<td>NR</td>
<td>NR</td>
<td>&gt;0.15</td>
</tr>
<tr>
<td></td>
<td>Kaneva, 1995</td>
<td>60</td>
<td>State Anxiety</td>
<td>8</td>
<td>24</td>
<td>50.3 (11.8)</td>
<td>-7.80</td>
<td>26</td>
<td>45.80 (11.40)</td>
<td>-1.30</td>
<td>-6.50</td>
<td>0.0004</td>
</tr>
<tr>
<td>Fluoxetine + self-help (BN)</td>
<td>Mitchell, 2001</td>
<td>60</td>
<td>EDI</td>
<td>16</td>
<td>21</td>
<td>58.11 (15.1)</td>
<td>NR</td>
<td>22</td>
<td>72.11 (14.59)</td>
<td>NR</td>
<td>NR</td>
<td>&gt;0.15</td>
</tr>
<tr>
<td></td>
<td>Mitchell, 2001</td>
<td>60</td>
<td>HAM-D</td>
<td>16</td>
<td>21</td>
<td>8.1 (6.6)</td>
<td>NR</td>
<td>22</td>
<td>10.91 (5.89)</td>
<td>NR</td>
<td>NR</td>
<td>&gt;0.15</td>
</tr>
</tbody>
</table>

Abbreviations: BDI=Beck Depression Inventory; BED=binge eating disorder; BES=Binge Eating Scale; BMI=body mass index; BITE=Bulimia Investigatory Test Edinburgh; BN=bulimia nervosa; CBT=cognitive behavioral therapy; EAT=Eating Attitudes Test; EDE-Q=Eating Disorder Examination Questionnaire; EDI=Eating Disorder Inventory; HAM-A=Hamilton Depression Rating Scale--Anxiety; HAM-D=Hamilton Depression Rating Scale--Depression; HDRS=Hamilton Depression Rating Scale; HDRS-17=Hamilton Depression Rating Scale--17; IDS=Inventory of Depressive Symptomatology; KQ=key question; MADRS=Montgomery-Åsberg Depression Rating Scale; NR=not reported; NS=not significant; SD=standard deviation; SDS=Zung Self-Rating Depression Scale; SE=standard error; SF-12=12-Item Short Form Survey; STAI=State-Trait Anxiety Inventory; T=treatment; YBOCS-BE=Yale–Brown Obsessive Compulsive Scale modified for binge eating.
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Time Point (weeks)</th>
<th>Intervention (N)</th>
<th>Psychiatric/Mood-Related Adverse Effects</th>
<th>Neurologic/Sleep-Related Adverse Effects</th>
<th>GI Adverse Effects</th>
<th>Respiratory-Related Adverse Effects</th>
<th>Other Adverse Outcomes (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnold, 2002</td>
<td>6</td>
<td>G1: Fluoxetine (30) G2: Placebo (30)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>No significant differences between treatment groups in the incidence of adverse events*</td>
</tr>
</tbody>
</table>
## Appendix E Table 3. Harms

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Time Point (weeks)</th>
<th>Intervention (N)</th>
<th>Psychiatric/Mood-Related Adverse Effects</th>
<th>Neurologic/Sleep-Related Adverse Effects</th>
<th>GI Adverse Effects</th>
<th>Respiratory-Related Adverse Effects</th>
<th>Other Adverse Outcomes (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, Year Quality</td>
<td>Time Point (weeks)</td>
<td>Intervention (N) Control (N)</td>
<td>Psychiatric/Mood-Related Adverse Effects</td>
<td>Neurologic/Sleep-Related Adverse Effects</td>
<td>GI Adverse Effects</td>
<td>Respiratory-Related Adverse Effects</td>
<td>Other Adverse Outcomes (N)</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Fluoxetine Bulimia Nervosa Collaborative Study Group, 1992</td>
<td>8</td>
<td>G1: Fluoxetine 20 mg (129) G1: Fluoxetine 60 mg (129) G3: Placebo (129)</td>
<td>NR</td>
<td>AE that occurred significantly (p&lt;0.05) more with G1 + G2 vs. G3: Insomnia G1: 23 G2: 30 G3: 10 p&lt;0.001 Asthenia G1: 16 G2: 23 G3: 11 p=0.039 Tremor G1: 4 G2: 12 G3: 0 p&lt;0.001</td>
<td>AE that occurred significantly (p&lt;0.05) more with G1 + G2 vs. G3: Nausea G1: 20 G2: 28 G3: 14 p=0.021</td>
<td>NR</td>
<td>AE that occurred significantly (p&lt;0.05) more with G1 + G2 vs. G3: Sweating G1: 4 G2: 7 G3: 1 p=0.036 Urinary frequency G1: 0 G2: 5 G3: 2 p=0.012 Palpitation G1: 1 G2: 5 G3: 2 p=0.017 Yawn: G1: 1 G2: 5 G3: 1 p=0.017 Mydriasis G1: 0 G2: 3 G3: 0 p=0.018 Vasodilatation G1: 4 G2: 1 G3: 0 p=0.029</td>
</tr>
</tbody>
</table>
### Appendix E Table 3. Harms

