National Guideline Alliance

Version 1.0

Eating Disorders: recognition and treatment

Full guideline

NICE Guideline

Methods, evidence and recommendations

December 2016

Draft for Consultation

Commissioned by the National Institute for Health and Care Excellence

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Contents

1

2

3

1 Guideline summary

1.1 Preface

Background

There are over 700,000 individuals in the UK with an eating disorder (Beat, 2015). While the prevalence rate is relatively stable, the number of cases identified in clinical settings is increasing as clinicians become more aware of these disorders and patients come forward more readily. However, many cases remain unidentified.

Those who suffer from eating disorders (anorexia nervosa, bulimia nervosa, binge eating disorder, other specified feeding and eating disorders) experience physical complications, psychological comorbidity, poor quality of life, disrupted relationships, emotional distress, social isolation and economic disadvantage (Beat, 2015). The risk of early death in this population is among the highest among patients with psychiatric disorders, whether due to physical complications (e.g., starvation) or mental health issues (e.g., suicide). Their families, carers and loved ones also suffer as a result of the eating disorder, experiencing high levels of stress (Treasure et al., 2001).

About 90% of cases in the UK are female and most are of normal weight or above (only about 15-20% meet criteria for anorexia nervosa) (Fairburn and Harrison, 2003). There is little evidence of any difference in prevalence rates among different ethnic groups. Some groups are at greater risk of developing eating disorders (e.g., those who work of study in areas where there is a strong focus on physical appearance, such as dancers, models or athletes). Others are at particularly high risk of complications if they develop an eating disorder (e.g., those with type 1 diabetes). Therefore, these groups are the subject of particular consideration in this guideline. Risk management should always be seen as the first consideration.

Eating disorders commonly have their origins in adolescence, but are often not identified or picked up by services until adulthood. Early intervention is strongly advocated (Treasure and Russell, 2011). However, the necessary early identification and prevention or treatment of eating disorders is a difficult task, due to the low base rate of such cases and limitations in tools suggested for early identification and prevention. Even very underweight patients routinely go unidentified or their needs are not responded to by clinicians and non-underweight patients are likely to be missed by clinicians and families alike. Such patients are unlikely to self-report in the early years of the disorder. Furthermore, most eating disorders have low rates of spontaneous remission. Therefore, this guideline focuses substantially on the evidence as to how clinicians can effectively treat and manage eating disorders that might have been present for many years.

The causes of eating disorders are not fully understood, with evidence of a mixture of social, biological, psychological and interpersonal causes. However, given that these disorders develop across their course (often several years), it is more important to consider the role of the factors that maintain the eating problems when planning and delivering treatment. Those maintaining factors are interlinked, but include:

- starvation and semi-starvation;
- social isolation and avoidance;
- emotional responses (particularly anxiety and depression);
- cognitive difficulties (concentration, narrowed thinking, memory, attentional biases, etc.);
- body image disturbance (negative body self-esteem, overestimation of body size);
- behaviours that maintain the problem (avoidance of food; purging behaviours, body checking, etc.)

Therefore, treatment needs to address these maintaining factors, while ensuring that risks are managed appropriately.

Key principles of evidence-based treatments for eating disorders

The first principle behind any treatment is to maintain life and avoid doing harm. The first of these is an important issue in a disorder with such a high mortality rate and the importance of medical care is undoubted. The risk of doing harm is addressed below – the danger of undertaking ineffective treatment, resulting in patients suffering a loss of engagement or belief in their ability to change.

The second principle is the need to engage the patient, family and carers in the process of change. Family and carers should be kept informed as far as possible and can be active agents in treatments. However, there is evidence that the process of engaging patients in the process of treatment differs according to the type of therapy (Graves et al., 2016). Developing the therapeutic alliance is a primary task when working with younger cases receiving family therapy, but with cognitive and behavioural treatments in adults it is the process of behavioural change that results in the development of positive therapeutic alliance (Brown et al., 2013). In short, clinicians need to work with the relevant evidence to get the best outcomes.

It is important to use all of the evidence-based approach that is advocated, rather than omitting elements that the clinician dislikes or feels uncomfortable delivering (Turner et al., 2014). Clinicians need to use protocols in a non-rigid, patient-centred way (Wilson, 1996), rather than omitting key requirements.

Nutritional restoration is a key part of most of the effective approaches to food, whether considering patient's physical health or psychotherapy outcomes (e.g., family meals; weight restoration; cessation of bingeing and purging). Its impact is: biological (e.g., resumption of menstruation and repair of bone structure); cognitive (e.g., flexibility of thought); emotional (e.g., stabilisation of mood); and interpersonal (e.g., restoration of social skills). While it is often contended that anorexia nervosa in particular might have some neuropsychological underpinning that would limit such benefits, it is important to remember that the evidence for this contention is highly limited (studies that are poorly designed, have too small a number, or they fail to replicate). Therefore, clinicians should not regard anorexia nervosa as a 'diagnosis of despair'.

The need for this guideline

The most recent NICE guideline for eating disorders was published in 2004. That guideline had relatively few recommendations based on level 'A' and 'B' evidence, reflecting the lack of high quality evidence at that time. Most noticeably, there were no evidence-based recommendations regarding the treatment of most atypical cases and the treatment of anorexia nervosa was based on clinical opinion rather than strong evidence. Two important strands of evidence have emerged since 2004, relating to:

- Enhanced evidence of treatment outcomes
- Evidence about problems regarding how therapies for eating disorders are delivered

A key consideration is the importance of focusing on treatments that have strong evidence at the end of the development of this guideline, in order to ensure that training demands for clinical services are manageable. For example, it is more beneficial and manageable for a service to train its staff in a limited number of effective therapies, rather than to train them in a wide range of therapies that are not proven to be effective.

Recent evidence of treatment outcomes

Since the 2004 NICE guideline, there has been a substantial increase in this empirical literature, including studies that might consolidate or amend the recommendations that were made at that time. In particular, the high-quality studies to be considered in this guideline include:

- Further studies of treatments (psychotherapies and medications) for bulimia nervosa and binge eating disorder
- Studies comparing treatments (both psychological and biological) for anorexia nervosa
- Studies on the impact of treatments on atypical cases (OSFED)
- Studies of the relevance (or otherwise) of comorbidities on the outcome of treatment for eating disorders.

These studies will be addressed directly in the guideline. Some therapies will be shown to be stronger in their evidence base, some new treatments will be considered, and some therapies that were included as possible courses of action in the 2004 guideline will be omitted as a result of lacking evidence.

These studies have been carried out with individuals across the young people-adult age span, involving parents and carers where appropriate to the cases in question. However, ARFID is not included in this guideline, as the evidence relating to treatment of this newly established diagnosis (APA 2013(APA, 2013)) is too limited to allow for recommendations to be made at this stage. Similarly, obesity in the absence of an eating disorder is not considered in this review, as it is the subject of a separate NICE guideline.

Evidence of treatment delivery

This guideline is also needed in light of other evidence that has emerged in the past decade. The relevant studies address:

- Continuing gaps in case identification in non-specialist primary care settings (Waller, Micali & James, 2014(Waller et al.)).
- The need for strategic use of relatively intensive treatments (e.g., day and in-patient) where appropriate (Gowers 2007).
- The need for appropriate implementation of evidence-based therapies, to ensure they are delivered at all (Waller 2012(Waller et al., 2012)) and competently (Fairburn & Cooper 2011(Fairburn and Cooper, 2011)).
- Clear evidence that evidence-based therapies can be delivered effectively in routine NHS and comparable settings (Byrne 2011(Byrne et al., 2011); Couturier 2010(Couturier et al., 2010); Turner 2015(Turner et al., 2015); Turner 2016(Turner et al., 2016); Waller 2014(Waller et al., 2014a)).
- The need to deliver treatment strategically (e.g., whether to maintain in-patient admission; whether to enhance out-patient treatment), considering whether it should be continued or adapted when there is early evidence of a lack of progress (Lock 2015(Lock et al., 2015); Turner 2015).

The reasons for this guideline address a clear issue in routine practice when working with eating disorders – therapist drift (Waller 2009(Waller, 2009); Waller 2016(Waller, 2016)). Clinicians commonly dismiss evidence-based practice as not relevant to their clinical setting and therefore routinely employ unevidenced therapeutic approaches (Kosmerly 2015(Kosmerly et al., 2015); Tobin 2007(Tobin et al., 2007)). Those approaches commonly include using relatively intensive treatments for far longer than is necessary or useful, or delivering full courses of therapy when an early lack of response was a strong indicator that this would not work.

Each of these clinical practices has the potential to be very costly and to cause negative outcomes (e.g., patients who lose belief that they could recover, or who come to be dependent on being in treatment despite the lack of progress).

Summary

In summary, eating disorders are relatively common in the UK population but are poorly identified in non-specialist NHS settings. These disorders are usually long-lasting and have serious implications, including risk of death, impaired health, psychiatric comorbidity and poor quality of life for the patient and those around them. A number of principles of care have been outlined.

Since the 2004 NICE guideline, two strands of evidence have emerged that necessitate a new eating disorders guideline in 2017. First, there is now far more evidence of efficacious treatments (both physical and psychotherapeutic), allowing for firmer guidelines to be developed. Second, it has become clear that clinicians vary substantially in their identification of cases and their delivery of the evidence-based treatments that are recommended (for reasons that are connected with therapists' unawareness of the evidence, preferences, or lack of training and competence).

Consequently, NICE and the Department of Health have concluded that it is necessary to produce this new guideline, to inform clinical practice and policy. The aim is to ensure that patients receive the best treatments possible, from clinicians who are knowledgeable and well trained.

1.2 Committee membership, National Guideline Alliance (NGA) staff and acknowledgements

Table 1: Committee members

Name	Role
Anthony Bateman (Chair)	Visiting Professor in the Psychoanalysis Unit, University College London. Consultant Psychiatrist and Psychotherapist and Honorary Senior Lecturer, University College and Royal Free Medical Schools, Barnet, Enfield and Haringey Mental Health NHS Trust and St Ann's Hospital, London.
Jane Dalgliesh	Mental Health Nurse, Nurse Practitioner/Team Manager/Head of Service, Eating Disorders Service, South Essex University Foundation Trust (SEPT).
Ivan Eisler	Emeritus Professor of Family Psychology and Family Therapy, Kings College Institute of Psychiatry, Psychology and Neuroscience Consultant Clinical Psychologist and Joint Head of Child and Adolescent Eating Disorders Service, South London and Maudsley NHS Foundation Trust. Lead for Psychological Treatments, CAMHS, South London and Maudsley NHS Foundation Trust.
Christopher Fairburn	Principal Research Fellow and Professor of Psychiatry, University of Oxford. Honorary Consultant Psychiatrist, Oxford Health NHS Foundation Trust. Governor, MQ: Transforming Mental Health through Research, London. Governor, Oxford Mindfulness Foundation.
Lee Hudson	Consultant General Paediatrician at Great Ormond Street Hospital and Elern Meade Eating Disorder Unit, London.
Mike Hunter	Consultant Psychiatrist; Clinical Director (Inpatient Services); Associate Medical Director (Research & Strategy), Sheffield Health and Social Care NHS Foundation Trust.
Dasha Nicholls	Consultant Child and Young people Psychiatrist and Honorary Senior Lecturer, Great Ormond Street Hospital, London. Honorary Senior Lecturer, UCL Institute of Child Health, London

Name	Role
Jessica Parker	Lay member
Daniel Perry	Lay member
Ursula Philpot	Senior Lecturer in Nutrition and Dietetics, Leeds Beckett University. Freelance Consultant Dietician.
Susan Ringwood	Lay member
Mandy Scott	Mental Health Nurse, CAMHS Case Manager, NHS England, East Anglia Area Team.
Lucy Serpell	Clinical Psychologist and Senior Lecturer in Psychology of Eating Disorders, UCL. Clinical Lead for Eating Disorders, North East London NHS Foundation Trust.
Phillip Taylor	Consultant Dentist, Clinical Director of Dentistry and Oral and Maxillofacial Surgery and Clinical Lead in Restorative Dentistry, Barts and the London NHS Trust.
Dominique Thompson	GP and Director of the University of Bristol's Students' Health Service. Lead GP for Bristol for eating disorders.
Janet Treasure	Psychiatrist, Director of Eating Disorders Unit and Professor of Psychiatry. Chief Medical Advisor, Beat.
Hannah Turner	Consultant Lead Clinical Psychologist, Southern Health NHS Eating Disorders Service.
Christine Vize	Consultant Psychiatrist, Cotswald House Eating Disorders Service, Savemake Hospital, Marlborough, Wiltshire.
Glenn Waller	Professor of Psychology, Department of Psychology, University of Sheffield.

Table 2: NGA staff

Name	Role
Annabel Flint	Project Manager (from August 2016)
Linyun Fou	Research Assistant
Timothy Kendall	Clinical Advisor and Guideline Lead
Stephen Pilling	Clinical Advisor
Elena Marcus	Research Assistant
Leanne Saxon	Senior Systematic Reviewer
Eric Slade	Health Economist
Sarah Stockton	Information Scientist
Ifigeneia Mavranezouli	Senior Health Economist
Jo Wolfreys	Project Manager (until August 2016)
Katrina Blears	Project Manager (from September 2016)

Additional support was received from Nuala Ernest, Sabrina Naqvi and Katherine Andrea, Dr Sofia Dias TSU Director, University of Bristol and Edna Keeney, TSU Scientific Coordinator, University of Bristol.

1.3 Other versions of the guideline

NICE produce a number of versions of this guideline:

- The 'short guideline' lists the recommendations, context and recommendations for research
- 'Information for the public' is written using suitable language for people without specialist medical knowledge

- NICE Pathways brings together all connected NICE guidance.

1.4 Schedule for updating the guideline

For the most up-to-date information about guideline reviews, please see the latest version of the NICE guidelines manual available from the NICE website

2 Introduction

2.1 What is an eating disorder?

Eating disorders have been described as “a persistent disturbance of eating behaviour or behaviour intended to control weight, which significantly impairs physical health or psychosocial functioning” (Fairburn & Walsh, 2002), although more recent definitions have reduced the emphasis on ‘intent’. The relevant behaviours include: restriction of dietary intake; overeating with a sense of loss of control; and compensatory behaviours (e.g., vomiting, exercise, laxative abuse). These behaviours are accompanied by cognitive disturbances (e.g., overvaluation of weight; body image disturbance), emotional triggers and consequences (e.g., anxiety, shame) and social difficulties (e.g., isolation). The majority of individuals with eating disorders (80-85%) are not underweight, as detailed in following sections. However, regardless of weight status, patients with eating disorders are at increased physical risk as a result of starvation (e.g., cardiac problems; bone deterioration), binge-eating (e.g., physical damage; complications of excess weight, such as diabetes), purging (e.g., electrolyte imbalance) and mood (e.g., suicidality).

2.1.1 What do we know about the causes of eating disorders?

While there are many theories regarding the causes of eating disorders, the evidence to date is weak for any one causal factor. There is some evidence suggesting roles of genetic, neurobiological and sociocultural factors though their specificity and generalisability are limited at present. The limitations of the research are due in part to the low prevalence rate of the disorders, making early screening for risk factors impractical. However, the time between onset of the disorders (which is not always a clear point) and identification of the disorders is often several years (due to issues such as control and shame), meaning that determining causality is very difficult. Consequently, there is a greater focus on the maintaining factors, which can be identified much more readily. These include the cognitive, emotional, physical and social consequences of starvation, binge-eating, purging, etc., which have been built into models that have been tested in empirical studies and treatments.

2.1.2 The natural course of eating disorders

The majority of eating disorders have their origins in adolescence and young adulthood, though a substantial number of cases begin at younger or older ages. If left untreated (or if treated inadequately), the maintaining factors mean that many cases continue for decades, though severity can vary over time and there can be temporary periods of remission. Many cases will change from one diagnostic status to another, usually away from low-weight presentations (e.g., an individual whose diagnosis changes over several years from restrictive anorexia nervosa to binge/purge anorexia nervosa to bulimia nervosa).

2.1.3 Special issues regarding children with eating disorders

Children and adults can each have the full range of eating disorders (see following sections). However, there are important differences in how treatment should be focused at different age points. First, because child cases are usually those that are identified earlier in the process, the patients are more commonly underweight. Second, while early intervention is to be encouraged at all ages, the long-term physical complications of malnutrition that are specific at this age (e.g., growth, pubertal delay, dental problems, osteoporosis, fertility problems) makes such early intervention critical in children with eating disorders. Finally, the role of the family in addressing the disorders tends to be emphasised in younger cases, while the focus is more commonly on the individual in adult cases.

2.1.4 The cultural context of eating disorders in Western societies

Eating disorders are found globally, though they have been identified for longer in Western cultures. It can be argued that this is a result of such societies placing stress on the value of specific body types, particularly among women. There is some evidence that encroaching 'westernisation' is followed by a greater level of eating disorders in 'non-Western' societies, as those values are spread via a range of media and individuals strive to fit to the newly incorporated standards.

It is also important to consider whether there are cultural differences within societies. While it is clear that females are more likely to experience eating disorders than males, there has been a noticeable increase in identification of male sufferers in recent years. There is also a greater prevalence among those in 'at-risk' professions where there is a focus on body shape and weight (e.g., athletes, dances, models). Despite this, there is no evidence that other differences (e.g., ethnicity; socioeconomic status; sexuality) make it more likely that a member of any one group will have an eating disorder. However, cultural factors do play one important role. For example, professionals are less likely to identify an eating disorder if the sufferer does not fit the stereotype for such cases (e.g., non-Caucasian or male).

2.2 Epidemiology

2.2.1 Anorexia nervosa

The prevalence of anorexia nervosa in young females is around 0.3% (range = 0.2 – 0.8%; van Hoeken 2003 (van Hoeken et al., 2003)). Incidence rates range from 4.2 to 12.6 per 100,000 person-years (Currin 2005 (Currin et al., 2005); Micali 2013 (Micali et al., 2013); Steinhausen & Jensen, 2015 (Steinhausen and Jensen, 2015)). The incidence of anorexia nervosa among males is lower, at 1 per 100,000 person-years (Currin 2005). Anorexia nervosa is relatively rare in children under 13 years, with a reported incidence rate of 1.1 per 100,000 person-years (Nicholls 2011). It is also relatively rare among middle-aged and elderly women (Lapid 2010). The incidence of anorexia nervosa has remained stable over the past decade, with a peak age of onset of 15-19 years (Micali 2013).

Anorexia nervosa has the highest rate of mortality among all mental disorders. Its weighted crude mortality rate (CMR) is approximately 5.1 deaths per 1,000 person-years. The most common causes of death are suicide (20%) and cardiac complications (Arcelus et al., 2011).

2.2.2 Bulimia nervosa

The prevalence of bulimia nervosa is 1% in women and 0.1% in men (van Hoeken 2003). Incidence studies suggest an increase in diagnoses in the 1980s and mid-1990s, followed by a decrease in incidence in the late 1990s (Currin 2005), with stability since that time (Micali 2013). The incidence of bulimia nervosa showed a similar pattern of increase and decrease over that time period, peaking at 12.2 per 100,000 person-years in 1993 but reducing to 6.6 per 100,000 in 2000 (Currin 2005). Age at identification also appears to be decreasing, currently sitting among 15 – 24-year-old females (Smink et al., 2012), though it is not clear whether this reflects earlier detection or earlier age of onset. Bulimia nervosa has a weighted crude mortality rate of 1.74 per 1,000 person years and an overall standardised mortality ratio of 1.93 (Arcelus 2011).

2.2.3 Binge-eating disorder

The lifetime prevalence of BED is around 1.9% for women and 0.3% for men (Preti 2009). Compared with the other eating disorders, BED is more common in males and older individuals. BED is commonly associated with obesity, which in turn is associated with increased risk of mortality.

2.2.4 Atypical cases

Such cases are labelled as eating disorder not otherwise specified (EDNOS) or other specified feeding or eating disorder (OSFED) in different editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM) (APA 1994(APA, 1994); APA 2013). They form the largest single category of cases of eating disorders (e.g. Fairburn & Harrison 2003 (Fairburn and Harrison, 2003)). Lifetime prevalence of EDNOS is around 4.8% in young people and 4.6% in adults (Le Grange 2012). More cases have been identified in recent years (Micali 2013), from 17.7 per 100,000 in 2000 to 28.4 per 100,000 in 2009. Those authors showed that EDNOS is the most commonly diagnosed eating disorder in males (4.2 per 100,000) and its female to male ratio for EDNOS was 7.7:1 in 2009.

2.3 Anorexia Nervosa

Anorexia nervosa is characterised by the individual maintaining their weight as low as possible by controlling their overall energy balance. Primarily this is done through restricting food, but some patients exhibit other behaviours that both reduce energy intake and increase energy expenditure, such as exercise, taking medications (e.g., metabolic stimulants, laxatives, diuretics), exposing themselves to cold, purging, or chewing and spitting food.

Low weight is defined as less than minimally normal in adults, or minimally expected in children and young people. For adults, this would typically mean a body mass index (BMI) of less than 18.5. For children and young people, it would mean a BMI-for-age at less than the 5th percentile. While hormonal disturbance is a common feature, amenorrhea is no longer treated as a criterion among females and is only one of a number of markers of hormonal insufficiency due to low weight. Physical complications are detailed below.

Many individuals have poor insight into their condition and do not consider themselves to be ill. In other cases, there is a level of secrecy about the symptoms, for example hiding weight loss. This lack of clarity can have implications for treatment, as it is likely to delay the time between the onset of the illness and contact with medical professionals. This duration of untreated illness might indicate a marker of poor prognosis. Family or friends often play a role initiating the pathway into care, as they note that the individual loses weight becomes irritable and withdraws (especially around meal times or events that involve food). Patients might become selective about food and choose to prepare their own meals, absenting themselves at family meal times, saying they aren't hungry or that they have already eaten, or expressing a dislike for food they once enjoyed. They might develop unusual habits around meal times, for example only eating at certain times, always using the same cutlery or breaking the food into small pieces. Other behaviours that might be observed are regular monitoring of their shape or weight, with persistent weighing, measuring and mirror checking.

The preoccupation with food and weight is often related to a pursuit of thinness, or later in the development of the disorder, a fear of gaining weight. Patients often have low self-esteem and a drive for perfection, resulting in a desire for control. Some individuals believe that they do not deserve to eat, or that their behaviours will result in increased happiness and self-worth and/or positively influence how others perceive them. However, in some cases it is not possible to ascertain any reason for the poor intake of food, especially in children and young people who may find it difficult to articulate why they are restricting their eating.

2.4 Bulimia Nervosa

Bulimia nervosa is characterised by recurrent binge eating, extreme weight-control behaviour and overconcern about body shape and weight. ["Binges" are episodes of eating in which large amounts of food are consumed and there is a sense of loss of control at the time.] Bulimia nervosa is substantially more common than anorexia nervosa in the population, though services are often more focused on the care of anorexia nervosa.

The disorder generally starts in late adolescence or early adulthood. It usually begins in much the same way as anorexia nervosa (weight loss, experienced positively) but after some months or years the dietary restriction becomes punctuated by repeated binges (resulting in fear of weight gain). In most cases, these binges are followed by self-induced vomiting or the misuse of laxatives in an attempt to minimise the impact on body weight and between the binges there are continuing attempts to restrict eating. Despite this, any weight lost tends to be gradually regained and further weight gain is a common outcome. Sufferers are able to keep the problem secret for many years, as their appearance is generally unremarkable and they can eat normally in public. Most delay seeking help because of the shame associated with this way of eating.

Once fully developed, bulimia nervosa tends to be highly self-perpetuating, with adverse effects on mood, self-esteem and relationships. It also has carries substantial physical risks (see below). Most cases are in their 20s and about one in 10 is male.

2.5 Binge Eating Disorder

Individuals with binge eating disorder regularly binge on large amounts of food in a discrete period, with an accompanying sense loss of control. However, they do not fast or use other compensatory behaviours to a significant degree. Bingeing is accompanied by significant distress and can involve high levels of guilt and shame, eating in secret and eating despite not being hungry or until feeling uncomfortably full. Recurrent binges might occur against a background of a general tendency to overeat, or the individual might eat normally between binges. As a result, many (but not all) people with binge eating disorder are overweight or obese. Binge eating disorder is particularly common among individuals referred for bariatric surgery. The demographic distribution of binge eating disorder is distinctive compared to anorexia nervosa or bulimia nervosa, in that the majority of patients are middle-aged and about a third are male. The course of binge eating disorder is also quite different from other eating disorders. Rather than being persistent, it tends to remit and recur, with extended periods, often lasting many months, free of the eating disorder. It is generally recognised that treatment should be focused around reducing or eliminating bingeing rather than on weight loss, as a target of weight loss is likely to result in greater levels of binge eating.

2.6 Atypical eating disorders (eating disorders not otherwise specified)

Many patients present to eating disorder services with the features of anorexia nervosa or bulimia nervosa combined in such a way (or at a level of severity) that makes it impossible to make either diagnosis fully. There is no consensus over how to denote these presentations and they are often referred to as ‘atypical’ eating disorders, even though they are more common than the ‘typical’ states. The terminology and criteria have changed over the years in the DSM framework, including a change in labels from EDNOS to OSFED.

The atypical eating disorders resemble anorexia nervosa and bulimia nervosa, with an absence of some features or emphasis on some rather than others. For example, binge eating disorder used to be classified as an atypical eating disorder, while ‘purging disorder’ is currently classified as an example of an atypical eating disorder. All such disorders share the same overconcern about eating, shape and weight and the same tendency to engage in persistent and extreme dieting and other forms of weight control or disordered eating behaviour, for example binge eating, purging and restriction. Body weight tends to be low if the dietary restriction is marked. The atypical eating disorders do not include ‘avoidant restrictive food intake disorder’ in which weight and shape concerns are not a feature.

Most people with an atypical eating disorder are female and in their 20s. Many have a history of anorexia nervosa, bulimia nervosa or both, reflecting the diagnostic migration that is common among the eating disorders. The atypical eating disorders are as impairing as

bulimia nervosa, with a similar duration and impact on everyday functioning. Atypical eating disorders are also common in those with other primary mental health diagnoses.

2.7 Physical Complications

The physical complications of eating disorders are common, especially in anorexia nervosa, where physical causes are implicated in around half of patients who die from an eating disorder. Such complications arise as a result of malnutrition, binge-eating and compensatory behaviours (including vomiting, diuretic and laxative misuse), misuse of other drugs or alcohol and excessive exercise. Physical complications can be considered as either acute and longer term/chronic complications.

Acute complications may relate to how underweight the individual is, but also to the rate of weight loss. They include effects in a range of physiological systems:

- Cardiovascular effects include low heart rate, low blood pressure, postural hypotension, poor peripheral circulation or conduction problems (e.g., prolonged QTc interval, arrthymias). Pericardial effusion and heart failure may also occur. Myocardial fibrosis may be a cause of sudden cardiac death in these patients.
- Haematological effects due to a reduction in bone marrow activity and the quantity and quality of blood cells, in particular, reduced amounts of platelets and neutrophils. This can result in an impaired immune response, anaemia and a higher the risk of stroke.
- Metabolic effects include electrolyte abnormalities (particularly low potassium, sodium, phosphate and magnesium levels), vitamin and iron deficiencies and a reduction in bone mineral density, which can be severe enough to cause osteoporosis even in young patients) This results in an increased risk of fractures. There may be a loss of stature in fully grown patients. Patients may suffer from hypothermia and temperature regulation is impaired. Skin may be dry or easily bruised and pressure sores can develop, which are slow to heal.
- Muscular effects are found throughout the body as fat stores are depleted and protein is used as fuel, producing complications such as muscle weakness and proximal myopathy, pain in skeletal muscles and joints, cardiomyopathy and delayed gut transit, causing discomfort, bloating and constipation. Lung function may be compromised (e.g., emphysematous changes). Pulmonary oedema can develop. Hepatitis and more rarely pancreatitis, can occur.

Effects on other systems are widespread. With severe weight loss, the brain is reduced in size with widened ventricles and sulci. In such cases, cognitive deficits become noticeable as weight drops, with memory and concentration impairment. Peripheral neuropathy can occur. Frequent vomiting by people with an eating disorder can cause both short- and long-term damage to dental health and in particular to appearance, which exacerbates body image related psychological issues. Dental damage can occur quickly, after as little as only a few months of frequent vomiting and is caused mainly by stomach acids washing over the tooth surfaces thereby causing dissolution of the dental enamel and dentine, which is often termed ‘erosion’. Eating too much fruit and drinking carbonated drinks can also cause similar problems. Another common appearance in patients with eating disorders, in particular the bulimic group, is an enlargement of the salivary glands, particularly the parotid gland (up to 36% enlarged) which gives an appearance of swollen cheeks.

Chronic/longer term complications include:

- Growth and development in children and young people. Growth may be slowed or cease, so that the person does not reach their potential height. Puberty can be delayed, incomplete or not start.

- Reproductive system effects include amenorrhoea in females and loss of potency in males. Fertility may be sub-optimal, even in recovered patients. There is a higher rate of obstetric complications in women who do conceive.
- Bone mineral density is commonly affected in males and females with longer standing low weight (generally longer than six months), especially when being underweight coincides with key periods of bone mineral acquisition (e.g., adolescence). Patients with anorexia nervosa have a greater risk of contemporary and lifetime fractures.
- Weight gain and mechanical effects of bingeing, weight gain and obesity are also important to note in many cases. In particular, type 2 diabetes can develop with associated complications. There can also be pain and mobility problems associated with such consequences of disturbed eating patterns. There can also be other mechanical complications associated with bingeing and purging behaviours, including damage and tears to the gastrointestinal tract.

2.8 Comorbidities

In addition to the physical complications outlined above, psychiatric comorbidity is common in eating disorders. These include:

- a range of anxiety disorders (particularly social anxiety/phobia and obsessive compulsive disorder), which commonly predate the eating disorders but which are exacerbated by the eating pathology
- mood disorders, such as depression (with attendant suicidal thoughts and behaviours), which are often secondary to the effects of malnutrition (including weight loss, but also resulting from low carbohydrate intake resulting in low serotonin levels), loss of control over eating behaviours, shame, etc.
- Compulsive behaviours, such as skin-picking, hair-pulling and compulsive exercise (even in the absence of obsessive-compulsive disorder)
- Impulsive behaviours, such as self-harm, alcohol use, drug use and aggression

Other disorders are less common, for example psychotic disorders, but are still present on occasion.

Personality-level pathology is also commonly comorbid with eating disorders. This can include:

- pathological levels of perfectionist traits, which again commonly predate the eating disorder but are exacerbated by the consequences of the eating pathology
- autistic spectrum disorder and attention deficit hyperactivity disorder appear to be over-represented in patients with eating disorders and might be vulnerability factors
- personality disorders, particularly obsessive-compulsive personality disorder, borderline personality disorder and avoidant personality disorder

However, given that these personality-level comorbidities are often alleviated by successful treatment for the eating disorder (particularly dietary and weight normalisation), their role as triggers or maintaining factors for eating disorders is uncertain.

2.9 The treatment and management of eating disorders in the NHS

Variation in existing provision

There is wide variation in how eating disorders are treated and managed in the NHS. This variation can be seen across the whole care pathway from the initial referral, through primary psychological services, outpatient child and young people mental health services (CAMHS),

as well as across adult services. Some teams provide generic support, while others offer more specialist input. This variation is also applicable to inpatient services, with some service users being treated in specialist eating disorder units and others being admitted to generic mental health units.

Significant geographical inconsistencies exist, with different areas providing widely contrasting services. This pattern of geographical difference is likely to be influenced by variations in funding across the country, as well as differences in referral criteria to specialist services. For example, NHS England Specialised Commissioning has contracts with both NHS and private inpatient units, with admission criteria largely dependent on where the service user lives.

Primary and secondary care pathways

Most patients with eating or feeding symptoms initially present in primary care, though not necessarily with these symptoms as their presenting problem. The nature of an eating disorder and the inherent ambivalence (e.g., control issues, shame) often make it very unlikely that the individual with the eating disorder will seek help, and so treatment may initially be sought by a family member or carer, particularly in the case of children and young people. Clinicians working in primary care will have very different levels of experience, skill and confidence in diagnosing, assessing and managing such patients. Some primary care doctors and nurses lack the experience to offer the robust initial assessment that is needed to ensure that appropriate onward referral is made. Furthermore, referral pathways from primary care into secondary services will vary for adults, although these pathways are now being standardised for patients under 18. In some areas, much more specialised primary care assessment and triage services have developed in response to increasing demand from patients and link into very well developed secondary care pathways. Following an initial assessment GPs will consider referral to specialist secondary services where these are available locally. They will also consider medication, if relevant, and follow up as appropriate. Physical investigations may be carried out in primary care if indicated.

Dentists are often the first professionals to recognise eating problems, as a result of identifying patterns of dental erosion. Recognition of the effects of eating disorders is taught by all dental schools in the UK. Dental treatment can affect outcomes and appearance is important for these patients. Current additive techniques with tooth coloured composite resins provide simple, aesthetic, effective and protective treatment.

Access pathways for children and young people

In relation to children and young people, there have historically been inconsistencies in service provision across the country, some young people being treated in generic CAMHS teams rather than specialist eating disorder teams. In areas where there are specialist services for younger patient, some will provide treatment and management of eating disorders up until the age of 18 years. In some areas, the treatment and management of 16 – 18-year-olds will be delivered within adult services, whilst a small number of specialist services take patients aged 14 years and over. In 2015, NHS England published a commissioning guide to Access and Waiting Times, designed to redress some of this variation in care for younger patients by setting standards for access and waiting times to receive an evidence based psychological intervention (NCCMH 2015(NCCMH., 2015)).

Very few age inclusive teams that cover the full period of biological maturation (12-25 years) exist across the country. Service access criteria related to age are likely to impact important aspects of care, such as the extent to which parents/family are involved and the level of responsibility the young person is encouraged to take. It is best practice for CAMHS and adult services to have a clear transition protocol thus enabling a smooth transition between services when required.

Access pathways for adults

Access to adult services tends to be either directly via primary care or through secondary adult mental health services. As in CAMHS, referral criteria vary, often influenced by funding and historical practice. In some areas, services will accept referrals of a wide range of eating disorder presentations, but in areas where funding is more limited, services may have strict referral criteria, such as a BMI cut-off or therapy only being offered to those within a specific BMI range.

A further transition issue is the care of University students with eating disorders, as they are often based at home for part of the year and at University for the remainder. Therefore, these individuals require particularly high levels of coordination between services (and with funders) to ensure that their access to care and their therapy is not compromised.

Treatment diversity

Waiting times for treatment vary significantly. Alongside offering out-patient treatment, some community services will also provide more intensive day-care treatment, but again there is significant variation across the country. Specialist in-patient treatment for children and young people, as well as adults, is currently funded by NHS England. In-patient beds tend to be accessed via specialist community services.

There is great variety in current therapeutic choices (interventions) offered and in the skill-set of the clinicians delivering these interventions. At times this can leave both patients and healthcare professionals in some confusion as to how best to manage these clinical cases. It is a clinical area in which a diverse range of therapies have evolved over the last few years, adding to the potential for challenging medical, nursing and psychotherapeutic decision-making. However, it is clearly the case that evidence-based therapies are substantially under-used, at the child and adult levels. The Access and Waiting Times initiative aims to address this variation in the type of therapy as well as the referral pathway. At present this only applies to those under age of 18.

Inpatient care

Inpatient care, particularly for anorexia nervosa, was central to the management of eating disorders in the twentieth century. However, there has been a move to more community-centred care with hospital care restricted to the patient group with severe medical risk and/or a failure to respond to outpatient care. The main goal of inpatient treatment is to reduce medical risk by improving nutrition. This usually involves meals supervised by nurses. Facilities with less than 24 hour per day care such as partial hospitalisation or day care are used as an alternative to inpatient care, or as the second phase in a form of stepped care. As the effect of treatment is often transient, inpatient care is less often used for bulimia nervosa unless comorbidity with problems such as diabetes increase the medical risk.

The history of inpatient treatment

Hospital admission so that nurses could support eating and restore weight to normal was advocated as treatment for anorexia nervosa by Sir William Gull in the 19th century. A century later, care was mainly provided within specialised psychiatric units. Relapse post discharge was common but a seminal study of aftercare found that, in young people in the early stage of illness (less than three years), family-based psychotherapy reduced relapse more effectively than individual therapy. However, neither form of therapy reduced relapse in those with a longer or later-onset form of illness.

Theoretical Background

The theoretical underpinning of the nursing approach to refeeding has changed over time. Gull described nurses using "moral authority". Later behavioural principles were followed with a reward given or privileges re-instated contingent on weight gain. However, evidence from a cohort study reported that a strict behavioural approach was less acceptable to nurses and patients and produced an equivalent amount of weight gain to a more lenient approach. Meal support can be problematic, with meal times reported as taxing and confusing. In a recent survey of nurses in specialised UK units, some patients and nurses described meals as a battle. Such a conflictual meal environment may deepen negative memories associated with food and embed eating disorders behaviours more deeply. The importance of staff skills was emphasised by a study that found teaching nurses to use an intervention devised to train family members in the skills of meal support led to a more cost effective inpatient stay. Moreover such a joint approach may bridge the transition period and allow changes to be sustained post-discharge. Thus, recent trends have seen inpatient teams moving away from taking over ‘in loco parentis’ to working in collaboration with parents/carers to transfer skills from the inpatient to the home environment. This is particularly important in the care of young people, where responsibility often lies with adults to take primary responsibility for managing meal times. The location of responsibility shifts with age and duration of illness and underpins the theory behind the psychological therapies utilized at different stages of illness.

Clinical Practice

Levels of anxiety are high before, during and after meals. This may be marked by intense emotional displays, but more often patients have a ‘poker face’ with restricted facial expression of emotions. This blocks an empathic reaction from staff who can become frustrated and hostile. On the other hand, if others recognise the terror associated with food, they may be drawn into accommodating the illness enabling eating disorder behaviours to persist. Thus, careful planning and supervision is needed to achieve a balance between avoidance and coercion. Eating is non-negotiable. On the other hand, the form and content of food-related activities can be individualised to a degree. Advance planning and review and a rule of no negotiations during meals themselves are helpful strategies. Implementation interventions, i.e., “if ...then...” plans, can be helpful. For example, "If you are not able to eat all of your meal, then you will have liquid food replacement. If you are taking too long, then I will give you a reminder at half time and every 10 minutes". The skills of motivational interviewing (warmth, open reflections, side stepping resistance) are particularly helpful for managing the ambivalence and resistance that meals evoke.

Tube Feeding

Nasogastric (NG) feeding is recommended over other enteral routes or parenteral nutrition when nutrition cannot be taken orally. NG feeding is relatively common in children and young people with anorexia nervosa but not so often in adults in UK, although the quality commission as well as clinical guidance aim to ensure all units are equipped and competent to do it when needed. It is important that patients, parents and carers are involved and understand the rationale for its use as a way to provide adequate and safe quantities of calories where patients appear unable to do this orally. Conversely, efforts must be given to avoid either explicit or implicit punitive application.

The lower threshold for use in young people is multifactorial and includes the fact that the impact of malnutrition in young people can be more acute and have lasting consequences on growth and development. The law also puts emphasis on adults being responsible for the care of young people up to the age of 18, taking into consideration increasing autonomy and capacity. Additionally, there is evidence that early weight restoration has an impact on outcome, justifying an aggressive approach to refeeding in the early stages of the illness. However, randomised trials looking at how this weight gain is achieved (NG versus oral) have not been undertaken in young people. A case series of young people fed by NG tube

found that, at follow up two thirds of patients thought the intervention had been necessary, while the remaining third still had negative views. A proportion of young people (and adults) want to be tube fed as it can be preferable not to feel responsible for eating. Clinical practice varies widely on use of NG feeding in children and young people, with some units using it universally and others only as a last resort.

In adults, case-control designs suggested that nasogastric feeding does not have a negative impact on outcome and may be superior to oral feeding, particularly with the group with binge/purge symptoms (Rigaud & Brayer 2007 (Rigaud et al., 2007a), Rigaud & Brondel 2007 (Rigaud et al., 2007b), Rigaud & Pennacchio 2009 (Rigaud et al., 2009), Rigaud & Brayer 2011 (Rigaud et al., 2011)).

Rate of refeeding

The NICE guidelines produced to manage refeeding syndrome (not in the eating disorder context) suggested that people should start on a very low calorie intake supplemented by vitamins. There is now concern in the field that if these guidelines are applied to people with eating disorder they may be underfed. Several studies have been examining this question (Madden 2015, Redgrave 2015 (Redgrave et al., 2015)). The only randomized controlled trial of refeeding to date found that young people randomized to high energy intake (1200 kilocalories versus 500 kilocalories) had greater weight gain than those on a more conservative regime, without a statistically different increase in risk (O'Connor, 2016 (O'Connor et al., 2016a)). There is considerable variation in the amounts of starting calories internationally and between units.

Supplemental Treatments

Multidisciplinary interventions are regarded as a necessary part of a high quality eating disorder inpatient program. In addition to nutrition, these target the range of eating disorder features, for example cognitive and emotional style or body image work. A recent systematic review examined supplementary treatments to inpatient care. Four randomised controlled trials compared an antidepressant with placebo and found minimal effects. Four studies compared antipsychotic drugs with placebo and, again, the overall effect was small; the largest effect was found in young people with anorexia nervosa. Minimal effects were also found in the four studies using randomised controlled designs to compare psychological treatments; possibly, the detrimental effects of starvation on brain plasticity and functioning may reduce the experiential benefits from psychotherapy.

For young people, the inpatient environment can be a place to practice family meals and develop a structure for mealtime management with staff support.

It is challenging to develop a strong evidence base regarding interventions on inpatient settings, as randomisation to different forms of intervention within the inpatient setting can be problematic and it is difficult to find effects over and above those resulting from standard care. On the other hand, an admission can provide time, space and motivation for psychological change to begin. Home leave is important aspect of care, providing opportunities for skills to be put into practice. The value of inpatient care in those with severe and enduring anorexia nervosa is more questionable, other than as a life saving measure when appropriate.

Admission Criteria

There is no international agreement on the admission criteria for intensive care and the thresholds specified in national guidelines vary. Healthcare settings also differ internationally; in some paediatrics/medical models of care predominate, in others eating disorders form part of generic mental health services, and yet others specialist eating disorder units are the norm. In part, admission criteria depend on the facilities available and the amount of risk they

are able to manage. Patients with extreme medical risk and multiple organ failure are usually admitted to general medical hospitals. In the UK, the "Management of Really Sick Patients with Anorexia Nervosa MARSIPAN" (adult and junior) protocol has been developed to describe the care pathway for such cases and to optimise the liaison between physical and psychiatric care (Royal College of Psychiatrists 2012(Psychiatry, 2012)).

Discharge Criteria

The traditional goal of inpatient care was to restore weight to normal. The underlying assumption was that normal physiology and eating habits would then resume. Indeed, low weight at discharge increases the likelihood of relapse and readmission. However, the outcome of inpatient care is confounded by many factors such as the level of motivation and randomised controlled trials (such as those described below) are essential to interpret findings. Shorter periods of inpatient stay and lower discharge BMIs are part of current practice. For example, in the US, over a 14 year period, the average length of inpatient admission decreased from 149.5 days to just 23.7 days, and discharge BMIs decreased over the years from 19 kg/m2 (full restoration) to 17 kg/m2 (medical stabilisation). Results from two recent randomised trials validate this practice for young patients, as short admissions for medical stabilisation, rather than normalisation of weight, produced a similar improvement in symptoms. Moreover, the shorter (medical stabilisation) admissions were associated with less service use in the year following discharge. It remains to be seen whether such an approach is useful for people in the enduring stage of illness.

Aftercare

The relapse rate following inpatient care ranges from 20-50% and approximately a third of patients are readmitted in the following year. Psychological interventions for patients and/or carers delivered face-to-face or through various form of technology have been found to reduce the rate of relapse. On the other hand, dietary advice or medication had no impact. A systematic review of the longer term follow up of inpatient/aftercare treatment in two RCTs (Eisler & Dare 1997 (Eisler et al., 1997), Godart & Berthoz 2012 (Godart et al., 2012)) have found that involving families in the post intensive treatment aftercare of young people improves outcomes. One pilot RCT has found that involving carers in the aftercare of adults improves outcomes (Hibbs 2015(Hibbs et al., 2015), Magill 2016 (Magill et al., 2016)).

2.10 Use of health service resources

In keeping with the variable patterns of service provision and delivery, the level of resource allocation differs across settings in the UK. There are also substantial differences across countries.

UK resources

Beat (2015) has reported on the costs associated with eating disorders in the UK. Their calculation included three cost categories:

- treatment costs (including both NHS and private providers) amount to £8,850 per individual per annum (in likely 2011 prices);
- direct financial burden to sufferers and carers (excluding any payments for private treatment) amount to £4,300 per annum;
- indirect financial burden on sufferers and carers, resulting from disruption to education, employment and professional development amount to £10-15.4 thousand per annum (according to educational and employment status of the sufferer).

Thus, the cost to health service resources of the treatment element was £8,850 per individual per annum. However, that cost will be much higher for in-patients and lower for out-patients.

Using a broader basis for assessing treatment costs, including costs to the educational and voluntary sectors, Byford and colleagues (2007) showed that the cost per annum of treating young people was £26,738 in a specialist outpatient service and £34,531 in a specialist in-patient unit (in 2003/2004 prices). However, both were less expensive than treating them in general CAMHS services (annual cost of £40,794).

Comparison with resource use in other countries

Comparative resource costs are provided for other countries, with the provision that the baseline costs of care vary across countries, for example, the level of staffing and staff costs and the greater use of in-patient resources in some countries, reducing immediate comparability. In addition, the following figures need to be corrected upwards to allow for the year in which they were assessed.

- In Germany, the cost of a treatment episode was €5,251 for anorexia nervosa and €3,265 for bulimia nervosa (Haas et al., 2012); in likely 2011 Euros.
- In an earlier German study, the annual the cost was €5,300 for anorexia nervosa and €1,300 for bulimia nervosa (Krauth and Buser, 2002), possibly indicating that the length of treatment episode for bulimia nervosa has increased or become more hospital-based over the decade between the studies (in likely 2001 Euros).

Overall resource costs have been shown to be higher than the comparable costs for non-eating disordered individuals:

- Overall, in the US, those with eating disorders spent $1,869 per annum more on healthcare costs than those without eating disorders (Samnaliev et al., 2015) in likely 2014 US dollars.
- In the US, healthcare costs for eating disorders were substantially higher for eating disorders ($37,125 over five years) than for a non-disordered group ($13,725 over the same time period), though the figures for people with eating disorder were similar to those for people with depression (Mitchell et al., 2009) in likely 2008 US dollars.
- In the US, one year healthcare costs are higher for people with binge-eating disorder than for people with EDNOS (by as much as $5,589) and higher than non-eating disordered controls (by as much as $18,152) (Bellows et al., 2015) in likely 2014 US dollars.

Inpatient care is particularly costly. For example:

- In Germany, mean costs over three months of care for anorexia nervosa patients was costed at €5,866, but most of this accounted by the cost of hospitalisations (Stuhldreher et al., 2014) in likely 2014 prices.
- Health service costs in the US for acute inpatient care in a general hospital were $12,141 for anorexia nervosa and $8,697 for bulimia nervosa (O'Brien and Ward, 2003) in likely 2002 US dollars.
- Considering different eating disorders (Striegel-Moore et al., 2004), inpatient care has been shown to be considerably more expensive than outpatient care:
  - anorexia nervosa: $16,740 (inpatient) versus $2,242 (outpatient);
  - bulimia nervosa: $9,380 (inpatient) versus $1,848 (outpatient);
  - EDNOS: $12,748 (inpatient) versus $2,146 (outpatient); in likely 1999 US dollars.
- Among children and young people with eating disorders in the US, the mean cost of an inpatient stay (mean duration of 18.4 days) was $10,019 (Robergeau et al., 2006) in likely 2005 US dollars.
- In a further study of young people treated in the US, the annual cost of treatment was $33,105 (Lock et al., 2008) in likely 2007 US dollars. However, the medical element (inpatient and outpatient monitoring) accounted for 81% of that cost.

- Again in the US, Lock and colleagues (2003) costed a mean period of 23.2 days of inpatient treatment for young people with anorexia nervosa was $25,750; in likely 2002 US dollars.
- In 22 residential eating disorder treatment programmes (including anorexia nervosa and bulimia patients) in the US, the mean length of stay was 83 days and the mean cost per person was $79,348 (Frisch et al., 2006) in likely 2005 US dollars.

Non-health-service costs are also significant. For example:

- Disability payments to individuals with anorexia nervosa in British Columbia, Canada were calculated to be $101.7 million per year, which is approximately 30 times the cost of all tertiary care services of eating disorders in the province (Su and Birmingham, 2003) in likely 2002 Canadian dollars.

Additional resources

There are also substantial healthcare resource costs that are related to the use of non-eating disorder services. For example:

- In the US, people with binge-eating disorder had higher generic healthcare costs ($1,379 in six months) than age- and gender-based norms (Grenon et al., 2010) in likely 2009 US dollars.
- In the US, substantial amounts of herbal and alternative medications were used, with an estimated cost of $33.88 per individual per month (Steffen et al., 2006) in likely 2005 US dollars.

Conclusion

To summarise, the health service costs per eating-disordered patient in the UK are approximately £8,850 per individual per annum year. This cost refers to people with eating disorder, but do not distinguish across the type of eating disorder. However, the cost of anorexia nervosa treatment is likely to be higher than that of other eating disorders as this relatively small group of cases receives substantially more in-patient care. These figures are not directly comparable with those from other countries, though such figures do support the conclusion that eating disorders in particular anorexia nervosa result in substantial economic burden on the healthcare resources. Efficient use of available healthcare resources will maximise the health benefit for people with eating disorders and can potentially reduce costs to the healthcare system and society as a whole.

3 Methods used to develop this guideline

3.1 What is a NICE clinical guideline?

National Institute for Health and Care Excellence (NICE) clinical guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the NHS – from prevention and self-care through primary and secondary care to more specialised services. We base our clinical guidelines on the best available research evidence, with the aim of improving the quality of healthcare. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific review questions.

NICE clinical guidelines can:

- provide recommendations for the treatment and care of people by healthcare professionals
- be used to develop standards to assess the clinical practice of individual healthcare professionals
- be used in the education and training of healthcare professionals
- help patients to make informed decisions
- improve communication between patients and healthcare professionals.

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:

- The guideline topic is referred to NICE from the Department of Health.
- Stakeholders register an interest in the guideline and are consulted throughout the development process.
- The scope is prepared by the former National Collaborating Centre for Mental Health (NCCMH (National Guideline Alliance (NGA) from April 2016)).
- The NGA establishes a Guideline Committee.
- A draft guideline is produced after the group assesses the available evidence and makes recommendations.
- There is a consultation on the draft guideline.
- The final guideline is produced.

The NGA and NICE produce a number of versions of this guideline:

- The ‘full guideline’ contains all the recommendations, together with details of the methods used and the underpinning evidence.
- The ‘short guideline’ lists the recommendations, context and recommendations for research.
- ‘Information for the public’ is written using suitable language for people without specialist medical knowledge.
- NICE Pathways brings together all connected NICE guidance.

3.2 Remit

NICE received the remit for this guideline from the Department of Health. It commissioned the NGA to produce the guideline.

The remit for this guideline is to develop a clinical guideline on eating disorders: Core interventions in the treatment and management of anorexia nervosa, bulimia nervosa and related eating disorders (update).

3.3 Who developed the guideline?

A multidisciplinary committee comprising healthcare professionals, researchers and lay members developed this guideline (see the list of group members and acknowledgements).

NICE funds the NGA and thus supported the development of this guideline. The committee was convened by the NGA and chaired by Anthony Bateman in accordance with guidance from NICE.

The group met every four to six weeks during the development of the guideline. At the start of the guideline development process all group members declared interests including consultancies, fee-paid work, shareholdings, fellowships and support from the healthcare industry. At all subsequent group meetings, members declared arising conflicts of interest.

Members were either required to withdraw completely or for part of the discussion if their declared interest made it appropriate. The details of declared interests and the actions taken are shown in Appendix C.

Staff from the NGA provided methodological support and guidance for the development process. The team working on the guideline included a guideline lead, a project manager, systematic reviewers, health economists and information scientists. They undertook systematic searches of the literature, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate and drafted the guideline in collaboration with the committee.

3.4 What this guideline covers

3.4.1 Groups that will be covered

This guideline covers the following groups:

- Children, young people and adults with an eating disorder, including atypical presentations, or a suspected eating disorder.

3.4.2 Key clinical issues that will be covered

The following clinical issues will be covered in this guideline:

- Identification, assessment and monitoring:
  - recognition and early identification of eating disorders (including formal recognition tools)
  - assessment in people with an eating disorder (including formal assessment tools)
  - monitoring in people with an eating disorder.
- Interventions to treat eating disorders through all phases of the disorder including:
  - psychological interventions, including low-intensity interventions such as self-help and Internet-based therapies, high-intensity interventions such as family therapy and family-based treatments and individual therapies such as psychodynamically informed therapies, cognitive behavioural therapy (CBT), interpersonal psychotherapy and behavioural interventions
  - pharmacological interventions (note that guideline recommendations will normally fall within licensed indications; exceptionally, and only if clearly supported by evidence, use outside a licensed indication may be recommended. The guideline will assume that

prescribers will use a drug's summary of product characteristics to inform decisions made with individual patients)
    - nutritional interventions, including tube feeding
    - other physical interventions, such as transcranial magnetic stimulation and physiotherapy.
- The management of physical health problems caused by an eating disorder.
- Interventions for eating disorders in the context of common physical and psychological comorbidities.
- Interventions to support families and carers.
- Organisation and delivery of services to support practitioners in the effective and competent delivery of interventions.
- Consent and compulsory treatment.
- Note that guideline recommendations will normally fall within licensed indications. Exceptionally, and only if clearly supported by evidence, use outside a licensed indication may be recommended. This guideline will assume that prescribers will use a drug's summary of product characteristics to inform decisions made with individual patients.
- For further details please refer to the scope in Appendix A and review questions in Appendix D.

3.5 What this guideline does not cover

3.5.1 Groups that will not be covered

- People with disordered eating because of a separate physical or other primary mental health problem of which a disorder of eating is a symptom
- People with feeding disorders, such as avoidant restrictive food intake disorders
- People with obesity without an eating disorder.

3.5.2 Clinical areas that will not be covered

- The diagnosis or treatment of people with disordered eating in the context of a separate physical or other primary mental disorder of which a disorder of eating is a symptom (such as loss of appetite in depression)
- The management of loss of appetite, psychogenic disturbance of appetite or other conditions that involve significant weight loss but which are due to known physical illness.
- The management of the wider range of eating disturbances typically but not exclusively occurring in children (for example, Pica or avoidant restrictive food intake disorders such as food avoidance emotional disorder or picky/selective eating).
- Obesity in the absence of an eating disorder.

3.6 Relationships between the guideline and other NICE guidance

3.6.1 Related NICE guidance

- Patient experience in adult NHS services (2012) NICE guideline CG138
- Service user experience in adult mental health (2011) NICE guideline CG136
- Medicines Adherence (2009) (2009) NICE guideline CG76
- Nutritional Support in Adults (2006) NICE guideline CG32
- Behaviour change: individual approaches (2014) NICE guideline PH49

- Behaviour change: the principles for effective interventions (2007) NICE guideline PH6

3.7 Methodology

The development of this guideline followed The Guidelines Manual (NICE, 2014(NICE, 2014a)). A team of health and social care professionals, lay representatives and technical experts known as the Guideline committee with support from the NGA staff, undertook the development of a person-centred, evidence-based guideline. There are eight basic steps in the process of developing a guideline:

- Define the scope, which lays out exactly what will be included (and excluded) in the guidance.
- Define review questions that cover all areas specified in the scope.
- Develop a review protocol for each systematic review, specifying the search strategy and method of evidence synthesis for each review question.
- Synthesise data retrieved, guided by the review protocols.
- Produce evidence profiles and summaries using the GRADE system.
- Consider the implications of the research findings for clinical practice and reach consensus decisions on areas where evidence is not found.
- Consider the economic costs for each review question.
- Answer review questions with evidence-based recommendations for clinical practice.

The clinical practice recommendations made by the committee are therefore derived from the most up-to-date and robust evidence for the clinical and cost effectiveness of the interventions and services covered in the scope. Where evidence was not found or was inconclusive, the committee discussed and attempted to reach consensus on what should be recommended, factoring in any relevant issues. In addition, to ensure a service user and carer focus, the concerns of service users and carers regarding health and social care have been highlighted and addressed by recommendations agreed by the whole committee.

3.8 Developing the scope

Clinical guideline topics are referred from the Department of Health or the NHS Commissioning Board and the letter of referral defines the remit, which defines the main areas to be covered; see The Guidelines Manual (NICE, 2014) for further information. The NGA developed a scope for the guideline based on the remit (see Appendix 1). The purpose of the scope is to:

- provide an overview of what the guideline will include and exclude
- identify the key aspects of care that must be included
- set the boundaries of the development work and provide a clear framework to enable work to stay within the priorities agreed by NICE and the NGA and the remit from the Department of Health
- inform the development of the review questions and search strategy
- inform professionals and the public about expected content of the guideline
- keep the guideline to a reasonable size to ensure that its development can be carried out within the allocated period.

An initial draft of the scope was sent to registered stakeholders who had agreed to attend a scoping workshop. The workshop was used to:

- obtain feedback on the selected key clinical issues
- identify which population subgroups should be specified (if any)
- seek views on the composition of the Guideline Committee

- encourage applications for committee membership.

The draft scope was subject to consultation with registered stakeholders over a four week period. During the consultation period, the scope was posted on the NICE website. Comments were invited from stakeholder organisations The NGA and NICE reviewed the scope in light of comments received and the revised scope was signed off by NICE.

3.9 The Guideline Committee

During the scope consultation phase, members of the committee were appointed by an open recruitment process. Committee membership consisted of: professionals in psychiatry, clinical psychology, nursing, social work, general practice; academic experts in psychiatry and psychology; and service users and carers. The guideline development process was supported by staff from the NGA, who undertook the clinical and health economic literature searches, reviewed and presented the evidence to the committee, managed the process and contributed to drafting the guideline.

3.9.1 Guideline committee meetings

There were 12 committee meetings, held between May 2015 and July 2016. During each day-long committee meeting, in a plenary session, review questions and clinical and economic evidence were reviewed and assessed and recommendations formulated. At each meeting, all committee members declared any potential conflicts of interest (see Appendix B) and service user and carer concerns were routinely discussed as a standing agenda item.

3.9.2 Service users and carers

The committee included one carer member and two service users who contributed as full committee members to writing the review questions, providing advice on outcomes most relevant to service users and carers, helping to ensure that the evidence addressed their views and preferences, highlighting sensitive issues and terminology relevant to the guideline and bringing service user research to the attention of the committee. Input from both service users and carers was central to the development of the guideline and they contributed to writing the guideline's introduction and the recommendations from the service user and carer perspective.

3.9.3 Expert advisers

No Expert Advisors were used in the development of this guideline.

3.9.4 National and international experts

National and international experts in the area under review were identified through the literature search and through the experience of the committee members. These experts were contacted to identify unpublished or soon-to-be published studies, to ensure that up-to-date evidence was included in the development of the guideline. They informed the committee about completed trials at the pre-publication stage, systematic reviews in the process of being published, studies relating to the cost effectiveness of treatment and trial data if the committee could be provided with full access to the complete trial report. Appendix E lists researchers who were contacted.

3.10 Review protocols

Review questions drafted during the scoping phase were discussed by the committee at the first few meetings and amended as necessary. The review questions were used as the

starting point for developing review protocols for each systematic review (described in more detail below). The final list of review questions can be found in Appendix F.

For questions about interventions, the PICO (Population, Intervention, Comparison and Outcome) framework was used to structure each question (see Table 3).

Table 3: Features of a well-formulated question on the effectiveness of an intervention – PICO

Population:	Which population of service users are we interested in? How can they be best described? Are there subgroups that need to be considered?
Intervention:	Which intervention, treatment or approach should be used?
Comparison:	What is/are the main alternative/s to compare with the intervention?
Outcome:	What is really important for the service user? Which outcomes should be considered: intermediate or short-term measures; mortality; morbidity and treatment complications; rates of relapse; late morbidity and readmission; return to work, physical and social functioning and other measures such as quality of life; general health status?

Questions relating to case identification and assessment tools and methods do not involve an intervention designed to treat a particular condition and therefore the PICO framework was not used. Rather, the questions were designed to pick up key issues specifically relevant to clinical utility, for example their accuracy, reliability, safety and acceptability to the service user.

In some situations, review questions related to issues of service delivery are occasionally specified in the remit from the Department of Health. In these cases, appropriate review questions were developed to be clear and concise.

For each topic, addressed by one or more review questions, a review protocol was drafted by the technical team using a standardised template (based on the PROSPERO database of systematic reviews in health), review and agreed by the committee (all protocols are included in Appendix F).

To help facilitate the literature review, a note was made of the best study design type to answer each question. There are five main types of review question of relevance to NICE guidelines. These are listed in Table 4. For each type of question, the best primary study design varies, where ‘best’ is interpreted as ‘least likely to give misleading answers to the question’. For questions about the effectiveness of interventions, where randomised controlled trials (RCTs) were not available, the review of other types of evidence was pursued only if there was reason to believe that it would help the committee to formulate a recommendation.

However, in all cases, a well-conducted systematic review (of the appropriate type of study) is likely to always yield a better answer than a single study.

Table 4: Best study design to answer each type of question

Type of question	Best primary study design
Effectiveness or impact of an intervention	RCT; other studies that may be considered in the absence of RCTs are prospective and retrospective cohort studies

Type of question	Best primary study design
Diagnostic accuracy	Comparing the information against a valid gold standard in a cohort or case-control study
Prognostic reviews	Prospective cohort studies or case-control
Prevalence of disease, rare side-effects	Prospective cohort, registry, cross-sectional study, case-control
Experience of care	Qualitative research (for example, grounded theory, ethnographic research)

(a) RCT = randomised controlled trial.

3.11 Clinical review methods

The aim of the clinical literature review was to systematically identify and synthesise relevant evidence from the literature in order to answer the specific review questions developed by the Committee. Thus, clinical practice recommendations are evidence-based, where possible, and, if evidence is not available, either formal or informal consensus methods are used to try and reach general agreement between committee members and the need for future research is specified.

3.12 The search process

3.12.1 Scoping searches

A broad preliminary search of the literature was undertaken in January 2015 to obtain an overview of the issues likely to be covered by the scope and to help define key areas. The searches were restricted to clinical guidelines, Health Technology Assessment (HTA) reports, key systematic reviews and RCTs. A list of databases and websites searched can be found in Appendix H.

3.12.1.1 Systematic literature searches

After the scope was finalised, a systematic search strategy was developed to locate as much relevant evidence as possible. The balance between sensitivity (the power to identify all studies on a particular topic) and specificity (the ability to exclude irrelevant studies from the results) was carefully considered and a decision made to utilise a broad approach to searching to maximise retrieval of evidence to all parts of the guideline. Searches were restricted to certain study designs if specified in the review protocol and conducted in the following databases:

- Cochrane Database of Abstracts of Reviews of Effects
- Cochrane Database of Systematic Reviews
- Cochrane Central Register of Controlled Trials Excerpta Medica Database (Embase)
- HTA database (technology assessments)
- Medical Literature Analysis and Retrieval System Online (MEDLINE)/MEDLINE In-Process
- Psychological Information Database (PsycINFO)

The search strategies were initially developed for MEDLINE before being translated for use in other databases/interfaces. Strategies were built up through a number of trial searches and discussions of the results of the searches with the review team and committee to ensure that all possible relevant search terms were covered. In order to assure comprehensive coverage, search terms for mental health and learning disabilities were kept purposely broad to help counter dissimilarities in database indexing practices and thesaurus terms and

imprecise reporting of study populations by authors in the titles and abstracts of records. The search terms for each search are set out in full in Appendix H.

3.12.1.2 Reference Management

Citations from each search were downloaded into reference management software and duplicates removed. Records were then screened against the eligibility criteria of the reviews before being appraised for methodological quality (see below). The unfiltered search results were saved and retained for future potential re-analysis to help keep the process both replicable and transparent.

3.12.1.3 Search filters

To aid retrieval of relevant and sound studies, filters were used to limit a number of searches to systematic reviews, RCTs and observational. The search filters for systematic reviews and RCTs are adaptations of validated filters designed by the Health Information Research Unit (HIRU) at McMaster University. The search filter for observational studies is an in-house development. The filters have been recorded and can be found in Appendix H.

3.12.2 Date and language restrictions

Systematic database searches were initially conducted in May 2015 up to the most recent searchable date. Search updates were generated on a six monthly basis, with the final re-runs carried out in July 2016 ahead of the guideline consultation. After this point, studies were only included if they were judged by the committee to be exceptional (for example, if the evidence was likely to change a recommendation).

Although no language restrictions were applied at the searching stage, foreign language papers were not requested or reviewed, unless they were of particular importance to a review question.

Date restrictions were not applied, except for searches of systematic reviews which were limited to research published from 2001. The search for systematic reviews was restricted to the last 15 years as older reviews were thought to be less useful.

3.12.2.1 Other search methods

Other search methods involved: (a) scanning the reference lists of all eligible publications (systematic reviews, stakeholder evidence and included studies) for more published reports and citations of unpublished research; (b) sending lists of studies meeting the inclusion criteria to subject experts (identified through searches and the committee) and asking them to check the lists for completeness and to provide information of any published or unpublished research for consideration (see Appendix E); (c) checking the tables of contents of key journals for studies that might have been missed by the database and reference list searches; (d) tracking key papers in the Science Citation Index (prospectively) over time for further useful references; (e) conducting searches in ClinicalTrials.gov for unpublished trial reports; (f) contacting included study authors for unpublished or incomplete datasets. Searches conducted for existing NICE guidelines were updated where necessary. Other relevant guidelines were assessed for quality using the AGREE (Appraisal of Guidelines for Research and Evaluation Instrument) instrument (AGREE Collaboration, 2003(Collaboration, 2003)). The evidence base underlying high-quality existing guidelines was utilised and updated as appropriate.

Full details of the search strategies and filters used for the systematic review of clinical evidence are provided in Appendix H.

3.12.2.2 Study selection and assessment of methodological quality

All primary-level studies included after the first scan of citations were acquired in full and re-evaluated for eligibility at the time they were being entered into the study information database. Eligible systematic reviews and primary-level studies were critically appraised for methodological quality (risk of bias) using a checklist (NICE, 2012a) for templates). However, some checklists that were recommended in the 2014 manual update (NICE, 2014) were used (for example, for qualitative studies, for systematic reviews [Assessing the Methodological Quality of Systematic Reviews, AMSTAR, checklist] and for cross-sectional and cohort studies [the Newcastle Ottawa checklist for observational studies was used (Wells) for the epidemiological review on incidence and prevalence).

The Quality Assessment of Diagnostic Accuracy Studies – Revised (QUADAS-II) (Whiting et al., 2011) was used for evaluating risk of bias and indirectness of diagnostic and assessment tool studies.

For some review questions, it was necessary to prioritise the evidence with respect to the UK context (that is, external validity). To make this process explicit, the committee took into account the following factors when assessing the evidence:

- participant factors (for example, gender, age and ethnicity)
- provider factors (for example, model fidelity, the conditions under which the intervention was performed and the availability of experienced staff to undertake the procedure)
- cultural factors (for example, differences in standard care and differences in the welfare system).

It was the responsibility of the committee to decide which prioritisation factors were relevant to each review question in light of the UK context.

3.12.2.3 Double-sifting

Titles and abstracts of identified studies were screened by two reviewers against inclusion criteria specified in the protocols, until a good inter-rater reliability was observed (percentage agreement ≥90% or Kappa statistics, K>0.60). Any disagreements between raters were resolved through discussion. Initially 10% of references were double-screened. If inter-rater agreement was good then the remaining references were screened by one reviewer.

Once full versions of the selected studies were acquired for assessment, full studies were usually checked independently by two reviewers, with any differences being resolved. For some review questions a random sample of papers was checked for inclusion. Any studies that failed to meet the inclusion criteria at this stage were excluded.

3.12.2.4 Unpublished evidence

Stakeholders were invited to submit any relevant unpublished data using the call for evidence process set out in the 2014 edition of The Guidelines Manual. The committee used a number of criteria when deciding whether or not to accept unpublished data. First, the evidence must have been accompanied by a trial report containing sufficient detail to properly assess risk of bias. Second, the evidence must have been submitted with the understanding that data from the study and a summary of the study's characteristics would be published in the full guideline. Therefore, in most circumstances the committee did not accept evidence submitted 'in confidence'. However, the committee recognised that unpublished evidence submitted by investigators might later be retracted by those investigators if the inclusion of such data would jeopardise publication of their research.

3.12.3 Data extraction

3.12.3.1 Quantitative analysis

Study characteristics, aspects of methodological quality and outcome data were extracted from all eligible studies, using Review Manager Version 5.3.5 (Collaboration, 2014) and an Excel-based form.

In most circumstances, for a given outcome (continuous and dichotomous), where more than 50% of the number randomised to any group were missing or incomplete, the study results were excluded from the analysis (except for the outcome 'leaving the study early', in which case, the denominator was the number randomised). Where there were limited data for a particular review, the 50% rule was not applied. In these circumstances the evidence was downgraded.

Where possible, outcome data from an intention-to-treat analysis (that is, a 'once-randomised-always-analyse' basis) were used. Where intention-to-treat had not been used or there were missing data, the effect size for dichotomous outcomes were recalculated using worse-case scenarios for positive outcome and vice versa (for example, it was assumed that the person whose data was missing did not achieve remission). Results reported at 12 months follow up (after the end of treatment) or as close as possible to 12 months were extracted. However, this was not always possible and outcomes up to 5 years after treatment were sometimes reported.

All continuous outcomes were presented as a standardised mean difference (SMD) instead of a mean difference (MD). The final scores in each group were the preferred outcome for extraction. If final or change scores (from the baseline) were not reported, for example the study reported an F-value, p-value or t-value, the standard mean difference (SMD) was estimated if possible using a statistical calculator.

SMDs are typically used when different tools are used to measure the same outcome, for example if depression is measured using either the Becks Depression Inventory or the Hospital Anxiety and Depression Scale. However, in this guideline SMDs were also used to present the results of continuous outcomes when the same tool was used, for example eating psychopathology using the Eating Disorder Examination (EDE). The main reason for this is that the committee are apt at making decisions based on SMDs using the recommended interpretation of Cohen's effect size (d=0.2 small effect, d=0.5 moderate effect, d=0.7 large effect).

An outcome that had an SMD of ≥0.2 was considered clinically significant (or clinically important) and trends were discussed if the 95% confidence interval just crossed the line of no effect. This apparently low number of an SMD ≥0.2 was used as the threshold because of the small number of studies available and even small improvements on a scale that measures eating behaviour or mental health were considered clinically important for the person with an eating disorder.

The SMD results could have been converted back to MDs, however, no clinical consensus was made on what constitutes a minimally important difference (MID) and no published MIDs were found for body weight or for the various eating disorder scales reported. Granted, there are methods available for estimating whether an MD is clinically important and there are published MIDs for various depression scales, however, the committee acknowledged there are limitations with both approaches (SMD and MD) and in order to make decisions across many comparisons, SMDs was the preferred approach.

For dichotomous outcomes, clinical significance was considered anything that was +/- ≥10% difference. Trends were discussed if the difference was +/- ≥10% but just crossed the line of no effect.

Where the committee agreed that treatment effects were of sufficient magnitude to be clinically important they were described as 'favourable results' or being 'more effective' in the Linking Evidence to Recommendations (LETR) tables. Conversely, if the outcome favoured the control arm, the treatment was described as being 'less favourable' or 'less effective'. If an outcome showed an effect size that was clinically important but just crossed the line of no effect, it was considered clinically important with some uncertainty.

When calculating sensitivity and specificity for the case identification and assessment tools reviews using the diagnostic test accuracy data (i.e. data about the true and false positives and negatives yielded by the relevant test) a continuity correction of 0.5 was added to the numerator and denominator in the cases where the denominator was equal to zero.

Where necessary, standard errors were calculated from confidence intervals (CIs) or p value according to standard formulae; see the Cochrane Reviewers' Handbook 5.1.0 (Higgins and Green, 2011). Data were summarised using the generic inverse variance method using Review Manager.

Data from studies included in existing systematic reviews were extracted independently by one reviewer and cross-checked with the existing dataset. Where possible, two independent reviewers extracted data from new studies. Where double data extraction was not possible, data extracted by one reviewer was checked by the second reviewer. Disagreements were resolved through discussion. Where consensus could not be reached, a third reviewer or committee members resolved the disagreement. Masked assessment (that is, blind to the journal from which the article comes, the authors, the institution and the magnitude of the effect) was not used since it is unclear that doing so reduces bias (Berlin, 2001); (Jadad et al., 1996).

The analyses performed for existing systematic reviews incorporated into the guideline were not amended unless the committee considered that additional important aspects needed to be taken into consideration. For example, this could include stratifying data, conducting additional analyses, or using different results from the primary studies in a given analysis. Otherwise, the analyses were not amended.

3.12.4 Evidence synthesis

The method used to synthesise evidence depended on the review question and availability and type of evidence (see Appendix F for full details). Briefly, for questions about the psychometric properties of instruments, reliability, validity and clinical utility were synthesised narratively based on accepted criteria. For questions about test accuracy, bivariate test accuracy meta-analysis was conducted when there were data from four or more studies to calculate summary estimates of sensitivity and specificity for the relevant tool and threshold (if applicable). In the case where there was data from less than four studies, a narrative synthesis was presented. For questions about the effectiveness of interventions, standard meta-analysis was used where appropriate, otherwise narrative methods were used with clinical advice from the Committee. In the absence of high-quality research, formal and informal consensus processes were used.

3.12.5 Grading the quality of evidence

For questions about the effectiveness of interventions and the organisation and delivery of care, the GRADE approach was used to assess the quality of evidence for each outcome (Guyatt et al., 2011). The technical team produced GRADE evidence profiles (see below) using the GRADEpro guideline development tool, following advice set out in the GRADE handbook (Schünemann et al., 2013). All staff doing GRADE ratings were trained, and calibration exercises were used to improve reliability (Mustafa et al., 2013).

The analyses performed for existing systematic reviews incorporated into the guideline were not amended unless the committee considered that additional important aspects needed to be taken into consideration. For example, this could include stratifying data, conducting additional analyses, or using different results from the primary studies in a given analysis. Otherwise, the analyses were not amended.

For questions about what factors should be considered when admitting someone for compulsory treatment, a quality appraisal checklist of studies reporting correlations and associations was used. It is based on the appraisal step of the 'Graphical appraisal tool for epidemiological studies GATE' (Jackson et al., 2006). This checklist enables a reviewer to appraise a study's internal and external validity after addressing the following key aspects of study design: characteristics of study participants; definition of independent variables; outcomes assessed and methods of analyses. An estimate of the overall quality was based on the average answer given to each of the checklists and given either very low, low, moderate or high quality.

Heterogeneity was explored if the I2 test was greater than 50%. As described in the protocols, a sensitivity analysis was first conducted if see if studies that carried a high risk of bias explained the heterogeneity. If removing studies with a high risk of bias did not explain the results, then a subgroup analysis was conducted exploring the role duration of illness, severity of illness and presence of comorbidities. The full results of this are explained in the appendices and any subgroup analysis is shown in GRADE and explained in the LETR.

For observational studies included in any of the reviews, where randomised control trial evidence was not available, they were appraised using a quality appraisal checklist provided in the NICE manual 2012 (NICE, 2012b) This checklist assesses the study design, data collection, trustworthiness of the investigators, and the rigour of the analysis.

For questions about tools for case-identification and assessment of eating disorders (see Appendix M), a modified GRADE approach was used to produce an overall quality rating for the evidence according to the GRADE criteria of risk of bias, inconsistency, indirectness and imprecision. The default quality of evidence for cohort and cross-sectional studies was set as high quality; case-control studies were set as low quality since they overestimate the accuracy of tests due to spectrum bias (Kohn et al., 2013). The QUADAS-2 checklist was used to evaluate risk of bias and indirectness (Bossuyt et al., 2013). Whilst the QUADAS-2 framework does not provide an overall quality index for each study, such a rating was deemed important to assist the committee in interpreting the data on tools to augment assessment of mental health problems. We therefore adopted the terminology used within GRADE (high, moderate, low or very low quality evidence) (Schunemann et al., 2008); Bossuyt 2013). The approach taken to evaluating inconsistency and imprecision was discussed and agreed with the committee: inconsistency was evaluated either by visual inspection of the Summary Receiving Operating Characteristic (SROC) plot (where a meta-analysis was possible) or by inspection of the sensitivity and specificity forest plots (based on the primary measure of sensitivity or specificity, as appropriate), using the point estimates and confidence intervals of the identified studies. In the latter case, the evidence was downgraded by one increment if the individual studies varied across two areas (for example, 50–90% and 91–100%) and by two increments if the individual studies varied across three areas (for example, 0–50%, 51–90% and 91–100%). When a meta-analysis was possible, imprecision was evaluated by visual inspection of the confidence region on the SROC plot. In this case, particular attention when evaluating imprecision was given to whether the confidence region crossed the diagonal (which would indicate that the test was no better than chance at identifying or diagnosing the relevant condition). When there were less than four studies, imprecision was assessed according to the following criteria: a range of 0–20% of differences in point estimates of sensitivity (for the review on case identification) or specificity (for the review on assessment tools) was considered not imprecise, 21–40% serious imprecision and >40% very serious imprecision.

- The QUADAS-2 checklist uses signalling questions to evaluates the risk of bias across the fourdomans of patient selection (three questions), index test (two questions), reference standard (two questions)and the flow and timing (four questions) of the study; indirectness ('applicability concerns') is evaluated according to the first three domains by a single question. Each question can be answered as 'yes', 'no' or 'unclear'. Using the answers to the signalling questions, each domain is then evaluated for risk of bias and indirectness as 'Low', 'High' or 'Unclear'. The quality of evidence for risk of bias and indirectness was then downgraded by one increment (e.g. 'serious risk of bias') given the presence of one 'High' or 'Unclear' rating in a domain, and downgraded by two increments given the presence of more than two such ratings (e.g. 'very serious risk of bias').
- Overall quality of evidence - taking into account risk of bias, indirectness, inconsistency and imprecision - was then assessed using the GRADE terminology ('high', 'moderate', 'low' and 'very low'), with the quality rating downgraded by one increment given the presence of a 'serious' (e.g. 'serious imprecision') and by two increments given the presence of a 'very serious'. Thus a cohort study (which starts as high quality) that was evaluated as having serious risk of bias and very serious imprecision would be rated overall as very low quality evidence.

3.12.5.1 Evidence profiles

A GRADE evidence profile was used to summarise both the quality of the evidence and the results of the evidence synthesis for each 'critical' and 'important' outcome (see Table 5 for completed evidence profiles). The GRADE approach is based on a sequential assessment of the quality of evidence, followed by judgment about the balance between desirable and undesirable effects, and subsequent decision about the strength of a recommendation.

Within the GRADE approach to grading the quality of evidence, the following is used as a starting point:

- RCTs without important limitations provide high-quality evidence
- observational studies without special strengths or important limitations provide very low-quality evidence.

For each outcome, quality may be reduced depending on five factors: limitations, inconsistency, indirectness, imprecision and publication bias. For the purposes of the guideline, each factor was evaluated using criteria provided in Table 6.

For observational studies without any reasons for down-grading, the quality may be up-graded if there is a large effect, all plausible confounding would reduce the demonstrated effect (or increase the effect if no effect was observed), or there is evidence of a dose-response gradient (details would be provided under the 'other' column).

Each evidence profile includes a summary of findings: number of participants included in each group, an estimate of the magnitude of the effect, and the overall quality of the evidence for each outcome. Under the GRADE approach, the overall quality for each outcome is categorised into one of four groups (high, moderate, low, very low).

1 **Table 5: Example of a GRADE evidence profile**

Quality assessment							No. of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	**Intervention**	**Control group**	**Relative (95% CI)**	**Absolute**		
Outcome 1 (measured with: any valid method; better indicated by lower values)												
2	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious1	None	47	43	-	SMD 0.20 lower (0.61 lower to 0.21 higher)	⊕⊕⊕⊖ MODERATE	CRITICAL
Outcome 2 (measured with: any valid rating scale; better indicated by lower values)												
4	Randomised trials	Serious2	No serious inconsistency	No serious indirectness	Serious1	None	109	112	-	SMD 0.42 lower (0.69 to 0.16 lower)	⊕⊕⊖⊖ LOW	CRITICAL
Outcome 3 (measured with: any valid rating scale; better indicated by lower values)												
26	Randomised trials	No serious risk of bias	Serious[3]	No serious indirectness	No serious imprecision	None	521/5597 (9.3%)	798/3339 (23.9%)	RR 0.43 (0.36 to 0.51)	136 fewer per 1000 (from 117 fewer to 153 fewer)	⊕⊕⊕⊖ MODERATE	CRITICAL
Outcome 4 (measured with: any valid rating scale; better indicated by lower values)												
5	Randomised trials	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	503	485	-	SMD 0.34 lower (0.67 to 0.01 lower)	⊕⊕⊕⊕ HIGH	CRITICAL
Note.												

[1] OIS (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.
[2] Risk of bias across domains was generally high or unclear.
[3] There is evidence of moderate heterogeneity of study effect sizes.

CI = confidence interval; OIS = optimal information size; RR = risk ratio; SMD = standardised mean difference.

1

2

Table 6: Factors that decrease quality of evidence

Factor	Description	Criteria
Limitations	Methodological quality/ risk of bias.	Serious risks across most studies (that reported a particular outcome). The evaluation of risk of bias was made for each study using NICE methodology checklists (see Section 3.5.1).
Inconsistency	Unexplained heterogeneity of results.	Moderate or greater heterogeneity (using the methods suggested by GRADE1)
Indirectness	How closely the outcome measures, interventions and participants match those of interest.	If the comparison was indirect, or if the available evidence was substantially different from the population, intervention, comparator, or an outcome specified in the protocol for the question being addressed by the Committee.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect.	• the 95% confidence interval around the pooled or best estimate of effect included both (a) no effect and (b) appreciable benefit or appreciable harm (using default minimally important differences, MIDs). If a dichotomous outcome, the MIDs were 0.75 and 1.25, if a continuous outcome and SMD is reported the MIDs were -0.5 and 0.5. • If no MIDs were detected, the outcome was then checked to see if it met the optimal information size (OIS). • for dichotomous outcomes, OIS = 300 events; for continuous outcomes OIS = 400 participants
Publication bias	Systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.	Evidence of selective publication. This may be detected during the search for evidence, or through statistical analysis of the available evidence.

(b) 1 An I2 of 50% was used as the cut-off to downgrade for inconsistency. If heterogeneity was found, subgroup analysis was performed using the pre-specified subgroups in the protocol (see Appendix F); if subgroup analysis did not explain the heterogeneity, a random-effects model was used and the outcome was downgraded. GRADE = Grading of Recommendations Assessment, Development and Evaluation; NICE = National Institute for Health and Care Excellence; OIS = optimal information size.

3.12.6 Presenting evidence to the Guideline Committee

Study characteristics tables and, where appropriate, forest plots generated with Review Manager Version 5.3 and GRADE summary of findings tables (see Table 8) were presented to the Committee.

Where meta-analysis was not appropriate and/ or possible, the reported results from each primary-level study were reported in the study characteristics table and presented to the Committee. The range of effect estimates were included in the GRADE profile and, where appropriate, described narratively.

3.12.6.1 Summary of findings tables

Summary of findings tables generated from GRADEpro were used to summarise the evidence for each outcome and the quality of that evidence (Table 6). The tables provide anticipated comparative risks for dichotomous outcomes, which are especially useful when the baseline risk varies for different groups within the population.

Control group risks were not presented for SMDs as decisions on the clinical importance was based on the effect sizes independently of/ regardless of the control risk. This would obviously not be the case for MDs.

Table 7: Example of a GRADE summary of findings table

Outcomes	No of Participants (studies) Follow-up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with intervention (95% CI)
Global impression: 1. no improvement – short term	102 (1 study)	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, imprecision	RR 0.89 (0.69 to 1.16)	725 per 1000	80 fewer per 1000 (from 225 fewer to 116 more)
Behaviour: 1. average change score Adaptive Behaviour Scale – medium term	101 (1 study)	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, imprecision		The mean behaviour score was 1	0.60 SDs lower (1 to 0.21 lower)
Adverse effects: 1. extrapyramidal symptoms – medium term	243 (2 studies)	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, imprecision	RR 0.34 (0.05 to 2.1)	33 per 1000	21 fewer per 1000 (from 31 fewer to 36 more)

Note.
The basis for the assumed risk was the median control group risk across studies. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

[1] Generally unclear risk of bias and funded by manufacturer.
[2] OIS (for dichotomous outcomes, OIS = 300 events; for continuous outcomes, OIS = 400 participants) not met.

CI = confidence interval; GRADE = Grading of Recommendations Assessment, Development and Evaluation; OIS = optimal information size; RR = risk ratio; SD = standard deviation.

(c)

3.12.7 Evidence statements

Evidence statements provide a narrative of the results presented either in GRADE tables or other summary of evidence tables. For each outcome they describe what contributed to the overall result including the number of studies, the number of participants, the quality of the evidence, the direction of the effect and any uncertainty in the result. Subheading were used to describe the intervention and comparison and if the result was found at the end of treatment or long-term follow-up. The evidence statements were used by the guideline committee to formulate and prioritise recommendations.

3.12.8 Extrapolation

When answering review questions, if there was no direct evidence from a primary dataset, based on the initial search for evidence, data was extrapolated from another data set as indirect evidence. In this situation, the following principles were used to determine when to extrapolate:

- a primary dataset is absent, of particularly high risk of bias or is judged to be not relevant to the review question under consideration, and

- a review question is deemed by the committee to be important, such that in the absence of direct evidence, other data sources should be considered, and
- non-primary data source(s) is in the view of the committee available, which may inform the review question.

When the decision to extrapolate was made, the following principles were used to inform the choice of the non-primary dataset:

- the populations (usually in relation to the specified diagnosis or problem which characterises the population) under consideration share some common characteristic but differ in other ways, such as age, gender or in the nature of the disorder (for example, a common behavioural problem; acute versus chronic presentations of the same disorder), and
- the interventions under consideration in the view of the committee have one or more of the following characteristics:
  - share a common mode of action (for example, the pharmacodynamics of drug; a common psychological model of change – operant conditioning)
  - be feasible to deliver in both populations (for example, in terms of the required skills or the demands of the health care system)
  - share common side effects/harms in both populations, and
- the context or comparator involved in the evaluation of the different datasets shares some common elements which support extrapolation, and
- the outcomes involved in the evaluation of the different datasets shares some common elements which support extrapolation (for example, improved mood or a reduction in behaviour that challenges).

When the choice of the non-primary dataset was made, the following principles were used to guide the application of extrapolation:

- the committee should first consider the need for extrapolation through a review of the relevant primary dataset and be guided in these decisions by the principles for the use of extrapolation
- in all areas of extrapolation datasets should be assessed against the principles for determining the choice of datasets. In general the criteria in the four principles set out above for determining the choice should be met
- in deciding on the use of extrapolation, the committee will have to determine if the extrapolation can be held to be reasonable, including ensuring that:
  - the reasoning behind the decision can be justified by the clinical need for a recommendation to be made
  - the absence of other more direct evidence and by the relevance of the potential dataset to the review question can be established
  - the reasoning and the method adopted is clearly set out in the relevant section of the guideline.
  - If any data was extrapolated to help answer a review question, the results were downgraded in GRADE for indirectness.

3.12.9 Method used to answer a review question in the absence of appropriately designed, high-quality research

In the absence of appropriately designed, high-quality research (including indirect evidence where it would be appropriate to use extrapolation) an informal consensus was adopted.

3.12.9.1 Informal method of consensus

The informal consensus process involved a group discussion of what is known about the issues. The views of the committee were synthesised narratively by a member of the review team and circulated after the meeting. Feedback was used to revise the text, which was then included in the appropriate evidence review chapter.

3.13 Health economics methods

The aim of the health economics was to contribute to the guideline's development by providing evidence on the cost effectiveness of interventions and services covered in this guideline. This was achieved by a systematic literature review of existing economic evidence in all areas covered in the guideline.

Economic modelling was planned to be undertaken in areas with likely major resource implications, where the current extent of uncertainty over cost effectiveness was significant and economic analysis was expected to reduce this uncertainty, in accordance with The Guidelines Manual (NICE, 2014a). Prioritisation of areas for economic modelling was a joint decision between the Health Economist and the Committee. The rationale for prioritising review questions for economic modelling was set out in an economic plan agreed between NICE, the Committee, the Health Economist and the other members of the technical team. The following economic questions were selected as key issues to be addressed by economic modelling:

- Cost effectiveness of psychological therapies for adults with bulimia nervosa
- Cost effectiveness of psychological individual therapies for adults with binge eating disorder
- Cost effectiveness of psychological group therapies for adults with binge eating disorder

In addition, literature on the health-related quality of life (HRQoL) of people covered by this guideline was systematically searched to identify studies reporting appropriate utility scores that could be utilised in a cost-utility analysis.

In areas where modelling was not possible, the committee took into consideration resource implications and anticipated the cost effectiveness of interventions and services for people with eating disorders when making recommendations.

The methods adopted in the systematic literature review of economic evidence are described in the remainder of this section.

3.13.1 Search strategy for economic evidence

3.13.1.1 Scoping searches

A broad preliminary search of the literature was undertaken in January 2015 to obtain an overview of the issues likely to be covered by the scope and help define key areas. Searches were restricted to economic studies and HTA reports and conducted in the following databases:

- Embase
- MEDLINE/MEDLINE In-Process
- HTA database (technology assessments)
- NHS Economic Evaluation Database (NHS EED).

Any relevant economic evidence arising from the clinical scoping searches was also made available to the health economist during the same period.

3.13.1.2 Systematic literature searches

After the scope was finalised, a systematic search strategy was developed to locate all the relevant evidence. The balance between sensitivity (the power to identify all studies on a particular topic) and specificity (the ability to exclude irrelevant studies from the results) was carefully considered and a decision made to utilise a broad approach to searching to maximise retrieval of evidence to all parts of the guideline. Searches were restricted to economic studies and health technology assessment reports and conducted in the following databases:

- Embase
- HTA database (technology assessments)
- MEDLINE/MEDLINE In-Process
- NHS EED
- PsycINFO.

Any relevant economic evidence arising from the clinical searches was also made available to the health economist during the same period.

The search strategies were initially developed for MEDLINE before being translated for use in other databases/interfaces. Strategies were built up through a number of trial searches and discussions of the results of the searches with the review team and committee to ensure that all possible relevant search terms were covered. In order to assure comprehensive coverage, search terms for the guideline topic were kept purposely broad to help counter dissimilarities in database indexing practices and thesaurus terms and imprecise reporting of study interventions by authors in the titles and abstracts of records.

For standard mainstream bibliographic databases (Embase, MEDLINE and PsycINFO) search terms for the guideline topic combined with a search filter for health economic studies. For searches generated in topic-specific databases (HTA, NHS EED) search terms for the guideline topic were used without a filter. The sensitivity of this approach was aimed at minimising the risk of overlooking relevant publications, due to potential weaknesses resulting from more focused search strategies. The search terms are set out in full in Appendix F.

3.13.1.3 Reference Management

Citations from each search were downloaded into reference management software and duplicates removed. Records were then screened against the inclusion criteria of the reviews before being quality appraised. The unfiltered search results were saved and retained for future potential re-analysis to help keep the process both replicable and transparent.

3.13.1.4 Search filters

The search filter for health economics is an adaptation of a pre-tested strategy designed by CRD (2007). The search filter is designed to retrieve records of economic evidence (including full and partial economic evaluations) from the vast amount of literature indexed to major medical databases such as MEDLINE. The filter, which comprises a combination of controlled vocabulary and free-text retrieval methods, maximises sensitivity (or recall) to ensure that as many potentially relevant records as possible are retrieved from a search. A full description of the filter is provided in Appendix F.

3.13.1.5 Date and language restrictions

Systematic database searches were initially conducted in May 2015 up to the most recent searchable date. Search updates were generated on a six monthly basis, with the final re-runs carried out in July 2016. After this point, studies were included only if they were judged

by the committee to be exceptional (for example, the evidence was likely to change a recommendation).

Although no language restrictions were applied at the searching stage, foreign language papers were not requested or reviewed, unless they were of particular importance to an area under review. All the searches were restricted to research published from 2000 onwards in order to obtain data relevant to current healthcare settings and costs.

3.13.1.6 Other search methods

Other search methods involved scanning the reference lists of all eligible publications (systematic reviews, stakeholder evidence and included studies from the economic and clinical reviews) to identify further studies for consideration.

Full details of the search strategies and filter used for the systematic review of health economic evidence are provided in Appendix I.

3.13.2 Inclusion criteria for economic studies

The following inclusion criteria were applied to select studies identified by the economic searches for further consideration:

1. Only studies from Organisation for Economic Co-operation and Development countries were included, as the aim of the review was to identify economic information transferable to the UK context.
2. Only studies published from 2000 onwards were included in the review. This date restriction was imposed so that retrieved economic evidence was relevant to current healthcare settings and costs.
3. Selection criteria based on types of clinical conditions and service users as well as interventions assessed were identical to the clinical literature review.
4. Studies were included provided that sufficient details regarding methods and results were available to enable the methodological quality of the study to be assessed and provided that the study's data and results were extractable. Poster presentations of abstracts were excluded.
5. Full economic evaluations that compared two or more relevant options and considered both costs and consequences as well as costing analyses that compared only costs between two or more interventions were included in the review. Non-comparative studies were not considered in the review.
6. Economic studies were included if they used clinical effectiveness data from a clinical trial, a prospective or retrospective cohort study, or from a literature review. Studies with clinical effectiveness based on author's assumptions only were excluded.

3.13.3 Applicability and quality criteria for economic studies

All economic papers eligible for inclusion were appraised for their applicability and quality using the methodology checklist for economic evaluations recommended in The Guidelines Manual (NICE, 2014b). All studies that fully or partially met the applicability and quality criteria described in the methodology checklist were considered during the guideline development process. The completed methodology checklists for all economic evaluations considered in the guideline are provided in Appendix R.

3.13.4 Presentation of economic evidence

The economic evidence considered in the guideline is provided in the respective evidence chapters, following presentation of the relevant clinical evidence. The references to included studies and the respective evidence tables with the study characteristics and results are

provided in Appendix S. Characteristics and results of all economic studies considered during the guideline development process are summarised in economic evidence profiles provided in Appendix T. The full guideline includes only a brief summary of de-novo economic modelling undertaken. The detailed write up of de-novo economic models including the methods and full results are presented in the Appendix X.

3.13.5 Results of the systematic search of economic literature

The titles of all studies identified by the systematic search of the literature were screened for their relevance to the topic (that is, economic issues and information on HRQoL). References that were clearly not relevant were excluded first. The abstracts of all potentially relevant studies (16 references) were then assessed against the inclusion criteria for economic evaluations by the health economist. Full texts of the studies potentially meeting the inclusion criteria (including those for which eligibility was not clear from the abstract) were obtained. Studies that did not meet the inclusion criteria, were duplicates, were secondary publications of one study, or had been updated in more recent publications were subsequently excluded. All economic evaluations eligible for inclusion (12 studies in 13 publications) were then appraised for their applicability and quality using the methodology checklist for economic evaluations. Finally, those studies that fully or partially met the applicability and quality criteria set by NICE were considered at formulation of the guideline recommendations. Flow chart of studies for economic literature review is presented in the Appendix M. Excluded economic studies list is presented in the Appendix J.

3.14 From evidence to recommendations

Once the clinical and health economic evidence was summarised, the committee drafted the recommendations. In making recommendations, the committee took into account the trade-off between the benefits and harms of the intervention/instrument, as well as other important factors, such as the relative value of different outcomes reported in the evidence, quality of the evidence, trade-off between net health benefits and resource use, values and experience of the committee and society, current clinical practice, the requirements to prevent discrimination and to promote equality and the committee's awareness of practical issues.

Finally, to show clearly how the committee moved from the evidence to the recommendations, each chapter (or sub-section) has a section called 'recommendations and link to evidence'. Underpinning this section is the concept of the 'strength' of a recommendation. Some recommendations are 'strong' in that the committee believes that the vast majority of healthcare professionals and service users would choose a particular intervention if they considered the evidence in the same way that the committee has. This is generally the case if the benefits clearly outweigh the harms for most people and the intervention is likely to be cost effective. However, there is often a closer balance between benefits and harms and some service users would not choose an intervention whereas others would. This may happen, for example, if some service users are particularly averse to some side effect and others are not. In these circumstances the recommendation is generally weaker, although it may be possible to make stronger recommendations about specific groups of service users. The strength of each recommendation is reflected in the wording of the recommendation, rather than by using ratings, labels or symbols. For example a recommendation will use the words "consider" or "offer" a type of treatment, reflecting a weaker versus a stronger recommendation respectively.

Where the committee identified areas of uncertainty or where robust evidence was lacking, they developed research recommendations. Those that were identified as 'high priority' were developed further in the NICE version of the guideline and presented in Appendix G.

3.15 Stakeholder contributions

Professionals, service users and companies have contributed to and commented on the guideline at key stages in its development. Stakeholders for this guideline include:

- service user and carer stakeholders: national service user and carer organisations that represent the interests of people whose care will be covered by the guideline
- local service user and carer organisations: but only if there is no relevant national organisation
- professional stakeholders' national organisations: that represent the healthcare professionals who provide the services described in the guideline
- commercial stakeholders: companies that manufacture drugs or devices used in treatment of the condition covered by the guideline and whose interests may be significantly affected by the guideline
- providers and commissioners of health services in England
- statutory organisations: including the Department of Health
- Government, NHS Quality Improvement Scotland, the Care Quality Commission and the National Patient Safety Agency
- research organisations: that have carried out nationally recognised research in the area.

NICE clinical guidelines are produced for the NHS in England, so a 'national' organisation is defined as 1 that represents England, or has a commercial interest in England.

Stakeholders have been involved in the guideline's development at the following points:

- commenting on the initial scope of the guideline and attending a scoping workshop held by NICE
- commenting on the draft of the guideline

3.16 Validation of the guideline

This guidance is subject to a six week public consultation and feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders are responded to in turn and posted on the NICE website when the pre-publication check of the full guideline occurs.

3.17 Disclaimer

Healthcare providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by practitioners in light of individual patient circumstances, the wishes of the patient, clinical expertise and resources.

The National Guideline Alliance (NGA) disclaims any responsibility for damages arising out of the use or non-use of these guidelines and the literature used in support of these guidelines.

4 Identification and management of eating disorders

4.1 Introduction

The early signs of an eating disorder (ED) can be behavioural, psychological or physical. Common ED behaviours include extreme dieting and cutting out specific food groups, avoiding meal times and compulsive exercise. Psychological signs may include increased preoccupation with eating, weight and body shape, distorted body image and the adoption of strict food-related rules. Common physical signs include rapid or frequent weight change, tiredness, low energy and poor concentration. In some cases, vomiting can cause damage to teeth.

Key groups for screening and possible identification of an ED include those who are underweight compared with age norms, those who are disproportionately concerned about their weight, are dieting when underweight, women with menstrual disturbances, those with unexplained gastrointestinal symptoms and those who present with the physical signs of starvation or repeated vomiting. EDs are also more common in those with other mental health problems and those with type 1 diabetes and poor treatment adherence should also be screened. In children, poor growth or a sudden change in eating habits can be indictors of an ED.

Individuals with an ED may present in a range of settings in the NHS including primary care and secondary services such as gastroenterology, reproductive medicine and general mental health services. Because of the emphasis on physical appearance including weight and body shape in some sub-groups, they can be more vulnerable to developing an ED (for example, ballet dancers and fashion models).

Whilst some people can talk openly about their ED, others might be unaware that they have an ED or find it too difficult to disclose. People with EDs often feel ashamed of their symptoms and many are ambivalent about seeking treatment. It is therefore important to take a supportive, non-judgemental stance when talking with someone about whether they might have an ED.

Clinical change and the level of risk in mental and physical health should be monitored throughout treatment. Changes in ED symptoms (including behaviours, cognitions and physical symptoms) should be monitored weekly during treatment. This provides important information about the progress and likely effectiveness of any intervention. It is commonly done using brief self-report measures that should be regularly discussed with patients. In some cases physical and/or mental health risk may increase – for example, continued weight loss in anorexia nervosa. This is why it is important to monitor levels of risk, so that treatment can be reviewed and changed as required.

4.2 Review Question: What are the utility, validity and reliability of the instruments, tools and methods used for case identification in eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 8. Further information about the search strategy can be found in Appendix I; the full review protocols can be found in Appendix F.

This review considers the utility of instruments used to identify cases of eating disorders in people who are suspected of having an eating disorder. Randomised control trials, cohort

and cross-sectional studies that assessed the accuracy of the DAWBA-eating disorders section, ESP or SCOFF in identifying whether an eating disorder (or specific category thereof) is present as indicated by a full diagnostic interview were searched for. Studies were categorised according to whether they were used to identify cases of any eating disorder or a specific type of eating disorder.

Table 8: Clinical review protocol for the review question of: What are the utility, validity and reliability of the instruments, tools and methods used for case identification in eating disorders?

Component	Description
Review question(s)	What are the utility, validity and reliability of the instruments, tools and methods used for case identification in eating disorders?
Population	Children, young people and adults with: • a suspected eating disorder (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). Strata: • children (≤12), young people (13-≤17 years), adults ≥18 years.
Instruments, tools and methods	The following will be investigated: SCOFF (Sick-Control-One-Fat-Food) questionnaire DAWBA (self-assessment and parent/clinician component diagnostic and comorbidities) ESP (compared with SCOFF)
Reference tool	Reference tool (full diagnostic test for both clinical samples and population) • DSM • ICD-10
Critical outcomes	• Sensitivity (Se): the proportion of true positives of all cases diagnosed in the population • Specificity (Sp): the proportion of true negatives of all cases not-diagnosed in the population • Positive predictive value • Negative predictive value • Likelihood values
Important outcomes	VALIDITY • Concurrent validity, convergent validity, construct validity, content validity, predictive and discriminant validity RELIABILITY • Inter-rater reliability. Intra-rater reliability, test re-test reliability, , internal consistency
Study design	• RCTs • Cohort • Cross-sectional

4.2.1 Clinical Evidence for: What are the utility, validity and reliability of the instruments, tools and methods used for case identification in eating disorders?

No RCTs that satisfied the eligibility criteria were identified. Due to the paucity of evidence, it was decided to include case-control studies. Accordingly, 10 studies (seven cohort or cross-sectional studies; three case-control studies) met the eligibility criteria and were included in the review (Aoun et al., 2015), (Baudet et al., 2013), (Garcia-Campayo et al., 2005], Garcia-Campayo 2005{Garcia-Campayo, 2005 #16014), (House et al., 2008), (Liu et al., 2015), (Luck et al., 2002), (Morgan et al., 1999), (Moya et al., 2005), (Siervo et al., 2005). The majority of participants were adult females. Since case-control studies are likely to overestimate the accuracy of a test, data from such studies are presented separately from those of cohort and cross-sectional studies. For an overview of included studies see Table 9.

4.2.1.1 Development and Well-Being Assessment (DAWBA) – Eating Disorders Section

Two studies (n=231) on DAWBA, which is intended for use with children and young people, were included in the review (House 2008, Moya 2005). The vast majority of the participants were young females (there were only four young males in the total sample). The cohort study (n=57) examined the online version of DAWBA in a secondary care setting and used a clinical assessment and EDE or C-EDE (as appropriate) as the reference tool (House 2008). The case-control study (n=174) examined the interview version of DAWBA and used two groups of participants from a primary/secondary care setting (an eating disorders group and a non-eating disorders clinical control group) and a community control group (Moya 2005). There were not a sufficient number of studies to allow a meta-analysis of the diagnostic test accuracy data.

The quality of the evidence is presented for each study in the clinical evidence profiles below in Table 10, Table 11, Table 12 and Table 13. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

4.2.1.2 Eating Disorders Screen for Primary Care (ESP)

No relevant studies for the ESP in clinical samples or those at risk of an eating disorder were found. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

4.2.1.3 Sick-Control-One-Fat-Food (SCOFF)

For SCOFF, which is for use in adults, 1 cohort study (n=143), five cross-sectional studies (n=2513) and two case-control studies (n=438) were included in this review (Aoun 2015, Baudet 2013, Garcia 2011, Garcia-Campayo 2005, Liu 2015, Luck 2002, Morgan 1999, Siervo 2005).

All of the participants were adult females with the exception of one study, that assessed the accuracy of SCOFF in adult males and females (Liu 2015). Three of the cohort and cross-sectional studies were conducted in a primary care setting (Baudet 2013, Garcia-Campayo 2005, Luck 2002) whilst the remaining studies were conducted in secondary care settings (Aoun 2015, Liu 2015, Siervo 2005). In the two case-control studies, one study compared an eating disorder group with a healthy control group (Garcia 2011), whilst the other study compared a secondary care group with a group confirmed as not having an eating disorder (Morgan 1999).