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Time Point (weeks)</th>
<th>Intervention (N)</th>
<th>Psychiatric/Mood-Related Adverse Effects</th>
<th>Neurologic/Sleep-Related Adverse Effects</th>
<th>GI Adverse Effects</th>
<th>Respiratory-Related Adverse Effects</th>
<th>Other Adverse Outcomes (N)</th>
</tr>
</thead>
</table>
| Guerdjukova, 20081    | 12                | G1: Escitalopram (21) G2: Placebo (21) | Nervousness:  
G1: 2  
G2: 1 | Headache:  
G1: 2  
G2: 4  
G1: 3  
G2: 3 | Diarrhea:  
G1: 5  
G2: 5  
G1 flu:  
G1: 3  
G2: 2  
Nausea:  
G1: 1  
G2: 3 | URI:  
G1: 2  
G2: 1  
Cold/pharyngitis:  
G1: 1  
G2: 4 | Dry mouth:  
G1: 7  
G2: 6  
Fatigue:  
G1: 3  
G2: 5  
Increased urinary frequency:  
G1: 3  
G2: 0  
Sweating:  
G1: 3  
G2: 0  
Sexual dysfunction:  
G1: 3  
G2: 0  
Yawning:  
G1: 1  
G2: 2  
Edema:  
G1: 1  
G2: 3 |
| McElroy, 200767       | 16                | G1: Topiramate (202) G2: Placebo (202) | Difficulty with Concentration/attention (%)  
G1: 26  
G2: 5  
p<0.001 | Paraesthesia (%)  
G1: 113  
G2: 25  
p<0.001  
Somnolence (%)  
G1: 34  
G2: 26  
p=0.327  
Difficulty with memory NOS (%)  
G1: 25  
G2: 12  
p=0.037 | Nausea  
G1: 32  
G2: 25  
p=0.391 | URI (%)  
G1: 37  
G2: 20  
p=0.022 | Taste perversion (%)  
G1: 28  
G2: 2  
p<0.01  
Dry mouth (%)  
G1: 27  
G2: 22  
p=0.543 |
Appendix E Table 3. Harms

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Time Point (weeks)</th>
<th>Intervention (N) Control (N)</th>
<th>Psychiatric/Mood-Related Adverse Effects</th>
<th>Neurologic/Sleep-Related Adverse Effects</th>
<th>GI Adverse Effects</th>
<th>Respiratory-Related Adverse Effects</th>
<th>Other Adverse Outcomes (N)</th>
</tr>
</thead>
</table>

Columbia-Suicide Severity Rating Scale
Only a single case of any active suicidal ideation in each arm at weeks 1 and 12
"No positive affirmations including preparatory acts, actual, interrupted, or aborted suicide attempts with either treatment"
## Appendix E Table 3. Harms

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Time Point (weeks)</th>
<th>Intervention (N) Control (N)</th>
<th>Psychiatric/Mood-Related Adverse Effects</th>
<th>Neurologic/Sleep-Related Adverse Effects</th>
<th>GI Adverse Effects</th>
<th>Respiratory-Related Adverse Effects</th>
<th>Other Adverse Outcomes (N)</th>
</tr>
</thead>
</table>
| McElroy, 2016b<sup>1</sup> | 12                 | G1: Lisdexamfetamine (181) G2: Placebo (185) | Irritability  
G1: 9  
G2: 6  
Columbia-Suicide Severity Rating Scale (per authors):  
“No positive affirmations including preparatory acts, actual, interrupted, or aborted suicide attempts with either treatment”  
“No positive affirmations of active suicidal ideation with either treatment” | Headache  
G1: 32  
G2: 16  
Insomnia  
G1: 19  
G2: 6  
Feeling jittery  
G1: 10  
G2: 0 | Decreased appetite  
G1: 11  
G2: 3  
Nausea  
G1: 16  
G2: 8  
Constipation  
G1: 10  
G2: 1 | NR | Any treatment-emergent AE  
G1: 140  
G2: 94  
Serious treatment-emergent AE  
G1: 1  
G2: 2  
Treatment-emergent AE related to study drug  
G1: 119  
G2: 56  
Dry mouth  
G1: 60  
G2: 11  
Fatigue  
G1: 17  
G2: 9  
Blood pressure increased  
G1: 9  
G2: 5  
Fatigue  
G1: 17  
G2: 9 |
### Appendix E Table 3. Harms