There were sufficient cohort and cross-sectional studies to conduct a meta-analysis of the SCOFF at thresholds of 2 and 3 for the case identification of any eating disorder. (Note that the case-controls studies were not included in this analysis). Since the majority of the

identified studies reported the critical outcomes of sensitivity and specificity for several thresholds it was not possible to estimate a Summary ROC curve. However the available data from six studies (n=2513) permitted the estimation of a summary sensitivity and specificity at each of the two thresholds (Aoun 2015, Baudet 2013, Garcia-Campayo 2005, Liu 2015, Luck 2002, Siervo 2005). The quality of evidence for the meta-analysis of SCOFF, that included data for both male and female participants, for the case identification of any eating disorder is presented in Table 14, Table 15 and Table 16, whilst the quality of evidence for the relevant case-control studies are presented in Table 19. To enable a comparison of the data for the SCOFF tool at different threshold, the consequences of using SCOFF in terms of the number of false positives and false negatives the test would yield, and the related likelihood ratios, were calculated given a one-year prevalence per 100,000 of 0.5%, 1% and 5% (see Table 17).

The one cross-sectional study (n=1541) that included both male and female participants, conducted in an outpatient psychiatric clinic (Liu 2015), examined whether there were gender differences in the optimal SCOFF threshold to identify eating disorder cases. The quality of evidence for SCOFF in male participants is presented in Table 18.

Table 9: Study information for review of case identification of eating disorders in people with suspected eating disorders

Study ID	Country	Index Test	Version	Reference Tool	Type of study: sample	Sample N	Age (years)	Female (%)	Eating disorder tested for
House 2008	UK	DAWBA	Online	Multidisciplinary team clinical Interview, DSM-IV	Cohort: secondary care	57	15.7 (1.5)	93	Any eating disorder Anorexia nervosa Bulimia nervosa EDNOS
Moya 2005	Brazil	DAWBA	Interview	Open Clinical Interview, DSM-IV/ICD-10	Case-control: (I) eating disorder (II) clinical controls (III) community controls	174	15.3 (2.2)	100	Any eating disorder Anorexia nervosa Bulimia nervosa EDNOS
Aoun 2015	Lebanon	SCOFF	Written	Arabic MINI, DSM-IV	Cross-sectional: primary care	123	Range 15-55	100	Any eating disorder
Baudet 2013	France	SCOFF	Written	DSM-IV-TR French MINI/EDE-Q	Cohort: primary care	143	32.9 (9.2)	100	Any eating disorder
Garcia 2011	France	SCOFF	Written	French MINI, DSM-IV	Case-control: (I) eating disorder (II) healthy controls	226	22.1 Range 18-35	100	Any eating disorder Anorexia nervosa Bulimia nervosa
Garcia-Campayo 2005	Spain	SCOFF	Written	Spanish SCAN	Cross-sectional: primary care	203	29.2 (7.9) Range 14-55	100	Any eating disorder Anorexia nervosa Bulimia nervosa EDNOS
Liu 2015	Taiwan	SCOFF	Written	Mandarin Chinese SCID-I Patient, DSM-IV-TR	Cross-sectional: secondary care	1541	30.5 (7.8)	61	Any eating disorder
Luck 2002	UK	SCOFF	Interview	CDI, DSM-IV	Cross-sectional: primary care	341	Range 18-50	100	Any eating disorder
Morgan 1999	UK	SCOFF	Interview	CDI, DSM-IV	Case-control: (I) secondary care (II) No eating disorder	212	Range 18-40	100	Any eating disorder Anorexia nervosa Bulimia nervosa
Siervo	Italy	SCOFF	Written	CDI, DSM-IV	Cross-sectional:	162	Range	100	Any eating disorder

Study ID	Country	Index Test	Version	Reference Tool	Type of study: sample	Sample N	Age (years)	Female (%)	Eating disorder tested for
2005					secondary care		16-35		

Abbreviations: CDI, Clinical Diagnostic Interview; DAWBA, Development & Well-Being Assessment; EDE-Q, Eating Disorder Examination-Questionnaire; EDNOS, Eating Disorder Not Otherwise Specified; MINI, Mini-International Neuropsychiatric Interview; SCAN, Schedules for Clinical Assessment in Neuropsychiatry; SCID-I, Structured Clinical Interview for Axis I disorders, DSM-IV; SCOFF, Sick-Control-One-Fat-Food

.

Table 10: Clinical evidence profile for cohort and cross-sectional studies on the Development and Well-Being Assessment-eating disorders section – Online

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
DAWBA-eating disorder section (Online version)									
DAWBA for any eating disorder	1	57	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.94 (0.84, 0.99)	0.33 (0.04, 0.78)	MODERATE
DAWBA for anorexia nervosa	1	57	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.9 (0.73, 0.98)	0.93 (0.76, 0.99)	LOW
DAWBA for EDNOS	1	57	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Very serious imprecision[d]	0.67 (0.43, 0.85)	0.83 (0.67, 0.94)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 12 and Table 13 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 12 and Table 13 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for sensitivity for decision-making.

Table 11: Clinical evidence profile for case-control studies on the Development and Well-Being Assessment-eating disorders section – Interview

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
DAWBA-eating disorder section (Interview version)									
DAWBA for any eating disorder	1	174	Very serious risk of bias[a]	Not applicable[b]	Very serious indirectness[c]	No serious imprecision[d]	1.0 (0.93, 1.0)	0.94 (0.88, 0.97)	VERY LOW
DAWBA for anorexia nervosa	1	174	Very serious risk of bias[a]	Not applicable[b]	Very serious indirectness[c]	No serious imprecision[d]	1.0 (0.93, 1.00)	1.0 (0.97, 1.00)	VERY LOW

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
DAWBA for bulimia nervosa	1	174	Very serious risk of bias[a]	Not applicable[b]	Very serious indirectness[c]	No serious imprecision[d]	0.94 (0.83, 0.99)	0.96 (0.91, 0.99)	VERY LOW
DAWBA for EDNOS	1	174	Very serious risk of bias[a]	Not applicable[b]	Very serious indirectness[c]	Serious imprecision[d]	0.71 (0.56, 0.83)	0.95 (0.9, 0.98)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 12 and Table 13 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 12 and Table 13 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for sensitivity for decision-making.

Table 12: Summary table of QUADAS-2 results for risk of bias and indirectness for DAWBA

Study ID	RISK OF BIAS				INDIRECTNESS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
House 2008	☺	☹	☺	☺	☺	☺	☺
Moya 2005	☹	☺	☺	☹	☹	☺	☹
Key ☺=Low Risk ☹=High Risk							

Table 13: Graphical representation of QUADAS-2 results for risk of bias and applicability concerns for DAWBA

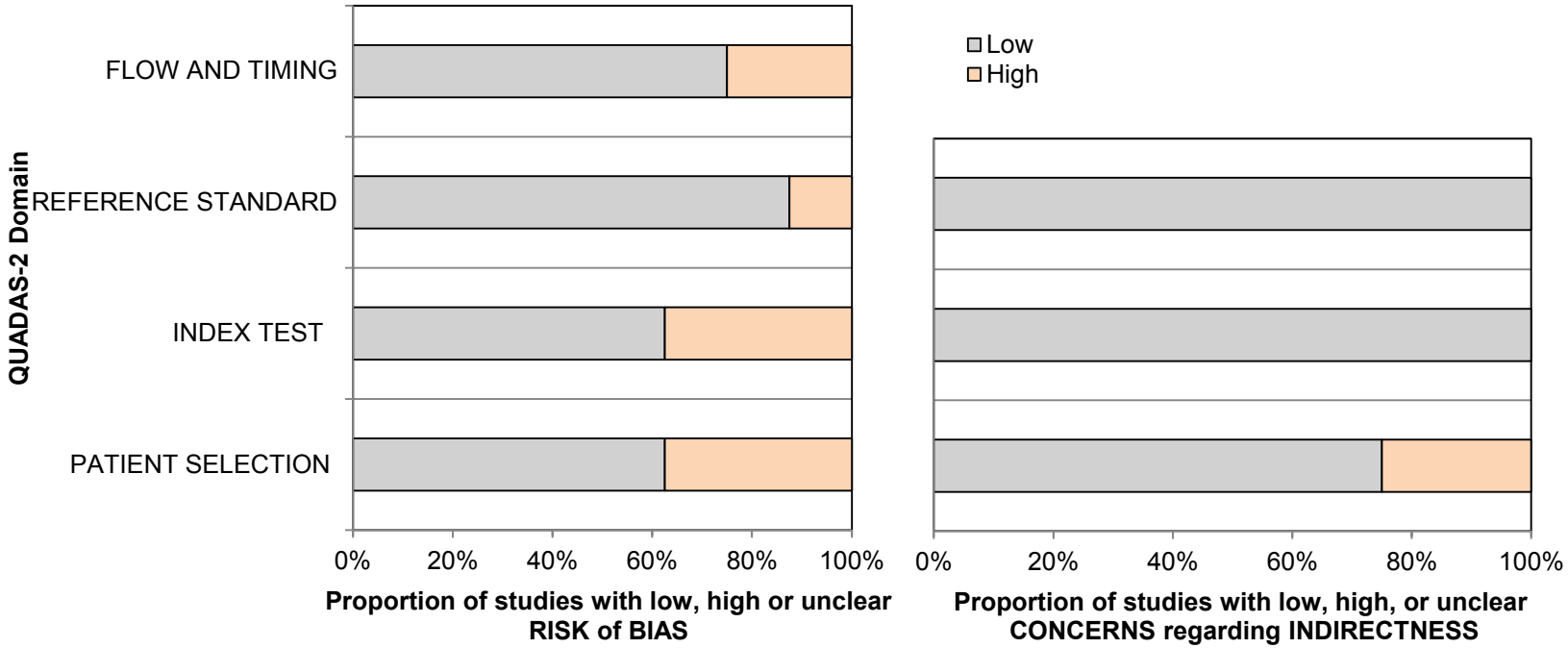

Table 14: Clinical evidence profile for cohort and cross-sectional studies of SCOFF for case identification of any eating disorder in adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for any eating disorder									
SCOFF at threshold ≥2	6	2513	No serious risk of bias[a]	Serious inconsistency[b]	No serious indirectness[c]	No serious imprecision[d]	Pooled 0.9 (0.81, 0.95)	Pooled 0.76 (0.56, 0.89)	MODERATE
SCOFF at threshold ≥3	6	2513	No serious risk of bias[a]	Very serious inconsistency[b]	No serious indirectness[c]	Serious imprecision[d]	Pooled 0.6 (0.46, 0.73)	Pooled 0.93 (0.82, 0.98)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 15 and Table 16 for more details; (b) Inconsistency was assessed by visual inspection of the summary Receiving Operating Characteristic (sROC) plots across studies, using the point estimates and confidence intervals. (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 15 and Table 16 for more details; (d) Judgement of precision was based on visual inspection of the confidence region in the diagnostic meta-analysis.

Table 15: Summary table of QUADAS-2 results for risk of bias and indirectness for cohort and cross-sectional studies on SCOFF

	RISK OF BIAS				INDIRECTNESS		
Study ID	**PATIENT SELECTION**	**INDEX TEST**	**REFERENCE STANDARD**	**FLOW AND TIMING**	**PATIENT SELECTION**	**INDEX TEST**	**REFERENCE STANDARD**
Aoun 2015	☺	☺	☺	☺	☺	☺	☺
Baudet 2013	☺	☺	☺	☺	☺	☺	☺
Garcia 2011	☹	☺	☺	☹	☹	☺	☺
Garcia-Campayo 2005	☺	☹	☺	☺	☺	☺	☺
Liu 2015	☺	☺	☺	☺	☺	☺	☺
Luck 2002	☺	☺	☺	☺	☺	☺	☺
Morgan 1999	☹	☹	☺	☹	☹	☺	☺
Siervo 2005	☹	☹	☹	☺	☺	☺	☺
Key ☺=Low Risk ☹=High Risk							

Table 16: Graphical representation of QUADAS-2 results for risk of bias and indirectness for cohort and cross-sectional studies on SCOFF

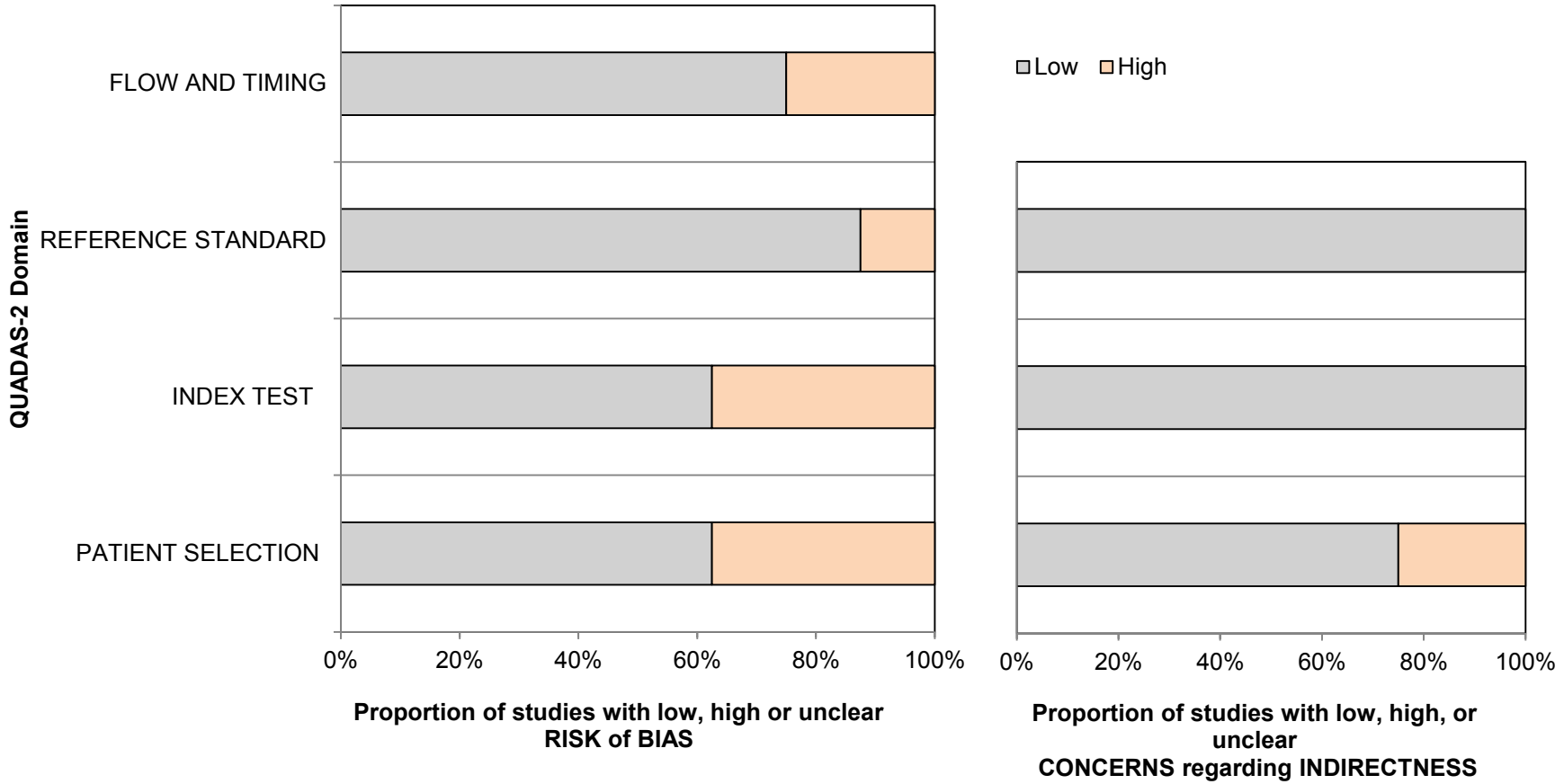

Table 17: Consequences of key findings on SCOFF for case identification of any eating disorder

	One-year prevalence per 100,000					
	0.5% prevalence		1% prevalence		5% prevalence	
	SCOFF Threshold		SCOFF Threshold		SCOFF Threshold	
Consequences	≥2	≥3	≥2	≥3	≥2	≥3

	One-year prevalence per 100,000					
Positive likelihood ratio	5.29	19.9	4.74	17.33	4.96	17.5
Negative likelihood ratio	0.38	0.21	0.12	0.31	0.07	0.31
False positive per 1000	189	40	188	40	180	38
False negative per 1000	0	1	1	3	3	15

Note: the number of false positives and false negatives per 1000 were calculated using the pooled sensitivity and specificity results from the meta-analysis of the cohort and cross-sectional data only. A positive likelihood ratio indicates how much more likely a person with a disease tests positive compared with a person without the disease, whilst a negative likelihood ratio indicates how less likely a person with a disease tests negative compared with a person without the disease. Likelihood ratios of <0.1 or >10 are typically interpreted as indicating that the relevant test is clinically very useful, 0.1-0.2 or 5-10 as moderately useful and 0.2-1 or 1-5 as not particularly useful.

Table 18: Clinical evidence profile for cohort and cross-sectional studies of SCOFF for case identification of any eating disorder in male adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for any eating disorder									
SCOFF at threshold ≥2	1								
Liu 2015		605	No serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.86 (0.70, 0.96)	0.74 (0.70, 0.78)	MODERATE
SCOFF at threshold ≥3	1								
Liu 2015		605	No serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.66 (0.48, 0.81)	0.91 (0.88, 0.93)	MODERATE

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 15 and Table 16 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 15 and Table 16 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for sensitivity for decision-making.

Table 19: Clinical evidence profile for case-control studies of SCOFF for case identification of any eating disorder in adult females

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for any eating disorder									
SCOFF at threshold ≥2	2	438							
Garcia 2011		226	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	No serious imprecision[d]	0.94 (0.88, 0.97)	0.94 (0.88, 0.97)	VERY LOW
Morgan 1999		212	Very serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	No serious imprecision[d]	1.00 (0.97, 1.00)	0.88 (0.79, 0.93)	VERY LOW
SCOFF at threshold ≥3	2	438							
Garcia 2011		226	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	No serious imprecision[d]	0.66 (0.57, 0.75)	1.00 (0.97, 1.00)	VERY LOW
Morgan 1999		212	Very serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[c]	No serious imprecision[d]	0.99 (0.95, 1.00)	0.96 (0.90, 0.99)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist; (b) Inconsistency was assessed by inspection of the sensitivity and specificity forest plots (based on the primary measure), using the point estimates and confidence intervals. The evidence was downgraded by 1 increment if the individual studies varied across 2 areas (for example, 50–90% and 90–100%) and by 2 increments if the individual studies varied across 3 areas (for example, 0–50%, 50–90% and 90–100%); See Table 20 for more details; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability; See Table 20 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for sensitivity for decision-making.

Table 20: Summary table of QUADAS-2 results for risk of bias and indirectness for case control studies on SCOFF

Study	RISK OF BIAS				INDIRECTNESS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Garcia 2011	☹	☺	☺	☹	☹	☺	☺
Morgan 1999	☹	☹	☺	☹	☹	☺	☺
Key ☺=Low Risk ☹=High Risk							

Table 21: Clinical evidence profile for cohort and cross-sectional studies of SCOFF for case identification of anorexia nervosa in adult females

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for anorexia nervosa									
SCOFF at threshold ≥2	1								
Garcia-Campayo 2005		195	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.93 (0.77, 0.99)	0.94 (0.90, 0.97)	LOW
SCOFF at threshold ≥3	1								
Garcia-Campayo 2005		195	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.9 (0.73, 0.98)	1.00 (0.98, 1.00)	LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist; (b) Inconsistency not applicable due to only one study. See Table 15 and Table 16 for more details; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for sensitivity for decision-making.

Table 22: Clinical evidence profile for case-control studies of SCOFF for case identification of anorexia nervosa in adult females

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for anorexia nervosa									
SCOFF at threshold ≥2	2	438							
Garcia 2011		226	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[d]	No serious imprecision[e]	0.96 (0.87, 0.99)	0.93 (0.82, 0.99)	VERY LOW
Morgan 1999		212	Very serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[d]	No serious imprecision[e]	1.00 (0.95, 1.00)	0.88 (0.81, 0.92)	VERY LOW
SCOFF at threshold ≥3	1								
Garcia 2011		226	Serious risk of bias[a]	Not applicable[c]	Serious indirectness[d]	Serious imprecision[e]	0.66 (0.53, 0.77)	1.00 (0.98, 1.00)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 15 and Table 16 for more details; (b) Inconsistency was assessed by inspection of the sensitivity and specificity forest plots (based on the primary measure), using the point estimates and confidence intervals. The evidence was downgraded by 1 increment if the individual studies varied across 2 areas (for example, 50–90% and 90–100%) and by 2 increments if the individual studies varied across 3 areas (for example, 0–50%, 50–90% and 90–100%); (c)

Inconsistency not applicable due to only one study; (d) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 15 and Table 16 for more details; (e) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for sensitivity for decision-making.

Table 23: Clinical evidence profile for cohort and cross-sectional studies of SCOFF for case identification of bulimia nervosa in adult females

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for bulimia nervosa									
SCOFF at threshold ≥2	1								
Garcia-Campayo 2005		195	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.98 (0.89, 1.00)	0.94 (0.89, 0.97)	MODERATE
SCOFF at threshold ≥3	1								
Garcia-Campayo 2005		195	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.73 (0.58, 0.85)	1.00 (0.98, 1.00)	LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 15 and Table 16 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 15 and Table 16 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for sensitivity for decision-making.

Table 24: Clinical evidence profile for case-control studies of SCOFF for case identification of bulimia nervosa in adult females

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for bulimia nervosa									
SCOFF at threshold ≥2	2	438							
Garcia 2011		226	Serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[d]	No serious imprecision[e]	0.93 (0.82, 0.99)	0.94 (0.89, 0.97)	VERY LOW
Morgan 1999		212	Very serious risk of bias[a]	No serious inconsistency[b]	Serious indirectness[d]	No serious imprecision[e]	1.00 (0.93, 1.00)	0.87 (0.81, 0.92)	VERY LOW
SCOFF at threshold ≥3	1								
Garcia 2011		226	Serious risk of bias[a]	Not applicable[c]	Serious indirectness[d]	Serious imprecision[e]	0.67 (0.51, 0.8)	1.00 (0.98, 1.00)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 15 and Table 16 for more details; (b) Inconsistency was assessed by inspection of the sensitivity and specificity forest plots (based on the primary measure), using the point estimates and confidence intervals. The evidence was downgraded by 1 increment if the individual studies varied across 2 areas (for example, 50–90% and 90–100%) and by 2 increments if the individual studies varied across 3 areas (for example, 0–50%, 50–90% and 90–100%); (c) Inconsistency not applicable due to only one study; (d) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 15 and Table 16 for more details; (e) Due to an insufficient number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for decision-making.

Table 25: Clinical evidence profile for cohort and cross-sectional studies of SCOFF for case identification of EDNOS in adult females

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SCOFF for EDNOS									
SCOFF at threshold ≥2	1								
Garcia-Campayo 2005		195	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	1.00 (0.94, 1.00)	0.95 (0.90, 0.98)	MODERATE
SCOFF at threshold ≥3	1								
Garcia-Campayo 2005		195	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.24 (0.13, 0.37)	1.00 (0.98, 1.00)	LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test sensitivity as this was the primary measure discussed in decision-making. (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 15 and Table 16 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 15 and Table 16 for more details; (d) Due to an insufficient number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure for decision-making.

4.2.2 Economic Evidence

No economic evidence on the tools for identification of eating disorders was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

4.2.3 Clinical evidence statements

4.2.3.1 DAWBA

The overall quality of evidence from the one identified cohort study on the use of the eating disorders section of DAWBA-online (n=57) to identify eating disorder cases in people suspected of having an eating disorder ranged from moderate for any eating disorder, to low for anorexia nervosa and very low for EDNOS. The quality of evidence for all three cases was downgraded due to concerns over risk of bias and concerns over imprecision in the case identification of anorexia nervosa and EDNOS.

The reported sensitivity was 0.94 for any eating disorder, 0.9 for anorexia nervosa and 0.67 for EDNOS; the related 95% CIs range were relatively consistent (i.e. narrow) and ranged from 0.84 to 0.99 for any eating disorder, 0.73 to 0.98 for anorexia nervosa, and from 0.43 to 0.85 for EDNOS. By contrast, the reported specificity of DAWBA-online was 0.33 for any eating disorder, 0.93 for anorexia nervosa and 0.83 for EDNOS; the related 95% CIs were much wider, ranging from 0.04 to 0.78 for any eating disorder, 0.76 to 0.99 for anorexia nervosa, and from 0.67 to 0.94 for EDNOS.

The quality of evidence from the one identified case control study on the eating disorders section of DAWBA-interview (n=174) was very low for the case identification of any eating disorder, anorexia nervosa, bulimia nervosa and EDNOS. The quality was affected by concerns over risk of bias, indirectness and/or imprecision. As expected, the estimated sensitivity and specificity of the eating disorders section of DAWBA-interview yielded by the case control studies was higher than that of the online version. The reported sensitivity for this study was 1.0 for any eating disorder and anorexia nervosa, and 0.71 for EDNOS; the related 95% CI ranged from 0.93 to 1.00 for any eating disorder and anorexia nervosa, and from 0.56 to 0.83 for EDNOS. The reported sensitivity of DAWBA-interview in identifying cases of bulimia nervosa was 0.94, with the 95% CI ranging from 0.83 to 0.99. Similarly, the reported specificity of the case control study on DAWBA-interview was higher than that reported by the cohort study on DAWBA-online. The reported specificity for this study was 0.94 for any eating disorder, 1.00 for anorexia nervosa and 0.95 for EDNOS; the related 95% CIs ranged from 0.88 to 0.97 for any eating disorder, 0.97 to 1.00 for anorexia nervosa, and from 0.9 to 0.98 for EDNOS. The reported specificity of DAWBA-interview in identifying cases of bulimia nervosa ranged from 0.91 to 0.99, with the 95% CI ranging from 0.91 to 0.99.

4.2.3.2 SCOFF

Any eating disorder

The overall quality of evidence from the six identified cohort and cross-sectional studies (n=2513) included in the meta-analysis was moderate for SCOFF at a threshold of 2 or more, but very low for SCOFF at a threshold of 3 or more. The quality of evidence for the use of SCOFF in adult populations suspected of having an eating disorder was downgraded due to concerns about inconsistency and/or imprecision.

The estimation of a summary sensitivity and specificity for SCOFF at a threshold of 2 yielded a pooled sensitivity of 0.90 with a 95% CI from 0.81 to 0.95 and a pooled specificity of 0.76

with a 95% CI from 0.56 to 0.89. The estimation of a summary sensitivity and specificity for SCOFF at a threshold of 3 yielded a pooled sensitivity of 0.6 with a 95% CI from 0.46 to 0.73 and a pooled specificity of 0.93 with a 95% CI from 0.82 to 0.98. Whilst the number of false positives per 1000 at the one-year prevalence rates per 100,000 of 0.5%, 1% and 5% were much higher for the SCOFF at a threshold of 2 compared to a threshold of 3, the number of false negatives was lower.

The quality of evidence from the one study that examined the use of SCOFF in male populations suspected of having an eating disorder was moderate (n=605) due to concerns about imprecision. In male participants at a SCOFF threshold of 2, the one identified study yielded a sensitivity of 0.86 with a 95% CI from 0.7 to 0.96 and a specificity of 0.74 with a 95% CI from 0.7 to 0.78. At a threshold of 3, the sensitivity of SCOFF was 0.66 with a 95% CI from 0.48 to 0.81, whilst the specificity was 0.91 with a 95% CI from 0.88 to 0.93.

The quality of evidence from the two identified case-control studies (n=438) that examined the use of SCOFF in populations with suspected eating disorders was very low due to concerns over risk of bias and indirectness. The reported sensitivity for SCOFF at a threshold of 2 or more ranged from 0.94 to 1.00 with the related 95% CIs ranging from 0.88 to 1.00. The reported specificity of SCOFF at a threshold of 3 or more ranged from 0.66 to 0.99 with the related 95% CIs ranging from 0.57 to 1.00, whilst the specificity ranged from 0.90 to 1.00 with the related 95% CIs also ranging from 0.90 to 1.00.

Overall, the evidence suggests that SCOFF is generally a more useful case identification tool when used with a threshold of 2 compared with a threshold of 3 at identifying any eating disorder.

Anorexia nervosa

The quality of evidence from the one cross-sectional study (n=195) that examined the use of SCOFF to identify cases of anorexia nervosa in adult female populations suspected of having an eating disorder was low for both thresholds examined due to concerns over risk of bias and imprecision. The reported sensitivity for SCOFF at a threshold of 2 in this study was 0.93 with 95% CI ranging from 0.77 to 0.99, whilst the specificity was 0.94 with the related 95% CI ranging from 0.90 to 0.97. The sensitivity for SCOFF at a threshold of 3 in this study was slightly lower at 0.9 with 95% CI ranging from 0.73 to 0.98, whilst the specificity was 1.00 with the 95% CI ranging from 0.98 to 1.00.

The quality of evidence from the two case control studies (n=438) that examined the use of SCOFF at a threshold of 2 or more to identify cases of anorexia nervosa in adult female populations was very low due to concerns over risk of bias and indirectness. The reported sensitivity for SCOFF at a threshold of 2 or more ranged from 0.96 to 1.00 with the related 95% CIs ranging from 0.87 to 1.00, whilst the specificity ranged from 0.88 to 0.93 with the related 95% CIs ranging from 0.81 to 1.00.

The quality of the one case control study (n=226) that examined the use of SCOFF at a threshold of 3 to identify cases of anorexia nervosa in adult female populations suspected of having an eating disorder was also very low due to concerns over risk of bias, indirectness and imprecision. The reported sensitivity was 0.66 (with 95% CI ranging from 0.53 to 0.77), whilst the specificity was 1.00 (with 95% CI ranging from 0.98 to 1.00).

Bulimia nervosa

The quality of evidence from the one cross-sectional study (n=195) that examined the use of SCOFF to identify cases of bulimia nervosa in adult female populations suspected of having an eating disorder was moderate for a threshold of 2 or more and low for a threshold of 3 or more, due to concerns over risk of bias and/or imprecision. The reported sensitivity for SCOFF at a threshold of 2 was 0.98 (with 95% CI ranging from 0.89 to 1.00), whilst the specificity was 0.94 (with 95% CI ranging from 0.89 to 0.97). At a threshold of 3, the reported

sensitivity was 0.73 (with 95% CI ranging from 0.58 to 0.85), whilst the specificity was 1.00 (with 95% CI ranging from 0.98 to 1.00).

The quality of evidence from the two case control studies (n=438) that examined the use of SCOFF at a threshold of 2 or more to identify cases of bulimia nervosa in adult female populations was very low due to concerns over risk of bias and indirectness. The reported sensitivity for SCOFF at a threshold of 2 or more ranged from 0.93 to 1.00 with the related 95% CIs ranging from 0.82 to 1.00, whilst the specificity ranged from 0.87 to 0.94 with the related 95% CIs ranging from 0.81 to 0.97.

The quality of the one case control study (n=226) that examined the use of SCOFF at a threshold of 3 to identify cases of bulimia nervosa in adult female populations suspected of having an eating disorder was again very low due to concerns over risk of bias, indirectness and imprecision. The reported sensitivity was 0.67 (with 95% CI ranging from 0.51 to 0.80), whilst the specificity was 1.00 (with 95% CI ranging from 0.98 to 1.00).

EDNOS

The quality of evidence from the one cross-sectional study (n=195) that examined the use of SCOFF to identify cases of EDNOS in adult female populations suspected of having an eating disorder was moderate for a threshold of 2 or more and low for a threshold of 3 or more, due to concerns over risk of bias and/or imprecision. The reported sensitivity for SCOFF at a threshold of 2 was 1.00 (with 95% CI ranging from 0.94 to 1.00), whilst the specificity was 0.95 (with 95% CI ranging from 0.90 to 0.98). At a threshold of 3, the reported sensitivity was no better than chance at 0.24 (with 95% CI ranging from 0.13 to 0.37), whilst the specificity was 1.00 (with 95% CI ranging from 0.98 to 1.00).

4.2.4 Economic Evidence statements

No economic evidence on the tools for the identification of eating disorders was available.

4.2.5 Recommendations and link to evidence for the review on: What are the utility, validity and reliability of the instruments, tools and methods used for case identification in eating disorders?

Initial assessments in primary and secondary mental health care

	1. Do not use screening tools (for example SCOFF) as the sole method to determine whether or not people have an eating disorder.
Relative value of different outcomes	For the review on the validity of tools that may be used for case-identification, assessment or monitoring eating disorders, the committee considered the critical outcomes were sensitivity, specificity, positive predictive value, and negative predictive value and likelihood ratio values. Other outcomes were considered important but studies were not included if they did not measure any of the critical outcomes. Important outcomes included numerous validity and reliability measures. The critical outcome of sensitivity was used as the primary measure for decision making given the need to minimise false negatives when seeking to identify eating disorder cases in people with a suspected eating disorder. That is, such a test needs to minimise the number of false negatives so that the test is more inclusive and ensures more people who are likely to have an eating disorder go on to receive the full diagnostic test (e.g. at the secondary care stage). Studies were excluded if they investigated how well the tool was at screening the general population for eating disorders because this would not be considered a

	good use of resources in an NHS setting. Instead, the usefulness of a case-identification tool in a clinical setting, such as when a person with a suspected eating disorder visits a general practitioner, was considered. The outcomes of positive and negative predictive value, and positive and negative likelihood ratios were presented although not considered by the committee (with the exception of the latter relating to the results of the meta-analysis of SCOFF for any eating disorder).
Trade-off between clinical benefits and harms	The review on what tests were effective at identifying people with an eating disorder (case identification) showed that the Development and Well-Being Assessment (online or interview) may be a better case identification tool in interview format than online for young people with any eating disorder, anorexia nervosa, bulimia nervosa or OSFED. Overall, although there was a similar number of false negatives for case identification of any eating disorder, anorexia nervosa, bulimia nervosa and EDNOS. However, the online version produced a higher number of false positives compared with the full interview. This in turn would result in a high number of young people needing to undergo a full diagnostic test, leading to a waste in resources and an increase in costs. Another tool where evidence was found was SCOFF (Sick-Control-One-Fat-Food). The review compared how well this tool identifies adults with a suspected eating disorder with different cut-off scores used. Overall, the evidence suggests that when a threshold of two or more is used, it is better at identifying the presence of an eating disorder (both generally and for the specific disorders of anorexia, bulimia and EDNOS) in populations who are suspected to have such and minimising the number of false negatives. When a cut-off score greater than three is used, the number of false negatives increases as the prevalence of eating disorders increases in the population. However, a cut-off score greater than two leads to a higher number of false positives compared to a cut-off score greater than three, which would both waste secondary care resources and increase the associated costs of assessment.
Trade-off between net health benefits and resource use	There was no evidence on the cost effectiveness of identification tools in people with eating disorders. The committee discussed the time it takes to administer such tools (for example, DAWBA can take up to 50 min and SCOFF only few minutes) and the consequence associated with eating disorders. The committee considered very limited clinical evidence and noted that that even though there are various tools available there is no convincing evidence that any of these tools are effective on their own in the identification of eating disorders. Based on the administration time SCOFF would be the preferred option. However, given the range in quality of the clinical evidence (which was mostly very low quality and conflicting) and the relatively low prevalence of eating disorders (especially anorexia nervosa) – both in the general population and in those people presenting in a primary care setting - the committee felt that even if the sensitivity of the reviewed case identification tools were higher, their utilisation would not be an efficient use of resources. As a result, the committee refrained from recommending any case identification tool, and noted that such tools should not be used as a sole method to determine whether people have an eating disorder.
Quality of evidence	The overall quality of evidence of the case-identification studies was assessed using a modified GRADE approach that used the QUADAS-2 checklist to evaluate the risk of bias and indirectness. The quality of evidence ranged from moderate quality to very low quality. Outcomes were downgraded for: i) risk of bias, ii) indirectness, iii) imprecision, and iv) inconsistency (if applicable). The evidence from case-control studies started at low quality because of the risk of spectrum bias. As expected, these studies generally yielded higher estimates of the sensitivity of the relevant tools compared with the identified cohort and cross-sectional studies. Few studies were identified for the case-identification review (only two for the Development and Well-Being Assessment), so an overall specificity and sensitivity score could not be estimated. Also, no evidence was found on the ESP test. The sensitivity of the interview and online versions of DAWBA yielded similar estimates for any eating disorder, anorexia nervosa and EDNOS, however, it was limited to one cohort study conducted in a UK secondary care setting and one

	case-control study conducted in Brazil. There was also substantial variability in its estimated specificity. There was substantially more evidence for the performance of SCOFF for any eating disorder, but only one cross-sectional study conducted in a Spanish primary care setting and two case-control studies that examined its performance in identifying cases of particular eating disorders. Whilst the quality of evidence for SCOFF at a cut-off score greater than two was moderate, the pooled estimate of its sensitivity was only 0.9. That is, if 100 people with an eating disorder were to take the test, only 90 of them would test positive. Any variability observed across the studies may have been due to a difference in the prevalence of the disease in the populations used. More generally, if there is a low prevalence in the population used, then there will be more people in whom the condition is barely present and fewer people in whom the condition is clearly present. As such, sensitivity may be lower (detecting true positives) because it will be more difficult to detect people with the target conditions. Conversely, if there is a higher disease prevalence, for example a sample from tertiary care or a case control study, there may be fewer participants with limited forms of the disease and more with the clear manifested forms. In such cases, it will be easier to clearly detect those with the condition and sensitivity will be higher. Indeed, the estimated sensitivity of SCOFF for both cut-off scores in the studies conducted in primary care were generally lower than for those conducted in secondary care. There are a number of other reasons that may explain test variability, including how similar the symptoms appear in a pool of participants. The more similar the underlying conditions appear in people, the more false positives are likely to be found. When underlying conditions appear very much alike, the target condition may also be recognised as a comorbidity, which may result in more false negatives. This could lead to a lower sensitivity and specificity in a study population that has been selected for case identification compared with the general population. Other reasons for variability across studies include reader expectation (what the person diagnosing typically sees), and study design such as case-controls versus a spectrum of participants that reflects what the clinician would typically see in practice. In the case of SCOFF, the estimated sensitivity in the case control studies was generally higher than that in the cohort and cross-sectional studies.
Other considerations	The committee discussed whether to recommend any of the case-identification tools. Although SCOFF was relatively good at identifying true cases and minimising false negatives and that it could be integrated into the early stages of identifying people with an eating disorder, the committee agreed it would be better for clinicians to use their judgment rather than one of the tools considered in this review. This was mostly due to the low prevalence of eating disorders, the variable likelihood ratios associated with the SCOFF for any eating disorder, and the wide range in the number of false negatives and positives that it would yield for the various types of eating disorders. The committee agreed that when a person with a suspected eating disorder presents for evaluation, there are better ways of determining whether s/he has an eating disorder (e.g. using a full diagnostic test such as the EDE). The Committee decided that no research recommendation was required. The tools investigated in this review were by no means a comprehensive list of all the tools available for case-identification or for a full-diagnosis. The committee were asked to provide a list of the most common and relevant tools that could be investigated in the time available for this review. It is important to note that results from the studies will vary depending on the population used to assess the tools. This is because sensitivity and specificity may vary with the prevalence of the disease. For this reason, clinicians are advised to base their decisions on studies that most closely match their own clinical situation. Although the committee discussed whether further guidance should be given, they concluded that clinical expertise is sufficient for identifying eating disorder cases.

4.3 Review Question: What is the validity and reliability of the instruments, tools and methods used to assess and monitor eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 26. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

Randomised control trials, cohort and cross-sectional studies were sought that assessed the accuracy of the Anorexia Nervosa Inventory for Self-rating, Bulimic Investigatory test Edinburgh, ED-15, Eating Attitudes Test, Eating Disorders Assessment for DSM-V, the Eating Disorders Examination-Questionnaire, Eating Disorder Inventory, the Structured Inventory for Anorexic and Bulimic Eating Disorders-Interview, Structured Inventory for Anorexic and Bulimic Eating Disorders-self-rating and the Short Evaluation for Eating Disorders in people already identified (e.g. in primary care) as having either an eating disorder or an early onset eating disorder, as defined by the DSM, ICD-10 or the semi-structured 'gold standard' EDE interview or the structured SCID-I-P were searched for. The studies were categorised according to the specific type of eating disorder assessed.

Table 26: Clinical review protocol for the review question: What is the validity and reliability of the instruments, tools and methods used to assess and monitor eating disorders?

Component	Description
Review question(s)	What is the validity and reliability of the instruments, tools and methods used to assess and monitor eating disorders?
Population	Children, young people and adults with: • early onset of eating disorders, e.g. people with body shape dissatisfaction • clinical samples (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). Strata: • children (≤12), young people (13-≤17 years), adults ≥18 years • Exclude: • People with disordered eating because of a physical health problem or another primary mental health problem of which a disorder of eating is a symptom (for example, depression). • People with feeding disorders, such as pica or avoidant restrictive food intake disorders (for example, food avoidance emotional disorder or picky/selective eating). • People with obesity without an eating disorder. • People from the general population where the tool would be used for screening.
Instruments, tools and methods	• The following will be investigated as a tool to use after a suspected index case has been raised: • EAT (Eating Attitudes Test; including different versions: EAT-40, EAT-26, ChEAT etc.). • EDI (Eating Disorder Inventory; distinguish between different versions)

Component	Description
	• BITE (Bulimic Investigatory Test, Edinburgh) • EDE-Q (Eating Disorder Examination-Questionnaire; distinguish between different versions) • SEED • ED-15 • The Structured Inventory for Anorexic and Bulimic Eating Disorders: available as a structured clinical interview for experts (SIAB-EX) and as a self-rated questionnaire (SIAB-S) • Munich ED-Quest (Munich Eating Disorder Questionnaire) • ANIS (Anorexia Nervosa Inventory for Self-rating) • EDA-5 (The Eating Disorder Assessment for DSM-5; for feeding or eating disorders or related conditions according to the DSM-5 criteria)
Reference	Gold standard, relevant ED definition as reported in: • DSM • ICD-10 • EDE –Interview • SCID (1)
Critical outcomes	• Sensitivity (Se): the proportion of true positives of all cases diagnosed in the population • Specificity (Sp): the proportion of true negatives of all cases not-diagnosed in the population • Positive predictive value • Negative predictive value • Likelihood values
Important outcomes	VALIDITY • Concurrent validity, convergent validity, construct validity, content validity, predictive and discriminant validity RELIABILITY • Inter-rater reliability. Intra-rater reliability, test re-test reliability, , internal consistency
Study design	• RCTs • Cohort • Cross-sectional

4.3.1 What is the validity and reliability of the instruments, tools and methods used to assess and monitor eating disorders

Due to the numerous studies that reported important outcomes only and the paucity of studies reporting critical outcomes on the aforementioned assessment tools, it was decided to include only studies that reported sensitivity and specificity (or reported data from which these could be derived) and moreover to include case-control studies. No RCTs that satisfied the eligibility criteria for this review were found. 15 cohort, cross sectional or case-control studies were identified (seven cohort or cross sectional studies, eight case-control studies), the majority of which were in adult females (Allen 2011 (Allen et al., 2011), Alvarez-Rayon 2004 (Alvarez-Rayon et al., 2004), Berg 2012 (Berg et al., 2012), Fichter 2000 (Fichter and

Quadflieg, 2000), Fichter 2001 (Fichter and Quadflieg, 2001), Fichter 2015 (Fichter et al., 2015), Henderson 1987 (Henderson and Freeman, 1987), Rivas 2010 (Rivas et al., 2010), Ro 2015 (Ro et al., 2015), Sysko 2015 (Sysko et al., 2015), Thurfjell 2003 (Thurfjell et al., 2003), Vander Wal 2011 (Vander Wal et al., 2011), Waller 1992 (Waller, 1992)). Only 1 study was found that evaluated an assessment tool specifically designed for use in children and young people (Thurfjell 2003). No studies that reported the critical outcomes of sensitivity and specificity were found for the ANIS, ED-15 and SEED assessment tools. An overview of the included studies can be found in Table 27, whilst an overview of risk of bias and indirectness can be found in Table 28. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

For the majority of assessment tools, only one or two studies were found that met the eligibility criteria. Whilst four studies were found for the EDE-Q, only two of these were cohort or cross-sectional studies. Since case-control studies overestimate the accuracy of a test, a meta-analysis was not conducted on the EDE-Q and data from such studies are presented separately from those of cohort and cross-sectional studies. In the case that a study reported data for more than one threshold for a given assessment tool, the data from the threshold recommended by the study was used. To enable visual comparisons between tests for any eating disorder or a specific type of disorder, the sensitivity and specificity of the assessment tools were plotted on a ROC curve.

4.3.1.1 Any eating disorder

Children and young people

One case-control study (n=2274) was identified (Thurfjell 2003) that examined an assessment tool specifically designed for children and young people. The majority of participants in this study were female. The quality of evidence is presented in Table 28 and Table 29. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Children, young people and adults

One cohort study (n=212) (Allen 2011) using the EDE-Q, which was conducted in a tertiary care setting, and three case-control studies using the EAT-40 (n=556) (Alvarez-Rayon 2004), EAT-26 (n=172) (Rivas 2010) and EDE-Q (n=2465) (Ro 2015) were identified. The majority of participants in the cohort study, and all the participants in the case-control studies, were female. The quality of evidence is presented in Table 30 and Table 31. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

4.3.1.2 Anorexia nervosa

Six cohort and cross-sectional studies using the EDE-Q for DSM-IV (n=429) (Allen 2001, Berg 2012), EDE-Q for DSM-V (n=217) (Berg 2012), Munich ED-Quest (n=195) (Fichter 2015), SIAB-EX (n=80) (Fichter 2001), EDA-5 (n=66) (Sysko 2015) and EDA-5 App (n=71) (Sysko 2015) and 3 case-control studies using the EAT-40 (n= 556) (Alvarez-Rayon 2004), EDE-Q (n= 2465) (Ro 2015) and BITE (n=81) (Waller 1992), were identified. The majority of cohort and cross-sectional studies were conducted in tertiary care settings, whilst the majority of participants were female. One case-control study that examined the BITE assessment tool, which was originally designed to assess binge eating, evaluated its utility in assessing the restricting and binge-purge subtypes of anorexia nervosa (Waller 1992).

The quality of evidence is presented in Table 32 and Table 33. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

4.3.1.3 Bulimia nervosa

Six cohort and cross-sectional studies using the EDE-Q for DSM-IV (n=429 (Allen 2001, Berg 2012), EDE-Q for DSM-V (n=217) (Berg 2012), Munich ED-Quest (n=195) (Fichter 2015), SIAB-EX (n=80) (Fichter 2001), EDA-5 (n=66) (Sysko 2015) and EDA-5 App (n=71) (Sysko 2015) and 3 case-control studies using the EAT-40 (n= 556) (Alvarez-Rayon 2004), EDE-Q (n= 2465) (Ro 2015) and BITE (n=81) (Waller 1992), were identified. The majority of cohort and cross-sectional studies were conducted in tertiary care settings, whilst the majority of participants were female. One case-control study that examined the BITE assessment tool, which was originally designed to assess binge eating, evaluated its utility in assessing bulimia nervosa with and without a history of anorexia nervosa (Waller 1992).

The quality of evidence is presented in Table 34 and Table 35. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

4.3.1.4 Anorexia and bulimia nervosa

Two cohort studies using the SIAB-EX (n=377) (Fichter 2000) and SIAB-S (n=80) (Fichter 2001) and 1 case-control study using the EAT-40 (n=556) (Alvarez-Rayon 2004) were identified. Both cohort studies were conducted in a secondary care setting and were in both adult males and females. The quality of evidence is presented in Table 36 and Table 37. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

4.3.1.5 Binge eating disorder

Three cohort and cross-sectional studies using EDA-5 (n=66) (Sysko 2015), EDA-5 App (n=71) (Sysko et al., 2015) (Sysko 2015) and EDE-Q (n=217) (Berg 2012) and 3 case-control studies using BITE (n=119) (Henderson 1987) and EDE-Q (n=41 (Vander Wal 2011) were identified. All three cohort and cross-sectional studies were conducted in a tertiary care setting and the majority of participants were female.

The quality of evidence is presented in Table 38 and Table 39. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

4.3.1.6 EDNOS

Four cohort or cross-sectional studies using the EDA-5 (n=66) (Sysko 2015), EDA-5 App (n=71) (Sysko 2015) and EDE-Q (n=429) (Allen 2001, Berg 2012) and 2 case-control studies using the EAT-40 (n=556) (Alvarez-Rayon 2004) and EDE-Q for DSM-IV (n=2465) (Ro 2015) were identified. All four cohort and cross-sectional studies were conducted in a tertiary care setting and the majority of participants were female.

The quality of evidence is presented in Table 40 and Table 41. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 27: Study information for review on assessment tools of eating disorders in early onset or clinical populations

Study ID	Country	Assessment tool	ED Reference definition	Sample	Sample N	Age (years)	Female (%)	Eating Disorder assessed
Henderson 1987 Study 1	UK	BITE	DSM-III	Case-control: (I) Binge eaters=15 (II) Controls=40	55	24.2 (5.5)	76	BED
Henderson 1987 Study 2	UK	BITE	DSM-III	Case-control: (I) BN=32 (II) Controls=32	64	25.1 (5.7)	100	BED
Waller 1992	UK	BITE	Interview, DSM-III-R	Case-control: (I) ED outpatients=81 (II) Controls=27	81	25.7 (7.6)	100	AN-R, AN-BP, BN with history of AN, BN without history of AN
Rivas 2010 Study 2	Spain	EAT-26	Spanish Q-EDD, DSM-IV	Case-control: (I) ED outpatients=86 (II) Controls=79	172	18.6 (4.4) Range 12-35	100	Any ED
Alvarez-Rayon 2004	Mexico	EAT-40	DSM-IV	Case-control: (I) ED outpatients=276 (ii) Controls=280	556	19.4 (3.9)	100	Any ED AN, BN, AN or BN, EDNOS
Sysko 2015 Study 1	USA	EDA-5	EDE, Version 16, DSM-IV	Cross sectional: individuals seeking or receiving treatment for ED at tertiary care centre	66	30.9 (11) Range 14-58	89	AN, BN, BED EDNOS or OSFED/USFED
Sysko 2015 Study 2	USA	EDA-5 App	Clinician interview, DSM-V	Cohort: receiving treatment for ED, tertiary care	71	32.7 (11.9) Range 18-65	94	AN, BN, BED EDNOS or OSFED/USFED
Allen 2011	Australia	EDE-Q	EDE,	Cohort:	212	26.18	99	Any ED

Study ID	Country	Assessment tool	ED Reference definition	Sample	Sample N	Age (years)	Female (%)	Eating Disorder assessed
			DSM-IV	ED outpatients, tertiary care		(9.21) Range 16-72		AN, BN, EDNOS
Berg 2012	USA	EDE-Q	EDE, DSM-IV	Cross-sectional: ED outpatients, tertiary care	217	19.6 (9.6) Range 9-61	90	AN, BN, BED, EDNOS
Ro 2015	Norway	EDE-Q	Clinical diagnosis, ICD-10 transforme d to DSM-IV	Case-control: (I) ED in- and out- patients=620 (II) Controls=1845	2465	30.7 (10.3) Range 16-66	100	Any ED AN, BN, EDNOS
Vander Wal 2011	USA	EDE-Q	Diagnostic items EDE, Version 12, DSM-IV	Case-control: Overweight or obese individuals with BED (n=15) and without BED (n=26)	41	52.0 (12.1)	71	BED
Thurfjell 2003	Sweden	EDI-C	Semi-structured interview adapted for young people, DSM-IV	Case-control: (I) ED patients from special ED unit=201 (II) Controls=2073	2274	15.7 (1.6)	100	Any ED
Fichter 2015 Sample 6	Germany	Munich ED-Quest	SIAB-EX, DSM-V	Cross-sectional: ED inpatients, tertiary care	195	21.7 (8.7)	100	AN, BN
Fichter 2000	Germany	SIAB-EX	EDE, DSM-IV	Cohort: ED inpatients, secondary care	377	29.1 (9.3)	97	AN or BN
Fichter 2001	Germany	SIAB-S	SIAB-EX	Cohort: ED inpatients, secondary care	80	28.8 (9.5)	96	AN, BN, AN or BN

Abbreviations: AN, anorexia nervosa; BN, bulimia nervosa; EDNOS eating disorder not otherwise specified: ED, eating disorder; EDE-Q, Eating Disorder Examination; EDI, Eating Disorder Inventory; BITE, Bulimic Investigatory Test, Edinburgh; DSM, diagnostic statistical manual of mental disorders; SIAB EX The Structured Inventory for Anorexic and Bulimic Eating Disorders: available as a structured clinical interview for experts; SIAB S The Structured Inventory for Anorexic and Bulimic Eating Disorders as a self-rated questionnaire; ANIS, Anorexia Nervosa Inventory for Self-rating; EDA-5, The Eating Disorder Assessment for DSM 5

Table 28: Summary table of QUADAS-2 results for risk of bias and indirectness for studies on assessmental tools of eating disorders

Study	RISK OF BIAS				INDIRECTNESS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
BITE							
Henderson 1987 Study 1	☹	☹	☺	☺	☺	☹	☺
Henderson 1987 Study 2	☹	☹	☺	☺	☺	☹	☺
Waller 1992	☹	☹	☹	☺	☺	☹	☺
EAT-26							
Rivas 2010 Study 2	☹	☹	☹	☺	☺	☹	☺
EAT-40							
Alvarez-Rayon 2004	☹	☹	☺	☹	☺	☹	☺
EDA-5							
Sysko 2015 Study 1	☺	☹	☺	☺	☺	☺	☺
EDA-5 App							
Sysko 2015 Study 2	☺	☹	☹	☺	☺	☺	☺
EDE-Q							
Allen 2011	☺	☺	☹	☺	☺	☺	☺
Berg 2012	☺	☹	☹	☹	☺	☺	☺
Ro 2015	☹	☹	☺	☹	☺	☹	☺
Vander Wal 2011	☹	☹	☺	☹	☹	☺	☺
EDI-C							
Thurfjell 2003	☹	☹	☹	☹	☺	☹	☺
Munich ED-Quest							

Study	RISK OF BIAS				INDIRECTNESS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Fichter 2015 Sample 6	☹	☹	☹	☺	☺	☺	☺
SIAB-EX							
Fichter 2000	☹	☺	☺	☺	☺	☺	☺
SIAB-S							
Fichter 2001 Study 2	☺	☹	☺	☺	☺	☺	☺
Key ☺=Low Risk ☹=High Risk							

Table 29: Clinical evidence profile for case-control studies on assessment tools for eating disorders in children and young people

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EDI-C	1								
Thurfjell 2003		2274	Very serious risk of biasa	Not applicableb	Serious indirectnessc	No serious imprecision[d]	0.74 (0.67, 0.79)	0.77 (0.75, 0.79)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 30: Clinical evidence profile for cohort and cross-sectional studies on assessment tools for eating disorders in children, young people and adults

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EDE-Q for DSM-IV	1									
Allen 2001		Cohort	232	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Not applicable[d]	0.64 (0.57, 0.7)	1.00[d]	LOW[d]

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) There were no false positives and no true negatives in the sample. A continuity correction was applied to allow calculation of specificity. The related confidence interval, and therefore imprecision, was not estimable. The overall quality of the evidence was therefore downgraded by one in this case.

Table 31: Clinical evidence profile for case-control studies on assessment tools for eating disorders in children, young people and adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EAT-26	1								
Rivas 2010		172	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.6 (0.48, 0.71)	0.95 (0.88, 0.99)	VERY LOW
EAT-40 ≥ 26	1								
Alvarez-Rayon 2004		556	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.83 (0.78, 0.87)	0.91 (0.87, 0.94)	VERY LOW
EDE-Q for DSM-IV ≥ 2.5	1								
Ro 2015		2465	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.86 (0.83, 0.89)	0.86 (0.84, 0.88)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 32: Clinical evidence profile for cohort and cross-sectional studies on assessment tools for anorexia nervosa in children, young people and adults

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EDA-5	1									
Sysko 2015 Study 1		Cross-sectional	66	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	1.0 (0.85 ,1.0)	0.83 (0.69, 0.93)	LOW
EDA-5 App	1									
Sysko 2015 Study 2		Cohort	71	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.83 (0.59, 0.96)	1.0 (0.93, 1.0)	LOW
EDE-Q for DSM-IV	2		429							
Allen 2001		Cohort	212	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	No serious imprecision[d]	0.71 (0.54, 0.85)	0.97 (0.93, 0.99)	MODERATE
Berg 2012		Cross-sectional	217	Very serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	No serious imprecision[d]	0.73 (0.39, 0.94)	0.99 (0.96, 1.0)	LOW
EDE-Q for DSM-V	1									
Berg 2013		Cross-sectional	217	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.73 (0.45, 0.92)	0.99 (0.96, 1.0)	LOW
Munich ED-Quest	1									
Fichter 2015		Cross-sectional	195	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.93 (0.89, 0.99)	0.98 (0.93, 1.0)	LOW
SIAB-EX	1									
Fichter 2001		Coh	80	Serious	Not applicable[b]	No serious	No serious	0.67	0.92	MODE

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
		ort		risk of bias[a]		indirectness[c]	imprecision[d]	(0.38, 0.88)	(0.83, 0.97)	RATE

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency was assessed by inspection of the sensitivity and specificity forest plots (based on the primary measure), using the point estimates and confidence intervals. The evidence was downgraded by 1 increment if the individual studies varied across 2 areas (for example, 50–90% and 91–100%) and by 2 increments if the individual studies varied across 3 areas (for example, 0–50%, 51–90% and 91–100%). Inconsistency not applicable if only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 33: Clinical evidence profile for case-control studies for anorexia nervosa in children, young people and adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
BITE for AN-R	1								
Waller 1992		81	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	1.0 (0.72, 1.00)	1.00 (0.95, 1.0)	VERY LOW
BITE for AN-BP	1								
Waller 1992		81	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.40 (0.12, 0.74)	1.0 (0.95, 1.0)	VERY LOW
EAT-40 ≥ 28	1								
Alvarez-Rayon 2004		556	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.85 (0.72, 0.93)	0.93 (0.90, 0.95)	VERY LOW
EDE-Q for DSM-IV ≥ 2.09	1								
Ro 2015		2465	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.8 (0.73, 0.86)	0.8 (0.78, 0.82)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 34: Clinical evidence profile for cohort and cross-sectional studies on assessment tools for bulimia nervosa in children, young people and adults

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EDA-5	1									
Sysko 2015 Study 1		Cross sectional	66	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.71 (0.44, 0.9)	0.96 (0.85, 0.99)	MODERATE
EDA-5 App	1									
Sysko 2015 Study 2		Cohort	71	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.96 (0.8, 1.0)	0.98 (0.88, 1.0)	LOW
EDE-Q for DSM-IV	2									
Allen 2001		Cohort	212	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	No serious imprecision[d]	0.53 (0.42, 0.63)	0.88 (0.81, 0.94)	MODERATE
Berg 2012		Cross-sectional	217	Very serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	No serious imprecision[d]	0.74 (0.58, 0.87)	0.91 (0.86, 0.95)	LOW
EDE-Q for DSM-V	1									
Berg 2012		Cross-sectional	217	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.73 (0.58, 0.85)	0.94 (0.89, 0.97)	LOW
Munich ED-Quest	1									
Fichter 2015		Cross-sectional	195	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.73 (0.57, 0.85)	0.97 (0.93, 0.99)	LOW
SIAB-EX	1									
Fichter 2001		Coh	80	Serious	Not applicable[b]	No serious	Serious	0.63	0.79	MODE

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
		ort		risk of bias[a]		indirectness[c]	imprecision[d]	(0.44, 0.79)	(0.65, 0.9)	RATE

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency was assessed by inspection of the sensitivity and specificity forest plots (based on the primary measure), using the point estimates and confidence intervals. The evidence was downgraded by 1 increment if the individual studies varied across 2 areas (for example, 50–90% and 90–100%) and by 2 increments if the individual studies varied across 3 areas (for example, 0–50%, 51–90% and 91–100%). Inconsistency not applicable if only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 35: Clinical evidence profile for case-control studies on assessment tools for bulimia nervosa in children, young people and adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
BITE for BN with history of AN	1								
Waller 1992		81	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.83 (0.61, 0.95)	1.0 (0.94, 1.0)	VERY LOW
BITE for BN without history of AN	1								
Waller 1992		81	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.9 (0.55, 1.0)	1.0 (0.95, 1.0)	VERY LOW
EAT-40 ≥ 28	1								
Alvarez-Rayon 2004		556	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.87 (0.79, 0.93)	0.93 (0.90, 0.95)	VERY LOW
EDE-Q for DSM-IV ≥ 2.62	1								
Ro 2015		2465	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.87 (0.82, 0.91)	0.87 (0.86, 0.88)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the

QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision.

Table 36: Clinical evidence profile for cohort and cross-sectional studies on assessment tools for anorexia or bulimia nervosa in children, young people and adults

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
SIAB-EX	1									
Fichter 2001		Cohort	80	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.68 (0.53, 0.81)	0.79 (0.61, 0.91)	LOW
SIAB-S	1									
Fichter 2000		Cohort	377	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.7 (0.64, 0.75)	0.8 (0.7, 0.87)	MODERATE

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details (b) Inconsistency not applicable due to only one study;(c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 37: Clinical evidence profile for case-control studies on assessment tools for anorexia and bulimia nervosa in children, young people and adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EAT-40 ≥ 28	1								
Alvarez-Rayon 2004		556	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.86 (0.79, 0.91)	0.94 (0.91, 0.96)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 38: Clinical evidence profile for cohort and cross-sectional studies on assessment tools for binge eating disorder in children, young people and adults

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EDA-5	1									
Sysko 2015 Study 1		Cross-sectional	66	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	1.0 (0.48, 1.0)	0.98 (0.91, 1.0)	MODERATE
EDA-5 App	1									
Sysko 2015 Study 2		Cohort	71	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.86 (0.57, 0.98)	0.96 (0.88, 1.0)	LOW
EDE-Q for DSM-IV	1									
Berg 2012		Cross-sectional	217	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.28 (0.1, 0.53)	0.97 (0.94, 0.99)	LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study;(c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 39: Clinical evidence profile for case-control studies on assessment tools for binge eating disorder in children, young people and adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
BITE	2	119							
Henderson 1987 (Study 1)		55	Serious risk of bias[a]	No serious consistency[b]	Serious indirectness[c]	No serious imprecision[d]	1.0 (0.89, 1.0)	1.0 (0.89, 1.0)	VERY LOW
Henderson 1987 (Study 2)		64	Serious risk	No serious	Serious	No serious	0.93	0.95	VERY

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
			of bias[a]	consistency[b]	indirectness[c]	imprecision[d]	(0.68, 1.0)	(0.83, 0.99)	LOW
EDE-Q for DSM-IV ≥ 3.2-3.3	1								
Vander Wal 2011		41	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	Serious imprecision[d]	0.73 (0.45, 0.92)	0.81 (0.61, 0.93)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study;(c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 40: Clinical evidence profile for cohort and cross-sectional studies on assessment tools for EDNOS in children, young people and adults

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EDA-5	1									
Sysko 2015 Study 1		Cross-sectional	66	Serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.65 (0.38, 0.86)	0.96 (0.85, 0.99)	MODERATE
EDA-5 App	1									
Sysko 2015 Study 2		Cohort	71	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	Serious imprecision[d]	0.73 (0.39, 0.94)	0.9 (0.79, 0.96)	VERY LOW
EDE-Q for DSM-IV	2		439							
Allen 2001		Cohort	212	Serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	No serious imprecision[d]	0.72 (0.61, 0.81)	0.63 (0.54, 0.71)	MODERATE
Berg 2012		Cross-secti	217	Very serious risk of bias[a]	No serious inconsistency[b]	No serious indirectness[c]	Serious imprecision[d]	0.84 (0.77, 0.9)	0.71 (0.6, 0.81)	VERY LOW

Index test	Number of studies	Type of study	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
		onal								
EDE-Q for DSM-V	1									
Berg 2012		Cross-sectional	217	Very serious risk of bias[a]	Not applicable[b]	No serious indirectness[c]	No serious imprecision[d]	0.82 (0.74, 0.88)	0.76 (0.67, 0.84)	LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency was assessed by inspection of the sensitivity and specificity forest plots (based on the primary measure), using the point estimates and confidence intervals. The evidence was downgraded by 1 increment if the individual studies varied across 2 areas (for example, 50–90% and 91–100%) and by 2 increments if the individual studies varied across 3 areas (for example, 0–50%, 51–90% and 91–100%). Inconsistency not applicable if only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

Table 41: Clinical evidence profile for case-control studies on assessment tools for EDNOS in children, young people and adults

Index test	Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Sensitivity (95% CI)	Specificity (95% CI)	Quality
EAT-40 ≥ 22	1								
Alvarez-Rayon 2004		556	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.83 (0.75, 0.89)	0.82 (0.78, 0.86)	VERY LOW
EDE-Q for DSM-IV ≥ 2.63	1								
Ro 2015		2465	Very serious risk of bias[a]	Not applicable[b]	Serious indirectness[c]	No serious imprecision[d]	0.88 (0.83, 0.92)	0.88 (0.87, 0.89)	VERY LOW

Notes: The assessment of the evidence quality was conducted with emphasis on test specificity as this was the primary measure discussed in decision-making; (a) Risk of bias was assessed using the QUADAS-2 checklist. See Table 28 for more details; (b) Inconsistency not applicable due to only one study; (c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. See Table 28 for more details; (d) Due to the small number of studies, a range of 0–20% of differences in point estimates of sensitivity was considered not imprecise, 21–40% serious imprecisions, and >40% very serious imprecision. Imprecision was assessed on the primary measure of specificity for decision-making.

4.3.2 Economic Evidence

No economic evidence on the tools for the assessment and monitoring of eating disorders was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

4.3.3 Clinical evidence statements

4.3.3.1 Any eating disorder

The quality of evidence from the one case control study (n=2274) that examined the use of an assessment tool (the EDI-C) in children and young people with eating disorders was very low due to concerns over risk of bias and indirectness. The reported specificity was 0.77 (with 95% CI ranging from 0.75 to 0.79), whilst the sensitivity was 0.74 (with 95% CI ranging from 0.67 to 0.79).

The quality of evidence from the one cohort study (n=212) that examined the use of the EDE-Q in an eating disorder population (i.e. children, young people and adults) was low due to concerns about risk of bias. The reported specificity was 1.00 (the 95% CI was not estimable) whilst the sensitivity was 0.64 (with 95% CI ranging from 0.57 to 0.70).

The quality of evidence from the three case control studies that examined the use of an assessment tool in an eating disorder population (i.e. children, young people and adults) was very low due to concerns about risk of bias and indirectness. The assessment tools examined were EAT-26 (n=172), EAT-40 at a threshold of 26 or more (n=556) and EDE-Q for DSM-IV at a threshold of 2.5 or more (n=2465). The specificity of the tools were relatively similar: the EAT-26 had the highest specificity of 0.95 (with 95% CI ranging from 0.88 to 0.99), followed by the EAT-40 with a specificity of 0.91 (with 95% CI ranging from 0.87 to 0.94) and the EDE-Q for DSM-IV with a specificity of 0.86 (with 95% CI ranging from 0.84 to 0.88). By contrast, the sensitivity of the tools were wider: the EDE-Q had the highest sensitivity of 0.86 (with 95% CI ranging from 0.83 to 0.89), followed by the EAT-40 with a sensitivity of 0.83 (with 95% CI ranging from 0.78 to 0.87) and the EAT-26 with a specificity of 0.6 (with 95% CI ranging from 0.48 to 0.71).

4.3.3.2 Anorexia nervosa

The quality of evidence from the six cohort or cross-sectional studies examining assessment tools for anorexia nervosa in eating disordered populations ranged from moderate to low due to concerns about risk of bias. The evaluated assessment tools included the interview version of EDA-5 (n=66), the electronic application version of EDA-5 (n=71), EDE-Q for DSM-IV (n=429), EDE-Q for DSM-V (n=217), Munich ED-Quest (n=195) and SIAB-EX (n=80). The reported specificity of these tools was relatively high and the related 95% CIs were also relatively narrow: the EDA-5 App had the highest specificity of 1.00 (with 95% CI ranging from 0.93 to 1.00), followed by the EDE-Q (specificity ranging from 0.97 to 0.99, with 95% CI ranging from 0.93 to 1.00 for DSM-IV, and specificity of 0.99 with 95% CI ranging from 0.96 to 1.00 for EDE-Q for DSM-V) and Munich ED-Quest (specificity=0.98 with 95% CI ranging from 0.93 to 1.00), the SIAB-EX (specificity= 0.92 with 95% CI ranging from 0.83 to 0.97) and the interview version of the EDA-5 (specificity=0.83 with 95% CI ranging from 0.69 to 0.93). However, the reported sensitivity of the tools was more variable and the 95% CIs were relatively wide: the interview version of the EDA-5 had the highest sensitivity of 1.00 (with 95% CI ranging from 0.85 to 1.00), followed by Munich ED-Quest (sensitivity=0.93 with 95% CI ranging from 0.89 to 0.99), EDA-5 App (sensitivity=0.83 with 95% CI ranging from 0.59 to 0.96), EDE-Q (sensitivity ranging from 0.71 to 0.73 with 95% CI ranging from 0.39 to

0.94 for DSM-IV; sensitivity=0.73 with 95% CI ranging from 0.45 to 0.92 for DSM-V) and SIAB-EX (sensitivity=0.67 with 95% CI ranging from 0.38 to 0.88).

The quality of evidence for the three identified case control studies was very low due to concerns over risk of bias and indirectness. The EAT-40 at a threshold of 28 had the highest specificity of 0.93 (with 95% CI ranging from 0.90 to 0.95) for the assessment of anorexia nervosa, followed by the EDE-Q at a threshold of 2.09 (specificity=0.8 with 95% CI ranging from 0.78 to 0.82). The EAT-40 at a threshold of 28 had the highest sensitivity of 0.85 (with 95% CI ranging from 0.72 to 0.93), followed by the EDE-Q at a threshold of 2.09 (sensitivity=0.8 with 95% CI ranging from 0.73 to 0.86).

The specificity of the BITE for both restricting and binge-purge subtypes of anorexia nervosa was 1.00 and the 95% CIs ranged from 0.95 to 1.0 for both subtypes. However, whilst the sensitivity of BITE for the restricting subtype was 1.0 (with 95% CI ranging from 0.72 to 1.00), the sensitivity for the binge-purge subtype was much lower at 0.4 (with 95% CI ranging from 0.12 to 0.74).

4.3.3.3 Bulimia nervosa

The quality of evidence from the six cohort and cross-sectional studies examining assessment tools for bulimia nervosa in eating disordered populations ranged from moderate to low due to concerns about risk of bias. The evaluated assessment tools included the interview version of EDA-5 (n=66), the electronic application version of EDA-5 (n=71), EDE-Q for DSM-IV (n=429), EDE-Q for DSM-V (n=217), Munich ED-Quest (n=195) and SIAB-EX (n=80). The reported specificity of these tools was relatively high and the related 95% CIs were also relatively narrow: the EDA-5 App had the highest specificity of 0.98 (with 95% CI ranging from 0.88 to 1.00), followed by Munich ED-Quest (specificity=0.97 with 95% CI ranging from 0.93 to 0.99), the interview version of EDA-5 (specificity=0.96 with 95% CI ranging from 0.85 to 0.99), EDE-Q (specificity ranging from 0.88 to 0.91 with 95% CI ranging from 0.81 to 0.95 for DSM-IV; specificity=0.94 with 95% CI ranging from 0.89 to 0.97 for DSM-V) and SIAB-EX (specificity=0.79 with 95% CI ranging from 0.65 to 0.90). The reported sensitivity of these tools and their related 95% CIs was much more wide-ranging, with the EDA-5 App also having the highest sensitivity of 0.96 (with 95% CI ranging from 0.80 to 1.00), followed by the Munich ED-Quest (sensitivity=0.73 with 95% CI ranging from 0.57 to 0.85) and EDE-Q (sensitivity ranging from 0.53 to 0.74 with 95% CI ranging from 0.42 to 0.87 for DSM-IV; sensitivity=0.73 with 95% CI ranging from 0.58 to 0.85 for DSM-V), EDA-5 (sensitivity=0.71with 95% CI ranging from 0.44 to 0.90) and SIAB-EX (sensitivity=0.63 with 95% CI ranging from 0.44 to 0.79).

The quality of evidence for the three identified case control studies was very low due to concerns over risk of bias and indirectness. The EAT-40 at a threshold of 28 had the highest specificity of 0.93 (with 95% CI ranging from 0.90 to 0.95) for the assessment of bulimia nervosa, followed by the EDE-Q at a threshold of 2.62 (specificity=0.87 with 95% CI ranging from 0.86 to 0.88). The EAT-40 at a threshold of 28 also had the highest sensitivity of 0.85 (with 95% CI ranging from 0.72 to 0.93), followed by the EDE-Q at a threshold of 2.62 (sensitivity=0.87 with 95% CI ranging from 0.82 to 0.91).

The specificity of the BITE for the assessment of people with bulimia nervosa with and without a history of anorexia nervosa was 1.0 with the related 95% CI ranging from 0.94 to 1.0 for the former, and from 0.95 to 1.0 for the latter. The sensitivity of BITE to bulimia nervosa with and without a history of anorexia nervosa was also similar at 0.83 (95% CI from 0.61 to 0.95) and 0.9 (95% CI from 0.55 to 1.0) respectively.

4.3.3.4 Anorexia or bulimia nervosa

The quality of evidence for the two cohort studies that examined an assessment tool for either anorexia or bulimia nervosa in eating disordered populations was moderate to low due

to concerns over risk of bias and/or imprecision. The assessment tools evaluated were the SIAB-EX (n=80) and SIAB-S (n=377). The SIAB-S had the highest specificity of these two studies (0.8, 95% CI ranging from 0.7 to 0.87) closely followed by the SIAB-EX (0.79, 95% CI ranging from 0.61 to 0.91). Similarly, the SIAB-S had the highest sensitivity (0.7, 95% CI ranging from 0.64 to 0.75), again closely followed by the SIAB-EX (0.68, 95% CI ranging from 0.51 to 0.81).

The quality of evidence for the one case control study that examined an assessment tool for the evaluation of anorexia or bulimia nervosa was very low due to concerns over risk of bias and indirectness. The specificity and sensitivity of the case control study on EAT-40 at a threshold of 28 or more (n=556) was 0.94 (with 95% CI ranging from 0.91 to 0.96) and 0.86 (with 95% CI ranging from 0.79 to 0.91), respectively.

4.3.3.5 Binge eating disorder

The quality of evidence from the three cohort or cross-sectional studies that examined assessment tools for binge eating disorder in eating disordered populations ranged from moderate to low. The tools evaluated were the interview (n=66) and electronic application versions (n=71) of the EDA-5 and the EDE-Q for DSM-IV (n=217). The interview version of the EDA-5 had the highest specificity (0.98, 95% CI ranging from 0.91 to 1.00), followed by EDE-Q for DSM-IV (0.97, 95% CI ranging from 0.94 to 0.99) and the electronic application version of EDA-5 (0.96, 95% CI ranging from 0.88 to 1.00). The reported sensitivity of these studies was much more wide ranging, with the EDA-5 having a high sensitivity (1.0, 95% CI ranging from 0.48 to 1.00), followed by the EDA-5 App (0.86. 95% CI ranging from 0.57 to 0.98). The EDE-Q for DSM-IV had a very low sensitivity of 0.28 (95% CI ranging from 0.1 to 0.53).

The quality of evidence for the three case control studies was very low. The reported specificity of the three case control studies were from 0.93 to 1.0 for BITE (with 95% CI ranging from 0.83 to 1.0) and 0.73 (95% CI ranging from 0.61 to 0.93) for EDE-Q for DSM-IV at a threshold of between 3.3 and 3.33. The reported sensitivity for these studies was from 0.93 to 1.00 for the BITE (95% CIs ranging from 0.68 to 1.0) and 0.73 for EDE-Q for DSM-IV (95% CI ranging from 0.45 to 0.92).

4.3.3.6 EDNOS

Very low to moderate quality evidence from five cohort or cross-sectional studies (n= 793) showed the EDA-5 interview has the highest specificity (0.96) for assessing EDNOS in clinical or early onset populations compared with EDE-5 App, EDE-Q for DSM-IV or DSM-V.

Very low to moderate quality evidence from five cohort or cross-sectional studies (n= 793) showed other assessment tools (EDE-5 App, EDE-Q for DSM-IV or DSM-V) were neither specific nor accurate.

Very low quality evidence from two case-control studies (n= 3021) showed EAT-40 at a threshold of 22, and EDE-Q for DSM-IV at a threshold of 2.63 were not particularly accurate.

The quality of evidence for the five cohort or cross-sectional studies that examined an assessment tool for EDNOS in eating disordered populations ranged from moderate to very low due to concerns over risk of bias and/or imprecision. The tools evaluated included the interview (n=66) and electronic application (n=71) versions of EDA-5 and the EDE-Q for DSM-IV (n=439) and DSM-V (n=217). The tool with the highest reported specificity of 0.96 was the EDA-5 (95% CI ranging from 0.85 to 0.99), followed by EDA-5 App (0.9, 95% CI ranging from 0.79 to 0.96), EDE-Q for DSM-V (0.76, 95% CI ranging from 0.67 to 0.84) and EDE-Q for DSM-IV (specificity ranging from 0.63 to 0.71, with 95% CI ranging from 0.54 to 0.81). The reported sensitivity of these studies was more wide ranging: the EDE-Q for DSM-IV had a reported sensitivity of 0.72 and 0.84 (with 95% CI ranging from 0.61 to 0.9), whilst the EDE-Q for DSM-V had a sensitivity of 0.82 (95% CI ranging from 0.74 to 0.88). The EDA-

5 App had a sensitivity of 0.73 (95% CI ranging from 0.39 to 0.94), whilst the EDA-5 had a sensitivity of 0.65 (95% CI ranging from 0.38 to 0.86). (Note that since both the EDA-5 and EDA-5 App are intended for use with DSM-V, it can be used to assess Other Specified Feeding and Eating Disorders [OSDED] and Unspecified Eating and Feeding Disorder [USFED] in addition to the DSM-IV category of EDNOS.)

The quality of evidence for the two case control studies was very low due to concerns over risk of bias and indirectness. The reported specificity of the EAT-=40 at a threshold of 22 or more was 0.82 (95% CI ranging from 0.78 to 0.86), whilst its sensitivity was 0.83 (95% CI ranging from 0.75 to 0.89). The specificity of the EDE-Q for DSM-IV at a threshold of 2.63 or more was 0.88 (95% CI ranging from 0.87 to 0.89), whilst its sensitivity was 0.88 (95% CI ranging from 0.83 to 0.92).

4.3.4 Economic Evidence statements

No economic evidence on the tools for the assessment and monitoring of eating disorders was available.

4.3.5 Recommendations and link to evidence for the review on: What is the validity and reliability of the instruments, tools and methods used to assess and monitor eating disorders?

Initial assessments in primary and secondary mental health care

2. **Be aware that eating disorders present in a range of settings, including:**
    - primary and secondary health care
    - social care
    - education

3. **Think about the possibility of an eating disorder in children and young people with poor growth (for example a low weight or height for their age).**

4. **Think about the possibility of an eating disorder in people with one or more of the following:**
    - an unusually low or high BMI or body weight for their age
    - dieting or restrictive eating practices (such as dieting when they are underweight) that are worrying them, their family members or carers, or professionals
    - family members or carers report a change in eating behaviour
    - other mental health problems
    - a disproportionate concern about their weight (for example, concerns about weight gain as a side effect of contraceptive medication)
    - problems managing a chronic illness that affects diet, such as diabetes
    - menstrual or other endocrine disturbances, or unexplained gastrointestinal symptoms
    - physical signs of:

	○ starvation, such as poor circulation, dizziness, palpitations, fainting or pallor ○ compensatory behaviours, such as laxative misuse, vomiting or excessive exercise ○ dental erosion • taking part in activities associated with a high risk of eating disorders (for example, professional sport, fashion, dance, or modelling). **5. When assessing for an eating disorder, think about all of the points in recommendation 3 regardless of the person's gender, ethnicity or socio-economic background.** **6. Think about the possibility of an eating disorder in children and young people with poor growth (for example a low weight or height for their age).** **7. Be aware that the risk of eating disorders is highest in young women (13 to 17 years), and that young men are also at greater risk between 13 and 17 years than at other ages.** **8. Do not use single measures such as BMI or duration of illness to determine whether to offer treatment for an eating disorder.** **9. Professionals in primary and secondary mental health settings should assess the following in people with a suspected eating disorder:** • their physical health, including checking for any physical effects of starvation or of compensatory behaviours such as vomiting • the presence of mental health problems commonly associated with eating disorders, including depression, anxiety, self-harm and obsessive compulsive disorder • the possibility of alcohol or substance misuse. • the need for emergency care in people whose physical health is compromised or who have a suicide risk.
Relative value of different outcomes	For the review on the utility and validity of tools that may be used for assessment and/or monitoring of eating disorders, the committee considered the critical outcomes were sensitivity, specificity, positive predictive value, and negative predictive value and likelihood ratio values. Other outcomes were considered important but studies were not included if they did not measure any of the critical outcomes. Important outcomes included numerous validity and reliability measures. The critical outcome of specificity was used as the primary measure for decision making given the need to minimise false positives when seeking to identify eating disorder cases in people with an eating disorder or in people who are in the early stages of developing one. That is, such a test needs to minimise the number of false positives so that the test is more exclusive and ensures people who do not have an eating disorder are not given unnecessary treatment (e.g. at the secondary care stage). The GC evaluated the performance of the relevant tests on the critical outcomes of

	sensitivity and specificity (and the related number of false positives and negatives) and decided that identification of eating disorders should be on the basis of clinical judgment via a full diagnostic 'gold standard' interview. Since such an interview is by definition valid and reliable in the UK, further analysis of the important outcomes for the relevant tools was not deemed necessary. The outcomes of positive and negative predictive value, and positive and negative likelihood ratios were presented although not considered by the committee. No review was conducted on where healthcare professionals may expect to find people with an eating disorder, or what they should consider when conducting an initial assessment including any safeguarding concerns.
Trade-off between clinical benefits and harms	For the review on what tools are effective at assessing those with an eating disorder, the critical outcome is specificity. Evidence was found on a number of tests including Eating Disorder Examination-Questionnaire, Munich-ED, the 40-item Eating Attitudes Test, Structured Expert Interview for Anorexic and Bulimic Syndromes and Structured Inventory for Anorexic and Bulimic Syndromes self-report, EDE-Q, EDA-5 (interview and electronic application) and Bulimic Investigatory Test, Edinburgh. To assess whether people have anorexia nervosa, the best tests appear to be EDA-5 App, EDE-Q (for both DSM-IV and DSM-5), and Munich-ED, which all had a specificity greater than 97%. To assess whether people have bulimia nervosa, the best tests appear to be EDA-5 App, Munich ED-Quest and EDA-5, which all had a specificity greater than 95%. To assess whether people have binge eating disorder, the best tests appear to be EDA-5, EDE-Q for DSM-IV and EDA-5 App, which all had a specificity over 95%. For OSFED, the best test appears to be EDA-5, which had a specificity over 95%. Overall, the EDA-5 appears to be most versatile for assessing anorexia nervosa, bulimia nervosa and OSFED in adults. Despite the relatively high specificity of the above tests, and given the low number of studies on each assessment tool and wide range in quality of evidence, the committee expressed the view that the risk of false positives (and hence inappropriate treatment) due to the use of these assessment tools outweighed the potential benefits, especially given that a full clinical diagnostic 'gold standard' interview, which by definition has 100% specificity, would be required. In lieu of any recommendation to use a particular assessment tool, the committee based their recommendations above on their clinical experience and current practice.
Trade-off between net health benefits and resource use	There was no evidence on the cost effectiveness of methods for the assessment and monitoring of eating disorders. The committee expressed the view that, in principle, if assessment and monitoring were to lead to the timely identification, and appropriate treatment, of an eating disorder then the additional costs associated with undertaking such assessment and monitoring would likely be outweighed by both the longer term improvements in health outcomes and the potential future cost savings to the healthcare system, given that delays in treatment exacerbate symptoms. Furthermore, providing timely assessment and monitoring may prevent the need of expensive secondary care. Whilst the committee agreed this in principle, they were of the opinion that in practice, there was little point in taking the time to administer the reviewed assessment tools when a full clinical diagnostic interview (such as the 'gold standard' EDE) would in any case be required.
Quality of evidence	The overall quality of evidence of the diagnostic studies was assessed using a modified GRADE approach that used the QUADAS-2 checklist to evaluate risk of bias and indirectness. The quality of evidence of the included studies ranged from moderate to very low quality. Outcomes were downgraded if: i) there was risk of bias, ii) for indirectness, iii) imprecision, and iv) inconsistency (if applicable). All case-control studies started at low quality because of the risk of spectrum bias. As expected, the case control studies generally yielded higher estimates of the sensitivity of the relevant tools compared to the cohort and cross-sectional studies. When reviewing the quality of evidence for assessment tools, the critical outcome is specificity. That is, considering how well the test is at reducing the number of false positives, so that people who do not have an eating disorder do not undergo unnecessary treatment. Few studies were identified that measured the effectiveness of a tool in different eating disorders, so interpretation of the data

	was based on few studies with small number of participants. Studies that used a case-control study design were highlighted to the committee as being at a high risk of spectrum bias given that it does not reflect a real-life situation and there is a greater chance of true positive and true negatives.
Other considerations	The Committee discussed the relevance of using an alternative tool to DSM-IV for assessing whether a person has an eating disorder. The consensus was that any other tool would not be used on its own, but could be used as part of clinical assessment at baseline, and throughout the treatment to monitor progress. The Committee did not agree on a particular tool to recommend, although the EDE-Q is probably the most commonly used. However, they did not infer that this was the preferred tool. The problem with EDE-Q is that some important outcomes may be missed, such as whether people are bingeing. In conclusion, the committee decided to generate their own recommendation on diagnosing eating disorders based on current best practice. No evidence was reviewed to develop the recommendations on where healthcare professionals may expect to find people with an eating disorder, or what they should consider when conducting an initial assessment including any safeguarding concerns. The committee used their knowledge, experience and expertise to generate these recommendations.

1. Research recommendation: How effective are the current guideline recommendations in improving symptoms and remission rates for men (aged over 18 years) with an eating disorder?

5 Coordinating care of eating disorders

5.1 Introduction

The coordination of care for those experiencing an eating disorder and for their families/ carers, varies widely across the country depending on a number of factors, including where the individual lives, allocated clinical commissioning/ specialised commission funding in that area, differences in thresholds (i.e. some services are restricted due to BMI criteria), variation in specialist skills and overarching inconsistencies across care pathways.

Coordination of care generally commences at the point of the individual accessing a service. This may be via a self-referral to the GP (not always for the initial eating disorder itself but for an associated problem) or primary care team. This is a crucial point for the coordination of care and access to an appropriate onward referral relies on the GP having received adequate training in the field of eating disorders. GPs are also under increasing time pressures, which can lead to the individual not sharing the extent of the problem. However, there are now a number of GPs across the country who have extensive expertise in working with such eating disorders. Appropriate referral also depends on the availability of specialist secondary care services, the provision of which varies greatly around the UK, with some regions having fully integrated primary to tertiary stepped services and other regions having a complete lack of eating disorder specialist support.

In some circumstances, involvement may start via another point of contact. This may include schools, dentists, pharmacies, sports clubs, and paediatricians. A parent/ carer may be the one who initiates an appointment with a professional as it is not uncommon for the individual experiencing the eating disorder to be lacking insight or ambivalent about accessing help.

Well established care pathways take into consideration the issue of consent and the ethical issues arising from eating disorders. Care coordination starts with initial assessment of the patient and assessment of risk, with referral being made to the relevant specialist service if available, ideally resulting in an agreed plan of care implemented.

Good practice emphasises a seamless approach from initial referral to treatment with targets met around waiting times. Historically, this has been due to lengthy wait list controls. New initiatives regarding standardising access and waiting times and narrowing the gap around inconsistencies that result due to geographical disparities continue to be addressed nationally and are being highlighted for children and young people with eating disorders in particular, via the Access and Waiting Time Standard for Children and Young People with an Eating Disorder, Commissioning Guidance published in August 2015, in collaboration with NHS England and the National Collaboration Centre for Mental Health (NCCMH., 2015).

Care provided to children, adults, their families and carers should be delivered by professionals experienced in the evidence based management and treatment of eating disorders.

One of the issues that often results in barriers to the coordination of care is poor communication between teams. The individual may be receiving support from a number of services, however if these are not in regular communication with each other, this often results in gaps in the provision of care. Engagement of the individual may also present as a barrier in the coordination of care and can require time and sensitivity from the professional in building up a therapeutic relationship with the service user.

The transition period from child to adult services can also lead to barriers in the continuity of care coordination and relies on effective transition protocols being delivered. A few areas are now served by community eating disorder teams for people across the age spectrum however these are rare.

5.1.1 Co-ordinating care

The co-ordination of care for patients with eating disorders varies across the country, depending on local commissioning arrangements and service design. Where specialist eating disorder services exist, a 'case manager' or care co-ordinator within the service commonly oversees the co-ordination and delivery of treatment. Case managers are responsible for ensuring that the package of care being delivered is appropriate for the individual's needs and for liaison with relevant professionals and agencies. Whilst for many this may involve the delivery of psychological therapies, for those with more severe eating difficulties it may include co-ordinating funding and access to either a specialist day or in-patient unit. Alongside discussing and agreeing an appropriate package of care, case managers are responsible for communicating the agreed treatment plan with all those involved in an individual's care. In many parts of the country this is done using the Care Programme Approach (CPA) and is likely to include a combination of CPA meetings and regular written updates to all relevant professionals, as well as to patients themselves. Whilst many specialist services will include medical monitoring as part of care, in some cases this responsibility will be held by the GP, a paediatrician or other physician. In these situations, excellent communication and clear guidance from the specialist team is needed, ideally under a shared care agreement and results of blood tests and other investigations fed back to the specialist team on a regular basis to inform further care planning. Such agreements need to specify who is responsible for taking action when results are abnormal or deteriorating and incorporate a shared understanding of the 'concern' and 'alert' ranges for blood results, physical observations, weight etc.

Care co-ordination may also include other health care professionals as required, such as gastroenterologists or diabetes specialists. The co-ordination of care for those who have an eating disorder as part of a broader psychiatric presentation, which may include depression, personality difficulties including emotionally unstable personality disorder, significant suicidality or substance misuse, is often overseen by a care-co-ordinator based in a local Community Mental Health Team (CMHT) or generic Child and Young people Mental Health Service (CAMHS) (although this may change in response to new recommendations (NCCMH., 2015)). This type of care co-ordination may involve integrating treatments from a number of services (e.g. CAMHS, general psychiatry, drug and alcohol teams). This way of working can enable 'joined-up' treatment packages to be drawn together that address the patient's broad ranging needs, thus reducing the possibility of individuals being 'bounced' between services. However, multiple professionals being involved in an individual's care can lead to differences in clinical opinion and tensions between teams as to what constitutes the most appropriate care package, so effective communication and supervision are particularly important.

Most referrals to specialist services are made directly from primary care or by self-referral, in some areas they are made via the local CMHT or CAMHS. This can result in a two-step process, introducing barriers to treatment and delaying diagnosis and access to care. In England the recently introduced Access and Waiting Times Standards for Young People with an Eating Disorder are intended to reduce duration of untreated illness by removing these barriers through direct access (NCCMH., 2015). Recent innovations in adult care have involved rolling the secondary care team out into the community with triage, assessment and therapy all being delivered in a primary care setting.

5.1.2 Transition of care

There are a number of reasons why patients might need to transition. They include transition between child and young people and adult services, between inpatient and outpatient where these are separate (typical for young people), between different geographical areas (common in the student population) or between different types of treatment, including back to primary care at the end of specialist treatment.

Most transitions between specialist CAMHS and adult services will follow a transition protocol which is likely to recommend a six month transitional period. During this time a therapist from the adult service will join meetings in CAMHS and get to know the patient and their family. This gives an opportunity for individuals to talk about any anxieties they might have about moving into adult services, as well as time to plan an appropriate on-going package of care. This transitional time can be difficult for parents as, once in adult services, their children are likely to have more say in relation to their treatment and whether they want to continue to share information about their care with their parents.

Another difficult time of transition relates to those moving between geographically separate specialist services. Whilst the ideal process would include a handover CPA with both the existing and new care teams, in reality local service procedures often hinder a smooth transition. This can be particularly difficult for students who may split their time between home and university. In these cases best practice might include services working together, although in reality there are often difficulties related to temporary registration with different GPs and services declining to get involved with patients whose addresses are out of their catchment area. This can lead to disjointed and disrupted treatment. Handover from specialist 'home' teams to university GP practices can aid the transition of care significantly.

Transitions also occur between types of care, most notably between in-patient and out-patient treatments. In some cases moving to day-care can provide a supportive 'step-down' to community care, which can reduce the risk of relapse and help patients to adjust to life outside of a hospital environment. This is important given that the risk of relapse is high following an episode of in-patient care.

As there are a limited number of beds in England, patients may be admitted to units geographically distant from home and from their community specialist or general services. As well as being challenging for sufferers and families to manage, this disconnect between inpatient and outpatient care can result in increased lengths of inpatient stay because of the need to re-establish therapeutic working relationships and, in the case of young people in particular, re-empower parents and carers to their caring role. A role of case managers is to oversee this process. It is important that the regular CPA reviews for these patients include home teams, to plan support during periods of home leave as well as at discharge.

Some patients will be treated in general hospitals and may then be transferred to a specialist eating disorder bed. The transition point must be agreed jointly by the clinical teams, so that patients are adequately medically stabilised at the point of transfer. The MARISPAN and Junior MARSIPAN reports (Psychiatry, 2014, Psychiatry, 2012) emphasise the importance of collaboratively developed protocols to facilitate shared care and admission and discharge processes and cite examples of poor outcome when these are not in place.

5.2 Coordinating care

5.2.1 Review questions: Do different ways of coordinating care produce benefits/harms for people with eating disorders? Does the setting (inpatient, outpatient or other specific setting) and different ways of coordinating, transitioning and integrating care for treating eating disorders produce benefits/harms in people with eating disorders?

The review protocol summary in Table 42 incorporated two review questions since the interventions overlapped and could be described as either a way of coordinating care or the setting in which treatment should be provided. Further information about the search strategy can be found in Appendix H and the full review protocols can be found in Appendix F.

Table 42: Clinical review protocol summary for the review of: Do different ways of coordinating care produce benefits/harms for people with eating disorders?

Does the setting (inpatient, outpatient or other specific setting) and different ways of coordinating, transitioning and integrating care for treating eating disorders produce benefits/harms in people with eating disorders?

Component	Description
Review question(s)	Do different ways of coordinating care produce benefits/harms for people with eating disorders? Does the setting (inpatient, outpatient or other specific setting) and different ways of coordinating, transitioning and integrating care for treating eating disorders produce benefits/harms in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) and a common comorbidity (e.g. diabetes, hypothyroidism). • **Strata:** • children (≤12), young people (13-≤17 years), adults ≥18 years • eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Case management (named person coordinates patient) vs. none • Specialist vs. non specialist • Inpatient vs. outpatient • Mental health vs. paediatric (physical health) practitioner • Teams vs. individual practitioners • Stepped care Compulsory vs. voluntary treatment
Comparison	• Note the comparison listed against the intervention.
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• All-cause mortality • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective

Component	Description
	cohort studies, (if no RCTs)

5.2.2 Clinical Evidence for: Do different ways of coordinating care produce benefits/harms for people with eating disorders? Does the setting (inpatient, outpatient or other specific setting) and different ways of coordinating, transitioning and integrating care for treating eating disorders produce benefits/harms in people with eating disorders?

5.2.2.1 Specialist vs. non-specialist care and inpatient versus outpatient treatment

10 RCTs (N=748) met the eligibility criteria for the two review questions, and mostly included people with anorexia nervosa and bulimia nervosa (Crisp 1991 (Crisp et al., 1991), Gowers 1994 (Gowers et al., 1994), Durand 2003 (Durand and King, 2003), Gowers 2007 (Gowers et al., 2007), Herpertz-Dahlmann 2014 (Herpertz-Dahlmann et al., 2014), Jager 1996 (Jager et al., 1996), Kong 2005 (Kong, 2005), Madden 2014 (Madden et al., 2014), Zeeck 2009a (Zeeck et al., 2009a), Zeeck 2008 (Zeeck et al., 2008)).

19 observational studies (n=2883) fulfilled the criteria for this review. The studies included people with anorexia nervosa and bulimia nervosa and some with any eating disorder (Arcelus 2008 (Arcelus et al., 2008), Birchall 2002 (Birchall et al., 2002), Carmen 2007 (Carmen et al., 2007), COSI-CAPS 2007 (Tulloch et al., 2008), Goddard 2013 (Goddard et al., 2013), Golan 2005 (Golan and Heyman, 2005), Hogdahl 2013 (Hogdahl, 2013), House 2012 (House et al., 2012),Hughes 2014 (Hughes et al., 2014), Kawai 2015 (Kawai et al., 2015), Kells 2012 (Kells et al., 2012), Meguerditchian 2010 (Meguerditchian et al., 2010), Milos 2004 (Milos et al., 2004), Olmsted 2002 (Olmsted et al., 2002), Tantillo 2009 (Tantillo et al., 2009), Vandereycken 2009 (Vandereycken and Vansteenkiste, 2009), Waller 2016 (Waller et al., 2016), Zeeck 2004 (Zeeck et al., 2004), Zeeck 2011 (Zeeck et al., 2011)).An overview of the trials included in the analysis can be found in Table 43 and Table 44. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

5.2.2.2 Stepped care interventions for anorexia nervosa

One RCT (n=45) met the eligibility criteria for the two review questions, that was on young people with anorexia nervosa (Lock 2015 (Lock, 2015)). The trial examined the effect of adding three sessions of intensive family coaching to participants who failed to gain more than 2.3 kg after four sessions of family-based treatment. An overview of the trials included in the analysis can be found in Table 43. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

5.2.2.3 Stepped care interventions for bulimia nervosa

Three RCTs (n=461) met the eligibility criteria for this review, that were all in adults (Davis 1999 (Davis et al., 1999), Mitchell 2011 (Mitchell et al., 2011), Crow 2013 (Crow et al., 2013), Treasure 1996 (Treasure et al., 1996)). An overview of the trials included in the analysis can be found in Table 46. Further information about both included and excluded studies can be found in Appendix J.

One study (n=58; Davis 1999) compared group psychoeducation then CBT-ED or wait list control. The study examined the effect of following six 90 min sessions of group psychoeducation over six weeks with either 16 weeks of CBT-ED (12 or 20 sessions depending on binge/purge frequency) or wait list control.

One study (n=110; Treasure 1996) compared self-help then CBT-ED versus CBT-ED. The study compared the effect of following eight weeks of self-help followed by eight weeks of CBT-ED (if the participants had experienced a binge or purge episode in the past 28 days) with 16 weekly sessions of CBT-ED.

One study (n=293; Mitchell 2011/Crow 2013) compared guided self-help followed by an antidepressant and CBT-ED versus CBT-ED then fluoxetine. The study participants were moved to the subsequent treatment if at the end of each treatment they had experienced a binge or purge episode in the past 28 days.

Summary of findings for stepped care intervention for bulimia nervosa can be found in Table 67, Table 68 and Table 69. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

5.2.2.4 Stepped care for binge eating disorder

No studies were found that examined a stepped care intervention for people with binge eating disorder.

5.2.2.5 Stepped care for EDNOS

No studies were found that examined a stepped care intervention for people with EDNOS.

Table 43: Study information on the RCTs included in the analysis of coordinating care and treating settings for people with an eating disorder.

Study	Eating Disorder	Age mean (SD)	BMI	N	Stage of illness	Intervention	Comparison	Duration
Crisp 1991/ Gowers 1994 UK	AN 100% female	23.2 (4.9) years	40.8 (6.1) kg	90	Duration of illness 41.0 (30.17) months 33.4 (25.9) months 27.5 (25.8) months 53.5 (52.9) months	Inpatient. Weight restoration to the mean matched-population weight at the age of onset of anorexia, including individual therapy, family therapy, group therapy, dietary counselling and occupational therapy. Followed by 12 sessions of out-patient psychotherapy sessions	Outpatient individual CBT-E and family psychotherapy + dietary counselling. Outpatient group therapy + dietary counselling No further treatment	12 months (unclear duration of therapy) 2 year FU
Durand 2003 UK	BN 100% female	24.5 (5.2) years	Not reported	68	5.9 (3.9) years 7.7 (4.6) years	Specialist Clinic Treatment A consultant psychiatrist managed each clinic. Other staff included psychiatrists, psychologists, nurse specialists, and dieticians. Each clinic offered similar forms of therapy, including a CBT and interpersonal psychotherapy.	General practice self-help CBT-ED. Received a copy of Bulimia Nervosa: a guide to recovery and advised to work through it while keeping in regular contact with their GP.	9 months
Gowers 2007/Gowers 2010 UK	AN 100% female	14.9 years	Weight below 85% of expected based on age and height	170	Mean length of illness 13 months	General community CAMHS a multidisciplinary, family-based approach with variable dietetic, individual supportive therapy and paediatric liaison	Specialist out-patient. It comprised an initial motivational interview, individual CBT plus parental feedback, parental counselling with dietary therapy. multi-modal feedback In patient (CBT + FT) The treatment was not manualised, but services all used a multidisciplinary psychiatric approach with the aim of normalising eating,	6 months 1 year FU (only data available)

Study	Eating Disorder	Age mean (SD)	BMI	N	Stage of illness	Intervention	Comparison	Duration
							restoring healthy weight and facilitating psychological (cognitive) change. Each participant received both individual supportive or cognitive therapies and family therapy	
Herpertz-Dahlmann 2014 Germany	AN 100% female	15.2 (1.5) years	15.1 (1.2) BMI	172	Duration of illness: 53.7 (39.6) weeks 42.4 (33.1) weeks	Continued inpatient care	Day patient treatment. At week 3 after inpatient All patients were provided with the same outpatient treatment programme. Typical day patient treatment in Germany offers a structured eating disorder programme from 0800 h to 1630 h on weekdays	12 months after admission (follow up),
Jager 1996 Germany	BN 100% female	23.8 (17-35) years	20.7 (16.6-29.3) BMI	61	Duration of illness: 4.7 (0.6 to 23) years	In patient therapy. Group therapy psychoanalytical group sessions integrated, structured groups which are presented with a problem-oriented task or topic. Treatment focuses on comorbidity with low self-esteem and interpersonal problems.	Outpatient treatment focuses on interrelationships. This form of therapy follows the Milan family therapy model.	1 year
Kong 2005 South Korea	AN: 32%: BN: 41%; EDNOS 25% 100% female	27.0 (7.2) years	21.5 (5.7) BMI	50	Duration of illness: 4.2 (1.8) years 3.4 (1.1) years	Modified day treatment Includes CBT, FT, includes meals supervised, group sessions, body image therapy, community meetings and exercise and nutrition education, cooking classes, dance and art therapy,	Traditional outpatient IPT, CBT and pharmacotherapy in an individual format. The treatment with traditional outpatient programme was continued after the study	10.7 (3.8) weeks

Study	Eating Disorder	Age mean (SD)	BMI	N	Stage of illness	Intervention	Comparison	Duration
						sexuality groups and weekend planning group	finished.	
Madden 2014 Australia	AN 95% female	14.9 (1.5) years	Not reported	82	Duration of illness: 7.4 (5.4) months	Hospitalization for medical stabilisation followed by FT out-patient if they had no markers of medical instability for 72 h after nasogastric feeds were ceased	Weight restoration Participants in the WR arm continued in hospital on supported meals without nasogastric feeding once they had no markers of medical instability for 72 hours, until they reached 90% ideal body weight before discharge to out-patient FBT.	In hospital 36.9 (17.1) days + 12 months FT
Zeeck 2009a/2008b Germany	BN 100% female	24.0 (7.6) years	21.5 (2.2) BMI	55	Duration of illness: 7 years (6.5) or 10.5 years (7.6)	Inpatient clinic program includes CBT and integrates a treatment contract, meal plans; 1 or 2 sessions with the family. Meals are supervised Inpatients are allowed to leave for the weekends during the last weeks of treatment.	Day clinic (treatment as usual). The day clinic and inpatient program comprise the same treatment components	12 weeks. Mean treatment duration: 86.7 days (23.6) 1 year FU

Abbreviations: AN – anorexia nervosa; BN – bulimia nervosa; ENDOS – eating disorder not otherwise specified; BMI – body mass index; CBT- cognitive behavioural therapy; FT – family therapy; FBT- family based therapy; FU – follow up; GP – general practice; h – hours.

Table 44: Study information on the observational studies included in the analysis of coordinating care and treatment settings for people with an eating disorder.