<table>
<thead>
<tr>
<th>Author, Year Quality</th>
<th>Time Point (weeks)</th>
<th>Intervention (N) Control (N)</th>
<th>Psychiatric/Mood-Related Adverse Effects</th>
<th>Neurologic/Sleep-Related Adverse Effects</th>
<th>GI Adverse Effects</th>
<th>Respiratory-Related Adverse Effects</th>
<th>Other Adverse Outcomes (N)</th>
</tr>
</thead>
</table>
| McElroy, 2015 
[9] | 11 | G1: Lisdexamfetamine 30 mg/day (66)  
G2: Lisdexamfetamine 50 mg/day (65)  
G3: Lisdexamfetamine 70 mg/day (65)  
G4: Placebo (63) | Irritability  
G1: 5  
G2: 3  
G3: 3  
G4: 4  
Anxiety  
G1: 4  
G2: 2  
G3: 1  
G4: 0  
Feeling jittery  
G1: 1  
G2: 3  
G3: 5  
G4: 0  
Sleep disorder  
G1: 1  
G2: 3  
G3: 4  
G4: 0 | Insomnia  
G1: 17  
G2: 13  
G3: 12  
G4: 4  
Headache  
G1: 6  
G2: 3  
G3: 5  
G4: 1  
Diabetes  
G1: 4  
G2: 5  
G3: 1  
G4: 0 | Decreased appetite  
G1: 8  
G2: 2  
G3: 3  
G4: 2  
Constipation  
G1: 1 | URI  
G1: 1  
G2: 3  
G3: 5  
G4: 4  
Nasopharyngitis  
G1: 8  
G2: 1  
G3: 3  
G4: 2 | Any treatment-emergent AE  
G1: 57  
G2: 56  
G3: 53  
G4: 37  
Serious treatment-emergent AE  
G1: 2  
G2: 0  
G3: 1  
G4: 0  
Dry mouth  
G1: 22  
G2: 22  
G3: 27  
G4: 5  
Pulmonary embolism  
G1: 4  
G2: 2  
G3: 3  
G4: 0 | Incidence of any treatment-emergent AE was significantly higher for the combined treatment (84.7%) vs. placebo group (58.7%) |

One participant died during the study. Postmortem toxicology analysis reported that methamphetamine/amphetamine levels were consistent with a methamphetamine overdose.

* Authors provided only the “most common” adverse effects in the fluoxetine group only (including dry mouth, headache, and nausea), but no rates from the placebo group.

**Abbreviations:** AE=adverse event; G=group; GI=gastrointestinal; NOS=not otherwise specified; NR=not reported; NS=not significant; URI=Upper Respiratory Infection.
Figure Notes: The 95% confidence region provides a visual estimate of the amount of variation around the pooled estimate that is due to sampling variation (i.e., chance). It is the region within which we expect the true pooled summary point to lie. It can be used to assess precision of the pooled estimate. The smaller the region, the more precise the estimate. In this figure, precision of the estimates for specificity is higher compared with the precision of the estimates for sensitivity. The 95% prediction region provides a visual estimate of the between-study variability that cannot be attributed to chance. It is the region within which we expect any future individual study estimate to lie. It can be used to assess the consistency of study findings. The larger the prediction region is within the SROC space and relative to the size of the confidence region, the more inconsistency (i.e., heterogeneity) is present.

Abbreviations: HSROC=hierarchical summary receiver operating characteristic; SROC=summary receiver operating characteristic.
Appendix F Figure 2. Summary Receiver Operating Characteristics Curve for Screening Test Accuracy of SCOFF (Cut Point ≥ 3) in Adults

**Figure Notes:** The 95% confidence region provides a visual estimate of the amount of variation around the pooled estimate that is due to sampling variation (i.e., chance). It is the region within which we expect the true pooled summary point to lie. It can be used to assess precision of the pooled estimate. The smaller the region, the more precise the estimate. In this figure, precision of the estimates for specificity is higher compared with the precision of the estimates for sensitivity. The 95% prediction region provides a visual estimate of the between-study variability that cannot be attributed to chance. It is the region within which we expect any future individual study estimate to lie. It can be used to assess the consistency of study findings. The larger the prediction region is within the SROC space and relative to the size of the confidence region, the more inconsistency (i.e., heterogeneity) is present.

**Abbreviations:** HSROC=hierarchical summary receiver operating characteristic
SROC=summary receiver operating characteristic.
Appendix References

Appendix References


Appendix References


Appendix References


Appendix References


Appendix References


Appendix References