Study ID	Eating Disorder	N	Age mean (SD)	BMI	Stage of illness	Intervention	Comparison	Duration
Arcelus 2008 UK	AN. BN, EDNOS, BED NR: number	206	Not reported	19.7 (10.7 to 43.1)	Age of onset of illness: 14.0 (1.8) years.	Patients who had previous involvement at CAMHS as out-patients and referred to a specialist Adult Eating Disorders Service (AEDS)	Adults referred to a specialist Adult Eating Disorders Service (AEDS)	4 years

Study ID	Eating Disorder	N	Age mean (SD)	BMI	Stage of illness	Intervention	Comparison	Duration
	of females					Previous treatment with CAMHS due to their ED (less than 5 years ago).	Specialist eating disorders service. Outpatient service after referral from primary care.	
Birchall 2002 UK	Severe AN NR: number of females	27	15.0 (1.9) years	Not reported	Not reported	Day programme Open 10.00 a.m. until 5.15 p.m. 5 days a week and has a maximum capacity of 10 places. The ethos is to provide whatever help patients who struggle towards weight restoration and recovery, and transitions between out-patient, day patient and in-patient treatment are designed to be as seamless as possible. The day programme is not a facility for patients 'stuck' with anorexia	Early programme (inpatient) Before the day programme was opened, the treatment of anorexia nervosa was dichotomous. In-patient treatment is intensive, lengthy and costly, but provides round-the-clock care	Not reported. Data available: 6 mo FU for BMI and 1.5-1.8 years FU for readmission
Carmen 2007 UK	Any ED 100% females	138	Not reported	Not reported	After first appointment with general practitioner	Opt-in letter. Outpatient The opt-in form was attached and they were invited to return it within 4 weeks from the date of the letter.	No opt-in letter. Outpatient	Not reported
COSI-CAPS 2007/Tulloch 2008 UK	AN NR: number of females	107	12-18 years	15.3 (2.2)	Inpatient care	Specialist eating disorder inpatient ward	General ward (non-specialist)	140 days
Goddard 2013 UK	AN. 95% females	166	26.5 (8.9) years	14.0 (1.7)	Length of illness 8.2 (8.3) years Length of illness 2.0	Inpatient care Patients were considered discharged when they ceased to receive intensive treatment for their eating disorder (that is, inpatient treatment	Day patient care (no details)	17.8 (10.4) weeks 29.0 (11.9)

Study ID	Eating Disorder	N	Age mean (SD)	BMI	Stage of illness	Intervention	Comparison	Duration
					(1.7) years	or day patient treatment ≥ 4 days a week).		
Golan 2005 Israel	AN (49%) and BN (51%) 100% females	123	AN 20.6 (4.1) years BN 22.1 (3.9) years	AN 16.2 (1.7) BN 21.4 (1.6) BMI	Duration of illness AN: 5.1 (3.4) years BN 6.0 (3.6) years	Extensive Outpatient Program Psychotherapy and nutrition counselling in addition to a variety of outreach services delivered by clinical mentors	Limited Outpatient program. Psychotherapy and nutrition counselling	1-5 hours /weeks
Hogdahl 2013 Sweden	BN, BED or EDNOS-BN 96% females	79	27.9 (7.5) years	24.0 (4.6)	Duration of illness: 11.3 (8.0) years	Guided self-help It contains facts about eating disorders, and a six-step self-help program with detailed instructions, assignments and checklists. Patient and therapist decide on a day of the week when the therapist answers the patient's messages.	Day patient psychodynamic intensive group treatment with group and individual therapy, meals, body knowledge, and art therapy.	16 weeks
House 2012 UK	AN 58%; BN:25%; EDNOS 17% 97% females	345	15.1 years	Mean weight for height 82.8%		Specialist eating disorders service. Outpatient service they first came face-to-face with after referral from primary care. A specialist service was defined as a minimum of a multidisciplinary team, a team with the expertise to deliver recommended treatments (assessment of physical risk and psychological therapies including family therapy); and the resources to offer routine outpatient treatment.	Specialist CAHMS or Specialist NHS eating disorder service. Assessment and treatment. (May have patients referred for treatment from non-specialist CAMHS). Non-specialist CAMHS. Assessment and at least initial treatment but ultimately transferred for specialist CAMHS for treatment.	12 months
Hughes 2014 Australia	AN NR: number of	161	Not reported	Not reported	Point of diagnosis	New model of care. Family therapy as first line. For patients diagnosed with AN or EDNOS-AN type. FBT is an outpatient intervention in which a	Old model of care. Mostly inpatient Inpatient admissions were routinely used in response	Not reported

Study ID	Eating Disorder	N	Age mean (SD)	BMI	Stage of illness	Intervention	Comparison	Duration
	females					mental health clinician helps parents become actively involved in supporting weight gain and normalizing eating patterns for their child.	to medical instability, failure of outpatient management. Mental health input was generally sought during inpatient admissions; however, there were inconsistencies in outpatient mental health care from community mental health services (e.g., individual, group, or family therapies).	
Kawai 2015 Japan	AN. 97% females	249	22.5 (8.6)	14.0 (1.8)	Duration 7.3 (7.4) years Duration 4.5 (5.4) years	Urgent hospitalisation. Patients hospitalized with disturbances of consciousness and/or difficulty walking on the day of consultation.	Planned inpatient admission. Patients hospitalized for AN up to one year after the day of consultation. For patients whose BMI was not expected to increase and/or eating behaviour abnormality was not expected to improve, inpatient treatment with CBT-ED was offered.	Not reported
Kells 2012 USA	AN NR number of females	52	15.9 (2.5)	15.9 (2.5)	Not reported	Meal supervision Goal of inpatient treatment was weight gain goal of 0.2 kg/day Group includes all patients receiving at least one supervised meal during course of admission. Meal supervision in hospitalized patients modifies the Maudsley method by using clinical staff, rather than parents, as active and	No meal supervision Patients who were not supervised at meal times during their IP stay.	Not reported

Study ID	Eating Disorder	N	Age mean (SD)	BMI	Stage of illness	Intervention	Comparison	Duration
						supportive observers during meal time.		
Meguerditchian 2010 France	AN 100% females	143	26.0 years	15.1 (12-20)	Disease duration: 4.1 (0.16 to 29) years	Hospitalization was prescribed only for life-threatening medical conditions due to severe malnutrition, patient incapacity to reach weight objectives patient request, or marked suicidal ideation. Discharge weight was negotiated with each patient	Ambulatory care. Outpatient Initial outpatient treatment was based on a weight contract (progressive return to normal BMI), and consisted of somatic and nutritional assessment, nutritional education by the dietician, monthly medical follow up by a physician specialized in nutrition, and weekly psychotherapy sessions of mixed cognitive-behavioural and analytical types.	4.8 year FU (duration of care 13.4 (1-68) months)
Milos 2004 Switzerland	AN (29.7%), BN (56.8%), EDNOS (13.5%) 100% females	222	26.8 (6.6) years	AN 15.2 (1.5) BN-21 (5.4) EDNOS/ BN 22.2 (5.4).	Not reported	Inpatient. Participants who had spent at least 1 day as an inpatient (in a psychiatric or psychotherapeutic inpatient unit or day hospital or in a somatic hospital) during that period were classified as inpatients.	No history of inpatient experience. All other participants were classified as not currently receiving inpatient treatment.	Not reported
Olmsted 2002 Canada	AN and BN 100% females	581	25.4 (6.6)	21.3 (5.1)	Not reported	5 day programme Intensive group therapy, nutritional rehabilitation, and pharmacotherapy when indicated. The program is predominantly cognitive-behavioural in orientation with some interpersonal and experiential components. Psychopharmacologic interventions include antidepressants and	4 day programme The intensity, goals, and modality of treatment did not change at this time. Rather, the aim was to be more efficient with less time.	8 hours a day. 8.6 weeks

Study ID	Eating Disorder	N	Age mean (SD)	BMI	Stage of illness	Intervention	Comparison	Duration
						selective serotonin reuptake inhibitors.		
Tantillo 2009 USA	AN (40%), BN (20%), EDNOS (40%) 100% females	35	22.1 (5.5)	20.4 (3.5)	Duration of illness: 6.1 (6.6) years	Combined treatment group Partial Hospitalization and Supported Housing, Sage House. The overall aim of supported housing is to help integrate individuals into the community and allow them to lead fulfilling and satisfying lives outside a hospital or institutional setting. In addition to a safe living environment, the individual residing in supported housing may receive individual counselling or case management to improve problem-solving and daily living skills, support groups, meals, and transportation to appointments/events in the community.	Partial Hospitalization. Inpatient Eating Disorders Partial Hospitalization programs are intensive multidisciplinary treatment venues in which patients spend 7–12 hours per day and receive at least two supervised meals and one snack. The majority time is spent in group and individual therapy, nutritional counselling, psychopharmacological evaluation medication monitoring, and case management. Patients and family members receive family and multifamily therapy, and parenting group is often avail-able for parents of young people and young adults.	Unclear
Vandereycken 2009 Belgium	AN (53%), BN (32%) EDNOS (15%) Not reporter number	174	21 (15 to 45) years	14.8 (7.9)	Duration of illness 4.2 years	New strategy. Inpatient the patient is proposed to come at least for an introductory week: a minimum commitment of 5 days (Monday to Friday) with the explicit promise that she will be free to leave the hospital even if her family would prefer her to stay. However, if her medical condition were at serious risk, she	Old strategy. Inpatient The staff took all decisions concerning further treatment, including the choice of the treatment group and the duration of treatment. Both staff and family used various ways to convince patients to stay in	6 months

Study ID	Eating Disorder	N	Age mean (SD)	BMI	Stage of illness	Intervention	Comparison	Duration
	of females					would be transferred to the internal medicine department of a general hospital nearby	the treatment, including medical arguments and psychological pressure with direct or indirect guilt-inducing messages. If patients ran away or refused to return, the family was supposed to bring them back to the hospital.	
Waller 2016 USA	AN NR number of females	29	14.5 (2.1)	74.1 % ABW	Duration of illness 12.4 (13.8) months	Continuum Care Program To make available additional levels of intensive care such as partial hospitalisation, day treatment, and especially intensive outpatient treatment so that inpatient can be used more selectively.	Historical exclusive inpatient Patients admitted exclusively inpatient treatment of AN	3 year FU
Zeeck 2004 (Zeek 2011) Germany	Severe BN 100% females	36	27.1 (6.8) years	23.9 (3.7)	Duration of illness 9.4 (7.8) years	Outpatient Monday to Friday from 08.00 to 16.00 hours. Patients stay for 3 months The nurses provide weekly sessions with a strong symptom orientation including cooking sessions. The initial focus is on symptomatology shifting later more and more to a focus on underlying conflicts or personality deficits. Main elements of treatment are a multidisciplinary team approach and treatment in a 'therapeutic community'.	Inpatient The inpatient treatment is comparable to the day clinic treatment in all treatment components other than that patients stay overnight, have their own room and nurses whom they can meet in the evenings and at weekends.	3-4 months

1 *Abbreviations: AN – anorexia nervosa; BN – bulimia nervosa; ABW- average body weight; CAMHS – child and adolescent mental health services: EDNOS – eating disorder not*
2 *otherwise specified; FU – follow up;*

Table 45: Study information on the RCT included in the analysis of stepped care in young people with anorexia nervosa.

Study ID	N Randomized	Female (%)	Sample	Intervention	Duration of Intervention	Comparison(s)	Duration of Comparison
Lock 2015	45*	92	AN excluding amenorrhea criterion	Family-based treatment (15 sessions) (+ 3 sessions of Intensive Parental Coaching if weight gain <4.8 lb)	6 months (+ 3 sessions IPC in-between FBT Session 4 and FBT Session 5)	Family-based treatment (15 sessions)	6 months

Notes: Participants initially randomised into two groups, one that received family-based treatment only (n=10), and one that received family-based treatment with the possibility of also receiving intensive family coaching if weight gain was less than 2.3 kg after 4 sessions of therapy (n=35). Data only included for participants in latter group.
Abbreviations: AN – anorexia nervosa; IFT – intensive family coaching; DSM- IV – Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; lb- pounds; < less than.

Table 46: Study information of the RCTs included in the review of stepped care interventions people with bulimia nervosa.

Study ID	N Random-ized	Female (%)	Sample	Intervention	Age at onset	Duration of illness (years)	Course of Intervention	Comparison(s)	Age at onset	Duration of illness (years)	Course of Comparison	Follow up
Davis 1999	58	100	BN	Group psychoeducation → CBT-ED (12 or 20 sessions)*	Not reported	7.6 (5.4) a	Week 1-6 → Week 7-22	Group psychoeducation → Wait list control	Not reported	7.6 (5.4)	Week 1-6 → Week 7-22	Not reported
Mitchell 2011/ Crow 2013	293	Not reported	BN-P or BN-NP	Guided self-help manual for BN → Fluoxetine (20, 40, 60 mg)** → CBT-ED (20 sessions)***	Not reported	Not reported	Week 1-18 → Week 10-70 → Week 18-44	CBT-ED (20 sessions) → Fluoxetine (20, 40, 60 mg)**	Not reported	Not reported	Week 1-18 → Week 5-70	Not reported
Treasure 1996	110	Not reported	BN or atypical BN	Self-help manual for BN → CBT-ED	17.5 (4.8)	8.0 (5.0)	8 weeks → 8 weeks	CBT-ED (16 sessions)	17 (4.4)	9.1 (6.5)	Week 1-16	18 months

Study ID	N Randomized	Female (%)	Sample	Intervention	Age at onset	Duration of illness (years)	Course of Intervention	Comparison(s)	Age at onset	Duration of illness (years)	Course of Comparison	Follow up
				(8 sessions)***								

*Notes: a, Whole sample; *20 sessions if ≥4 binge or purge episodes past 28 days. 12 sessions if ≤3 binge or purge episodes past 28 days. **, Patients assigned to treatment if they were predicted to be non-responders and consented to treatment; ***, Patients offered treatment if they had not achieved abstinence (no binge nor purge episodes in past 28 days). Arrow indicates the following treatment patients were stepped up to. Abbreviations: BN-NP, Bulimia nervosa non purging subtype; BN-P, Bulimia nervosa purging subtype; CBT- ED, cognitive behavioural therapy with an eating disorder focus.*

Randomised control trials for coordinating care and treatment settings for people with an eating disorder

Table 13: Summary table of findings for inpatient care versus another setting (other) for people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Inpatient vs. Other (AN) (95% CI)
BMI Adults - Inpatient vs. Day Clinic	43 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI adults - inpatient vs. day clinic in the intervention groups was 0.04 standard deviations higher (0.56 lower to 0.64 higher)
Bingeing - Adults - Inpatient vs. Day Clinic	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing - adults - inpatient vs. day clinic in the intervention groups was 0.45 standard deviations lower (1.05 lower to 0.16 higher)
Vomiting- Adults - Inpatient vs. Day Clinic	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting- adults - inpatient vs. day clinic in the intervention groups was 0.39 standard deviations lower (0.99 lower to 0.21 higher)

EDI-2 Bulimia - Adults- Inpatient vs. Day Clinic	43 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 bulimia - adults- inpatient vs. day clinic in the intervention groups was 0.12 standard deviations higher (0.48 lower to 0.72 higher)
Change in Global MR - In-patient vs. Outpatient Individual + FT_Adults	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in global mr - in-patient vs. outpatient individual + ft_adults in the intervention groups was 0.14 standard deviations lower (0.70 lower to 0.43 higher)
Change in Global MR - In-patient vs. Outpatient Group Adults	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in global mr - in-patient vs. outpatient group adults in the intervention groups was 0.06 standard deviations higher (0.5 lower to 0.63 higher)
Change in Global MR - In-patient vs. WLC Adults	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in global mr - in-patient vs. wlc adults in the intervention groups was 0.03 standard deviations higher (0.54 lower to 0.60 higher)
Change in MR: Menstruation - In-patient vs. Outpatient Individual + FT	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: menstruation - in-patient vs. outpatient individual + ft in the intervention groups was 0.02 standard deviations lower (0.59 lower to 0.55 higher)
Change in MR: Menstruation - In-patient vs. Outpatient Group	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: menstruation - in-patient vs. outpatient group in the intervention groups was 0.16 standard deviations lower (0.72 lower to 0.41 higher)
Change in MR: Menstruation - In-patient vs. WLC	50 (1 study)	⊕⊕⊖⊖ LOW1,2		Not	The mean change in mr: menstruation - in-patient vs. wlc in the intervention groups was

		due to risk of bias, imprecision		calculable for SMD values	0.02 standard deviations higher (0.55 lower to 0.58 higher)
Change in MR: Nutrition - In-patient vs. Outpatient Individual + FT	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: nutrition - in-patient vs. outpatient individual + ft in the intervention groups was 0.06 standard deviations lower (0.63 lower to 0.51 higher)
Change in MR: Nutrition - In-patient vs. Outpatient Group	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: nutrition - in-patient vs. outpatient group in the intervention groups was 0.2 standard deviations lower (0.77 lower to 0.36 higher)
Change in MR: Nutrition - In-patient vs. WLC	50 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: nutrition - in-patient vs. wlc in the intervention groups was 0.33 standard deviations higher (0.24 lower to 0.90 higher)
Change MR: Mental State - In-patient vs. Outpatient Individual + FT	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change mr: mental state - in-patient vs. outpatient individual + ft in the intervention groups was 0.29 standard deviations lower (0.86 lower to 0.28 higher)
Change MR: Mental State - In-patient vs. Outpatient Group	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change mr: mental state - in-patient vs. outpatient group in the intervention groups was 0.07 standard deviations higher (0.50 lower to 0.64 higher)
Change MR: Mental State - In-patient vs. WLC	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias,		Not calculable for	The mean change mr: mental state - in-patient vs. wlc in the intervention groups was 0.12 standard deviations lower (0.69 lower to 0.45 higher)

		imprecision		SMD values	
Change in MR: Sexual adjustment - In-patient vs. Outpatient Individual + FT	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: sexual adjustment - in-patient vs. outpatient individual + ft in the intervention groups was 0.11 standard deviations higher (0.65 lower to 0.87 higher)
Change in MR: Sexual adjustment - In-patient vs. Outpatient Group	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: sexual adjustment - in-patient vs. outpatient group in the intervention groups was 0.07 standard deviations lower (0.83 lower to 0.69 higher)
Change in MR: Sexual adjustment - In-patient vs. WLC	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: sexual adjustment - in-patient vs. wlc in the intervention groups was 0.05 standard deviations lower (0.81 lower to 0.71 higher)
Change in MR: Social economic adjustment - In-patient vs. Outpatient Individual + FT	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: social economic adjustment - in-patient vs. outpatient individual + ft in the intervention groups was 0.31 standard deviations lower (0.88 lower to 0.26 higher)
Change in MR: Social economic adjustment - In-patient vs. Outpatient Group	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: social economic adjustment - in-patient vs. outpatient group in the intervention groups was 0 standard deviations higher (0.57 lower to 0.57 higher)
Change in MR: Social economic adjustment - In-patient vs. WLC	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in mr: social economic adjustment - in-patient vs. wlc in the intervention groups was 0.13 standard deviations higher (0.43 lower to 0.70 higher)

Global Severity Index -_Adults - Inpatient vs. Day Clinic	43 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean global severity index -_adults - inpatient vs. day clinic in the intervention groups was 0.41 standard deviations higher (0.19 lower to 1.02 higher)
Remission - _Adults - Inpatient vs. Day Clinic_ITT	55 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR 1.81 (0.6 to 5.5)	143 per 1000	116 more per 1000 (from 57 fewer to 643 more)
BMI- _Adults FU - Inpatient vs. Specialist Outpatient	76 (1 study)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi- _adults fu - inpatient vs. specialist outpatient in the intervention groups was 0.00 standard deviations higher (0.47 lower to 0.47 higher)
BMI- Adultss FU - Inpatient vs. General Outpatient	74 (1 study)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi- adults fu - inpatient vs. general outpatient in the intervention groups was 0.25 standard deviations lower (0.73 lower to 0.23 higher)
BMI-Young people FU - Inpatient vs. Day patient	161 (1 study)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi-young people fu - inpatient vs. day patient in the intervention groups was 0.09 standard deviations lower (0.4 lower to 0.22 higher)
Bingeing - _Adults FU - Inpatient vs. Day Clinic	43 (1 study)	⊕⊕⊖⊖ LOW4,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing - _adults fu - inpatient vs. day clinic in the intervention groups was 0.36 standard deviations higher (0.24 lower to 0.97 higher)
Vomiting - _Adults FU - Inpatient vs. Day Clinic	44 (1 study)	⊕⊕⊖⊖ LOW3,10 due to risk of		Not calculab	The mean vomiting - _adults fu - inpatient vs. day clinic in the intervention groups was 0.31 standard deviations lower

		bias, imprecision		le for SMD values	(0.91 lower to 0.28 higher)
Menstruation regular - Young people FU	156 (1 study)	⊕⊖⊖⊖ VERY LOW6,9 due to risk of bias, imprecision	RR 0.81 (0.41 to 1.6)	198 per 1000	38 fewer per 1000 (from 117 fewer to 119 more)
EDI Total - _Adults FU - Inpatient vs. Specialist Outpatient	85 (1 study)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total - _adults fu - inpatient vs. specialist outpatient in the intervention groups was 0.28 standard deviations lower (0.7 lower to 0.15 higher)
EDI Total - _Adults FU - Inpatient vs. General Outpatient	83 (1 study)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total - _adults fu - inpatient vs. general outpatient in the intervention groups was 0.46 standard deviations lower (0.9 to 0.02 lower)
EDI Total - Young people FU - Inpatient vs. Day Patient	143 (1 study)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total - young people fu - inpatient vs. day patient in the intervention groups was 0.11 standard deviations higher (0.22 lower to 0.43 higher)
EDI-2 Bulimia - Young people FU - Inpatient vs. Day Clinic	43 (1 study)	⊕⊕⊖⊖ LOW4,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 bulimia - young people fu - inpatient vs. day clinic in the intervention groups was 0.58 standard deviations higher (0.03 lower to 1.19 higher)
MR: Total Outcome - FU - Inpatient vs. Specialist Outpatient	103 (1 study)	⊕⊕⊖⊖ LOW5,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean mr: total outcome - fu - inpatient vs. specialist outpatient in the intervention groups was 0.04 standard deviations lower (0.43 lower to 0.35 higher)

MR: Total Outcome - FU - Inpatient vs. General Outpatient	104 (1 study)	⊕⊕⊖⊖ LOW5,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean mr: total outcome - fu - inpatient vs. general outpatient in the intervention groups was 0 standard deviations higher (0.38 lower to 0.38 higher)
Global severity index - Young people FU - Inpatient vs. Day Patient	141 (1 study)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean global severity index - young people fu - inpatient vs. day patient in the intervention groups was 0.20 standard deviations higher (0.13 lower to 0.53 higher)
Global severity index - Adults FU - Inpatient vs. Day Patient (Copy)	43 (1 study)	⊕⊕⊖⊖ LOW5,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean global severity index - adults fu - inpatient vs. day patient (copy) in the intervention groups was 0.21 standard deviations higher (0.39 lower to 0.81 higher)
Serious adverse events - Young people FU	161 (1 study)	⊕⊖⊖⊖ VERY LOW6,9 due to risk of bias, imprecision	RR 1.31 (0.5 to 3.44)	81 per 1000	25 more per 1000 (from 41 fewer to 199 more)
Remission - Young people FU - Inpatient vs. Day patient_ITT (Copy)	172 (1 study)	⊕⊕⊖⊖ LOW9,11 due to risk of bias, imprecision	RR 0.91 (0.73 to 1.14)	671 per 1000	60 fewer per 1000 (from 181 fewer to 94 more)
Readmissions/Relapse for ED - Young people FU - Inpatient vs. Day patient	161 (1 study)	⊕⊕⊖⊖ LOW9,12 due to risk of bias, imprecision	RR 1.68 (0.89 to 3.16)	151 per 1000	103 more per 1000 (from 17 fewer to 327 more)
Remission _Adults FU - Inpatient vs. Specialist Outpatient_ITT	113 (1 study)	⊕⊕⊖⊖ LOW8,12 due to risk of bias, imprecision	RR 1.41 (0.77 to 2.57)	236 per 1000	121 more per 1000 (from 57 fewer to 473 more)

Remission - Adults FU - Inpatient vs.General Outpatient_ITT	113 (1 study)	⨁⊖⊖⊖ VERY LOW6,8 due to risk of bias, imprecision	RR 0.92 (0.55 to 1.52)	364 per 1000	55 fewer per 1000 (from 200 fewer to 229 more)
Remission - Adults FU - Inpatient vs. Day patient_ITT	55 (1 study)	⨁⨁⊖⊖ LOW10,11 due to risk of bias, imprecision	RR 0.52 (0.14 to 1.87)	214 per 1000	103 fewer per 1000 (from 184 fewer to 186 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: intention to treat; FU: follow up

1 Unclear how randomisation sequence was generated or if allocation concealment was conducted. Participants and investigators were not blind. It was unclear if assessor was blind. High dropout rates were detected in one arm >20%
2 95% CI crossed 2 MIDs (-0.5 and 0.5)
3 95% CI crossed 1 MID (-0.5)
4 95% CI crossed 1 MID (0.5)
5 For a continuous outcome, there were fewer than 400 participants.
6 95% CI crossed 2 MIDs (0.75 and 1.25)
7 In Gowers 2007, it was unclear how randomisation sequence was generated or if allocation concealment was conducted. It was unclear if participants, investigators were blind. Assessor were blind. High dropout rates were detected in one arm >20% In Herpertz-Dahlmann 2014 performed adequate randomisation and allocation concealment. Patients and investigators were not blind and assessors were only blind at baseline.
8 In Gowers 2007, it was unclear how randomisation sequence was generated or if allocation concealment was conducted. It was unclear if participants, investigators were blind. Assessor were blind. High dropout rates were detected in one arm >20%
9 In Herpertz-Dahlmann 2014 performed adequate randomisation and allocation concealment. Patients and investigators were not blind and assessors were only blind at baseline.
10 In Zeek 2009/2008b, it was unclear if adequate randomisation sequence was generated or if allocation concealment was performed. Participants and investigators were not blind but assessors were.
11 For a dichotomous outcome, there are fewer than 300 events.
12 95% CI crossed 1 MID (1.25)

1 **Table 47: Summary table of findings for specialised outpatient care versus general outpatient care for people with anorexia nervosa**

Outcomes	**No of Participants (studies) Follow up**	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with General outpatient (AN)	**Risk difference with Specialist outpatient (95% CI)**

BMI FU	98 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.29 standard deviations lower (0.69 lower to 0.11 higher)
EDI Total FU	82 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total fu in the intervention groups was 0.17 standard deviations lower (0.6 lower to 0.26 higher)
MR: Total Outcome FU	103 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean mr: total outcome fu in the intervention groups was 0.04 standard deviations higher (0.35 lower to 0.43 higher)
Subsequent admission to hospital FU	110 (1 study)	⨁⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 1.13 (0.63 to 2.03)	273 per 1000	35 more per 1000 (from 101 fewer to 281 more)
Remission FU_ITT	110 (1 study)	⨁⨁⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 0.65 (0.36 to 1.17)	364 per 1000	127 fewer per 1000 (from 233 fewer to 62 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up; MR: Morgan Russell

1 It is unclear how the randomisation sequence was generated and if allocation concealment was performed. It is unclear if participants and investigators were blind, however, the assessors were masked. High dropouts were reported >20%.
2 95% CI crossed 1 MID (-0.5)
3 For a continuous outcome, there were fewer than 400 participants.
4 95% CI crossed 2 MIDs (0.75 and 1.25)
5 95% CI crossed 1 MID (0.75)

1 **Table 48: Summary table of findings for inpatient care with group psychotherapy versus family therapy (FT) outpatient care for people**
2 **with bulimia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with FT Outpatient (BN)	Risk difference with Inpatient Group (95% CI)

Binges FU	71 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binges fu in the intervention groups was 0.06 standard deviations lower (0.53 lower to 0.41 higher)
Self-induced vomiting FU	71 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean self-induced vomiting fu in the intervention groups was 0.11 standard deviations lower (0.57 lower to 0.36 higher)
Depression FU	71 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.14 standard deviations higher (0.33 lower to 0.61 higher)
Bulimic severity score FU	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic severity score fu in the intervention groups was 0.07 standard deviations lower (0.55 lower to 0.42 higher)
Remission FU_ITT	71 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 0.79 (0.43 to 1.43)	436 per 1000	92 fewer per 1000 (from 248 fewer to 187 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 The study was only partially randomised, only 52% were assigned randomly. The investigators decided some patients needed to be allocated due to their clinical condition. It was unclear if either the participants, investigators and assessors were blind. High dropouts were detected in one arm >20%
2 95% CI crossed 1 MID (-0.5)
3 95% CI crossed 1 MID (0.5)
4 95% CI crossed 2 MIDs (0.75 and 1.25)

1 **Table 49: Summary table of findings for modified day treatment compared to traditional outpatient for any eating disorder**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with traditional outpatient (ANY ED)	Risk difference with modified day treatment (95% CI)
BMI	43 (1 study)	⊕⊕⊖⊖ LOW1,2		Not calculable for SMD	The mean BMI in the intervention groups was

		due to risk of bias, imprecision		values	0.57 standard deviations higher (0.12 to 1.02 higher)
Bingeing episodes	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing episodes in the intervention groups was 0.93 standard deviations lower (1.57 to 0.3 lower)
Purging episodes	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging episodes in the intervention groups was 1.21 standard deviations lower (1.87 to 0.56 lower)
Depression	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.83 standard deviations lower (1.45 to 0.2 lower)
EDI-2 Total score	43 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 total score in the intervention groups was 1.42 standard deviations lower (2.09 to 0.74 lower)
EDI-2 Drive for thinness	43 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 drive for thinness in the intervention groups was 1.88 standard deviations lower (2.61 to 1.15 lower)
EDI-2 Bulimia	43 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 bulimia in the intervention groups was 1.52 standard deviations lower (2.21 to 0.83 lower)
EDI-2 Body dissatisfaction	43 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 body dissatisfaction in the intervention groups was 1.2 standard deviations lower (1.86 to 0.55 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear if allocation concealment was performed. It was also unclear if either the participants, investigators and assessors were blind.
2 95% CI crossed 1 MID (0.5)
3 95% CI crossed 1 MID (-0.5)

4 For a continuous outcome, there were fewer than 400 participants.

Table 50: Summary table of findings for inpatient weight stabilisation (short) compared with weight restoration (long) for any eating disorder

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with weight restoration (longer) (AN)	Risk difference with Inpatient weight stabilisation (short) (95% CI)
Remission Young people_ITT	82 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.11 (0.5 to 2.45)	220 per 1000	24 more per 1000 (from 110 fewer to 318 more)
Change EDE Global score Young people FU	69 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change ede global score young people fu in the intervention groups was 0.12 standard deviations lower (0.59 lower to 0.36 higher)
Hospital readmission Young people FU	78 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 0.95 (0.53 to 1.72)	368 per 1000	18 fewer per 1000 (from 173 fewer to 265 more)
Remission Young people FU_ITT	82 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 0.92 (0.48 to 1.78)	317 per 1000	25 fewer per 1000 (from 165 fewer to 247 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Randomisation was adequate however it was unclear if allocation concealment was performed. Participants and investigators were not blind, however, the assessor was blind to treatment allocation.
2 95% CI crossed 1 MID (1.25)
3 95% CI crossed 1 MID (-0.5)
4 95% CI crossed 2 MIDs (0.75 and 1.25)

5.2.2.6 **Observational studies for coordinating care and the best setting for treating eating disorders (quality starts at very low)**

Table 51: Summary of findings table for inpatient care versus day patient care for people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with day patient - Adult - AN	Risk difference with Inpatient (95% CI)
Binge eating	152 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.41 (0.26 to 0.64)	667 per 1000	393 fewer per 1000 (from 240 fewer to 493 fewer)
Laxative use	152 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.66 (0.16 to 2.66)	133 per 1000	45 fewer per 1000 (from 112 fewer to 221 more)
Self-induced vomiting	152 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.57 (0.26 to 1.26)	333 per 1000	143 fewer per 1000 (from 247 fewer to 87 more)
Excessive Exercise	152 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.64 (0.35 to 1.17)	467 per 1000	168 fewer per 1000 (from 303 fewer to 79 more)
EDE- Total	152 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- total in the intervention groups was 0.25 standard deviations lower (0.79 lower to 0.28 higher)
BMI	179 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.55 standard deviations lower (0.99 to 0.1 lower)
Quality of life	152 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.08 standard deviations lower (0.62 lower to 0.45 higher)
BMI FU	27 (1 study)	⊕⊖⊖⊖ VERY LOW1,3		Not calculable for	The mean BMI fu in the intervention groups was

		due to risk of bias, imprecision		SMD values	0.35 standard deviations lower (1.11 lower to 0.42 higher)
Readmission FU	24 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 2 (0.45 to 8.94)	167 per 1000	167 more per 1000 (from 92 fewer to 1000 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 The day patients were heavier/had a higher BMI than inpatients at baseline and slightly lower duration of illness. The authors did not adjust for potential confounders. Length of stay was longer for inpatients vs. day patient. Investigators and participants were not blinded.
2 For a dichotomous outcome, there are fewer than 300 events.
3 For a continuous outcome, there were fewer than 400 participants

Table 52: Summary of findings table for inpatient care versus ambulatory care for people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with ambulatory care AN	Risk difference with Inpatient (95% CI)
BMI FU	143 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.13 standard deviations lower (0.48 lower to 0.22 higher)
Hospitalisation in last 6 months FU	143 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 2.67 (1.5 to 4.77)	155 per 1000	258 more per 1000 (from 77 more to 583 more)
Remission _ITT_FU	143 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 0.82 (0.37 to 1.82)	186 per 1000	33 fewer per 1000 (from 117 fewer to 152 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Patient in hospital had a lower BMI versus ambulatory care. Pure restrictive forms were overrepresented in the inpatient group. Prevalence of history of

suicide attempts in the last 24 months was also higher. This group underwent longer treatment (on average of 1.5 years) than the ambulatory group. Finally, a larger percentage of patients were still followed by specialists in nutrition and/or psychiatry at the time of the survey. Neither patients nor investigators were blind.
2 For a continuous outcome, there were fewer than 400 participants
3 For a dichotomous outcome, there are fewer than 300 events

Table 53: Summary of findings table for partial hospitalisation (PH) and support versus partial hospitalisation for people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with PH AN	Risk difference with Partial Hospitalisation + Support (95% CI)
Difference in Weight Gain	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in weight gain in the intervention groups was 1.02 standard deviations higher (0.13 to 1.91 higher)
Difference in BMI	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in BMI in the intervention groups was 0.4 standard deviations higher (0.26 lower to 1.06 higher)
Difference in Purging	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in purging in the intervention groups was 0.57 standard deviations higher (0.38 lower to 1.52 higher)
Difference in EDI-2 Total Risk	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in edi-2 total risk in the intervention groups was 0.92 standard deviations higher (0.12 to 1.72 higher)
Difference in EDI-2 Drive for thinness	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in edi-2 drive for thinness in the intervention groups was 0.68 standard deviations higher (0.12 lower to 1.48 higher)
Difference in EDI-2 Body dissatisfaction	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in edi-2 body dissatisfaction in the intervention groups was 0.51 standard deviations higher (0.31 lower to 1.33 higher)

Difference in EDI-2 Bulimia	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in edi-2 bulimia in the intervention groups was 1.31 standard deviations higher (0.51 to 2.11 higher)
Difference EDEQ: Restraint	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference edeq: restraint in the intervention groups was 0.39 standard deviations higher (0.38 lower to 1.16 higher)
Difference EDEQ: Eating concern	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference edeq: eating concern in the intervention groups was 0.33 standard deviations higher (0.44 lower to 1.1 higher)
Difference EDEQ: Shape concern	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference edeq: shape concern in the intervention groups was 0.33 standard deviations higher (0.47 lower to 1.13 higher)
Difference EDEQ: Weight concern	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference edeq: weight concern in the intervention groups was 0.83 standard deviations higher (0.03 to 1.63 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Patients were not matched at baseline. Those who needed supported housing to potentially ensure successful outcome, were initially encouraged to receive Sage House service. However, the investigators attempted to address this by controlling for age, duration of eating disorder, and EDPHP length of stay
2 For a continuous outcome, there were fewer than 400 participants.

1 **Table 54: Summary of findings table for family therapy compared with inpatient care for people with anorexia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Inpatient AN	Risk difference with Family therapy (95% CI)
Readmission	171	⊕⊖⊖⊖	RR 0.56	546 per 1000	240 fewer per 1000

	(1 study)	VERY LOW1,2 due to risk of bias, imprecision	(0.36 to 0.87)		(from 71 fewer to 350 fewer)
Readmission > 3 times	90 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.6 (0.2 to 1.77)	185 per 1000	74 fewer per 1000 (from 148 fewer to 143 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Likely to be a similar population seeking ED assessment. After 2008 patients were then allocated to FT compared with those historically who were not. However, no baseline data was provided. No adjustments were made to account for covariates. Neither participants nor investigators were blind.
2 For a dichotomous outcome, there were fewer than 300 events.

1 **Table 55: Summary of findings table for day patient care versus inpatient care for people with bulimia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Inpatient BN	Risk difference with Day patient (95% CI)
EDI - Drive for thinness	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness in the intervention groups was 0.22 standard deviations lower (0.87 lower to 0.44 higher)
EDI - Body dissatisfaction	33 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction in the intervention groups was 0.32 standard deviations higher (0.37 lower to 1.01 higher)
EDI - Bulimia	33 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 0.13 standard deviations higher (0.56 lower to 0.82 higher)
SCL -90R Global Severity Index	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean scl -90r global severity index in the intervention groups was 0.26 standard deviations lower (0.94 lower to 0.42 higher)

Depression	34 (1 study)	⊕⊖⊖⊖ VERY LOW1 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.27 standard deviations lower (0.94 lower to 0.41 higher)
Remission_ITT	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 0.83 (0.31 to 2.24)	333 per 1000	57 fewer per 1000 (from 230 fewer to 413 more)
EDI - Bulimia FU	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia fu in the intervention groups was 0.41 standard deviations lower (1.07 lower to 0.25 higher)
EDI - Drive for thinness FU	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness fu in the intervention groups was 0.49 standard deviations lower (1.15 lower to 0.18 higher)
SCL -90R Global Severity Index FU	36 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean scl -90r global severity index fu in the intervention groups was 0.35 standard deviations lower (1.01 lower to 0.3 higher)
Depression FU	36 (1 study)	⊕⊖⊖⊖ VERY LOW1 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.35 standard deviations lower (1.01 lower to 0.3 higher)
Bingeing FU	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.23 standard deviations lower (0.88 lower to 0.43 higher)
Vomiting Severity FU	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting severity fu in the intervention groups was 0.21 standard deviations higher (0.45 lower to 0.86 higher)
Remission FU_ITT	36 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 5 (1.27 to 19.68)	111 per 1000	444 more per 1000 (from 30 more to 1000 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%

confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 The day patient group were heavier in weight and the inpatient group had more general psychopathology in the SCL-90-R scale. That is inpatients were more severely ill. Differences were also detected for depression, and interpersonal sensitivity. The authors did not adjusted for these differences. Neither the participants nor investigators were blind to treatment. There was an unclear duration of follow up.
2 For a continuous outcome, there are fewer than 400 participants.
3 For a dichotomous outcome, there are fewer than 300 events.

1 **Table 56: Summary of findings table for 5 days versus 4 day care for people with anorexia nervosa and bulimia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with 4 days_AN_BN	Risk difference with 5 days (95% CI)
Bingeing	369 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.37 standard deviations lower (0.59 to 0.14 lower)
Vomiting	359 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.21 standard deviations lower (0.43 lower to 0.02 higher)
BMI	153 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.37 standard deviations lower (0.69 to 0.04 lower)
EDI - Drive for thinness	461 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness in the intervention groups was 0.64 standard deviations lower (0.85 to 0.42 lower)
EDI - Bulimia	461 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 0.49 standard deviations lower (0.71 to 0.28 lower)
EDI - Body dissatisfaction	461 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean edi - body dissatisfaction in the intervention groups was 0.55 standard deviations lower

		imprecision			(0.77 to 0.33 lower)
Depression	408 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.73 standard deviations lower (0.95 to 0.5 lower)
Remission_ITT	756 (1 study)	⨁⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 3.31 (2.29 to 4.78)	101 per 1000	233 more per 1000 (from 130 more to 381 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Patients in 5-day were older, lighter, had more binges, vomiting, had lower depression and self-esteem problems, EDI was also better. Pre-treatment scores were used as covariates. Neither patients nor participants were blind.
2 95% CI crossed 1 MID (-0.5)
3 For a continuous outcome, there were fewer than 400 participants.
4 For a dichotomous outcome, there were fewer than 300 events.

1 **Table 57: Summary of findings table for inpatient CAMHS versus outpatient CAMHS for people with any eating disorder**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Outpatient CAMHS ANY ED	Risk difference with Inpatient CAMHS (95% CI)
BMI FU	57 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.17 standard deviations lower (0.69 lower to 0.36 higher)
EDI Bulimia FU	57 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia fu in the intervention groups was 0.4 standard deviations higher (0.14 lower to 0.93 higher)
EDI Body dissatisfaction FU	57 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction fu in the intervention groups was 0.05 standard deviations lower (0.57 lower to 0.48 higher)

EDI Drive for thinness FU	57 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness fu in the intervention groups was 0.19 standard deviations lower (0.71 lower to 0.34 higher)
SCL-90 Global Severity Index FU	57 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean scl-90 global severity index fu in the intervention groups was 0.22 standard deviations lower (0.75 lower to 0.31 higher)
Rosenberg Self Esteem FU	57 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean Rosenberg self-esteem fu in the intervention groups was 3.1 standard deviations higher (2.31 to 3.89 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; FU: follow up

1 There were significant differences between the groups for maturity, age of onset and Self-Esteem score at baseline. Patients treated as in-patients had significantly higher scores in the RSES and MF subscale comparing to the other two groups. The difference in the age of onset was statistically significant between patients treated as outpatients and those not treated by CAMHS. The authors did not adjust for any confounders. CAHMS patients were likely to have gotten treatment for a longer period compared with those who entered AMHS. Neither participants nor investigators were blind to treatment.
2 For a continuous outcome, there were fewer than 400 participants.

1 **Table 58: Summary of findings table for guided self-help (SH) versus day patient care for people bulimia nervosa or EDNOS**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Day Patient BN or EDNOS	Risk difference with Guided SH (95% CI)
EDE-Q Total	66 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q total in the intervention groups was 0.15 standard deviations higher (0.34 lower to 0.63 higher)
Objective binge eating	66 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean objective binge eating in the intervention groups was 0.43 standard deviations higher (0.06 lower to 0.92 higher)
Vomiting	65 (1 study)	⊕⊖⊖⊖ VERY LOW1,2		Not calculable for SMD	The mean vomiting in the intervention groups was

		due to risk of bias, imprecision		values	0.24 standard deviations higher (0.25 lower to 0.73 higher)
Excessive Exercise	66 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean excessive exercise in the intervention groups was 0.22 standard deviations lower (0.71 lower to 0.26 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; FU: follow up					
1 The patients were well matched at baseline for illness duration and severity (based on BMI). However, the ED diagnosis was different: CBT_GSH had higher number of BED and EDNOS-BN. The authors did not adjust for confounders. Neither participants nor investigators were not blinded. 2 For a continuous outcome, there were fewer than 400 participants.					

Table 59: Summary of findings table for extensive program (community outreach combined with limited) compared with a limited program (psychotherapy and nutritional counselling) for people with any eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Limited Program ANY ED	Risk difference with Extensive Program (95% CI)
Remission	123 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.39 (0.21 to 0.73)	537 per 1000	328 fewer per 1000 (from 145 fewer to 424 fewer)
Remission - AN	60 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.41 (0.18 to 0.91)	455 per 1000	268 fewer per 1000 (from 41 fewer to 373 fewer)
Remission - BN	63 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.38 (0.16 to 0.95)	578 per 1000	358 fewer per 1000 (from 29 fewer to 485 fewer)
Remission FU	123 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias,	RR 0.5 (0.35 to 0.72)	761 per 1000	381 fewer per 1000 (from 213 fewer to 495 fewer)

		imprecision			
Remission FU - AN	60 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.42 (0.26 to 0.68)	818 per 1000	475 fewer per 1000 (from 262 fewer to 605 fewer)
Remission FU - BN	63 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.61 (0.35 to 1.05)	733 per 1000	286 fewer per 1000 (from 477 fewer to 37 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Patients were allocated depending on their physical status, symptom severity, comorbidity, and occupational functioning. Patients who did not respond to limited treatment or who needed structured eating and had no regular occupation were assigned to intensive treatment. Patients assigned to intensive treatment had a higher rate of comorbidity, a longer duration of illness, more previous treatments, lower scores in social and occupational adjustment than those offered limited treatment. The authors did not adjust for confounders. Neither participants nor investigators were blinded.
2 For a dichotomous outcome, there were fewer than 300 events.

Table 60: Summary of findings table for history of inpatient care compared with no history of inpatient care for people with any eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with No history ANY ED	Risk difference with History of Inpatient (95% CI)
EDI- Drive for thinness	222 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- drive for thinness in the intervention groups was 0.02 standard deviations higher (0.28 lower to 0.31 higher)
EDI- Bulimia	222 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- bulimia in the intervention groups was 0.07 standard deviations higher (0.22 lower to 0.36 higher)
EDI-Body dissatisfaction	222 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-body dissatisfaction in the intervention groups was 0.18 standard deviations lower (0.48 lower to 0.11 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It is not clear what the differences in severity were between those who had (historically) received inpatient vs not. No adjustments were made for confounders. Neither participants nor investigators were blinded.
2 For a continuous outcome, there were fewer than 300 events.

Table 61: Summary of findings table for specialist (Sp) compared to non-specialist (Non-Sp) assessment for people with any eating disorder

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with non-specialist assessment and treatment (ANY ED)	Risk difference with Specialist (95% CI)
Admitted to inpatient treatment - Sp to Sp vs. NonSp to Non Sp	69 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.81 (0.24 to 2.68)	188 per 1000	36 fewer per 1000 (from 142 fewer to 315 more)
Admitted to inpatient treatment - Sp to Sp vs. NonSp to Sp	68 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.38 (0.15 to 0.92)	400 per 1000	248 fewer per 1000 (from 32 fewer to 340 fewer)
Admitted to inpatient treatment - Non Sp to Non Sp vs. Non Sp to Sp	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.47 (0.14 to 1.55)	400 per 1000	212 fewer per 1000 (from 344 fewer to 220 more)
Continuity of care - Sp to Sp vs. NonSp to Sp	69 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.11 (0.81 to 1.51)	750 per 1000	83 more per 1000 (from 142 fewer to 382 more)
Continuity of care - Sp to Sp vs. NonSp to NonSp	68 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 2.08 (1.1 to 3.9)	400 per 1000	432 more per 1000 (from 40 more to 1000 more)
Continuity of care - Non Sp to Sp vs. Non Sp to Sp	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2	RR 1.88 (0.95 to	400 per 1000	352 more per 1000 (from 20 fewer to

		due to risk of bias, imprecision	3.71)		1000 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up; OR: Odds ratio; Sp- Specialist; Non Sp – non-specialist

1 Comparisons between PCT groups revealed no statistically significant differences in age, gender, ethnicity, weight for height percentage at assessment, or referrals. Thus no adjustments were needed. But unclear how they estimated predicted referrals and no data was provided on success rates. Neither participants nor investigators were blind.
2 For a dichotomous outcome, there were fewer than 300 events.

Table 62: Summary of findings table for inpatient treatment versus variation of day, inpatient and outpatient care for people with any eating disorder

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Variation (Day, Hospital, OutP) - AN	Risk difference with Inpatient (95% CI)
Body Weight (ABW)	29 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean body weight (abw) in the intervention groups was 0.75 standard deviations lower (1.51 lower to 0.01 higher)

Table 63: Summary of findings table for prior opt-in versus post opt-in for people with any eating disorder

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Post opt-in ANY ED	Risk difference with Prior opt-in (95% CI)
% attended their first appointment	138 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.1 (1.02 to 1.18)	618 per 1000	62 more per 1000 (from 12 more to 111 more)
Overall attrition rates	138 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.80 (0.77 to 4.25)	103 per 1000	82 more per 1000 (from 24 fewer to 335 more)
Did not attend	138 (1 study)	⨁⊖⊖⊖ VERY LOW1,2	RR 3.2 (1.04 to	44 per 1000	97 more per 1000 (from 2 more to 317 more)

		due to risk of bias, imprecision	8.18)		
No cancellations	138 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.97 (0.93 to 1.02)	0 per 1000	-

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 No demographic data so unable to know if there were any differences pre and post opt-in intervention.
2 For a dichotomous outcome, there were fewer than 300 events.

1 **Table 64: Summary of findings table for meal supervision versus no meal supervision for people with any eating disorder**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with No meal supervision	Risk difference with Meal Supervision (95% CI)
Length of Hospital Stay	51 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean length of hospital stay in the intervention groups was 0.51 standard deviations higher (0.13 lower to 1.15 higher)
Weight gain	47 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean weight gain in the intervention groups was 0.33 standard deviations higher (0.33 lower to 0.99 higher)
Bradycardia (HR <45 BPM) % days in treatment	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean bradycardia (hr <45 bpm) % days in treatment in the intervention groups was 0.62 standard deviations lower (1.28 lower to 0.04 higher)

2 **Table 65: Summary of findings table for eating disorder specialist ward versus general ward for people with anorexia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with General ward	Risk difference with Eating disorder unit (95% CI)
BMI	110	⊕⊖⊖⊖	Not calculable for	The mean BMI in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with General ward	Risk difference with Eating disorder unit (95% CI)
	(1 study)	VERY LOW1,2 due to risk of bias, imprecision	SMD values	1.29 standard deviations higher (0.87 to 1.72 higher)
Length of time in hospital	110 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean length of time in hospital in the intervention groups was 0.02 standard deviations higher (0.37 lower to 0.40 higher)
Morgan Russell Score	110 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean Morgan Russell score in the intervention groups was 0.68 standard deviations higher (0.28 to 1.07 higher)
General health	110 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean general health in the intervention groups was 0.19 standard deviations higher (0.19 lower to 0.57 higher)
Children's global assessment	110 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean children's global assessment in the intervention groups was 0.15 standard deviations lower (0.54 lower to 0.23 higher)

1 **Anorexia nervosa stepped care**

2 **Table 66: Summary table of findings for family-based treatment (FBT) then intensive family coaching versus family-based treatment**
3 **for young people with anorexia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with FBT	Risk difference with FBT->IPC (95% CI)
Recovered from AN (>=95% EBW)	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.12 (0.6 to 2.07)	522 per 1000	63 more per 1000 (from 209 fewer to 558 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with FBT	Risk difference with FBT->IPC (95% CI)
BMI	35 (1 study)	⊕⊖⊖⊖ VERY LOW2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.28 standard deviations higher (0.42 lower to 0.98 higher)
% Expected Body Weight	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean % expected body weight in the intervention groups was 0.22 standard deviations higher (0.48 lower to 0.92 higher)
EDE Global	35 (1 study)	⊕⊖⊖⊖ VERY LOW2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.92 standard deviations higher (0.18 to 1.65 higher)
Depression BDI	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.59 standard deviations higher (0.12 lower to 1.3 higher)
Yale-Brown-Cornell Eating Disorder Scale	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean Yale-Brown-Cornell eating disorder scale in the intervention groups was 0.71 standard deviations higher (0.01 lower to 1.43 higher)
Service user experience Helping Relationship Questionnaire	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias,		Not calculable for	The mean service user experience in the intervention groups was 0.86 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with FBT	Risk difference with FBT->IPC (95% CI)
		indirectness, imprecision		SMD values	(1.59 to 0.13 lower)
Number of Sessions attended	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean number of sessions attended in the intervention groups was 0.92 standard deviations higher (0.18 to 1.65 higher)
Suitability of therapy - child Therapy Suitability and Patient Expectancy	35 (1 study)	⊕⊖⊖⊖ VERY LOW2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean suitability of therapy - child in the intervention groups was 0.38 standard deviations lower (1.09 lower to 0.32 higher)
Child's expectations about therapy Therapy Suitability and Patient Expectancy	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean child's expectations about therapy in the intervention groups was 0.45 standard deviations lower (1.16 lower to 0.26 higher)
Suitability of therapy - Mother Therapy Suitability and Patient Expectancy	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean suitability of therapy - mother in the intervention groups was 0.64 standard deviations higher (0.08 lower to 1.35 higher)
Mother's expectations about therapy Therapy Suitability and Patient Expectancy	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean mother's expectations about therapy in the intervention groups was 0.54 standard deviations higher (0.17 lower to 1.25 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with FBT	Risk difference with FBT->IPC (95% CI)
Suitability of therapy - Father Therapy Suitability and Patient Expectancy	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean suitability of therapy - father in the intervention groups was 0 standard deviations higher (0.7 lower to 0.7 higher)
Father's expectations about therapy Therapy Suitability and Patient Expectancy	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean father's expectations about therapy in the intervention groups was 0.27 standard deviations lower (0.97 lower to 0.43 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Lock & Le Grange 2015: High risk of selection and performance bias.
2 Participants initially randomized into FBT only and FBT/IPC groups. Participants in FBT/IPC group subsequently divided into IPC (those <2.3 kg weight gain by week 4 of FBT) and No IPC groups (those >2.3 kg weight gain by week 4 of FBT). Data only for FBT+IPC vs FBT+No IPC groups.
3 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).
4 CI crosses either 0.75 or 1.25 (Risk ratio), or either 0.5 or -0.5 (SMD).

1 **Bulimia nervosa stepped care**

2 **Table 67: Summary table of findings for group psychoeducation then either CBT-ED or wait list control (WLC) in adults with bulimia**
3 **nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group Psychoeducation->WLC	Risk difference with Group Psychoeducation->CBT-ED (95% CI)
Not in Remission	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.67 (0.48 to 0.95)	842 per 1000	278 fewer per 1000 (from 42 fewer to 438 fewer)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group Psychoeducation->WLC	Risk difference with Group Psychoeducation->CBT-ED (95% CI)
Not in Remission from Bingeing	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.62 (0.41 to 0.92)	789 per 1000	300 fewer per 1000 (from 63 fewer to 466 fewer)
Not in Remission from Purging	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.58 (0.38 to 0.89)	789 per 1000	332 fewer per 1000 (from 87 fewer to 489 fewer)
Binge Frequency EDE 28 days	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.54 standard deviations lower (1.11 lower to 0.02 higher)
Purge Frequency EDE 28 days	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean purge frequency in the intervention groups was 0.7 standard deviations lower (1.27 to 0.13 lower)
EDE Global	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.08 standard deviations lower (0.63 lower to 0.48 higher)
Depression BDI	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.17 standard deviations lower (0.72 lower to 0.39 higher)
General Psychopathology Brief Symptom Inventory	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0.21 standard deviations lower (0.76 lower to 0.35 higher)
General Functioning SAS	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean social adjustment in the intervention groups was 0.3 standard deviations lower (0.86 lower to 0.25 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group Psychoeducation->WLC	Risk difference with Group Psychoeducation->CBT-ED (95% CI)
confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Davis 1999: unclear randomization method and allocation concealment. No participant blinding, unclear investigator and assessor blinding. Unclear whether baseline characteristics similar. 2 CI crosses either 0.75 or 1.25 (Risk Ratio).					

Table 68: Summary table of findings for self-help manual then CBT-ED versus CBT-ED at end of treatment and follow up in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT-ED	Risk difference with Self-Help Manual for BN -> CBT-ED (95% CI)
Remission Abstinence from bingeing, purging or other weight control behaviour in past month (or if not available: BITE Symptom score<=11 and BITE Severity score=0)	86 (1 study) 18 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.01 (0.53 to 1.93)	300 per 1000	3 more per 1000 (from 141 fewer to 279 more)
Remission 18-mo FU Abstinence from bingeing, purging or other weight control behaviour in past month (or if not available: BITE Symptom score<=11 and BITE Severity score=0)	64 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.97 (0.54 to 1.76)	412 per 1000	12 fewer per 1000 (from 189 fewer to 313 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Treasure 1996: inadequate randomization method and allocation concealment; No participant blinding, unclear investigator and assessor blinding; dropout rate of CBT-ED group>20%. 2 CI crosses both 0.75 and 1.25.					

Table 69: Summary table of findings for guided self-help CBT-ED then antidepressant then CBT-ED versus CBT-ED then antidepressant (AD) in adults with bulimia nervosa.

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with CBT-BN->AD	Risk difference with GSH CBT->AD->CBT-BN (95% CI)
Remission	293 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, indirectness, imprecision	RR 0.94 (0.67 to 1.33)	313 per 1000	19 fewer per 1000 (from 103 fewer to 103 more)
EDE Global	293 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values
EDE Restraint	293 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values
EDE Shape Concerns	293 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values
EDE Weight Concerns	293 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values
EDE Eating Concerns	293 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values
Yale-Brown-Cornell ED Scale - Preoccupation	293 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT-BN->AD	Risk difference with GSH CBT->AD->CBT-BN (95% CI)
Yale-Brown-Cornell ED Scale - Ritual	293 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values
Depression BDI	293 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values
Quality of Life Quality of Well Being Scale	293 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	Not calculable for SMD values

The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up; BDI: Becks Depression Index

1 Mitchell 2011/Crow 2013: Unclear allocation concealment. No participant nor investigator blinding. Dropout rates of both groups>20%, no details provided for reasons.

2 I2>50%.

3 Randomization was to different treatments. No randomisation to next level of stepped care.

4 CI crosses both 0.75 and 1.25 (Risk Ratio).

5 <400 participants.

5.2.3 Economic Evidence

5.2.3.1 Coordination of care and treatment setting

The systematic search of the economic literature undertaken for the guideline identified:

- One UK study on the cost effectiveness of inpatient psychiatric treatment versus specialist outpatient treatment and general outpatient treatment in young people with AN (Byford et al., 2007); follow up data in (Gowers et al., 2010);
- One German study on the cost effectiveness of day treatment in young people with AN (Herpertz-Dahlmann et al., 2014);
- One US study on the cost effectiveness of partial day hospital care in adults with AN or sub-threshold AN or BN or sub-threshold BN (Williamson et al., 2001);
- One US study assessing the cost effectiveness of an adequate care model (inpatient care, partial hospital care, psychotherapy and medication management) in people with AN (Crow and Nyman, 2004);
- One Australian study on the cost effectiveness of a best practice model (early intervention, a range of care from GPs, self-help, intensive outpatient and residential care, inpatient care, stepped care approach) in people with AN, BN, BED or EDNOS (Deloitte Access Economics, 2012).

References to included studies and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix P. Completed methodology checklists of the studies are provided in Appendix O. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix Q.

Byford and colleagues (2007) evaluated the cost effectiveness of an inpatient psychiatric treatment compared with specialist outpatient treatment and also with general outpatient treatment in young people aged between 11-17 years with AN in the UK. The economic evaluation was undertaken alongside an RCT (Byford 2007) (N=167 at baseline, N=160 at two years for the effectiveness data). Byford and colleagues (2007) report results over two years, whereas the publication by Gowers and colleagues (2010a) is based on the same RCT but reports cost over three to five years. Inpatient psychiatric treatment was provided within generic children's or young people psychiatric inpatient unit. It lasted six weeks and comprised a multidisciplinary psychiatric approach with the aim of normalising eating, restoring healthy weight and facilitating psychological (cognitive) change. Each participant received both individual supportive or cognitive therapies and family therapy. Specialist outpatient treatment comprised motivational interview, individual CBT plus parental feedback (12 sessions), parental counselling with the individual (minimum of four sessions, increasing to eight for younger people), dietary therapy (four sessions, with parental involvement as required), multi-modal feedback (weight, self-report and clinician-rated questionnaire) and monitoring (four sessions). The treatment was designed to last six months. General outpatient treatment adopted a multidisciplinary, family-based approach, with variable dietetic, individual supportive therapy and paediatric (medical) liaison. The analysis was conducted from a public sector perspective (health, social care and education). The study considered a range of costs including secondary health services (inpatient and outpatient visits, day care attendances, A&E visits), community health and social service contacts (GP, practice nurse, dietician, district nurse, health visitor, community paediatrician, community psychiatric nurse, clinical psychologist, counsellor, family therapist, dentist, school doctor, school nurse, social worker, eating disorders association, family therapy, foster care), education (state day school, independent day school, independent boarding school, hospital school, home tuition, school counsellor, education welfare officer). The resource use estimates were based on the RCT (N=135 at 2 years, N=71 3-5 years). The unit costs were obtained from national sources. The measure of outcome for the economic analysis was the

improvement measured on MRAOS scale. The time horizon of the analysis was five years. Results were reported at two years and then during three to five years. Costs beyond year one were discounted at 3.5%.

The specialist outpatient treatment resulted in higher MRAOS scores at two years follow up when compared with the other two treatment options. The scores were 8.3 (SD 2.6), 8.4 (SD 2.4) and 8.3 (SD 2.6) for inpatient treatment, specialist outpatient treatment and general outpatient treatment, respectively. The difference between inpatient and specialist outpatient treatment was not significant. There was no difference between inpatient and general outpatient treatment. Outcomes at five years were not reported.

The mean total cost per participant at two years follow up was £34,531 (SD £52,439) for inpatient treatment, £26,738 (SD £46,809) specialist outpatient treatment and £40,794 (SD £63,652) for the general outpatient treatment (in likely 2003/04 prices). The difference between inpatient treatment and specialist outpatient treatment was £7,793 (specialist outpatient treatment had lower cost) and the difference between inpatient treatment and general outpatient treatment was £6,262 (inpatient treatment had lower cost). However, none of the cost differences was statistically significant.

The mean total cost per participant during three to five years of follow up was £15,304 (SD £69,083) for the inpatient treatment, £15,636 (SD £46,545) for the specialist outpatient treatment and £15,203 (SD £61,275) for the general outpatient treatment. None of the cost differences was statistically significant.

Based on the above, at two years follow up specialist outpatient treatment dominated both inpatient and general outpatient treatment. At a WTP of £0 per additional point of improvement on MRAOS scale, the probability that specialist outpatient treatment is cost effective is 78%, the probability that inpatient treatment is cost effective is 16% and the probability that general outpatient treatment is cost effective is only 6%.

The findings were robust to changes in the discount rate and assumptions underlying analyses of missing data; also the exclusion of education costs had no impact of the conclusions.

The analysis was judged by the committee to be directly applicable to the NICE decision-making context. Even though the authors did not attempt to estimate QALYs this was not a problem in terms of interpretation of findings since the specialist outpatient treatment was found to be dominant at two years follow up. Overall, this was a well conducted study and was judged by the committee to have only minor methodological limitations.

Herpertz-Dahlmann and colleagues (2013) evaluated the cost effectiveness of a day treatment programme (following a short inpatient treatment) compared with continued inpatient treatment (following a short inpatient treatment) in young people females (11-18 years) with AN in Germany. The economic analysis was conducted alongside an RCT (Herpetz-Dahlmann 2014) (N=172).

The analysis was conducted from the health care provider perspective. The study considered a range of costs including psychiatrist visits, psychologist visits, admissions (including re-admissions) and outpatient visits. The resource use estimates were based on the RCT. The unit costs were obtained from hospital tariffs.

The measure of outcome for the economic analysis was the improvement in BMI (between the time of admission and follow up). The time horizon of the analysis was 12 months.

Day treatment resulted in a greater improvement in BMI when compared with the continued inpatient treatment (3.2 versus 2.7 points, a difference of 0.46 points; $p < 0.0001$). The mean total cost per participant at the 12 month follow up was €31,114 (SD €16,246) and €39,481 (SD €16,174) for day treatment and inpatient treatment, respectively; a difference of €8,367

(in favour of day treatment) in likely 2013 Euros, p = 0.002. Based on the above, day treatment was the dominant option (that is, it resulted in better outcomes and lower costs).

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it has been conducted in Germany. The authors did not attempt to estimate quality adjusted life years (QALYs). However, this was not a problem in terms of the interpretation of the findings, since day treatment was the dominant option. This was a well conducted study and was judged by the committee to have only minor methodological limitations.

Williamson and colleagues (2001) attempted to assess the cost effectiveness of partial day hospital compared with inpatient care in people with AN, sub-threshold AN or BN and sub-threshold BN in the US. People assigned to inpatient or partial day treatment attended the same psychological treatment programme. Inpatients stayed on an adult or young people unit. People receiving day hospital care lived at home or stayed in local hotels. The programme included supervised meals and group therapy, including special groups for body image, behaviour management, CBT, meal planning, nutrition education, activity therapy and exercise. Also, most people were prescribed psychotropic medication. The economic analysis was based on an observational cohort study (N=51). The analysis was conducted from a health care provider perspective. The study considered only costs associated with treatment and admissions. The clinical effectiveness data and resource use estimates were based on observational cohort study. Unit costs were obtained from local sources (hospital financial records). The measure of outcome for the economic analysis was improvement as measured by BMI and MAEDS. However, the authors did not report the effectiveness data for each arm of the study. The authors only reported that in both groups there was a significant improvement in BMI and on all MAEDS subscales at the end of treatment and 12 month follow up, p < 0.007. So, in effect, this was a cost analysis.

The total mean cost per participant at 12 month follow up was $12,740 (SD $16,414) and $22,385 (SD $18,024) for partial day hospital and inpatient care, respectively; a difference of $9,645 (in favour of partial day hospital), p < 0.02 (in likely 2000 US dollars).

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it has been conducted in the US. This study was judged by the committee to have potentially serious methodological limitations, including the study design (small observational cohort study), lack of consideration of wider health care and social care costs and the use of local unit costs.

Crow and Nyman (2004) evaluated the cost effectiveness of an adequate care model compared with SC in people with AN in the US. This was a modelling study. Adequate care was defined as 45 days of inpatient hospital treatment, 20 days of partial hospital, 50 sessions of psychotherapy (50 min per each session), medication management (20 sessions) and fluoxetine (60 mg per day) for two years. SC was defined as seven days of inpatient hospital treatment, 15 days of partial hospital, 25 sessions of psychotherapy (50 min per each session), medication management (20 sessions) and fluoxetine (60 mg per day) for two years.

The analysis was conducted from a health care payer perspective. The study considered a range of costs including inpatient treatment, partial hospitalisation, psychotherapy, outpatient visits, medication and medication management. The resource use estimates were based on charge data. Unit costs were obtained from local sources. The measure of outcome for the economic analysis was the number of life years saved (LYS). The time horizon of the analysis was life time.

The adequate care model resulted in 2.75 additional life years saved. The mean life time costs per person were $119,200 and $36,200 for the adequate care model and SC, respectively; a difference of $83,000 (in likely 2003 US dollars). Based on the above, the ICER of adequate care was $30,180 per additional LYS.

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it has been conducted in the US. The authors did not attempt to estimate quality adjusted life years (QALYs) which made it difficult to interpret the cost-effectiveness results and to compare the findings with other studies. This study was judged by the committee to have potentially serious methodological limitations, including the assumptions about mortality rates and treatment efficacy that were based on authors' opinion, no consideration of wider health care costs, use of local unit costs and lack of sensitivity analyses.

Deloitte Access Economics (2012) evaluated the cost effectiveness and cost benefit of a best practice model when compared with treatment as usual (TAU) in people with AN, BN, BED and EDNOS in Australia. This was a modelling study with effectiveness data derived from a systematic review of RCTs, other published sources and authors' assumptions. The best practice model focused on early intervention, a range of delivery options from general practitioners and online self-help, through intensive outpatient and residential programmes, to full inpatient hospitalisation; a stepped care approach, realising that people might need to progress both up and down (sometimes repeatedly) through delivery levels; and long-term follow up, to prevent relapse. TAU was defined as patchy services (largely untreated ED), no specialist ED inpatient services, no continuity in care and sub-optimal treatment dose. The analysis was conducted from a societal perspective. The study considered a range of costs including treatment provision and other health care costs, productivity, employment and welfare. The resource use estimates were from published sources. The source of unit costs was unclear. The measure of outcome for the economic analysis was disability adjusted life years (DALYs) and monetised DALYs. DALYs were converted into a dollar figure using an estimate of the value of a statistical life year (VSLY). The VSLY is an estimate of the value society places on an anonymous life year. The time horizon of the analysis was 10 years. A discount rate of 7% was applied to costs and monetised DALYs.

The best practice model resulted in fewer DALYs per participant (0.96 versus 2.25, respectively; a difference of 1.29 DALYs in favour of the best practice model. The monetised DALYs were equal to $161,346 and $353,647 (in likely 2013 AU dollars) with the best practice model and TAU, respectively; the net savings associated with the best practice model were $192,301 per participant. The best practice model also resulted in a reduction in the mean total costs over 10 years ($72,699 versus $130,390 for the best practice model and TAU, respectively; a difference of $57,690 in favour of the best practice model). Based on the above, the best practice model was found to be dominant (that is, it led to cost reductions and also fewer DALYs). When using monetised DALYs the savings amounted to $250,261 per participant over 10 years.

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it has been conducted in Australia. This study was judged by the committee to have potentially serious methodological limitations, including some of the clinical input parameters and resource use inputs were based on the authors' assumptions; the unclear source of unit cost data, discount rate of 7% for costs and outcomes and lack of sensitivity analyses.

5.2.3.2 Stepped care

The systematic search of the economic literature undertaken for the guideline identified:

- 1 US study and 1 Finnish study on the cost effectiveness of stepped care models in adults with BN (Crow et al., 2013, Pohjolainen et al., 2010).

References to included studies and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix P. Completed methodology checklists of the studies are provided in Appendix O. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix Q.

Crow and colleagues (2013) evaluated the cost effectiveness of a stepped care model compared with high intensity CBT treatment augmented as indicated with fluoxetine in adult women with purging or non-purging BN alongside an RCT (Crow 2013) (N=293) conducted in the US. The stepped care model involved a stepped series of interventions moving from less intensive and less expensive to more intensive and expensive interventions. Interventions included CBT, self-help, admissions, outpatient care and medication management.

The analysis was conducted from a health care provider perspective. The study considered a range of costs including CBT, self-help, medication, physician visits, emergency room, hospitalisation, individual therapy, group therapy and medication. The resource use estimates were based on the RCT (N=293). The unit costs were obtained from national sources (Medicare rates) and where necessary supplemented with other published sources. The measure of outcome for the economic analysis was the proportion of participants abstinent at 12 month follow up. The time horizon of the analysis was 12 months.

The stepped care model resulted in a greater proportion of participants abstinent at 12 month follow up (26% versus 18% for the stepped care model and high intensity CBT, respectively; a difference of 8%). The mean total costs per participant over 12 months were $3,158 for the stepped care model and $3,657 for a high intensity CBT, a difference of $499 (in favour of the stepped care model) in 2005 US dollars. Based on the above, the stepped care model was dominant (that is, it led to cost savings and also a greater proportion of participants abstinent at the 12 month follow up). Bootstrapping indicated that the stepped care model was both less expensive and more effective than high intensity CBT in 81% of the replications.

The results were robust to changing the assumptions pertaining to the unit cost estimates (that is, instead of using Medicare rates, actual fees charged were used).

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it has been conducted in the US. The authors did not attempt to estimate quality adjusted life years (QALYs). However, this was not a problem in terms of the interpretation of findings since the stepped care model was found to be dominant. Overall, this was a well conducted study and was judged by the committee to have only minor methodological limitations.

Pohjolainen and colleagues (2010) evaluated the cost utility of a stepped care model compared with ‘no treatment’ in adult females with BN in Finland. This was a modelling study with the effectiveness data derived from an observational cohort study (N=72), published studies and authors’ assumptions. The stepped care model was defined as psychoeducation that included elements of CBT, followed by group CBT (8 sessions) and then individual CBT (20 sessions), and followed by day hospital or inpatient treatment. Participants also received psychopharmacological treatment if needed, individual nutritional counselling and social skills training. The analysis was conducted from a health care payer perspective. The study considered only the costs associated with the intervention provision including admissions, outpatient visits, laboratory testing and radiology. The resource use estimates were from the observational cohort study. The unit costs were obtained from local sources. The measure of outcome for the economic analysis was the QALY with HRQoL weights derived using the 15D generic instrument with valuations provided by the general Finnish population. The time horizon of the analysis was 10 years. However, the costs and outcomes were measured only over six months. It was assumed by the authors that there was no difference in the costs between the groups over the study period between six months and 10 years. Also, it was assumed in the base case analysis that in untreated people their HRQoL improves linearly in 10 years to the same level as the treated people had after six months of treatment. For those treated, the authors assumed that the HRQoL gain by six months would persist until 10 years. Discounting was undertaken only in the sensitivity analysis using either 3% or 5% for outcomes.

The stepped care model resulted in 0.241 QALYs gained at the 10 year follow up. The incremental undiscounted cost of stepped care model at six month follow up was €3,972 (SD €5,518) per participant (in likely 2010 Euros). Based on the above, the mean undiscounted cost per QALY gained for stepped care model was €16,481 when compared with no treatment.

Deterministic sensitivity analyses indicated that the cost per QALY associated with the best practice model was €19,663 per QALY and €17,812 per QALY when using the discount rate for QALYs of 5% and 3%, respectively. Using the upper and lower 95% confidence intervals for incremental QALYs of 0.339 and 0.113 resulted in an ICER of €11,717 and €35,150 per QALY, respectively. Using upper and lower 95% confidence interval for the incremental costs of €5,269 and €4,702 resulted in an ICER of €21,863 and €19,510 per QALY. Using upper 95% confidence interval for the incremental costs and lower 95% confidence interval for incremental QALYs resulted in and ICER of €46,628 per QALY.

In the base case analysis, it was assumed that in untreated people the HRQoL improved linearly over 10 years to the same level as in the treated people after 6 months of treatment. In the best case scenario it was assumed that in people receiving 'no treatment' HRQoL did not improve at all. This best case scenario resulted in an ICER of €1,455 per QALY. Similarly, using the best case scenario, but discounting QALYs gained at 5% resulted in an ICER of €4,428 per QALY.

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it was conducted in Finland. The authors estimated QALYs; however, the HRQoL weights were derived using the 15D generic instrument with the valuations being obtained from a general Finnish population. This study was judged by the committee to have potentially serious methodological limitations, including the study design that provided the efficacy data (small observational cohort study), the consideration of intervention costs only, the various assumptions regarding future costs and benefits in both arms of the model and the use of local unit costs.

5.2.4 Clinical evidence statements

5.2.4.1 RCT evidence for coordination of care and treatment setting

Inpatient care versus day clinics for adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=43) showed no difference in the effect of inpatient care on BMI, vomiting, bingeing, EDI-bulimia, global severity index and remission compared with day clinics.

Inpatient care versus day clinic for adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=37) showed no difference in the effect of inpatient care on bingeing, vomiting, EDI-total, global severity index and remission, compared with day clinic.

Inpatient care versus outpatient individual and family therapy for adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=50) showed no difference in the effect of inpatient care on change in the following Morgan-Russell scores: global, menstruation, nutrition, mental state, sexual adjustment, social economic adjustment, compared with outpatient individual and family therapy.

Inpatient care versus outpatient group therapy for adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=50) showed no difference in the effect of inpatient care on change in the following Morgan-Russell scores: global menstruation, nutrition, mental state, sexual adjustment, social economic adjustment, compared with outpatient group therapy.

Inpatient care versus wait list control for adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=50) showed no difference in the effect of inpatient care on change in the following Morgan-Russell scores: global menstruation, nutrition, mental state, sexual adjustment, social economic adjustment, compared with wait list control.

Inpatient care versus specialist outpatient (CBT-ED) for adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=104) showed no difference in the effect of inpatient care on BMI, EDI-total, Morgan-Russell total score and remission compared with specialist outpatient.

Inpatient care versus general outpatient (CAMHS) for adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=100) showed no difference in the effect of inpatient care on BMI, EDI-total, Morgan-Russell total score and remission compared with general outpatient.

Inpatient care versus day clinic for young people with anorexia nervosa at follow up

Low quality evidence from one RCT (n=143 to 172) showed no difference in the effect of inpatient care on BMI, EDI-total, global severity index, remission, relapse, menstrual function and adverse events compared with day clinic

Low quality evidence from one RCT (n=143 to 172) showed inpatient care is less effective on EDI-bulimia compared with day clinic, but there was some uncertainty.

Specialist outpatient (CBT-ED) versus general outpatient (CAMHS) for adults with anorexia nervosa at follow up

Very low to low quality evidence from one RCT (n=98 to 110) showed no difference in the effect of specialist outpatient on BMI, EDI-total, Morgan-Russell total, remission and readmission to hospital compared with general outpatient care.

Inpatient group versus outpatient (family therapy) for adults with bulimia nervosa at follow up

Very low quality evidence from one RCT (n=71) showed no difference in the effect of inpatient group therapy on binges, vomiting, depression, bulimic severity score and remission compared with outpatient care

Specialist outpatient versus GP outpatient for adults with bulimia nervosa at the end of treatment

Very low quality evidence from one RCT (n=68) showed no difference in the effect of specialist outpatient treatment on binges, vomiting, EDE global or subscales, bulimic

investigatory test, depression or work, leisure, family life questionnaire compared with GP outpatient care.

Modified day treatment versus traditional outpatient for adults with any eating disorder at end of treatment

Low quality evidence from one RCT (n=43) showed modified day treatment is more effective on binges, vomiting, purging, depression, EDI-total score, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction compared with outpatient care.

Low quality evidence from one RCT (n=43) showed BMI increased in the modified day treatment compared with outpatient care but it is unclear if this is favourable outcome in this population.

Inpatient weight stabilisation versus inpatient weight restoration (long) for young people with any eating disorder at end of treatment

Low quality evidence from one RCT (n=69 to 82) showed no difference in inpatient weight stabilisation on remission and change in EDE-global score compared with weight restoration.

Inpatient weight stabilisation versus inpatient weight restoration (long) for young people with any eating disorder at follow up

Very low quality evidence from one RCT (n=78 to 82) showed no difference in inpatient weight stabilisation on remission and change in EDE-global score compared with weight restoration.

5.2.4.2 Observational evidence for coordination of care and treatment setting

Inpatient care versus day patient care for adults with anorexia nervosa at end of treatment

Very low quality evidence from one observational study (n=152) showed no difference in inpatient care on bingeing, laxative use, vomiting, excessive exercise, EDE-total, quality of life compared with day patient care.

Very low quality evidence from one study (n=152) showed inpatient care is less effective on BMI compared with day patient care.

Inpatient care versus outpatient care for adults with anorexia nervosa at follow up

Very low quality evidence from one observational study (n=143) showed no difference in inpatient care on BMI, hospitalisation and remission compared with outpatient ambulatory care.

Partial hospitalisation and support versus partial hospitalisation for adults with anorexia nervosa at end of treatment

Very low quality evidence from one observational study (n=35) showed partial hospitalisation and support is more effective on weight gain, change in EDI-total, EDI-bulimia and EDE-weight concern compared with partial hospitalisation.

Very low quality evidence from one observational study (n=35) showed no difference in partial hospitalisation and support on BMI, purging, EDI-drive for thinness, EDI-body dissatisfaction, EDE-restraint, EDE-eating concern and EDE-shape concern compared with partial hospitalisation.

Family therapy versus inpatient care for adults with anorexia nervosa at end of treatment

Very low quality evidence from one observational study (n=171) showed family therapy is more effective on remission and readmission compared with inpatient care.

Day patient versus inpatient care for adults with bulimia nervosa at end of treatment

Very low quality evidence from one observational study (n=33 to 36) showed no difference in the effect of day patient care on EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction, global severity index, depression and remission compared with inpatient care.

Day patient versus inpatient care for adults with bulimia nervosa at follow up

Very low to low quality evidence from one observational study (n=33 to 36) showed no difference in the effect of day patient care on EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction, global severity index and depression compared with inpatient care.

Low quality evidence from one observational study (n= 36) showed day patient care is more effective on remission compared with inpatient care.

Five day inpatient care versus four day inpatient care for adults with bulimia nervosa or anorexia nervosa end of treatment

Low quality evidence from one observational study (n= 153 to 756) showed 5-day inpatient care is more effective on bingeing, BMI, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction, depression and remission compared with 4 day inpatient care.

Low quality evidence from one observational study (n= 153 to 756) showed 5 day inpatient care is more effective on vomiting compared with 4-day inpatient care, but there was some uncertainty.

Inpatient CAMHS versus outpatient CAMHS for adults with any eating disorder at follow up

Very low quality evidence from one observational study (n=57) showed no difference in the effect of inpatient CAMHS on BMI, EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction and global severity index compared with outpatient CAMHS.

Very low quality evidence from one observational study (n=57) showed inpatient CAMHS is more effective on self-esteem compared with outpatient CAMHS.

Guided self-help versus day patient care for adults with bulimia nervosa and EDNOS at end of treatment

Very low quality evidence from one observational study (n=57) showed no difference in the effect of guided self-help on EDE-total, bingeing, vomiting, excessive exercise compared with day patient care.

Extensive programme versus limited programme for adults with anorexia nervosa at end of treatment

Very low quality evidence from one observational study (n=60) showed extensive programmes is less effective on remission compared with a limited programme.

Extensive programme versus limited programme for adults with bulimia nervosa at end of treatment

Very low quality evidence from one observational study (n=63) showed extensive programmes is less effective on remission compared with a limited programme.

Extensive programme (community outreach combined with limited programme) versus limited programme (psychotherapy and nutritional counselling) for adults with anorexia nervosa at follow up

Very low quality evidence from one observational study (n=60) showed extensive programmes is less effective on remission compared with a limited programme.

Extensive programme (community outreach combined with limited programme) versus limited programme (psychotherapy and nutritional counselling) for adults with bulimia nervosa at end of treatment

Very low quality evidence from one observational study (n=63) showed no difference in the effect of extensive programmes on remission compared with a limited programme.

History of inpatient care versus no history for adults with any eating disorder at end of treatment

Very low quality evidence from one observational study (n=222) showed no difference in the effect of treatment in those who had a history of inpatient care on EDI-drive for thinness, EDI-body dissatisfaction, EDI-bulimia compared those with no history.

Specialist care versus non-specialist care for adults with any eating disorder at end of treatment

Very low quality evidence from one observational study (n=69) showed no difference in the number who were admitted to inpatient treatment if the patient went via specialist assessment to specialist treatment compared with non-specialist assessment to non-specialist treatment.

Very low quality evidence from one observational study (n=31) showed no difference in the number who were admitted to inpatient treatment if the patient went via non-specialist assessment to non-specialist treatment compared with non-specialist assessment to specialist treatment.

Very low quality evidence from one observational study (n=68) showed a lower number admitted to inpatient treatment if the patient went via specialist assessment to specialist treatment compared with non-specialist assessment to specialist treatment.

Very low quality evidence from one observational study (n=69) showed no difference in the continuity of care if the patient went via specialist assessment to specialist treatment compared with non-specialist assessment to specialist treatment.

Very low quality evidence from one observational study (n=69) showed no difference in the continuity of care if the patient went via non-specialist assessment to non-specialist treatment compared with non-specialist assessment to specialist treatment.

Very low quality evidence from one observational study (n=69) showed better continuity of care if the patient went via specialist assessment to specialist treatment compared with non-specialist assessment to specialist treatment.

Inpatient care versus variation in care (day, hospital, outpatient) for adults with anorexia nervosa at end of treatment

Very low quality evidence from one observational study (n=29) showed inpatient care is less effective on body weight compared with any other type of care, but there was some uncertainty.

Prior to opt-in versus post opt-in in adults with any eating disorder at end of treatment

Very low quality evidence from one observational study (n=138) showed opt-in is less effective on attendance to the first appointment compared with no opt-in.

Very low quality evidence from one observational study (n=138) showed opt-in is more effective on reducing the number who do not attend compared with no opt-in.

Very low quality evidence from one observational study (n=138) showed no difference in the effect of opt-in on overall attrition rates compared with no opt-in.

Very low quality evidence from one observational study (n=138) showed no difference in the effect of opt-in on cancellations rates compared with no opt-in.

Meal supervision versus no supervision in adults with any eating disorder at end of treatment

Very low quality evidence from one observational study (n=47 to 51) showed no difference in the effect of meal supervision on length of hospital stay and weight gain compared with no meal supervision.

Very low quality evidence from one observational study (n=47 to 51) showed meal supervision is more effective on incidence of bradycardia compared with no meal supervision but there was some uncertainty.

Eating disorder ward versus general ward for adults with anorexia nervosa at end of treatment

Very low quality evidence from one observational study (n=110) showed an eating disorder ward is more effective on BMI and Morgan-Russell score compared with a general ward.

Very low quality evidence from one observational study (n=110) showed no difference in the effect of an eating disorder ward on time in hospital, general health and children's global assessment compared with a general ward.

5.2.4.3 Stepped care for anorexia nervosa

Intensive family coaching with family-based treatment versus family-based treatment for anorexia nervosa

Very low quality evidence from one RCT (n=35) showed no difference in the effect of adding intensive family coaching to family-based treatment on recovery, BMI, % expected body weight, depression, the family's (child, mother and father) expectations about therapy and the child and father's perceptions about the suitability of therapy compared with family-based treatment only.

Very low quality evidence from one RCT (n=35) showed that adding intensive family coaching to family-based treatment may be more effective on YBC-EDS scores and the mother's perceptions about the suitability of therapy compared with family-based treatment, although there was some uncertainty.

Very low quality evidence from one RCT (n=35) showed that adding intensive family coaching to family-based treatment is less effective on EDE-global, service user experience and number of therapy sessions attended compared with family-based treatment.

5.2.4.4 Stepped care for bulimia nervosa

Group psychoeducation then CBT-ED or wait list control in adults with bulimia nervosa

Low quality evidence from one RCT (n=56) showed that group psychoeducation followed by CBT-ED was more effective on the number of people not in remission, not in remission from bingeing and not in remission from purging and purge frequency compared with group psychoeducation followed by wait list control.

Low quality evidence from one RCT (n=56) showed that group psychoeducation followed by CBT-ED may be more effective on binge frequency compared with group psychoeducation followed by wait list control, although there was some uncertainty.

Low quality evidence from one RCT (n=56) showed no difference in the effect of group psychoeducation followed by CBT-ED on EDE-global, general psychopathology, depression and general functioning, compared with group psychoeducation followed by wait list control.

Self-help for BN then CBT-ED versus CBT-ED at end of treatment in adults with bulimia nervosa

Very low quality evidence from one RCT (n=86) showed no difference in the effect of a self-help manual for bulimia nervosa followed by CBT-ED on remission compared with CBT-ED only.

Self-help for BN then CBT-ED versus CBT-ED at follow up in adults with bulimia nervosa

Very low quality evidence from one RCT (n=64) showed no difference in the effect of a self-help manual for bulimia nervosa followed by CBT-ED on remission compared with CBT-ED only.

Guided self-help for BN then antidepressant then CBT-ED versus CBT-ED then antidepressant in adults with bulimia nervosa

Very low quality evidence from one RCT (n=293) showed no difference in the effect of guided self-help for BN then antidepressant then CBT-ED on remission, EDE-global, EDE-dietary restraint, EDE-shape concerns, EDE-weight concerns, EDE-eating concerns, YBC-EDS-preoccupation, YBC-EDS-ritual, depression and quality of life compared with CBT-ED followed by antidepressant.

5.2.4.5 Binge eating disorders

No clinical evidence on stepped care was found on binge eating disorder.

5.2.4.6 Eating disorders not otherwise specified

No clinical evidence on stepped care was found on EDNOS.

5.2.5 Economic Evidence statements

5.2.5.1 Coordination of care

There was limited UK evidence (N=167) showing that specialist outpatient care was dominated (that is, it was more effective and resulted in lower costs) when compared with both inpatient and general outpatient care in young people with AN. This evidence came from a directly applicable study that was characterised by minor methodological limitations.

There was evidence from one German study (N=172) showing that day treatment when compared with inpatient care was dominant in people with AN. This evidence came from a partially applicable study that was characterised by minor methodological limitations.

There was evidence from one US study (N=51) showing that partial day hospital care was cost saving when compared with inpatient care in people with AN or BN (or sub-threshold AN or BN). This evidence came from a partially applicable study that was characterised by potentially serious methodological limitations.

There was evidence from one US modelling study showing that adequate care model (inpatient care, partial hospital treatment, psychotherapy and medication treatment) when compared with standard care was potentially cost effective in people with AN. This evidence came from a partially applicable study that was characterised by potentially serious methodological limitations.

There was evidence from one AU modelling study showing that the best practice model (early intervention, a range of care from GPs, self-help, intensive outpatient and residential care, inpatient care and stepped care approach) when compared with treatment as usual was dominant in people with AN, BN, BED and EDNOS. This evidence came from a partially applicable study that was characterised by potentially serious methodological limitations.

5.2.5.2 Stepped care

There was evidence from one US study (N=293) showing that stepped care model was dominant when compared with high intensity CBT in people with BN. This evidence came from a partially applicable study that was characterised by minor methodological limitations.

There was evidence from one Finnish study (N=72) showing that stepped care model was potentially cost effective in people with BN. This evidence came from a partially applicable study that was characterised by potentially serious methodological limitations.

There was no economic evidence on the cost effectiveness of stepped care models for people with AN, BED or EDNOS.

5.2.6 Recommendations and link to evidence for the reviews on: Do different ways of coordinating care produce benefits/harms for people with eating disorders? Does the setting (inpatient, outpatient or other specific setting) and different ways of coordinating, transitioning and integrating care for treating eating disorders produce benefits/harms in people with eating disorders?

Improving access to services

	10. Be aware that people with an eating disorder may: • avoid contact with and find it difficult or distressing to interact with healthcare professionals, staff and other service users • be vulnerable to stigma and shame.

	11. Ensure that people with an eating disorder and their parents or carers (as appropriate) get equal access to treatments for eating disorders, regardless of: • gender or gender identity (including people who are transgender) • sexual orientation • religion, belief, culture or family origin • where they live and who they live with • any mental or physical health problems or disabilities.
Relative value of different outcomes	No formal review was conducted to address the barriers and facilitators for accessing treatment.
Trade-off between clinical benefits and harms	No formal review was conducted to address the barriers and facilitators for accessing treatment since this question was outside the scope. However, the committee agreed it was important to include general principles that healthcare professionals should incorporate when treating and managing people with an eating disorder. Moreover, the committee wanted to highlight how people with an eating disorder may feel vulnerable when accessing care and to ensure people of all backgrounds have equal access. The committee used their experience and knowledge to generate a group discussion about the issues, and the recommendations were developed using an informal method of consensus.
Trade-off between net health benefits and resource use	The committee expressed the view that eating disorders cause a significant burden on individual with an eating disorder and their parents or carers and also healthcare system in terms of increased health and social care costs and reduced quality of life. Given that there are effective treatments for eating disorders the committee stressed the importance of improving and ensuring equal access to treatment. It was noted that, for example, in rural areas certain type of therapies may not be possible, such as group therapies. In such cases, access to other effective treatments such as self-help (computer programmes that people can access from their homes) and individual therapies should be facilitated. The committee expressed the view that improving and facilitating equal access to treatments for eating disorders may incur additional resource use (for example, providing individual therapy versus group). However, if this results in timely treatment and management of eating disorders at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with improving and facilitating access to treatments is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	No quality assessment was conducted in the absence of a formal review.
Other considerations	In the absence of a formal review, the committee generated the recommendations based on their clinical and service-user experience. They considered a wide-range of ways that having an eating disorder could interfere with the process of accessing and receiving treatment. The barriers are not a comprehensive list or representative of all people with an eating disorder who may not access treatment. For people with an eating disorder, accessing care may be anxiety provoking and will involve discussing a number of issues, possibly for the first time that they may find humiliating or embarrassing. Mood and anxiety symptoms are very common, in addition to low self-esteem and low confidence. For people with anorexia nervosa, weight loss is experienced as a positive achievement and therefore, they may often deny the seriousness of their condition. Many will only access treatment when they reach crisis point. Typically, individuals will be persuaded to seek treatment by concerned family

	members, teaching staff or general practitioners with whom they consult about physical consequences. In some cases, they will seek treatment in their own right if they begin to see the damaging effects of the disorder. Children and young people rarely seek treatment independently and are often brought to treatment by parents or carers. For these reasons, the committee agreed that it was important that healthcare professionals are aware of the difficulties that people with an eating disorder may have when they seek help and that they understand the person may feel vulnerable and shame regarding their condition. A focus of all NICE guidelines is to ensure there is equal access to services and treatment. Therefore, the committee agreed it was important to highlight groups that may feel marginalised or reluctant to seek treatment. Such groups may include those who: are transgender; with a particular religion or belief; children who are looked after such as in foster care of they have another mental health or physical health problem. This is by no means an exhaustive list. One of the committee service users highlighted the reluctance of men to seek treatment or be considered at risk of an eating disorder by health professionals. They also mentioned people who are transgender may be prone to eating disorders because of the desire to fit the image of a new gender. For example in transgender men, anorexia nervosa may lead to the loss of breast tissue, no more menstrual cycles and a smaller frame. Thus, they are a group who may be particularly vulnerable to eating disorders and that we need to ensure they have equal access to care and treatment.

1 **Referral and coordination of care**

	12. Take particular care to ensure services are well coordinated when: • a young person moves from children's to adult services (see the NICE guideline on transition from children's to adults' services) • more than one service is involved (such as inpatient and outpatient services, or when a comorbidity is being treated by a separate service) • people need care in different places at different times of the year (for example, university students). **13. If an eating disorder is still suspected after the initial assessment, refer without delay to:** • a community based, age-appropriate eating disorders service for an assessment and treatment (if possible) **or** • day patient or inpatient services for people with clinical signs in the concern or alert ranges (see recommendations 36 and 48).
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of different treatment settings or coordinating care for children, young people and adults with an eating disorder. For those with anorexia nervosa, body weight or BMI and remission are of greatest concern. For bulimia nervosa, binge eating and remission are the most critical outcomes. Service user experience is also a critical outcome. For any eating disorder, other outcomes that are important but considered rare

	events, or rarely measured in RCTs for eating disorders, include all-cause mortality, adverse events, quality of life, resource use and relapse. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning and family functioning.
Trade-off between clinical benefits and harms	**Randomised control trials** **Anorexia nervosa:** A number of RCTs were identified that compared the effect of care within in different settings on weight and or remission in young people or adults with anorexia nervosa. Comparing an inpatient versus a day clinic setting for adults, no difference in BMI was found at the end of treatment or in remission rates at the end of treatment or at one year follow up. All other outcomes were similar between the two settings. They included: bingeing, vomiting, EDI-bulimia and global severity index. No data was available for BMI at follow up, or at either time point for family functioning, service user experience, all-cause mortality quality of life, resource use, or general psychopathology. Comparing inpatient treatment with an outpatient psychotherapy group that included individual and family therapy in adults with anorexia nervosa showed no difference in Morgan–Russell global or subscale scores between the two groups at the end of treatment. No data was available for the critical outcomes BMI or remission, or any other important outcomes at the end of treatment or follow up. They included general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. No differences were found between these same outcomes when inpatient treatment was compared with an outpatient group therapy or wait list controls. In young people, follow up data 12 months after admission was available for those who attended an inpatient versus a day clinic setting. The results show that inpatient care is equally effective as a day clinic on remission and BMI. EDI-total, EDI-bulimia and global severity index at follow up. However, relapse or readmission rates are higher in the inpatient-treated group compared with the day clinic. No data was reported at the end of treatment and no follow up data was available for family functioning, service user experience, all-cause mortality, quality of life, resource use, or general psychopathology. A study on young people with anorexia nervosa in the UK showed no difference at one year follow up in remission or BMI in those who attended an inpatient setting versus a specialised outpatient clinic (offered CBT-ED). EDE-total and Morgan–Russell scores were also similar at follow up. No data was reported at the end of treatment and no follow up data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, or general psychopathology. Similar results were found in the same study when comparing inpatients with those randomised to a general outpatient treatment (community child and young people mental health service [CAMHS]). At one year follow up, no difference was found in remission rates or BMI. EDI-total favoured the inpatient arm but no difference was found in the Morgan–Russell score. No data was reported at the end of treatment and no follow up data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, or general psychopathology. When comparing the two outpatient interventions described above (specialised treatment versus general CAMHS treatment), no difference was found in any of the outcomes at follow up, including BMI, remission, EDI- total, Morgan-Russell score and readmission to hospital. No data was reported at the end of treatment and no follow up data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, or general psychopathology. **Bulimia nervosa:** One RCT was identified in adults with bulimia nervosa and showed at 12 to 14 months' follow up, there was no difference in remission or binge eating between

	those who were treated as inpatients and received group psychoanalytical therapy versus those who received outpatient family therapy. Other outcomes were also similar they included vomiting, depression and bulimic severity score. No data was available at the end of treatment and no outcomes were reported for general functioning, family functioning, service user experience, all-cause mortality, quality of life or resource use. Evidence from one study on adults with bulimia nervosa showed no difference at the end of treatment in the effect of specialist outpatient treatment on binges, vomiting, EDE global or subscales, bulimic investigatory test, depression or work/leisure/family life questionnaire compared with GP outpatient care. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality or resource use. **Any eating disorder:** One study randomised participants to a modified day clinic and compared the outcome with those who attended traditional outpatient therapy. At the end of treatment, the results favoured the day clinic. Bingeing episodes, purging episodes, depression, EDE-total, EDI(2)-drive for thinness, EDI(2)-bulimia, EDI(2)-body dissatisfaction were all in favour of the day clinic, except for BMI. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life or resource use. One study compared different durations and aims of inpatient treatment. In one arm the aim was for medical stabilisation (mean 22 days), the other was for weight restoration (mean 38 days). The outcomes were similar. Remission, hospital readmission and change in EDE-global score were the same at the end of treatment. Remission was also similar at long-term follow up. No data was available for family functioning, service user experience, all-cause mortality, quality of life, resource use, general functioning or general psychopathology. **Observational studies** **Anorexia nervosa:** A cohort study comparing outcomes of adults with anorexia nervosa who were in inpatient care with day patient care showed that day patient care was favourable for improving BMI. Results regarding binge eating favoured inpatient care, but otherwise all other outcomes were similar, including laxative use, vomiting, excessive exercise and quality of life. At follow up, no differences were found between BMI (after 6 months) and readmission rates (after 1.7 years). No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. Another cohort study compared the outcomes of people with anorexia nervosa who were admitted to a general admissions unit with those who admitted to an eating disorder unit. At the time of discharge, there was a similar improvement in symptoms using the Morgan–Russell score and Children's Global Assessment Scale (CGAS). However, those in the specialist unit did achieve a higher BMI, but they had a longer hospital stay. No data was available for remission, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. At 4.8 years follow up, another inpatient versus outpatient cohort study in adults showed that remission and BMI were similar. However, hospitalisation rates were higher in the inpatient-treated group. No data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. Inpatient care compared with family therapy also showed less favourable results in adults, with higher readmission rates but no difference for readmission rates of more than three times. No data was available for body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. When comparing inpatient care with a hybrid of outpatient treatments (including day clinic and outpatient care) in young people with anorexia nervosa the findings favoured hybrid treatment since they showed greater gains in body weight. No data was available for remission, general functioning, family functioning, service user

experience, all-cause mortality, quality of life, resource use or general psychopathology.

One study in adults compared partial hospitalisation and support (community housing with counselling and case management) with partial hospitalisation alone. The findings favoured the additional support for improvements in weight gain, EDI (2)-total, EDI (2)-bulimia and EDE-weight concern. No other differences were found in other EDI (2) or EDE-subscales, BMI or purging. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology.

Bulimia nervosa:

One study compared day patient care with inpatient care in adults with bulimia nervosa and showed no difference in the outcomes for remission or global severity, along with depression and the EDI-subscales. Interestingly at 3 years follow up, remission rates favoured the day patient group, while all other outcomes showed no difference between the two arms. They included: bingeing, vomiting, depression, global severity and EDI-bulimia and EDI-drive for thinness. No data was available for family functioning, service user experience, all-cause mortality, quality of life or resource use.

An inpatient study with a mix of people with bulimia nervosa and anorexia nervosa compared a five day versus a four day hospital programme and showed all outcomes favoured the five day treatment. The outcomes included remission, bingeing, BMI, depression, EDI-bulimia, EDI-drive for thinness and EDI-body dissatisfaction. Vomiting also favoured the five day treatment, however there was some uncertainty. No data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life or resource use.

Any eating disorder:

Comparing inpatient with outpatient CAMHS programmes for those with any eating disorder showed at follow up no difference in most outcomes including BMI, EDI-bulimia, EDI-body dissatisfaction, EDI-drive for thinness, and global severity. Self-esteem scores were higher in the inpatient group. No data was available for remission, family functioning, service user experience, all-cause mortality, quality of life or resource use.

A day hospital programme compared with guided self-help showed bingeing improved more in the day hospital programme, while other outcomes were similar including EDE-total, vomiting and excessive exercise. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology.

A study that compared an extensive programme (that included an additional community out-reach programme) with a limited programme (that included combined psychotherapy with nutritional counselling) identified it will improve remission rates in adults with anorexia nervosa and bulimia nervosa at the end of treatment and at follow up. No data was available for body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology.

One study compared the long-term (unclear duration) outcomes of patients who had a history of inpatient care compared with those who had no history and found no difference in scores for EDI-drive for thinness, EDI-body dissatisfaction and EDI-bulimia. No critical outcomes were reported. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology.

One study compared the outcomes in young people with any eating disorder who had progressed through different pathways of care in the UK. Those who were referred and treated in an eating disorder-specialised CAMHS or private eating disorder service (Sp in GRADE) had better outcomes compared with those who were referred and treated in a non-eating disorder specialised CAMHS service or referred to a non-eating disorder specialist CAMHS (Non Sp in GRADE), but then treated at a specialist eating disorder setting. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology.

	Compared with those who stayed in a non-specialised setting, the eating disorder-specialised treated group had better continuity of care and were more likely to receive care from a specialist, but no difference in admission to hospital. Compared with those who ultimately received specialised care, the continuous specialised group showed lower rates of admission to hospital but no difference in continuity of care. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One study investigated the effectiveness of an opt-in intervention on attendance to first appointments. Opt-in systems require the patient to respond in some way to the offer of an appointment. Those who do not respond are ineligible to attend. The results showed after the opt-in programme was introduced it was less effective on ensuring people attended their first appointment. However, although the number of people failing to attend a first appointment was reduced, more people were seen. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One study compared inpatient care with a variety of other care settings (including day, hospital and outpatient) and showed inpatient care was less effective on body weight compared with any other type of care, but there was some uncertainty. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One inpatient study compared the effectiveness of meal supervision versus not in adults with any eating disorder. The results showed the length of hospital stay and weight gain was no different but bradycardia results were better in the meal supervised group. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. The evidence in both young people and adults with anorexia nervosa, bulimia nervosa and any eating disorder clearly showed that inpatient care does not result in better outcomes than for those treated as outpatients.
Trade-off between net health benefits and resource use	The existing economic evidence on the cost effectiveness of specialist eating disorder service is very sparse. The existing UK study indicated that at two years follow up specialist outpatient treatment dominated both inpatient and general outpatient treatment. Also, specialist outpatient treatment had a high probability of being cost effective. The existing limited evidence is characterised by minor methodological limitations. The committee also took into account the psychological and financial burden associated with eating disorders both for people with eating disorders and for their families, as well as the benefits associated with the specialist eating disorder service. The committee considered the substantial costs associated with delayed diagnosis and management of unrecognised eating disorders and recognised that early diagnosis of eating disorders which is most likely to be facilitated by a specialist eating disorder service offers a benefit to the individuals who receive appropriate treatment, and may also result in a considerable reduction in healthcare resource use. Regarding assessment, the committee acknowledged that appropriate assessment of people with eating disorders enables them to receive suitable treatment according to their needs, thus ensuring efficient use of available healthcare resources. Generally, the committee considered that coordinated approach to the management of eating disorders may have resource implications in terms of the extra time required to facilitate such approach to care. However, the committee expressed the view that if such service structures lead to prompt identification of needs and this results in subsequent treatment and management of an eating disorder (and potentially of any comorbidities) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating coordinated approach is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of	**Randomised controlled trials**

evidence	The majority of the evidence was very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as unclear randomisation, lack of clarity on whether allocation concealment was performed and if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm The study sizes were mostly small (fewer than 400 participants or 300 events) and very few studies were available for each comparison, so imprecision was detected in a lot of outcomes. Remission was not always measured, and some studies did not provide data at the end of the treatment, only at follow up. Service user outcomes were not reported either, so the preference of participants was not an outcome that the committee could consider. **Observational studies** The quality of the evidence was all very low quality. In GRADE all observational studies start at very low quality and can only be scored up if the effect size is large, there is a dose-response and the possible effect of confounders have been taken into account. The majority of studies did not adjust the data for potential confounders, and in many cases the cohorts were not matched for factors such as severity of illness. In a lot of cases remission was not measured. The majority of the studies were from outside the UK, so applying the findings to the NHS is difficult and may be considered indirect evidence. For instance, one study in the USA looked at the benefit of adding community housing to people who were partially hospitalised. This is not something the NHS would recommend, even though it showed some benefit. When including studies from other countries, it is important to consider the different pathways of care, who pays for the treatment (insurance versus nationalised health service), the culture of inpatient admissions, the availability of beds and the costs of treatment. To include in the recommendation refer people to a community-based eating disorder service (depending on risk), the committee focused on a study that showed referral of young people to such services results in fewer people being admitted to inpatient care compared with people referred to a general CAMHS services. This study was very low quality because it is an observational study and no adjustments were made to the data, although no statistically significant differences were found between the groups in age, gender, ethnicity or weight for height percentage at the assessment stage or when being referred. Nevertheless, this study was conducted in the UK so the findings and setting are pertinent to this guideline.
Other considerations	**Referral to and delivery of care in a community-based eating disorder service** The committee agreed that the inpatient care does not result in better outcomes compared with those treated as outpatients, and in some instances inpatient care may result in worse outcomes. Community based eating disorder service is considered a type of outpatient care thus the evidence on outpatient care versus inpatient care can be used to justify the recommendation. The committee wanted to be specific and not just offer outpatient care because there was evidence in the UK that showed specialist care results in better outcomes than those who are treated in non-specialist settings. For clarification, community based eating disorder service is not part of the primary care team and clinics take place in the community that offer an assessment, triage and treatment. According to the committee most places in the UK have a community based eating disorder service for adults and most also provide care for young people. Thus, it was recommended that people with any eating disorder be referred to and treated in an age-appropriate community-based (outpatient) specialist eating disorder service. Although some committee members said such places are not available across the whole of the NHS, they wanted to recommend it in the hope that it would improve services across the country. **Background** In the Access and Waiting Time Standard for Children and Young People with an Eating Disorder (2015), the current referral pathways in the NHS are described.

Once an eating disorder has been identified, the first and most common referral pathway is from primary care to local community CAMHS that have varying levels of expertise in eating disorders and the treatments available.

Some community CAMHS have invested in developing eating disorder expertise and have eating disorder mini-teams that are able to offer specialist assessment and treatment. As a result they have the necessary skills to provide full community eating disorder service, but they may be limited in the number of cases they can care for. Some mini-teams provide home treatment or intensive outreach services to support cases in their homes for a limited period of time. A limitation of these mini-teams is that they cover a smaller geographical area compared with larger community eating disorder services.

The alternative route is to refer from primary care, generic CAMHS or an eating disorder mini-team to a community child and young people eating disorder service. The latter are multidisciplinary services that cover a large geographical area. They offer community-based treatment with the possibility of intensive community-based or day patient treatment. Although not many of these services are available across the NHS they are growing in number.

The committee highlighted that referral to general mental health services may have significant implications on the recovery of the person with an eating disorder because they may then need to be referred to a specialist eating disorder service.

The committee raised concerns over whether the recommendation may be misinterpreted as saying that only a specialist eating disorder service would be able to diagnose eating disorders. However, the recommendation is based on evidence and again it describes what should be happening wherever possible. The committee again acknowledged the caveat on the provision and (limited) availability of community-based eating disorder services across the NHS.

The committee emphasised that referral should be without delay and, according to the Access and Waiting Time Standard for Children and Young People with an Eating Disorder (2015), it is recommended that those who require urgent care should begin treatment within one week and those who do not require urgent care should begin treatment within three weeks.

The committee discussed the relevance of the term “without delay”, saying that it highlights to GP’s that they should not “wait and see” if the symptoms progress. If the person is seeing a GP about their eating disorder, it generally means it is bad enough for them to warrant a referral. “Without delay” will also mitigate instances where patients are told that they are not ill enough or that they need to lose more weight before they are eligible for treatment. Also, GPs may delay referral because they think waiting lists are long and they are not sure how serious the eating disorder is. The committee preferred the term without delay instead of immediately because the latter may result in GP’s referring patients the same day or a four hour triage assessment, which is not usually required.

The CAMHS Tier 4 Report (2014) on eating disorders highlights the need for the development of community-based eating disorder services, to reduce the need for admissions and improve service-user outcomes (CAMHS Tier 4 Report Steering Group, 2014).

No evidence for this review was found in children with an eating disorder. Nevertheless, the committee ensured that the recommendation was clear that people should be treated in age-appropriate community-based eating disorder settings.

Coordinated approach

The committee discussed how many people with an eating disorder will require treatment for a long period of time, often a number of years. During this time they may move from a young person’s to an adult service, or have different services involved in the treatment, especially if they have a comorbid condition or are living in different places throughout the year (for example, university students).

Transfer of care is likely to involve a requirement to establish new relationships as well as a shift in treatment approach. Where clear transition protocols are not in place, with adequate preparation for transfer of care, recovery may be hampered.

Based on their experience the committee highlighted that a lack of collaboration

	causes confusion, adds to the burden on children and young people and their parents or carers and has the potential to delay recovery. Therefore, effective collaboration to manage the treatment of the eating disorder, any coexisting mental health or physical health problems and the physical consequences of severe eating disorders is essential. However, the service users on the committee said coordinated care is often not in place and talked about their own experience of long delays in treatment. Although there was no evidence available to generate the recommendation on using a coordinated approach, the committee agreed that it is critical for the interface between different care providers to be managed effectively. This should include good communication, clear lines of responsibility and ensuring a transition protocol is in place.

1 **Communication and information**

	14. When assessing a person with a suspected eating disorder, find out what they and their family members or carers (as appropriate) know about eating disorders and address any misconceptions. **15. Ensure that people with an eating disorder and their parents or carers (as appropriate) understand the purpose of any meetings and the reasons for sharing information about their care with others.** **16. Offer people with eating disorders and their family members or carers (as appropriate) education and information on:** • the nature and risks of their eating disorder and how it is likely to affect them • the treatments available and their likely benefits and limitations. **17. When communicating with people with an eating disorder and their family members or carers (as appropriate):** • check that they understand what is being said • be sensitive when discussing a person's weight and appearance • be aware that family members or carers may feel guilty and responsible for the eating disorder • show empathy, compassion and respect
Relative value of different outcomes	No formal review was conducted to address the barriers and facilitators for accessing treatment.
Trade-off between clinical benefits and harms	No formal review was conducted to address the barriers and facilitators for accessing treatment since it was outside the scope. However, the committee agreed it was important for the guideline to include general principles that healthcare professionals should incorporate when treating and managing people with an eating disorder. The committee used their experience and knowledge to generate a group discussion about the issues and the recommendations were developed using an informal method of consensus.
Trade-off between net health benefits and	The committee expressed the view that offering people with eating disorders and their family members or carer's education and information is an integral part of most eating disorder specific psychological interventions and providing such supplementary advice would not incur significant extra resource implications.

resource use	
Quality of evidence	No quality assessment was conducted in the absence of a formal review. The committee used their knowledge and expertise to generate the recommendations.
Other considerations	Because it is a challenging and significant step for people with an eating disorder to seek help from others, the committee discussed the importance of making sure their experience was positive and met with care, compassion and understanding; as many barriers and triggers as possible should be removed from their pathway to recovery. Otherwise there is a risk that the person will not seek treatment, or will disengage soon after starting. The committee agreed that if people with an eating disorder do not seek help then further problems, such as those commonly comorbid with the disorder, will develop. The committee agreed that effective communication with the person and their family is a big part of ensuring their experience is positive. Prior to giving information to the person with an eating disorder and their parents or carers (if appropriate), healthcare professionals should first establish existing knowledge and take the opportunity to correct any misconceptions. Once existing knowledge has been established, the healthcare professional should then offer the person with the eating disorder and their family or carers information on the nature and risks of their eating disorder and how it may affect them, treatments that are available and their likely benefits and limitations. The risks are pertinent given the physical health problems typically associated with having an eating disorder, such as cardiac problems, delayed growth and development. Communicating with children and young people with an eating disorder and their parents or carers was regarded by the committee as particularly challenging. The committee drew upon the recommendations from the NICE guideline Service User Experience in Adult Mental Health and was mindful that healthcare professionals should take into account the child or young person's developmental level, emotional maturity and cognitive capacity. The use of plain language and the explanation of any clinical terms were considered to be very important, as was the use, where necessary, of communication aids (such as pictures, symbols, braille, different language, interpreters or sign language). One way to ensure the information is appropriate is to allow the person and or their family and carers to ask questions and to encourage them to do. Also explain any clinical terms and simply check that they understand what is being said. Given the person is likely to be self-conscious about their body weight and body image, it is important that sensitivity is used when discussing a person's weight and appearance. It is also likely that the parents or carers feel guilty about the causes and effects of the eating disorder, therefore it is important that is detected and addressed where necessary (see recommendations on working with family and carers). It is important to make sure that the person with the eating disorder has support around them when they are undergoing treatment, to contribute to them having a positive experience of the treatment. In some cases a long-term focus on eating and dieting may lead to social withdrawal and a compromise in time engaged in leisure activities, daily living and productivity in employment and/or education. Thus, healthcare professionals should assess how much support is available and encourage this support network to help where possible during the treatment. This network may extend beyond immediate family to peers, work colleagues and teachers, to name a few. The committee were aware of some observational studies that showed informal, supportive, social relationships outside of treatment and a supportive relationship with a partner are important in the recovery process for people with an eating disorder. It is also important that if, for instance, a child or young person who is in education and receiving treatment misses a day of classes they can be kept up to date with any homework. Given the importance of the role of parents or carers in the treatment of children and young people with eating disorders, especially anorexia nervosa, the committee agreed it was important for them to understand early on the purpose of

	any meetings and the reasons for sharing information about the care with other professionals

1 **Support and treatment for children and young people**

	18. For children and young people, assess the impact of their home, education, work and wider social environment on their eating disorder. Ensure that their emotional, education and social needs are met throughout treatment. **19. If appropriate, encourage family members, carers, teachers, and peers of children and young people to support them during their treatment.** **Psychological treatment for children** **20. For children with an eating disorder, consider using the treatments recommended for young people with the same eating disorder.**
Critical and important outcomes	The committee discussed the importance and relevance of various outcomes for the review on whether setting and different ways of coordinating, transitioning and integrating care for treating eating disorders produce benefits or harms in people with eating disorders. For any eating disorder, remission is of greatest concern and for those with anorexia nervosa body weight or BMI is critical and binge frequency for those with bulimia or binge eating disorder. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, family functioning and service user experience.
Trade off benefits and harms	No relevant published RCT or observational evidence was identified to generate these recommendations. However, the committee agreed that it was important to develop a recommendation focusing on the social and environmental factors that may positively or negatively influence the treatment for children and young people with an eating disorder. The committee used their experience and knowledge to generate a group discussion about the issues, and a recommendation was developed using an informal method of consensus.
Trade-off between net health benefits and resource use	The recommendations 15-17 relate to the principles of care and factors that directly impact on the treatment outcomes for children and young people with an eating disorder. The committee expressed the view that these recommendations may have modest resource implications, which are justifiable as these principles and factors are deemed essential in ensuring the success of treatment. Also, given the lack of evidence on psychological interventions for children with an eating disorder the GC expressed the view that if psychological treatments are cost effective in young people they are also likely to be cost effective in children.
Quality of the evidence	No quality assessment was conducted in the absence of a formal review.
Other considerations	The committee discussed how children and young people may experience a wide range of social and emotional difficulties that may lead to developing an eating disorder. Moreover, the resilience of a child or young person to cope with adversity or stress is affected by their own characteristics, as well as the support they receive from others, the environment they live and learn in and their

	opportunities for positive engagement and success. The committee therefore considered it important that healthcare professionals: • think about how the home, education, work and wider social environment affects a child's or young persons' eating disorder and the treatment they are having. • that the emotional, education and social needs of children and young people are not neglected during their treatment • that a planned and supportive environment is at the heart of supporting children and young people throughout their treatment It was discussed that children and young people with eating disorders may require additional support to improve engagement in learning at school. This may be the result of any number of behaviours including: reduced motivation, social withdrawal or isolation; obsessive behaviours; or emotional distress. Support may come in a variety of ways including modification to curriculum content and delivery, enhanced pastoral support and access to individual or group activities to promote emotional wellbeing and social development. The committee agreed in the absence of evidence on how to treat children with an eating disorder that health care professionals should follow the recommendations for young people with the same eating disorder. Children make up a relatively small proportion of the total number of people with eating disorders, however, they agreed that the treatments for young people should work equally well.

1 **Consent and confidentiality**

	21. When working with people with an eating disorder and their family members or carers (as appropriate): • hold discussions in places where confidentiality, privacy and dignity can be respected • explain the limits of confidentiality (that is, which health and social care professionals have access to information about their care, and when this may be shared with others). **22. When seeking consent for assessments or treatments for children or young people under 16, respect Gillick competence if they do not want their family members or carers involved.**
Relative value of different outcomes	No review was conducted on the concerns of children and young people surrounding confidentiality.
Trade-off between clinical benefits and harms	No evidence was reviewed to develop recommendations surrounding confidentiality when treating children and young people with eating disorder. This was considered outside of the scope, however, the committee agreed that it was important that a recommendation was made on consent and confidentiality because of the high prevalence of eating disorders in young people. The committee used their experience and knowledge to generate a group discussion about the issues, and a recommendation was developed using an informal method of consensus.
Trade-off between net health benefits and resource use	The recommendations 18-19 relate to the principles of care and factors that directly impact on the treatment outcomes for children and young people with an eating disorder. The committee expressed the view that these recommendations may have modest resource implications (for example finding a private room), which are justifiable as these principles and factors are deemed essential in ensuring the success of treatment.
Quality of evidence	No quality assessment was conducted in the absence of a formal review.

Other considerations	The committee discussed how it is often the case when treating children and young people that their parents or carers are involved in the treatment. For this reason they generated recommendations that would ensure discussions are held in places where confidentiality, privacy and dignity can be respected. The committee highlighted that consultation between those with an eating disorder and healthcare professionals are bound to generic rules of confidentiality that should only be breached if the person with an eating disorder or others are at significant risk, and that a breach of confidentiality is likely to reduce that risk (Department of Health, 2001). They also discussed how the person with an eating disorder should be informed of any breach of that confidentiality. Situations may arise where healthcare professionals have a statutory obligation to let other people know of health and safety issues that need to be considered. Therefore, while services should aim to keep treatment confidential, they also need to ensure that the person with an eating disorder's safety is considered. For this reason, the committee agreed it was important to explain the limits of confidentiality. That is, which health and social care professionals have access to information about the person's care, and when this may be shared with other professionals. Some older children and young people may be mature enough to make informed decisions about their own care and might therefore want to discuss and negotiate the extent to which their parents are involved. In such cases, the committee wanted to remind healthcare professionals to consider the child or young person's Gillick competence. That is, a term used in medical law to decide whether a child (16 years or younger) is able to consent to his or her own medical treatment, without the need for parental permission or knowledge. If the child or young person does not want family members or carers involved in their treatment and the Gillick competence has been approved, then it is important that healthcare professionals who are assessing and/or treating the person respect this and are reminded of this when communicating with family member or carers if the child or young person has asked that they not be involved. Some of the recommendation was adapted from the Service User Experience in Adult Mental Health NICE Guideline.

1 **Training and competencies**

	23. Health, social care and education professionals working with children and young people with an eating disorder should be trained and skilled in: • negotiating and working with parents and carers • managing issues around information sharing and confidentiality • safeguarding • working with multidisciplinary teams **24. Professionals who assess and treat eating disorders should be competent to do this for the age groups they care for.** **25. Base the content, structure and duration of psychological treatments on relevant manuals that focus on eating disorders.** **26. Professionals who provide interventions for treating eating disorders should:** • receive appropriate supervision • use standardised outcome measures, for example the

	Eating Disorder Examination Questionnaire (EDE-Q), bulimic behaviours or weight • monitor their competence (for example, by using recordings and external audit and scrutiny) • monitor treatment adherence in people who use their service. **Safeguarding** **27. Healthcare professionals assessing children and young people with eating disorders should be alert throughout assessment and treatment to signs of bullying, teasing, abuse (emotional, physical and sexual) and neglect. For guidance on when to suspect child maltreatment, see the NICE guideline on child maltreatment.**
Relative value of different outcomes	No review was conducted on what the training and competencies should be of healthcare professionals who manage and deliver care for those with an eating disorder.
Trade-off between clinical benefits and harms	No evidence was formally reviewed to develop recommendations on the training and competencies healthcare professionals should have when managing and delivering care for those with an eating disorder people since this question was outside of the scope. However, the committee agreed that a recommendation was needed to ensure a high standard of care is delivered and to increase the likelihood of people recovering from an eating disorder. The committee used their experience and knowledge to generate a group discussion about the issues, and a recommendation was developed using an informal method of consensus.
Trade-off between net health benefits and resource use	The committee expressed the view that training and competency monitoring is routinely undertaken in professionals working with children and young people and offering it in line with the principles outlined in the recommendations 20-24 would not incur significant extra resource implications. The committee expressed the view that the cost of providing training and monitoring/supervision of professional is relatively small, taking into account that it has the potential to significantly change the behaviour of professionals in meaningful and positive ways (for example, improved staff ability to recognise eating disorders through better sharing of information, better ability to communicate with the family and carers and the potential to reduce their burden) and make their overall interactions more efficient when dealing with people with eating disorders, parents, carers and other professionals and as a result lead to timely and appropriate care, improved health outcomes and the overall cost savings to the healthcare system. The committee expressed the view that supervision and monitoring of professionals is essential in ensuring that staff are competent in how to deal with people who have eating disorders.
Quality of evidence	No quality assessment was conducted in the absence of a formal review.
Other considerations	The committee discussed how the involvement of parents and carers in the treatment of eating disorders can be complex. For this reason, staff should receive training in the skills needed to negotiate with parents and carers in managing issues relating to information sharing and confidentiality. There may be instances where the eating disorder has developed in response to neglect or abuse in the family home. For children and young people who need protection, healthcare professionals should be skilled and trained in safeguarding. Safeguarding is defined as: • protecting children from maltreatment • preventing impairment of children's health and development

- ensuring children grow up in circumstances consistent with the provision of safe and effective care
- taking action to ensure all children have the best outcomes.

People with eating disorders who need safeguarding or treatment for a comorbidity are just two examples of instances where healthcare professionals need to communicate well with other services. Thus, the committee agreed it was important that they are trained to work with multidisciplinary teams.

The committee were resolute that professionals who assess and manage people with eating disorders must be competent in delivering interventions to the age group for which they provide care. There are a number of reasons why simply extrapolating care delivered to adults cannot be offered to children and young people. For instance, the goals of treatment may be different because in children it may need to address the completion of puberty and growth. Unlike the treatment of adults with anorexia nervosa where recovery usually involves returning to a health state, in young people it is more about entering early adulthood in a new healthy state. Also, a healthy weight will change as the young person grows. Attention should also be paid to the social and educational needs of young people. Along with providing additional care and support to the carers and parents of young people with an eating disorder. These are just some examples of why it is important that those who deliver care are competent in the relevant age group.

The committee highlighted that every member of an eating disorder team should have an appropriate qualification before delivering the NICE-recommended therapy. They agreed that only evidence-based treatments manuals should be used, since they were concerned that psychotherapies that are not specific to eating disorders may be delivered. The highest quality evidence included in the reviews (on what are the most effective psychotherapies for treating eating disorders) was from studies that used recognised manuals specifically designed to address eating disorders. For more on this please refer to the relevant LETRs.

As in all areas of mental and physical health, the healthcare professional should receive appropriate supervision. Supervision is a requirement in the UK by the British Association for Counselling and Psychotherapy but it is also seen as an ethical imperative. Supervision is needed to protect people receiving therapy and to improve the ability of the therapists to provide care.

Routine measures should be used throughout treatment to monitor progress and success. Measures such as the Eating Disorder Examination Questionnaire (EDE-Q), bulimic behaviours and weight are relatively quick and easy to measure.

A number of people with an eating disorder will be nonadherent at some point during their treatment. People with anorexia nervosa are often described as being ambivalent about seeking treatment. Treatment can lead to a feeling of loss of control (one of the central characteristics of having an eating disorder) and can result in a reluctance to engage fully in the intervention, high levels of treatment refusal and premature dropouts. Thus, prolonging recovery and increasing healthcare costs.

Given the concerns surrounding nonadherence, the committee agreed it is important that healthcare professionals monitor adherence throughout the treatment. The committee said the following methods could be used: recordings, external audits and general scrutiny.

The committee provided some anecdotal evidence where body weight or body mass index (BMI) had been incorrectly used to decide if treatment should be offered (for example, when a person with an eating disorder had a BMI that was considered too low to be offered binge eating disorder treatment). Duration of illness has also been misused as an indicator that the person is unlikely to respond to treatment. The committee therefore agreed that single measures should not be used to decide whether or not to offer treatment and that it should be based on a comprehensive psychological and physical assessment.

The setting for psychological treatment will depend on the medical stability and

	physical health of the person with eating disorder. As described in other recommendations relating to the coordination of care and inpatient care, psychological treatments should be offered in dedicated, age-appropriate, community-based eating disorder services unless physical health is compromised. In such cases, they should be offered inpatient or day patient care. The Access and Waiting Times Standard for Children and Young People with an Eating Disorder (2015) commissioning guide discussed how raising the awareness of professionals in primary care, education and other services will improve early identification of people either at risk of developing or currently experiencing an eating disorder. **Safeguarding** The committee wanted to stress that at any point throughout assessment and treatment, healthcare professionals should be alert to any signs of bullying, teasing, abuse (emotional, physical and sexual) and neglect. In such cases, they should ensure safeguarding of these individuals and refer to the NICE guideline (CG89) on child maltreatment when appropriate.

1 **Pregnancy and eating disorders**

	28. Provide advice and education to women with an eating disorder who plan to conceive, to increase the likelihood of conception and to reduce the risk of miscarriage. This may include information on the importance of: • maintaining good mental health and wellbeing • ensuring adequate nutrient intake and a healthy body weight • stopping behaviours such as bingeing, vomiting, laxatives and excessive exercise. **29. Nominate a dedicated professional (such as a GP or midwife) to monitor and support pregnant women with an eating disorder during pregnancy and in the post-natal period, because of:** • concerns they may have specifically about gaining weight • possible health risks to the mother and child. • the high risk of mental health problems in the perinatal period **30. For guidance on providing advice to pregnant women about healthy eating and feeding their baby, see the NICE guideline on maternal and child nutrition.** **31. Consider more intensive prenatal care for pregnant women with current or remitted anorexia nervosa, to ensure adequate prenatal nutrition and foetal development.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem. In pregnancy, health risks to the mother and child were considered critical outcomes. The other critical outcomes depended on the eating disorder included in the study. Remission is of greatest concern for any eating disorder. For those with anorexia nervosa, body weight or BMI are of greatest concern. For bulimia nervosa and binge eating disorder, binge eating is a

	critical outcome. For any eating disorder, other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	No relevant published RCTs or observational studies were identified on whether treatment for eating disorders need to be modified if the person is pregnant. The committee used their experience and knowledge to generate a group discussion about the issues, and the recommendations were developed using an informal method of consensus.
Trade-off between net health benefits and resource use	The committee considered that pregnancy can be extremely challenging time for women with eating disorders and providing advice and education to women who are thinking of conceiving is crucial. Also, pregnancy can trigger eating disorder flare outs and other severe problems and the use of a prompt multidisciplinary approach to monitor and support women is particularly important. Providing advice and education, and the use of multidisciplinary approach may have resource implications in terms of the extra time required to facilitate these. However, the committee expressed the view that any problems and complications should be treated as soon as possible since if the mother's life is compromised the life of the baby is compromised too. Any delays in care may require very expensive secondary care for mother and baby. The committee expressed the view that if such service structures lead to timely and better identification of health needs and this results in appropriate subsequent treatment and management of underlying health problems and complications at an earlier stage, before women (and potentially the baby, for example complications associated with the small gestational age) require more resource intensive management, then the additional costs associated with facilitating such service structures is expected to result in improved health outcomes in the longer term (both for mother and the baby) and potential future cost savings to the healthcare system.
Quality of evidence	No quality assessment was conducted in the absence of relevant published evidence for this review.
Other considerations	In the absence of relevant published evidence for this review the committee discussed how an eating disorder prior to or during pregnancy may be a cause of concern among women of reproductive age. A growing foetus requires adequate nutrition for normal development and growth and vital nutrients may not be available to the foetus in a woman who binge and subsequently purges, uses laxatives and/or diuretics, fasts and/or engages in excessive exercise before or during pregnancy. Moreover, complications such as oesophagitis, oesophageal and stomach bleeding and ruptures, dehydration, acid-base imbalances and cardiac arrhythmias if occurring during pregnancy have great potential for harming the foetus. The committee highlighted that advice and education to women with an eating disorder who are thinking of conceiving is important because in early stages of pregnancy the mother may be depriving herself of vital nutrients. Moreover, for women with irregular menstrual cycles, it is important for them to understand how having a regular menstrual cycle can help aid conception. For women with anorexia nervosa, regular menstrual cycles can be achieved by restoring body weight to healthy levels. For these reasons it is unusual for women with anorexia nervosa to conceive, although some may seek fertility treatment. Women with an eating disorder are overly concerned with body weight and body image, thus weight gain and shape changes during pregnancy may not be well accepted and may increase the risk of compensatory behaviours during pregnancy. Given that pre-pregnancy weight and maternal weight gain during pregnancy are the best predictors of infant birth weight and birth outcome, it was recommended that more advice is sought from the NICE guideline on maternal and child nutrition. This guideline was also considered helpful for providing guidance on

	feeding the baby A multidisciplinary team is critical for the care of person with an eating disorder who is pregnant because the aetiology is more complex than simply an obsession with body weight and body image. Open and active communication among all members of the team (including the person who is pregnant and has an eating disorder, the obstetrician, dietician, psychologist, psychiatrist and others) is important. Women with anorexia nervosa are at greater risk of premature offspring and those that are small for gestational age. There is evidence from case-series that women with anorexia nervosa also have difficulty feeding their children and that the child's growth can be abnormal. For these reasons, the committee agreed that it was important to recommend that pregnant women with anorexia nervosa may need more intensive prenatal care to ensure adequate prenatal nutrition and foetal development. The committee discussed the length of time that may be needed to monitor the mother. In the absence of evidence it was decided to not make a specific recommendation although there was some suggestion one year after the birth of the baby may be sufficient.

1 **Medication risk management**

	32. When prescribing medication for people with an eating disorder and comorbid mental or physical health conditions, take into account the impact malnutrition and compensatory behaviours can have on the effectiveness and the risk of side effects. **33. When prescribing for people with an eating disorder and a comorbidity, assess how the eating disorder will affect medication adherence (for example, for medication that can affect body weight).** **34. When prescribing for people with an eating disorder, take account of the risks of medication that can compromise physical health because of prexisting medical complications.** **35. Offer ECG monitoring for people with an eating disorder who are taking medication that can compromise cardiac functioning (for example, bradycardia below 50 beats per minute or a prolonged QT interval).**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of treating people with an eating disorder and a comorbidity. For binge eating disorder and bulimia nervosa, it was agreed binge eating frequency and remission are of greatest concern. For anorexia nervosa, body weight/BMI and remission are critical and for OSFED, remission and either binge eating or body weight/BMI depending on the eating disorder they most closely resemble. The other outcomes that are critical are the primary outcomes that are relevant to the physical or mental health comorbidity being treated. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, family functioning and service user experience.
Trade-off between	No relevant published RCT or observational evidence was identified where they treated people with an eating disorder and a comorbid condition with medication.

clinical benefits and harms	The committee used their experience and knowledge to generate a group discussion about the issues, and the recommendations were developed using an informal method of consensus.
Trade-off between net health benefits and resource use	There was no existing economic evidence on the costs and benefits associated with the medication risk management strategies. Medication risk management in line with the principles outlined in the recommendations may incur additional resources. However, the committee noted the importance of the appropriate use of the medication given that people with eating disorders have a high rate of comorbid health problems (such as cardiovascular problems, osteoporosis, kidney dysfunction, etc.) and this in turn may affect how medications work. Fragmented and inappropriate prescribing, sub-optimal dosing and poor adherence, affects health outcomes and overall healthcare costs (for example, sub-optimal dosing leads to poorer outcomes, which then increase healthcare utilisation and overall healthcare costs). Overall, the committee expressed the view that if medication risk management results in appropriate treatment, before individuals require more resource intensive care, then the additional costs associated with facilitating such service structures is expected to result in improved health outcomes (for example, prevention of liver or kidney damage) in the longer term and potential future cost savings to the healthcare system. Similarly, offering regular ECG tests for people with eating disorders who are taking medication that can compromise their cardiac functioning may incur additional resources. However, the committee expressed the view that the cost of EGC monitoring is very small relative to the costs associated with managing future cardiac problem (expensive secondary care and high cost surgical interventions).
Quality of evidence	In the absence of relevant published evidence no quality assessment was conducted.
Other considerations	Although medication is not being recommended for the sole treatment of any eating disorder, people with severe eating disorders have elevated rates of physical illnesses, psychiatric disorders and suicide. Thus, they may need medication for the treatment of morbid mental and physical health conditions such as antidepressants, antipsychotics and treatments for gastroenterological problems However, when medication is used to treat people with severe eating disorders, the side effects of the drugs (in particular, cardiac side effects) should be carefully considered because of the compromised cardiovascular function of many people with anorexia nervosa. For example, extremely malnourished people and those with electrolyte abnormalities are at risk of cardiac complications. Reported mortality rates of anorexia nervosa a highly variable and typically range from 5 to 20%. This risk appears to increase the longer the person suffers from the eating disorder. Causes of death range from suicide to sudden death. Case studies on people with anorexia nervosa who died suddenly showed prolonged QT intervals (the time between the start of the heart's electrical Q wave and the end of the T wave) in the electrocardiogram (ECGs) days before death occurred. Ventricular tachyarrhythmias have also been associated with prolonged QT intervals. These extended QT intervals may be the result epileptic and non-epileptic seizures associated with poor nutrition and malnutrition that result from starvation and liquid-protein diets in anorexia nervosa. Thus, these people are at risk for arrhythmia-related syncopal (fainting) attacks and sudden death. Healthcare professionals should be aware of the risk of drugs that prolong the QT interval on the ECG; for example, antipsychotics, tricyclic antidepressants, macrolide antibiotics and some antihistamines. (Antipsychotic drugs or antihistamines are frequently used to symptomatically to reduce the high levels of anxiety). If these medications are prescribed to people with severe anorexia nervosa, the committee agreed that ECGs should be offered. The committee also mentioned that other medications that can compromise physical health should also be taken into account before prescribing to person with a severe eating disorder whose health is already compromised. Other concerns surrounding prescribing medication for people with an eating disorder is that compromised nutritional status may affect the mechanism of drug

action and is rarely considered in studies. For example, there is some evidence that antidepressants may be less effective if the person has low oestrogen levels and if tryptophan levels are altered.

Compensatory behaviours may also affect the effectiveness of the medication. Starvation, vomiting, dehydration and over-hydration may influence pharmacokinetics (drug absorption and toxicity). For example, if someone vomits soon after taking medication, absorption is reduced. For this reason, the committee agreed that when prescribing medication it is important to consider how malnutrition and compensatory behaviours can affect the effectiveness of the medication, in addition to the side-effects as discussed above.

Medication adherence can also be problem in people with anorexia nervosa given the concerns they may have with weight gain. This may increase the desire for additional control of eating and weight and shape. Thus, the committee agreed that it was important for the health professionals to be aware of which medications will affect medication adherence.

An additional concern raised was that people with bulimia nervosa are at risk of self-harm and so risks of overdose need to be considered.

Very few drugs are recommended for children and young people under 18 years old.

1 **Health monitoring of all eating disorders**

36. GPs should assess fluid and electrolyte balance in people with an eating disorder who are using compensatory behaviours, such as vomiting, taking laxatives or diuretics, or water or salt loading.

37. GPs, paediatricians or psychiatrists should think about the need for acute medical care (including emergency admission) for people with severe electrolyte imbalance, dehydration or signs of incipient organ failure.

38. For people with continued unexplained electrolyte imbalance, GPs, eating disorder specialists, paediatricians or dieticians should assess whether it could be caused by another condition.

39. For people who need supplements to restore electrolyte balance, GPs, eating disorder specialists or dieticians should offer these orally unless the person has problems with gastrointestinal absorption.

40. GPs, eating disorder specialists, paediatricians, psychiatrists or cardiologists should assess whether ECG monitoring is needed, based on the following risk factors:

- rapid weight loss
- excessive exercise
- severe purging behaviours, such as laxative or diuretic use or vomiting
- bradycardia
- hypotension
- excessive caffeine (including from energy drinks)
- prescribed or non-prescribed medications
- muscular weakening
- electrolyte imbalance

	• previous abnormal heart rhythm. **41. GPs, eating disorder specialists or dieticians should encourage people who are vomiting to:** • have regular dental and medical reviews • avoid brushing teeth immediately after vomiting • rinse with non-acid mouthwash after vomiting • avoid highly acidic foods and drinks. **42. GPs, eating disorder specialists or dieticians should advise people who are misusing laxatives:** • that laxatives do not reduce calorie absorption and so do not help with weight loss. • to gradually reduce and stop laxative use. **43. For guidance on identifying, assessing and managing overweight and obesity, see the NICE guideline on obesity.**
Relative value of different outcomes	For the review on how to manage, treat or reduce the short and long-term physical health conditions associated with eating disorders, the committee agreed that the critical outcomes will depend on the health condition under review. Other outcomes that are important include quality of life, weight or BMI, compensatory behaviours, side-effects, remission and service user experience
Trade-off between clinical benefits and harms	No relevant published RCT or observational evidence was identified on how to manage, treat or reduce the short and long-term physical health conditions associated with eating disorders. The committee used their experience and knowledge to generate a group discussion about the issues, and the recommendations were developed using an informal method of consensus.
Trade-off between net health benefits and resource use	There was no existing economic evidence on the costs and benefits associated with health monitoring strategies of people with eating disorders. Health monitoring in line with the principles outlined in the recommendations may incur additional health care resources. However, the committee noted the importance of close health monitoring of people with eating disorders given that they have a very high rate of comorbid health problems (such as cardiovascular problems, osteoporosis, kidney dysfunction, etc.). The committee expressed the view that a large proportion of people with eating disorders die due to the cardio-metabolic risk and this is partially due to poor health monitoring. The committee considered that the costs of ECT, electrolytes tests, etc., are very low compared with the consequences of the potential health complications such as, kidney and liver damage and cardiovascular problems. Overall, if such careful monitoring leads to better identification of health needs and this results in timely and appropriate subsequent medical intervention for the underlying health problem at an earlier stage, before an individual requires more resource intensive management, then the additional costs associated with facilitating such health monitoring is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system. The committee noted that the aim of health monitoring should be prevention of the complications and not the treatment of accumulated problems that require very expensive multidisciplinary management.
Quality of evidence	In the absence of relevant published evidence no quality assessment was conducted.
Other considerations	The recommendations are based on good practice but provide clear advice on electrolyte imbalance and dehydration, ECG monitoring, dental care, laxative misuse and when to refer to other NICE guidelines.

Fluid and electrolyte imbalance is detected in approximately 10% of those with bulimia nervosa. It is often the result of laxative and/or diuretics use or water or salt loading, and may be detected with routine screening. For this reason, the committee recommended that fluid and electrolyte balance is assessed in those who are using such compensatory behaviours.

The concern of fluid and electrolyte imbalance is that it may lead to metabolic alkalosis and is generally accompanied by hypochloraemia and hypokalaemia. Metabolic alkalosis can also occur in people who are abusing laxatives as a result of the loss of bicarbonate from the bowel. Less often, hyponatraemia, hypocalcaemia, hypophosphataemia and hypomagnesaemia may develop. Constipation is also common, mainly due to dehydration. Abnormal electroencephalographic (EEG) may also be found as a result of fluid and electrolyte abnormalities. Causes of death in people with anorexia nervosa have been attributed to dehydration, electrolyte imbalance (particularly hypokalaemia) and metabolic complications.

When electrolyte imbalance is detected, it is usually sufficient to focus on eliminating the behaviour responsible. However, in people with severe electrolyte imbalance or dehydration or where there are signs of incipient organ failure the committee recommended that health professional consider the need for acute medical care, including emergency admission.

The committee also highlighted that electrolyte imbalance may not necessarily be due to vomiting, taking laxatives or diuretics, or water or salt loading. In such cases, it was recommended that health professional assess whether it is caused by another condition.

Because of the risk of death due to cardiac complications in people with anorexia nervosa, the committee agreed that health professionals should consider whether ECG monitoring is needed based on a number of risk factors including: rapid weight loss, excessive exercise, severe purging behaviours, bradycardia, hypotension, muscular weakening, electrolyte imbalance, and previous abnormal heart rhythm.

The committee discussed what rapid weight loss might be defined as, for instance 1kg or more per week. But this was controversial and potentially complicated because it may depend on those who already very underweight. So it was decided to not be specific about how much weight loss. The committee also questioned over how long ECG monitoring is needed but it was decided to best leave it up to the expertise of the health professional.

Because of the risk associated with vomiting on: 1) erosion of tooth enamel potentially leading to destruction of the whole dentition; 2) tooth pain and 3) having unattractive teeth that will also affect self-esteem, the committee agreed that it was important to include recommendations to reduce the acidic environment in the mouth by having regular dental reviews, avoid brushing after vomiting, rinse with non-acidic mouthwash and avoiding highly acidic foods and drinks.

Because of the risks discussed above on laxative misuse and the misconception associated with laxative misuse the committee suggested that people are advised that: laxatives do not reduce calorie absorption and that they should gradually reduce and stop their use.

The committee highlighted that it is important to note that people who have an eating disorder may not be as easy to engage with or have them comply with physical health monitoring so it is important clinicians are aware of this.

There was discussion on the need to review the literature on the treatment of binge eating disorder in presence of obesity in the context of it being a barrier to bariatric surgery. However, this was considered outside of the scope of the guideline. A committee member said there was evidence that bariatric surgery reduces binge eating and that it could be considered as a treatment for binge eating disorder in the presence of obesity. Without reviewing this evidence it was considered not something the group could recommend.

Committee members also discussed the complexities of treating obesity in people with eating disorder. There is the possibility that treatments may be less, equally or more effective (for example bariatric surgery may be equally effective in people with binge eating disorders, but behavioural weight loss may be less effective), but

	without reviewing moderators on the response to treatment, the committee were not able to be explicit on this and it was out of the scope. It was decided to refer to NICE guidance on identifying, assessing and managing overweight and obesity.

1 **Health monitoring for anorexia nervosa**

	44. GPs should offer a physical and mental health review at least annually to people with anorexia nervosa who are not receiving ongoing treatment for their eating disorder. The review should include: • weight or BMI • blood pressure • relevant blood tests • mood • any problems with daily functioning • assessment of risk (related to both physical and mental health) • an ECG, for people with purging behaviours and/or significant weight changes • discussion of treatment options. **45. Monitor physical and mental health (including weight and indicators of increased risk) in people who are having psychological interventions for anorexia nervosa.**
Relative value of different outcomes	For the review on how to manage, treat or reduce the short and long-term physical health conditions associated with eating disorders, the committee agreed that the critical outcomes will depend on the health condition under review. For treating delayed physical development or stunted growth, the committee agreed that the critical outcome is growth. Other outcomes that are important include quality of life, weight or BMI, compensatory behaviours, side-effects, resumption of menses, remission and service user experience.
Trade-off between clinical benefits and harms	No relevant published RCT or observational evidence was identified on how to manage, treat or reduce the short and long-term physical health conditions associated with eating disorders. The committee used their experience and knowledge to generate a group discussion about the issues, and the recommendations were developed using an informal method of consensus.
Trade-off between net health benefits and resource use	There was no existing economic evidence on the costs and benefits associated with health monitoring strategies of people with anorexia nervosa. Health monitoring in line with the principles outlined in the recommendations may incur additional health care resources. However, the committee noted the importance of close health monitoring of people with eating disorders and in particular people with anorexia nervosa given that they have a very high rate of comorbid health problems (such as cardiovascular problems, osteoporosis, kidney dysfunction, etc.). The committee expressed the view that most people with eating disorders die due to the cardio-metabolic risk and this is partially due to poor health monitoring. The committee considered that the costs of ECT, electrolytes tests, etc., are very low compared with the consequences of the potential health complications, such as, kidney and liver damage and cardiovascular problems. Overall, if such careful monitoring leads to better identification of health needs and this results in timely and appropriate subsequent medical intervention for the underlying health problem at an earlier stage, before individual requires more resource intensive

	management, then the additional costs associated with facilitating such health monitoring is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system. The committee noted that the aim of health monitoring should be prevention of the complications and not the treatment of accumulated problems that require very expensive multidisciplinary management.
Quality of evidence	In the absence of relevant published evidence no quality assessment was conducted.
Other considerations	The committee discussed how people with anorexia nervosa who are not currently receiving psychological treatment should be offered an annual review with a general practitioner. This is to ensure that either their long-term recovery is on track or that they are not relapsing. The committee listed parameters that should be taken into account in the review, including BMI or weight, blood pressure, bloods, mood, impairment of daily functioning, assessment of physical and mental health risk, an ECG and possible treatment options if needed. It was decided to separate the two recommendations, although similar, because the committee wanted to be clear that the GP should take responsibility for those who are no longer receiving active treatment. While for those who are undergoing treatment, a physical and mental health review should be part of the on-going treatment and management. It will be part of their original assessment but it should also be reviewed over the course of treatment. In the treatment plan, it must be clear who is responsible for the physical assessment and how any risk identified is to be managed. This often requires effective communication between primary and secondary or tertiary care services.

1 **Inpatient care**

	46. For people with an eating disorder and compromised physical health, consider inpatient treatment or appropriate day patient care for medical stabilisation and to initiate refeeding if these cannot be done in an outpatient setting. **47. Children and young people with an eating disorder who need inpatient treatment or day patient care should be admitted to age-appropriate facilities that are as near to their home as possible and that have the capacity to provide appropriate educational activities.** **48. For people with acute mental health risk (such as suicide risk), consider psychiatric crisis care or inpatient treatment** **49. When deciding whether to use day-patient or inpatient care, take the following into account:** • the person's BMI or weight, and whether either of these are below the safe range and rapidly dropping (for example more than 1 kg per week; be aware that there is no absolute weight or BMI threshold for admission) • whether several medical risk parameters (such as blood tests, physical observations and ECG [for example bradycardia below 50 beats per minute or a prolonged QT interval]) have values and/or rates of change in the concern or alert ranges (refer to Box 1 in MARSIPAN or Guidance 1 and 2 in junior MARSIPAN). • the person's current physical health and whether this is

	declining • whether the parents or carers of children and young people can support them and keep them from significant harm. **50. If a person is admitted for physical health problems caused by an eating disorder, start or continue psychological treatments for the eating disorder if appropriate.** **51. Do not use inpatient care solely to provide psychological treatment for eating disorders.** **52. Inpatient services should collaborate with other teams (including the community team) and the person's family members or carers (as appropriate), to help with treatment and transition.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of different treatment settings or coordinating care for children, young people and adults with an eating disorder. For those with anorexia nervosa, body weight or BMI and remission are of greatest concern. For bulimia nervosa, binge eating and remission are the most critical outcomes. Service user experience would also be a critical outcome. For any eating disorder, other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning and family functioning.
Trade-off between clinical benefits and harms	**Randomised control trials** **Anorexia nervosa:** A number of RCTs were identified that compared the effect of care within in different settings on weight and or remission in young people or adults with anorexia nervosa. Comparing an inpatient versus a day clinic setting for adults, no difference in BMI was found at the end of treatment or in remission rates at the end of treatment or at follow up. All other outcomes were similar between the two settings. They included: bingeing, vomiting, EDI-bulimia and global severity index. No data was available for BMI at follow up, or at either time point for family functioning, service user experience, all-cause mortality quality of life, resource use, or general psychopathology. Comparing inpatient treatment with an outpatient psychotherapy group that included individual and family therapy in adults with anorexia nervosa showed no difference in Morgan–Russell global or subscale scores between the two groups at the end of treatment. No data was available for the critical outcomes BMI or remission, or any other important outcomes at the end of treatment or follow up. They included general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. No differences were found between these same outcomes when inpatient treatment was compared with an outpatient group therapy or wait list controls. In young people, follow up data 12 months after admission was available for those who attended an inpatient versus a day clinic setting. The results show that inpatient care is equally effective as a day clinic on remission and BMI. EDI-total, EDI-bulimia and global severity index at follow up. However, relapse or readmission rates are higher in the inpatient-treated group compared with the day clinic. No data was reported at the end of treatment and no follow up data was available for family functioning, service user experience, all-cause mortality, quality

of life, resource use, or general psychopathology.

Another study on young people with anorexia nervosa in the UK showed no difference at follow up in remission or BMI in those who attended an inpatient setting versus a specialised outpatient clinic (offered CBT-ED). EDE-total and Morgan–Russell scores were also similar at follow up. No data was reported at the end of treatment and no follow up data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, or general psychopathology.

Similar results were found in the same study when comparing inpatients with those randomised to a general outpatient treatment (community child and young people mental health service [CAMHS]). At follow up, no difference was found in remission rates or BMI. EDI-total favoured the inpatient arm but no difference was found in the Morgan–Russell score. No data was reported at the end of treatment and no follow up data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, or general psychopathology.

When comparing the two outpatient interventions described above (specialised treatment versus general CAMHS treatment), no difference was found in any of the outcomes at follow up, including BMI, remission, EDI- total, Morgan-Russell score and readmission to hospital. No data was reported at the end of treatment and no follow up data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, or general psychopathology.

Bulimia nervosa:

One RCT was identified in adults with bulimia nervosa and showed at 12 to 14 months' follow up, there was no difference in remission or binge eating between those who were treated as inpatients and received group psychoanalytical therapy versus those who received outpatient family therapy. Other outcomes were also similar they included vomiting, depression and bulimic severity score. No data was available at the end of treatment and no outcomes were reported for general functioning, family functioning, service user experience, all-cause mortality, quality of life or resource use.

Evidence from one study on adults with bulimia nervosa showed no difference at the end of treatment in the effect of specialist outpatient treatment on binges, vomiting, EDE global or subscales, bulimic investigatory test, depression or work/leisure/family life questionnaire compared with GP outpatient care. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality or resource use.

Any eating disorder:

One study randomised participants to a modified day clinic and compared the outcome with those who attended traditional outpatient therapy. At the end of treatment, the results favoured the day clinic. Bingeing episodes, purging episodes, depression, EDE-total, EDI(2)-drive for thinness, EDI(2)-bulimia, EDI(2)-body dissatisfaction were all in favour of the day clinic, except for BMI. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life or resource use.

One study compared different durations and aims of inpatient treatment. In one arm the aim was for medical stabilisation (mean 22 days), the other was for weight restoration (mean 38 days). The outcomes were similar. Remission, hospital readmission and change in EDE-global score were the same at the end of treatment. Remission was also similar at long-term follow up. No data was available for family functioning, service user experience, all-cause mortality, quality of life, resource use, general functioning or general psychopathology.

Observational studies

Anorexia nervosa:

A cohort study comparing outcomes of adults with anorexia nervosa who were in inpatient care with day patient care showed that day patient care was favourable for improving BMI. Results regarding binge eating favoured inpatient care, but otherwise all other outcomes were similar, including laxative use, vomiting,

	excessive exercise and quality of life. At long-term follow up, no differences were found between binge eating and readmission rates. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. Another cohort study compared the outcomes of people with anorexia nervosa who were admitted to a general admissions unit with those who admitted to an eating disorder unit. At the time of discharge, there was a similar improvement in symptoms using the Morgan–Russell score and Children's Global Assessment Scale (CGAS). However, those in the specialist unit did achieve a higher BMI, but they had a longer hospital stay. No data was available for remission, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. At follow up, another inpatient versus outpatient cohort study in adults showed that remission and BMI were similar. However, hospitalisation rates were higher in the inpatient-treated group. No data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. Inpatient care compared with family therapy also showed less favourable results in adults, with higher readmission rates but no difference for readmission rates of more than three times. No data was available for body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. When comparing inpatient care with a hybrid of outpatient treatments (including day clinic and outpatient care) in young people with anorexia nervosa the findings favoured hybrid treatment since they showed greater gains in body weight. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One study in adults compared partial hospitalisation and support (community housing with counselling and case management) with partial hospitalisation alone. The findings favoured the additional support for improvements in weight gain, EDI(2)-total, EDI(2)-bulimia and EDE-weight concern. No other differences were found in other EDI(2) or EDE-subscales, BMI or purging. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. **Bulimia nervosa:** One study compared day patient care with inpatient care in adults with bulimia nervosa and showed no difference in the outcomes for remission or global severity, along with depression and the EDI-subscales. Interestingly at follow up, remission rates favoured the day patient group, while all other outcomes showed no difference between the two arms. They included: bingeing, vomiting, depression, global severity and EDI-bulimia and EDI-drive for thinness. No data was available for family functioning, service user experience, all-cause mortality, quality of life or resource use. An inpatient study with a mix of people with bulimia nervosa and anorexia nervosa compared a five day versus a four day hospital programme and showed all outcomes favoured the five day treatment. The outcomes included remission, bingeing, BMI, depression, EDI-bulimia, EDI-drive for thinness and EDI-body dissatisfaction. Vomiting also favoured the five day treatment, however there was some uncertainty. No data was available for general functioning, family functioning, service user experience, all-cause mortality, quality of life or resource use. **Any eating disorder:** Comparing inpatient with outpatient CAMHS programmes for those with any eating disorder showed at follow up no difference in most outcomes including BMI, EDI-bulimia, EDI-body dissatisfaction, EDI-drive for thinness, and global severity. Self-esteem scores were higher in the inpatient group. No data was available for remission, family functioning, service user experience, all-cause mortality, quality of life or resource use. A day hospital programme compared with guided self-help showed bingeing improved more in the day hospital programme, while other outcomes were similar

	including EDE-total, vomiting and excessive exercise. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. A study that compared an extensive programme (that included an additional community out-reach programme) with a limited programme (that included combined psychotherapy with nutritional counselling) identified it will improve remission rates in adults with anorexia nervosa and bulimia nervosa at the end of treatment and at follow up. No data was available for body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One study compared the long-term outcomes of patients who had a history of inpatient care compared with those who had no history and found no difference in scores for EDI-drive for thinness, EDI-body dissatisfaction and EDI-bulimia. No critical outcomes were reported. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One study compared the outcomes in young people with any eating disorder who had progressed through different pathways of care in the UK. Those who were referred and treated in an eating disorder-specialised CAMHS or private eating disorder service (Sp in GRADE) had better outcomes compared with those who were referred and treated in a non-eating disorder specialised CAMHS service or referred to a non-eating disorder specialist CAMHS (Non Sp in GRADE), but then treated at a specialist eating disorder setting. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. Compared with those who stayed in a non-specialised setting, the eating disorder-specialised treated group had better continuity of care and were more likely to receive care from a specialist, but no difference in admission to hospital. Compared with those who ultimately received specialised care, the continuous specialised group showed lower rates of admission to hospital but no difference in continuity of care. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One study investigated the effectiveness of an opt-in intervention on attendance to first appointments. Opt-in systems require the patient to respond in some way to the offer of an appointment. Those who do not respond are ineligible to attend. The results showed after the opt-in programme was introduced it was less effective on ensuring people attended their first appointment. However, although the number of people failing to attend a first appointment was reduced, more people were seen. No data was available for remission, body weight, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One study compared inpatient care with a variety of other care settings (including day, hospital and outpatient) and showed inpatient care was less effective on body weight compared with any other type of care, but there was some uncertainty. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. One inpatient study compared the effectiveness of meal supervision versus not in adults with any eating disorder. The results showed the length of hospital stay and weight gain was no different but bradycardia results were better in the meal supervised group. No data was available for remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or general psychopathology. The evidence in both young people and adults with anorexia nervosa, bulimia nervosa and any eating disorder clearly showed that inpatient care does not result in better outcomes than for those treated as outpatients.
Trade-off between net health	The limited economic evidence from the UK suggests that specialist outpatient treatment dominates (that is, results in better outcomes and lower costs) when compared with inpatient psychiatric treatment and general outpatient treatment in

benefits and resource use	young people with anorexia nervosa. The cost per admitted individual to the adult specialist eating disorder service is £450.82 per day versus £182.91 and £185.58 per community contact and outpatient attendance, respectively (DoH, 2015). The committee considered clinical benefits and high costs associated with the inpatient care and expressed the view that inpatient treatment should only be used for medical stabilisation and initiation of refeeding and that it should not be used solely for the psychological treatment of eating disorders.
Quality of evidence	**Randomised controlled trials.** The majority of the evidence was very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as unclear randomisation, it was unclear if allocation concealment was performed and if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm The study sizes were mostly small (fewer than 400 participants or 300 events) and very few studies were available for each comparison so imprecision was detected in a lot of outcomes. Remission was not always measured and in some studies they did not provide data at the end of the treatment only at follow up. Service user outcomes were not reported either, thus the preference of the person receiving treatment was not an outcome that the committee could consider. **Observational studies** The quality of the evidence was all very low quality. In GRADE, all observational studies start at very low quality and can only be scored up if the effect size is large, there is a dose-response and the possible effect of confounders have been taken into account. The majority of studies did not adjust the data for potential confounders, and in many cases the cohorts were not matched for factors such as severity of illness. In a lot of cases remission was not measured. The majority of the studies were from outside the UK, so applying the findings to the NHS is difficult and may be considered indirect evidence. For instance, one study in the USA looked at the benefit of adding community housing to those who are partially hospitalised. This is not something the NHS would recommend even though it showed some benefit. It is important to consider when including studies from other countries the different pathways of care, who pays for the treatment (insurance versus nationalised health service), the culture of inpatient admissions, the availability of beds and the costs of treatment.
Other considerations	The committee had an extensive discussion about the role, importance and effectiveness of inpatient care. A number of committee members agreed that inpatient care may be effective for some people with a severe eating disorder. For example, those who are unresponsive to outpatient care may need more intensive inpatient treatment. However, the evidence did not support such a recommendation. The RCT and observational evidence generally showed no difference in the outcomes in young people or adults with anorexia nervosa or bulimia nervosa who are treated as inpatients compared with outpatient treatment whether it be day patient or community based CBT-ED, group or family-based therapy. In some cases, remission rates may be lower if treated as an inpatient compared with a day clinic. Finally, the costs of inpatient treatment compared with an outpatient setting do not justify such a recommendation. The committee agreed the only time inpatient care is a viable option for people with an eating disorder is if their physical health is compromised and inpatient treatment is needed for medical stabilisation and refeeding. Additionally, only if this cannot be achieved in an outpatient setting. It is important that if children and young people are admitted to inpatient care, that it is to an age-appropriate facility that has the capacity to provide appropriate educational and related activities, especially if they are admitted for a number of weeks or more. In the case of suicide risk or other acute mental health risks, the committee recommended what is considered appropriate treatment for anyone with any mental health problem, not just an eating disorder. The committee agreed that it was important to list some of the most important

	factors to consider when deciding whether to use day-patient or in-patient care. They included outcomes that show: a BMI or body weight below a safe range; blood tests, physical observations or ECG results in the alert range; an overall ill health or rapid decline to ill health; and if the parents or carers cannot support the child or young person. The committee agreed that it was important to pay attention not just to a single measure but measures over time that indicate a rapid decline. Paying attention to the latter may ensure the person gets help early before they slip into a critical state. The committee discussed how people admitted for inpatient care, especially those who are in for a number of weeks or more, should start or continue with psychological therapy is possible. This should be offered in conjunction with the physical health treatment and not as a sole treatment alone in hospital. To ensure ongoing psychological treatment the inpatient services should collaborate with other teams, such as the community based eating disorder services and the person's parents or carers.

1 **Discharge with an appropriate care plan**

	53. Make a care plan for each person with an eating disorder, to cover the care they need after discharge. **54. Within one month of admission, review with the referring team, the person with an eating disorder and their parents or carers (as appropriate) whether inpatient care should be continued, stepped down to a less intensive setting, or stopped.** **55. As part of the review:** • assess whether enough progress has been made towards the goals agreed at admission (such as medical progress) • take into account the risk that people with an eating disorder can become institutionalised, and that a lack of change in their condition could indicate that inpatient treatment is harmful • consider seeking an independent second opinion. **56. Reaching a healthy weight should not be used as the only reason for discharging people with an eating disorder.**
Relative value of different outcomes	No review was conducted to consider what the optimal discharge plan should be if a person is admitted for inpatient care. The recommendation was generated out of the committee experience and expertise.
Trade-off between clinical benefits and harms	No evidence was formally reviewed to develop recommendations on what the appropriate discharge plan should be for a person with an eating disorder who has been admitted to hospital since this question is outside the scope. However, the committee agreed that a recommendation is needed to ensure people are not kept in hospital longer than they should be and are discharged only when appropriate. The committee used their experience and knowledge to generate a group discussion about the issues and the recommendations were developed using an informal method of consensus.
Trade-off between net health benefits and	There was no evidence on the cost effectiveness of care plan arrangements for people with eating disorders. However, the committee expressed the view that if such care plans lead to a timely identification of relapse and appropriate care then the additional costs associated with facilitating such service structures are likely to

resource use	be outweighed by the improvements in the health outcomes in the longer term and potential future cost savings to the healthcare system. Also, providing comprehensive care plan and reviews may prevent the need of expensive secondary care.
Quality of evidence	No quality assessment was conducted in the absence of a formal review.
Other considerations	The committee had an in-depth discussion about how to best manage a person who has been admitted to inpatient care and ensure that it is does not continue beyond the point where outpatient or day patient treatment could be safely reinstated. It was agreed that long-term admissions should be avoided, but at the same time ensure the person is not discharged too early and end up readmitted soon after. To help make a decision whether the person should be discharged, the committee agreed it should be a joint decision between the referral team, the person with the eating disorder and the parents and carers (if appropriate). The committee agreed a review on the need for ongoing treatment should be conducted within the first month of admission. RCT evidence from another review showed that similar outcomes can be achieved if the person is admitted for short-term medical stabilisation (mean 22 days) compared with longer-term weight restoration (mean 38 days). Thus, short-term treatment may be equally effective if the aim of the treatment is medical stabilisation rather than weight restoration. It was noted by the committee that an overemphasis on weight and weight restoration can be unhelpful or harmful for adults receiving inpatient treatment (Button & Warren, 2001) and lead to increased risk of drop out from treatment. For these reasons, the committee agreed that weight should not be the sole criterion for discharging people with an eating disorder. That is not to say that restoring weight to improve medical stability and reduce risk is not a goal of inpatient treatment, but it should not be a sole criterion. The committee used their knowledge and experience to agree on a number of factors that should be considered when discharging a patient from hospital. They included: whether the person has made enough progress towards the goals agreed at admission; the risk that the person may become dependent on inpatient care and reluctant to be discharged; that inpatient treatment is not working and that the person may be better placed at home for outpatient therapy; and seek an independent second opinion. When discharged, they committee agreed that the person with the eating disorder should be stepped down to less intensive treatment such as a day clinic or return home to receive ongoing psychological treatment in a community based eating disorder service (as described in the recommendation).

1 **Stepped care**

	No recommendation was made on stepped care. Instead a research recommendation was made.
Relative value of different outcomes	The GC discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating children, young people and adults with an eating disorder. For people with anorexia nervosa, body weight or BMI and remission are of greatest concern. For those with bulimia nervosa or binge eating disorder, binge eating and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in randomised controlled trials for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. Thus, they were extracted where possible, but did not factor strongly in the decision making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, family functioning and service user experience.
Trade-off between clinical	**Anorexia nervosa** An RCT on stepped care in young people with anorexia nervosa showed if family

	No recommendation was made on stepped care. Instead a research recommendation was made.
benefits and harms	therapy is stepped up to intensive parental coaching compared with continued family based therapy, there was no difference at the end of treatment on remission, body weight and depression and may be less effective on EDE-global scores and service use experience. The study did not report data on the important outcomes of general functioning, all-cause mortality, relapse or quality of life. **Bulimia nervosa** An RCT on stepped care in adults with bulimia nervosa showed that group psychoeducation stepped-up to CBT-ED may lead to better remission rates compared with group psychoeducation and wait list control but no other outcomes favoured the stepped care approach as they all showed no difference between the two arms, including binge frequency and EDE-global. The study did not report data on the important outcomes of family functioning, resource use, service user experience, all-cause mortality or relapse. Another study showed self-help stepped up to CBT-ED compared with CBT-ED alone had a similar effect on remission at the end of treatment and at follow up in adults with bulimia nervosa. The study did not report data on the critical outcome of binge eating, nor the important outcomes of family functioning, resource use, service user experience, all-cause mortality or relapse. Guided self-help stepped up to an antidepressant followed by CBT-ED showed no difference on remission rates compared with CBT-ED alone in adults with bulimia nervosa. The study did not report data on the important outcomes of general functioning, family functioning, resource use, service user experience, all-cause mortality or relapse. No published stepped-care evidence was found in people with binge eating disorder or EDNOS.
Trade-off between net health benefits and resource use	There was evidence from two studies showing that stepped care may potentially be cost effective in people with BN. However, both studies were non-UK and were only partially applicable to the NICE decision making context. One study was characterised by minor methodological limitations and one by potentially serious methodological limitations. The committee considered the above evidence. However, they could not draw any firm conclusions from it. Given that the overall existing evidence was positive the committee noted that there is a need for future well-conducted UK studies comparing the effectiveness and cost effectiveness of such care arrangements for people with eating disorders.
Quality of evidence	The quality of the evidence on stepped care was mostly very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as unclear randomisation, it was unclear if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded and high drop outs were detected >20%. Studies were included in the review if the higher level of care was the result of the participants' non-responsiveness to the previous intervention. However, in the study on people with anorexia nervosa, the non-responders were not randomised to a higher level of care (after failing to show a response to six months of family therapy). Therefore, at the end of treatment, the two arms of the study included two very different populations (one that showed a response early one versus those who did not). For this reason, the results are difficult to decipher and it is closer to an observational study. Only one of the studies on adults with bulimia nervosa (Davies 1999) randomised the participants after showing no response to the initial treatment, so the groups were more comparable at the end of treatment. The other two studies did not randomise participants after assessing whether they responded to the first-line treatment, they only randomised at the beginning of the study.
Other considerations	The committee agreed that the evidence was too limited to make a recommendation on stepped care and instead a research recommendation was generated that is relevant for any eating disorder to: "evaluate the effectiveness of stepped care for psychological treatment of eating disorders for people of all-ages."

2. Research recommendation: Evaluate the effectiveness of stepped care for psychological treatment of eating disorders for people of all-ages.

6 Treatment and management of anorexia nervosa

6.1 Introduction

People with Anorexia Nervosa (AN) induce weight loss and maintain low body weight. In psychological terms, this is often described as related to an underlying 'fear of fat' and an altered body image, where individuals see themselves as overweight whilst beneath the normal weight range. Anorexia Nervosa is more common in females than males, although there is recognition that it's incidence in males has been persistently underestimated. The onset of AN is characteristically in the young people years but may also occur in children and adults. Anorexia Nervosa usually runs a sustained course and persists into adulthood for most sufferers.

A number of behaviours are characteristically associated with inducing and maintaining low body weight. These include dietary restriction, excessive exercise and purging behaviours, for example, diuretic use or induced vomiting. Low body weight is an essential feature of AN and is frequently accompanied by other indicators of inadequate nutrition, including electrolyte imbalance, vitamin deficiency and secondary endocrine effects. Although the best known secondary endocrine effect is amenorrhoea in women, similar effects may be manifest as a loss of sexual interest in men and by delayed or arrested pubertal development in both sexes.

Therapeutic strategies in AN usually involve psychological approaches to beliefs and behaviour, with concurrent interventions to address the impact of low body weight and inadequate nutrition on people's general physical health.

6.2 Psychological interventions

6.3 Review question: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 70. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all psychological interventions that may be delivered to children, young people and adults with an eating disorder with or without a pharmacological intervention. The interventions were categorised according to their mode of delivery, i.e. individual, group or self-help, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to any other intervention or to wait list controls.

Table 70: Clinical review protocol summary

Component	Description
Review question(s)	Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with

Component	Description
	eating disorders compared with any other intervention or controls?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder. • Strata: • children (<12), young people (13-17 years), adults ≥18 years • eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder and atypical eating disorder) • mode of delivery (i. individual ii. family iii. group iv. self-help)
Intervention(s)	Psychological intervention including: • Dialectical behaviour therapy (DBT) • Counselling (Nutritional/Other) • Integrative Cognitive-Affective Therapy for Binge Eating (ICAT) • Maudsley model for treatment of adults with anorexia nervosa (MANTRA) • Cognitive remediation therapy (CRT) • Specialist supportive clinical management for anorexia nervosa (SSCM) • Behavioural therapy (BT) • CBT (General or ED specific) • Dynamic (IPT, Psychodynamic General or ED specific) • Guided Self Help w therapist guidance • Pure self help • E-therapies Psychological in combination with any pharmacological intervention.
Comparison	wait list control treatment as usual another other intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission (if symptoms were measured over a minimum 2 week period) • Binge eating for bulimia nervosa and binge eating disorder; and weight/body mass index (adjusted for age) for anorexia nervosa
Important outcomes	• Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Family functioning • Service user experience

Component	Description
	• Quality of life • All-cause mortality • Relapse • Adverse events • Resource use
Study design	• Systematic reviews • RCTs

6.3.1 Individual psychotherapy

16 RCTs (n=1181) were identified as relevant studies that investigated the effects of individual psychotherapy in people with anorexia nervosa, the majority of these were on adults and two were follow up studies (Carter 2011 (Carter et al., 2011), Dalle Grave 2013 (Dalle Grave et al., 2013), Dare 2001 (Dare et al., 2001), Eisler 1997 (Eisler et al., 1997), Gowers 2007 (Gowers et al., 2007), Hall 1987 (Hall and Crisp, 1987), Lock 2010 (Lock et al., 2010), McIntosh 2005 (McIntosh et al., 2005), Pike 2003 (Pike et al., 2003), Robin 1999 (Robin et al., 1999), Russel 1987 (Russell et al., 1987), Schmidt 2012 (Schmidt et al., 2012), Schmidt 2015 (Schmidt et al., 2015), Touyz 2013 (Touyz et al., 2013), Treasure 1995 (Treasure et al., 1995), Zipfel 2014 (Zipfel et al., 2014)). An overview of the trials included in the meta-analysis can be found in Table 71. Further information about both included and excluded studies can be found in Appendix J.

Summary of findings can be found in Table 74. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

6.3.2 Group therapy

No papers on group therapy in people with anorexia nervosa were identified.

6.3.3 Self-help

One RCT (n=221) on self-help was identified in people with anorexia nervosa (Fichter 2012 (Fichter et al., 2012)). An overview of the trials included in the meta-analysis can be found in. Summary of findings can be found in Table 73. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

6.3.4 Family therapy for anorexia nervosa

One of the 13 studies (n=60; Godart 2012) compared family therapy (ED) and treatment as usual with treatment as usual in young inpatients with anorexia nervosa. Summary of findings can be found in Table 86.

Two of the 13 studies (n=73; Geist 2000; Whitney 2012) compared family therapy-ED versus any other type of family-ED intervention in inpatients with anorexia nervosa. One study (Geist 2000) compared family therapy-ED with family group psychoeducation in young people, whilst Whitney 2012 compared family therapy-ED with a 3-day family day workshop intervention. Summary of findings can be found in Table 87, Table 100 and Table 101.

One of the 13 studies (n=164, Agras 2014) compared family therapy-ED with general family therapy in young people with anorexia nervosa. Summary of findings can be found in Table 88 and Table 89.

One of the 13 studies (n=169; Eisler 2016) compared multi-family therapy-ED with family therapy-ED in young people with anorexia nervosa. Summary of findings can be found in Table 90 and Table 91.

Four of the 13 studies (n=263) compared family therapy with any individual therapy in people with anorexia nervosa. The comparison interventions included young people-focused individual therapy, individual supportive therapy and cognitive analytic therapy. Three of these (n=179) were for anorexia nervosa in young people (Lock 2010, Robin 1999, Russell 1987, Eisler 1997) and one study (n=84) was for adults (Dare 2001). Summary of findings can be found in Table 92, Table 93 and Table 103.

Two of the 13 studies (n=147; Eisler 2000, le Grange 2016) compared family therapy-ED with an alternative family therapy-ED in people with anorexia nervosa. Both studies compared family therapy-ED in which the patient and carer are seen together by the therapist with family therapy-ED in which they are seen separately. Summary of findings can be found in Table 94 and Table 95.

One of the 13 studies (n=86, Lock 2005, Lock 2006) compared long-term family therapy-ED (12 months) with short-term family therapy-ED (6 months) in young people with anorexia nervosa. Summary of findings can be found in Table 96 and Table 97.

One of the 13 studies (n=23, Herscovici 2015) compared family therapy-ED and family meal with family therapy without family meal in young people with anorexia nervosa. Summary of findings can be found in Table 98 and Table 99.

One of the 13 studies (n=30; Hall 1987) compared a combined course of general family therapy and individual psychodynamic therapy with nutritional counselling. Summary of findings can be found in Table 102.

Table 71: Study information for trials included in the meta-analysis of individual psychotherapy versus any other intervention or wait list controls in people with anorexia nervosa

Study ID	Mean Age (SD) years	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison	Sessions N	Treatment Length	Long-term FU
Anorexia nervosa: individual therapy										
Dare 2001	26.3 (6.7)	15.4 (1.6)	98%	Most will have had > 3 years	84	Psychodynamic - General x 2	Psychiatric Counselling Family therapy	30	1 year	NA
Dalle Grace 2012	23.4 (6.9)	14.3 (1.8)	78%	Require inpatient treatment.	80	CBT-ED.1	CBT-ED.2 (variation in content)	Unclear	20 weeks	12 months FU
Gowers 2007	14.9	15.5 (1.6)	NR	<1 year	170	CBT-ED	Treatment as usual Inpatient care	NR	6 weeks to 6 months	18 months FU
Hall 1987	19.6 (14 to 25)	Deviation from matched population mean weight 25%	100%	Mean duration of illness 29.7 months	30	Psychodynamic	Nutritional counselling	12	12-24 weeks	6 months FU
Lock 2010	14.4 (1.6)	16.1 (1.1)	91%	Early most <3 years	121	Young people focused therapy	Family therapy	32	12 months	12 months FU
McIntosh 2005/ Carter 2011	17-40	17.3 (1.1)	100%	Unclear	56	CBT-ED	IPT SSCM	20	20 weeks	6.7 year FU
Pike 2003	26.1 (6.2)	16.0 (2.1)	100%	1 year after hospitalisation	33	CBT-ED	Nutritional counselling	50	1 year	NA
Robin 1999	14.9	15.86 (2.05)	100%	Developed AN less than	37	Young people focused therapy	Family therapy	Variable	15.9 months	12 months

Study ID	Mean Age (SD) years	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison	Sessions N	Treatment Length	Long-term FU
				one year						FU
Russell 1987/Eisler 1997	15.3 (1.8)	65.9 (8.0)% of average body weight	91%	Post hospital. Duration of illness 1.2 (0.7) years	21	Supportive therapy	Family therapy	Variable	12 months	5 year FU
Schmidt 2012	25.5 (6.9)	16.3 (1.3)	91.20 %	Most > 3y Duration of illness 77.3 months (70.8)	72	MANTRA	SSCM	26	10.6 months	2 months FU
Schmidt 2015	26.7 (7.7)	16.6 (1.2)	98.6%	Most > 3y Duration of illness 8.3 (7.3) years	142	MANTRA	SSCM	20	20 weeks	6 months FU
Touyz 2013	34.6 (9.0)	16.3 (1.3)	100%	Chronic AN (least 7 years)	63	CBT-ED	SSCM	30	8 months	12 months FU
Treasure 1995	25.3 (7)	15.0 (1.0)	100%	Poor prognosis	30	SSCM	Psychodynamic –General	20	20 weeks	32 weeks FU
Zipfel 2014	28.0 (8.6)	16.57 (1.0)	100%	40-60% AN>6 years	242	CBT-ED	Psychodynamic Treatment as usual	40	10 months	12 months FU

Abbreviations: AN – anorexia nervosa; BT – behavioural therapy; CBT-ED – cognitive behavioural therapy with an eating disorder focus; ED – eating disorder; ESM - emotional and social mind training; FU – follow up; ICAT – integrative cognitive affective therapy; IPT – interpersonal psychotherapy; N – number; NR – not reported; MANTRA – The Maudsley model of anorexia nervosa treatment for adults; SSCM – specialist supportive clinical management; WLC – wait list control; < - less than; > - greater than.

Table 72: Study information for trials included in the meta-analysis of self-help versus any other intervention or wait list control for people with anorexia nervosa.

Study ID	Mean age years (SD)	Mean BMI, kg/m2 (SD)	Female (%)	Sample	N random-ised	Intervention	Comparison	Number of sessions	Treatment Length	Long-term FU
Fichter 2012	23.8 (6.5)	17.8 (1.4)	100%	Discharged from inpatient care	258	Internet guided self-help	Treatment as usual (dependent on patient)	9	9 months	9 months FU

Abbreviation: FU – follow up

Table 73: Study information for trials included in the meta-analysis of family therapy versus any other intervention or wait list control for people with anorexia nervosa.

Study ID	Mean age years (SD)	Mean BMI, kg/m2 (SD)	Female (%)	Sample	N random-ised	Intervention	Comparison	Number of sessions	Treatment Length	Long-term FU
Agras 2014	15.3 (1.8)	Not reported	89	DSM-IV criteria for AN except for amenorrhea criterion	164	FBT-ED Duration of illness: 11.6 (9.8) months	SFT-ED Duration of illness: 15.4 (16.9) months.	16	9 months	12 months FU
Dare 2001	26.3 (6.7)	15.4 (1.6)	98	DSM-IV AN	84	FT-ED Duration of illness: 5.8 (4.9) years.	FPP Duration of illness: 6.7 (5.9) years. CAT Duration of illness: 6.7 (7.6) years Counselling Duration of illness: 6.1 (5) years	Mean 13.6 (8.6)	12 months	na
Eisler 2000	15.5	NA	98	DSM-IV or ICD-10 criteria for	40	Conjoint FT-ED	Separated FT-ED	Mean 16.4	12 months	na

Study ID	Mean age years (SD)	Mean BMI, kg/m2 (SD)	Female (%)	Sample	N random-ised	Intervention	Comparison	Number of sessions	Treatment Length	Long-term FU
	(1.6)			AN		Duration of illness: 13 months (range 2-36 mo)	Duration of illness: 12 months (range 2-36 mo)	(8.9)		
Eisler 2016	15.7 (1.6)	15.7 (1.2)	91	DSM-IV AN-R or EDNOS	169	FT-ED	Multi-FT-ED	FT-ED group, median =19 [IQR 12-27]; Multi-FT-ED, median =18.5 [IQR 11-24]	12 months	6 months FU
Geist 2000	14.6 (1.6)	na	100	Inpatients, DSM-IV except <90% IBW	25	FT-ED	Family Group PE	~64	4 months	na
Godart 2012	16.6 (1.6)	16.9 (1.1)	100	Inpatients, DSM-IV criteria for AN, ≤3 years since hospital admission for AN.	60	FT and treatment as usual	Treatment as usual	FT: Mean 11.8 (5.7). Overall (consultations, FT, individual therapy) mean: 33.7	18 months	18 months FU

Study ID	Mean age years (SD)	Mean BMI, kg/m2 (SD)	Female (%)	Sample	N random-ised	Intervention	Comparison	Number of sessions	Treatment Length	Long-term FU
								(24.6)		
Hall 1987	19.6 (13.5)	na	100	AN, <85% matched population mean weight + amenorrheic	30	General FT and IPT Duration of illness: 29.7 months (range: 6-72 months)	Nutritional Duration of illness: 24.5 months (range: 6-72 months)	12	12-24 weeks	12 months FU
Herscovici 2015	17.1 (2.3)	Weight (kg)=42.9 (7.3)	96	GOSH operational definition of AN	23	FT-ED with family meal	FT-ED without family meal	Intervention: mean 18 (14-25) Comparison mean: 14 (range 10-19)	6 months	6 months FU
Lock 2005/ Lock 2006	15.2	17.3 (1.5)	90	DSM-IV AN, though some (i) were partially weight restored or (ii) had only missed one menstrual period	86	Long-term FT-ED Duration: 12 (9.9) months	Short-term FT-ED Duration: 11.3 (10.4) months	20	12 months	3 years FU
Lock 2010	14.4 (1.6)	16.1 (1.1)	91	Early most <3 years	121	Family therapy	Young people focused therapy	32	12 months	12 months FU
Robin 1999	14.17	na	100	DSM-III-R criteria for AN.	37	BFST	AFT	variable	variable, 12-18 months	12 months FU

Study ID	Mean age years (SD)	Mean BMI, kg/m2 (SD)	Female (%)	Sample	N random-ised	Intervention	Comparison	Number of sessions	Treatment Length	Long-term FU
				Developed AN<=1 year.					[Average 15.9 mo]	
Russell 1987/ Eisler 1997	15.3 (1.8)	na	91	DSM-III criteria for AN. Just before discharge from ED unit. Illness for less than 3 years.	21	FT	Individual therapy	variable	12 months	5 years FU
Whitney 2012	na	13.3 (1.6)	98	Inpatient unit for AN. Duration of illness from <1 year, to >20 years (median 5-10 years)	48	Individual family work	Family day workshops	18 hours of treatment in 1-2 hr sessions (weekly or fortnightly) + 3 FU sessions	18 hours	36 months FU

Abbreviations: AFT, Young people-Focussed Therapy; BFST, Behavioural Family Systems Therapy; CAT, Cognitive Analytic Therapy; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; FBT-AN, Family-based Treatment for anorexia nervosa; FDW, Family Day Workshops; FPP, Focal Psychodynamic Psychotherapy; FT, Family Therapy; IFW, Individual Family Work; IPT, Individual Psychodynamic Therapy; IT, Individual Therapy; PE, Psychoeducation; SFT-AN, Systematic Family Therapy for anorexia nervosa; na, not applicable; TAU, treatment as usual.

6.3.51 Individual therapy

Table 74: Summary of findings table for CBT-ED versus any other intervention at the end of treatment and at follow up (FU) in adults and children and young people with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Another intervention	Risk difference with AN CBT-ED (95% CI)
Weight - Adults	298 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight - adults in the intervention groups was 0.17 standard deviations higher (0.07 lower to 0.42 higher)
EDE-Restraint - Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-restraint - adults in the intervention groups was 0.13 standard deviations lower (0.69 lower to 0.44 higher)
EDE-Eating concerns-Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concerns- adults in the intervention groups was 0.31 standard deviations lower (0.87 lower to 0.25 higher)
EDE-Weight concerns-Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concerns- adults in the intervention groups was 0.39 standard deviations higher (0.17 lower to 0.95 higher)
EDE-Shape concerns-Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concerns- adults in the intervention groups was 0.09 standard deviations lower (0.65 lower to 0.46 higher)
EDI - Drive for thinness-Adults	56 (1 study)	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness- adults in the intervention groups was 0.07 standard deviations lower (0.63 lower to 0.48 higher)
EDI - Body dissatisfaction- Adults	56 (1 study)	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction- adults in the intervention groups was 0.2 standard deviations lower (0.76 lower to 0.35 higher)

EDI - Bulimia- Adults	56 (1 study)	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia- adults in the intervention groups was 0.21 standard deviations lower (0.76 lower to 0.35 higher)
EDI Total - Adults	242 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total - adults in the intervention groups was 0.08 standard deviations lower (0.35 lower to 0.19 higher)
General psychopathology- Adults	242 (1 study)	⊕⊖⊖⊖ VERY LOW5,8,9 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general psychopathology-adults in the intervention groups was 0.25 standard deviations lower (0.52 lower to 0.02 higher)
Depression Adults	56 (1 study)	⊕⊕⊖⊖ LOW5,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression adults in the intervention groups was 0.20 standard deviations lower (0.76 lower to 0.35 higher)
Relapse	33 (1 study)	⊕⊕⊖⊖ LOW11,12 due to risk of bias, imprecision	RR 0.42 (0.16 to 1.12)	533 per 1000	309 fewer per 1000 (from 448 fewer to 64 more)
Remission ITT- Adults	275 (2 studies)	⊕⊖⊖⊖ VERY LOW1,13,14,15 due to risk of bias, inconsistency, indirectness, imprecision	RR 1.97 (0.67 to 5.80)	102 per 1000	99 more per 1000 (from 34 fewer to 488 more)
BMI-Adolescents FU	98 (1 study)	⊕⊕⊖⊖ LOW5,16 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi-adolescents fu in the intervention groups was 0.29 standard deviations lower (0.69 lower to 0.11 higher)
BMI - Adults FU	285 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,13,15 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean bmi - adults fu in the intervention groups was 0.05 standard deviations lower (0.29 lower to 0.2 higher)
EDE-Shape concerns - Adults FU	44 (1 study)	⊕⊕⊖⊖ LOW3,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concerns - adults fu in the intervention groups was 0.31 standard deviations lower (1.33 lower to 0.71 higher)

EDE-Eating concerns- Adults FU	43 (1 study)	⊕⊕⊖⊖ LOW3,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concerns- adults fu in the intervention groups was 0.16 standard deviations lower (0.78 lower to 0.45 higher)
EDE-Restraint - Adults FU	43 (1 study)	⊕⊕⊖⊖ LOW5,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-restraint - adults fu in the intervention groups was 0.36 standard deviations lower (0.97 lower to 0.26 higher)
EDE-Weight concerns - Adults FU	43 (1 study)	⊕⊕⊖⊖ LOW6,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concerns - adults fu in the intervention groups was 0.02 standard deviations lower (0.63 lower to 0.59 higher)
EDI - Body dissatisfaction- Adults FU	43 (1 study)	⊕⊕⊖⊖ LOW5,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction- adults fu in the intervention groups was 0.32 standard deviations lower (0.94 lower to 0.29 higher)
EDI - Bulimia - Adults FU	43 (1 study)	⊕⊕⊖⊖ LOW6,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia - adults fu in the intervention groups was 0.43 standard deviations higher (0.19 lower to 1.06 higher)
EDI - Drive for thinness - Adults FU	43 (1 study)	⊕⊕⊖⊖ LOW6,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness - adults fu in the intervention groups was 0.25 standard deviations higher (0.37 lower to 0.87 higher)
EDI Total Adults - FU	242 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi total adults - fu in the intervention groups was 0.07 standard deviations higher (0.19 lower to 0.34 higher)
EDI-Total Adolescents FU	82 (1 study)	⊕⊕⊖⊖ LOW5,16 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-total adolescents fu in the intervention groups was 0.17 standard deviations lower (0.6 lower to 0.27 higher)
Depression Adults FU	43 (1 study)	⊕⊕⊖⊖ LOW6,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression adults fu in the intervention groups was 0.13 standard deviations lower (0.48 lower to 0.75 higher)
General Function Adults	43	⊕⊕⊖⊖		Not calculable for	The mean general function adults fu in

FU	(1 study)	LOW5,17 due to risk of bias, imprecision		SMD values	the intervention groups was 0.04 standard deviations lower (0.65 to 0.57 lower)
General psychopathology Adults FU	242 (1 study)	⊕⊖⊖⊖ VERY LOW3,8,9 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general psychopathology adults fu in the intervention groups was 0.03 standard deviations higher (0.24 lower to 0.3 higher)
Remission- Adolescents FU ITT	110 (1 study)	⊕⊕⊖⊖ LOW15,16 due to risk of bias, imprecision	RR 1.25 (0.53 to 2.93)	145 per 1000	36 more per 1000 (from 68 fewer to 281 more)
Remission -Adults FU ITT	242 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,13,18,19 due to risk of bias, inconsistency, indirectness, imprecision	RR 0.85 (0.51 to 1.43)	235 per 1000	35 fewer per 1000 (from 115 fewer to 101 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was performed. High drop outs >20% were reported. Only assessors were blind in all studies.
2 In Ziphel, between baseline and end of treatment, the following had hospital study longer than 28 days for weight restoration: 5/ 80 (6%) focal psychodynamic, 8/80 (10%) CBT-ED and 9/82 (11%) TAU.
3 For a continuous outcome, there were fewer than 400 participants.
4 Heterogeneity present, I2>80%
5 95% CI crossed 1 MID (-0.5)
6 95% CI crossed 1 MID (0.5)
7 Unclear if allocation concealment was performed or how randomisation was conducted. Neither patients or investigators were blind, assessor was blind. High dropout >20% was reported.
8 Unclear if allocation concealment was performed. Participants were not blind, unclear if investigators were blind, Assessors were blind. High drop outs were detected >20%
9 High number of participants spent time in hospital: 23% Focal Psychodynamic, 34% CBT, 41% TAU had periods of hospitalisation
10 Unclear how randomisation was performed or if allocation concealment was performed. High drop outs were reported >20% in most studies. Only assessors were blind.
11 Unclear how randomisation was performed or if allocation concealment was conducted. Unclear if assessors, participants or investigators were blind.
12 95% CI crossed 1 MID (0.75)
13 Heterogeneity, I2 >50%
14 In Pike, participants were assigned to therapy within 1 week of successful completion of hospitalization. Different population to other studies.
15 For a dichotomous outcome, there were fewer than 300 events.

16 Unclear methods of randomisation. It was unclear if either participants, investigators or assessors were blind. High drop outs were reported >20%,
17 Unclear if allocation concealment was performed. Neither patients or investigators were blind, assessor was blind. High drop outs reported >20%...
18 95% CI crossed 2 MIDs (0.75 and 1.25)
19 Inconsistency was between 2 different comparisons from the same study that included 3-arms.

Table 75: Summary of findings table for psychiatric counselling versus any other intervention at the end of treatment in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with AN Psychiatric Counselling (95% CI)
Remission Adults	104 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.10 (0.95 to 1.28)	106 per 1000	11 more per 1000 (from 5 fewer to 30 more)
All-cause mortality Adults	84 (1 study)	⊕⊕⊖⊖ LOW1 due to risk of bias, imprecision	RR 1.01 (0.9 to 1.13)	16 per 1000	0 more per 1000 (from 2 fewer to 2 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how random sequence was generated or if sealed envelopes were opaque. Neither the investigators, assessors nor participants were blinded. High dropouts were reported >20%.
2 95% CI crossed 1 MID (1.25).

Table 76: Summary of findings table for supportive therapy versus another intervention at the end of treatment and at follow up in young people with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Another intervention Young people	Risk difference with AN Supportive therapy (95% CI)
Weight (percentile) Young people	21 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (percentile) young people in the intervention groups was 0.98 standard deviations lower (1.90 to 0.07 lower)

Did not achieve remission ITT Young people	21 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 2.27 (1.04 to 4.97)	600 per 1000	762 more per 1000 (from 24 more to 1000 more)
Weight (percentile) Young people FU	19 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (percentile) young people fu in the intervention groups was 0.57 standard deviations lower (1.50 lower to 0.35 higher)
Remission ITT- Young people FU	21 (3 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	See comment	400 per 1000	144 more per 1000 (from 184 fewer to 984 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Russel/Eisler. Unclear if allocation concealment was performed. High dropout rates >20% were reported. Assessors were blind, but it was unclear if participants were but investigators were not blind. *2 95% CI crossed 1 MID (-0.5)3 95% CI crossed 1 MID (0.75)*					

Table 77: Summary of findings table for young people focused therapy versus another intervention (other) at the end of treatment and at follow up in young people with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with AN Young people focused therapy (95% CI)
BMI Young people	139 (2 studies)	⊕⊕⊖⊖ LOW1,2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI young people in the intervention groups was 0.43 standard deviations lower (0.77 to 0.09 lower)
Remission ITT Young people	158 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, inconsistency, imprecision	RR 0.79 (0.61 to 1.01)	700 per 1000	147 fewer per 1000 (from 273 fewer to 7 more)

BMI Young people FU	129 (3 studies)	⨁⨁⊖⊖ LOW1,2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI young people fu in the intervention groups was 0.18 standard deviations lower (0.53 lower to 0.16 higher)
Remission ITT- Young people FU	158 (3 studies)	⨁⨁⊖⊖ LOW1,2,5 due to risk of bias, imprecision	See comment	588 per 1000	41 more per 1000 (from 100 fewer to 217 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Robin 1999. Unclear if allocation concealment was performed. It was unclear if either the participants, investigators or assessors were blind.
2 Lock 2010. Unclear if allocation concealment was performed. Assessors were blind, but participants and investigators were not blind.
3 95% CI crossed 1 MID (-0.5).
4 95% CI crossed 1 MID (0.75).
5 95% CI crossed 1 MID (1.25).

Table 78: Summary of findings table for focal psychodynamic general therapy versus another intervention at the end of treatment and follow up in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention Adults	Risk difference with AN Psychodynamic General (95% CI)
BMI Adults	242 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI adults in the intervention groups was 0.17 standard deviations lower (0.44 lower to 0.09 higher)
EDI Total - Adults	242 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi total - adults in the intervention groups was 0.02 standard deviations lower (0.29 lower to 0.24 higher)
All-cause mortality- Adults	84 (2 studies)	⨁⊖⊖⊖ VERY LOW4,5,6 due to risk of bias,	RR 1.05 (0.94 to 1.18)	49 per 1000	2 more per 1000 (from 3 fewer to 9 more)

		indirectness, imprecision			
General psychopathology- Adults	242 (1 study)	⊕⊖⊖⊖ VERY LOW1,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general psychopathology- adults in the intervention groups was 0.08 standard deviations higher (0.19 lower to 0.35 higher)
Remission_Adults_ITT	326 (2 studies)	⊕⊖⊖⊖ VERY LOW2,5,6,8 due to risk of bias, indirectness, imprecision	RR 1.73 (0.95 to 3.14)	89 per 1000	65 more per 1000 (from 4 fewer to 190 more)
Weight (BMI and kg)- Adult FU	293 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,5,8 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight (bmi and kg)- adult fu in the intervention groups was 0.09 standard deviations higher (0.14 lower to 0.33 higher)
EDE Bulimia- Adults FU	30 (1 study)	⊕⊖⊖⊖ VERY LOW8,9 due to risk of bias, imprecision	RR 0.76 (0.15 to 3.92)	188 per 1000	45 fewer per 1000 (from 159 fewer to 548 more)
EDI-Total- Adults FU	242 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi - total- adults fu in the intervention groups was 0.07 standard deviations lower (0.35 lower to 0.19 higher)
Morgan Russell ED- Adults FU	30 (1 study)	⊕⊕⊖⊖ LOW7,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean Morgan Russell ed- adults fu in the intervention groups was 0.32 standard deviations higher (0.4 lower to 1.04 higher)
General psychopathology - Adults FU	242 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general psychopathology - adults fu in the intervention groups was 0.00 standard deviations lower (0.27 lower to 0.27 higher)
Remission FU_- Adults ITT	272 (2 studies)	⊕⊖⊖⊖ VERY LOW2,9,10	RR 2.00 (1.33 to	174 per 1000	174 more per 1000 (from 57 more to 354 more)

		due to risk of bias, indirectness, imprecision	3.03)		

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear if allocation concealment was performed. Participants were not blind, it was unclear if investigators were, however, and assessors were blind to treatment allocation. High dropouts reported >20%.
2 In Zipfel, between baseline and end of treatment, the following had hospital study longer than 28 days for weight restoration: 5/ 80 (6%) focal psychodynamic, 8/80 (10%) CBT-ED and 9/82 (11%) TAU.
3 For a continuous outcome, there were fewer than 400 participants.
4 Unclear methods of randomisation and if allocation concealment was performed. High dropouts reported >20%. Unclear if either patient, investigator or assessor were blind.
5 In Dare, a number of patients were hospitalised during the treatment: 10% Family therapy, 14% focal psychodynamic, 9% focal psychodynamic CAT, 26% treatment as usual – counselling.
6 95% CI crossed 1 MID (1.25).
7 95% CI crossed 1 MID (0.5).
8 Unclear if allocation concealment was performed or if assessors were blind. High dropouts reported >20%.
9 95% CI crossed 2 MIDs (0.75 and 1.25).
10 Unclear if allocation concealment was performed or if participants, investigators or assessors were blind. High dropouts reported >20%.

Table 79: Summary of findings table for interpersonal psychotherapy (IPT) versus any other intervention at the end of the treatment and follow up in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with another intervention	Risk difference with AN IPT (95% CI)
BMI- Adults	56 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean bmi- adults in the intervention groups was 0.13 standard deviations lower (0.68 lower to 0.41 higher)
EDE-Restraint- Adults	56 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-restraint- adults in the intervention groups was 0.99 standard deviations higher (0.41 to 1.57 higher)
EDE-Eating concerns- Adults	56 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias,	Not calculable for SMD values	The mean ede-eating concerns- adults in the intervention groups was 0.49 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with another intervention	Risk difference with AN IPT (95% CI)
		imprecision		(0.06 lower to 1.04 higher)
EDE-Weight concerns- Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-weight concerns- adults in the intervention groups was 0.2 standard deviations lower (0.75 lower to 0.34 higher)
EDE-Shape concerns- Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-shape concerns- adults in the intervention groups was 0.25 standard deviations higher (0.29 lower to 0.8 higher)
General Function (GAF)- Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean general function (gaf)- adults in the intervention groups was 0.5 standard deviations lower (1.06 lower to 0.05 higher)
Depression (Hamilton)- Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression (Hamilton)- adults in the intervention groups was 0.4 standard deviations higher (0.15 lower to 0.95 higher)
EDI-Drive for thinness- Adults	48 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi - drive for thinness- adults in the intervention groups was 0.17 standard deviations lower (0.76 lower to 0.43 higher)
EDI-Bulimia- Adults	48 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi - bulimia- adults in the intervention groups was 0.36 standard deviations higher (0.24 lower to 0.96 higher)
EDI-Body dissatisfaction- Adults	48 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi - body dissatisfaction- adults in the intervention groups was 0.01 standard deviations higher (0.59 lower to 0.6 higher)
BMI - Follow up- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean BMI - follow up- adults in the intervention groups was 0.10 standard deviations higher (0.54 lower to 0.75 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with another intervention	Risk difference with AN IPT (95% CI)
EDE-Shape concerns Follow up- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-shape concerns follow up- adults in the intervention groups was 0.18 standard deviations higher (0.47 lower to 0.82 higher)
EDE-Eating concerns Follow up- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-eating concerns follow up- adults in the intervention groups was 0.17 standard deviations lower (0.81 lower to 0.47 higher)
EDE-Restraint Follow up- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-restraint follow up- adults in the intervention groups was 0.28 standard deviations lower (0.93 lower to 0.37 higher)
EDE-Weight concerns Follow up- Adults	43 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-weight concerns follow up- adults in the intervention groups was 0.1 standard deviations lower (0.74 lower to 0.54 higher)
EDI-Drive for thinness - FU- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi - drive for thinness - fu- adults in the intervention groups was 0.54 standard deviations lower (1.19 lower to 0.11 higher)
EDI-Bulimia - FU- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi - bulimia - fu- adults in the intervention groups was 0.21 standard deviations lower (0.85 lower to 0.44 higher)
EDI-Body dissatisfaction - FU- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi - body dissatisfaction - fu- adults in the intervention groups was 0.14 standard deviations higher (0.5 lower to 0.78 higher)
Depression (Hamilton) Follow up- Adults	43 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression (Hamilton) follow up- adults in the intervention groups was 0.08 standard deviations lower (0.72 lower to 0.56 higher)
General Function (GAF) Follow	43	⊕⊖⊖⊖	Not calculable for	The mean general function (gaf) follow up- adults

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with another intervention	Risk difference with AN IPT (95% CI)
up- Adults	(1 study)	VERY LOW1,2,3 due to risk of bias, imprecision	SMD values	in the intervention groups was 0.08 standard deviations higher (0.56 lower to 0.72 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear how randomisation was performed or if allocation concealment was conducted. Assessors were blind. High dropout rates were reported >20%
2 95% CI crossed 1 MID (-0.5)
[3] 95% CI crossed 1 MID (0.5)

Table 80: Summary of findings for SSCM versus any other intervention at the end of treatment and at follow up in adults with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with AN SSCM (95% CI)
BMI- Adults	269 (2 studies)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi- adults in the intervention groups was 0.04 standard deviations lower (0.28 lower to 0.21 higher)
EDE-Restraint- Adults	198 (2 studies)	⨁⨀⨀⨀ VERY LOW1,2,9 due to risk of bias, imprecision, inconsistency		Not calculable for SMD values	The mean ede-restraint- adults in the intervention groups was 0.58 standard deviations lower (1.41 lower to 0.24 higher)
EDE-Eating concerns- Adults	198 (2 studies)	⨁⨁⨀⨀ LOW1,2,3 due to risk of bias,		Not calculable for	The mean ede-eating concerns- adults in the intervention groups was 0.04 standard deviations higher

		imprecision		SMD values	(0.33 lower to 0.24 higher)
EDE-Weight concerns- Adults	198 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concerns- adults in the intervention groups was 0.07 standard deviations lower (0.36 lower to 0.22 higher)
EDE-Shape concerns- Adults	198 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concerns- adults in the intervention groups was 0.11 standard deviations lower (0.39 lower to 0.18 higher)
EDE - Global- Adults	213 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - global- adults in the intervention groups was 0.00 standard deviations lower (0.27 lower to 0.27 higher)
EDI - Drive for thinness- Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness- adults in the intervention groups was 0.29 standard deviations lower (0.88 lower to 0.29 higher)
EDI - Body dissatisfaction- Adults	56 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction- adults in the intervention groups was 0.14 standard deviations higher (0.44 lower to 0.72 higher)
EDI - Bulimia- Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia- adults in the intervention groups was 0.09 standard deviations lower (0.67 lower to 0.49 higher)
Depression - Adults	269 (3 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - adults in the intervention groups was 0.15 standard deviations lower (0.4 lower to 0.09 higher)

General Function (GAF)- Adults	56 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean general function (gaf)- adults in the intervention groups was 0.83 standard deviations higher (0.22 to 1.43 higher)
Remission_ ITT- Adults	216 (2 studies)	⊕⊖⊖⊖ VERY LOW1,6,7 due to risk of bias, imprecision	RR 1.22 (0.52 to 2.82)	83 per 1000	18 more per 1000 (from 40 fewer to 150 more)
BMI - Follow-up- Adults	286 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi - follow-up- adults in the intervention groups was 0.09 standard deviations lower (0.32 lower to 0.15 higher)
EDE-Weight concerns Follow-up- Adults	189 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concerns follow-up- adults in the intervention groups was 0.16 standard deviations higher (0.13 lower to 0.46 higher)
EDE-Shape concerns Follow-up- Adults	185 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concerns follow-up- adults in the intervention groups was 0.04 standard deviations higher (0.25 lower to 0.34 higher)
EDE-Restraint Follow-up- Adults	185 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-restraint follow-up- adults in the intervention groups was 0.20 standard deviations higher (0.09 lower to 0.5 higher)
EDE-Eating concerns Follow-up- Adults	185 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concerns follow-up- adults in the intervention groups was 0.24 standard deviations higher (0.06 lower to 0.53 higher)
EDE-Global FU- Adults	213 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias,		Not calculable for SMD	The mean ede-global fu- adults in the intervention groups was 0.13 standard deviations higher

		imprecision		values	(0.14 lower to 0.4 higher)
EDI - Body dissatisfaction - FU- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction - fu- adults in the intervention groups was 0.2 standard deviations higher (0.47 lower to 0.87 higher)
EDI - Bulimia - Follow-up- Adults	43 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia - follow-up- adults in the intervention groups was 0.15 standard deviations lower (0.82 lower to 0.52 higher)
EDI - Drive for thinness - Follow-up- Adults	43 (1 study)	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness - follow-up- adults in the intervention groups was 0.44 standard deviations higher (0.24 lower to 1.12 higher)
Depression Follow-up- Adults	256 (3 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression follow-up- adults in the intervention groups was 0.02 standard deviations lower (0.27 lower to 0.023 higher)
Bulimia- Adults	30 (1 study)	⊕⊖⊖⊖ VERY LOW7,8 due to risk of bias, imprecision	RR 1.31 (0.25 to 6.76)	143 per 1000	44 more per 1000 (from 107 fewer to 823 more)
General Function (GAF) Follow-up- Adults	43 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean general function (gaf) follow-up- adults in the intervention groups was 0.05 standard deviations lower (0.72 lower to 0.62 higher)
Remission FU_ITT- Adults	243 (3 studies)	⊕⊖⊖⊖ VERY LOW1,7 due to risk of bias, imprecision	RR 0.80 (0.49 to 1.3)	233 per 1000	47 fewer per 1000 (from 119 fewer to 70 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; FU: Follow up

1 Unclear if allocation concealment was performed. High dropout rates were reported >20% for McIntosh2005 and Schmidt 2015. It was unclear in McIntosh how randomisation was conducted. Across studies it was either unclear if participants and investigators were blind or they were not blind.
2 95% CI crossed 1 MID (-0.5)
3 For a continuous outcome, there were fewer than 400 participants.
4 95% CI crossed 1 MID (1.25)
5 95% CI crossed 1 MID (0.5)
6 Unclear if allocation concealment was performed. Across studies it was either unclear if participants and investigators were blind.
7 95% CI crossed 2 MIDs (0.75 and 1.25)
8 Unclear if allocation concealment was performed. It was unclear if participants, assessors and investigators were blind. High drop outs were reported >20%
9. Heterogeneity >50%

Table 81: Summary of findings table for MANTRA versus any other intervention at the end of treatment and at follow up in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with AN MANTRA (95% CI)
BMI Adults	213 (2 studies)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI adults in the intervention groups was 0.08 standard deviations higher (0.18 lower to 0.35 higher)
EDI-Total Adults	213 (2 studies)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - total adults in the intervention groups was 0.00 standard deviations higher (0.27 lower to 0.27 higher)
Depression- Adults	213 (2 studies)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression- adults in the intervention groups was 0.01 standard deviations lower (0.28 lower to 0.26 higher)
Remission ITT- Adults	213 (2 studies)	⨁⨁⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 0.82 (0.35 to 1.91)	103 per 1000	19 fewer per 1000 (from 67 fewer to 94 more)

BMI FU- Adults	213 (2 studies)	⨁⨀⨀⨀ VERY LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu- adults in the intervention groups was 0.11 standard deviations higher (0.16 lower to 0.37 higher)
Depression FU- Adults	213 (2 studies)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu- adults in the intervention groups was 0.01 standard deviations higher (0.25 lower to 0.28 higher)
EDI-Total Adults FU	213 (2 studies)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - total adults fu in the intervention groups was 0.13 standard deviations lower (0.4 lower to 0.14 higher)
Remission ITT FU- Adults	215 (2 studies)	⨁⨁⨀⨀ LOW1,4 due to risk of bias, imprecision	RR 1.22 (0.7 to 2.14)	165 per 1000	36 more per 1000 (from 50 fewer to 188 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 In Schmidt 2015, it was unclear if allocation concealment was performed. In both studies, the participants were not blinded, it was unclear in one if the investigators were blind, but in the other they were not. In both studies the assessors were blind, High dropouts were reported in one group >20%.
2 95% CI crossed 1 MID (0.5).
3 For a continuous outcome, there were fewer than 400 participants.
4 For a dichotomous outcome, there were fewer than 300 events.
5 95% CI crossed 2 MIDs (-0.5 and 0.5).

Table 82: Summary of findings table for CBT-ED (1) compared with another CBT-ED (2) inpatient program at the end of treatment and at follow up for adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with CBT-ED (2)	Risk difference with AN Inpatient CBT-ED (1) (95% CI)
BMI Adults	72 (1 study)	⨁⨁⨀⨀ LOW1,2	Not calculable for SMD values	The mean BMI adults in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with CBT-ED (2)	Risk difference with AN Inpatient CBT-ED (1) (95% CI)
		due to risk of bias, imprecision		0.09 standard deviations lower (0.56 lower to 0.37 higher)
EDE-Restraint Adults	72 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-restraint adults in the intervention groups was 0 standard deviations higher (0.46 lower to 0.46 higher)
EDE-Eating concerns Adults	72 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-eating concerns adults in the intervention groups was 0.09 standard deviations higher (0.37 lower to 0.56 higher)
EDE-Weight concerns Adults	72 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-weight concerns adults in the intervention groups was 0.07 standard deviations lower (0.54 lower to 0.39 higher)
EDE-Shape concerns Adults	72 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-shape concerns adults in the intervention groups was 0.06 standard deviations higher (0.4 lower to 0.52 higher)
General psychiatric features Adults	72 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean general psychiatric features adults in the intervention groups was 0.3 standard deviations higher (0.16 lower to 0.77 higher)
BMI - Adults FU	68 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean BMI - adults fu in the intervention groups was 0.04 standard deviations higher (0.43 lower to 0.52 higher)
General psychiatric features - Adults FU	68 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean general psychiatric features - adults fu in the intervention groups was 0.14 standard deviations higher (0.33 lower to 0.62 higher)
EDE-Restraint Adults FU	68 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,	Not calculable for SMD values	The mean ede-restraint adults fu in the intervention groups was 0.06 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with CBT-ED (2)	Risk difference with AN Inpatient CBT-ED (1) (95% CI)
		imprecision		(0.54 lower to 0.42 higher)
EDE-Eating concerns Adults FU	68 (1 study)	⨁⨁⨀⨀ LOW3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-eating concerns adults fu in the intervention groups was 0 standard deviations higher (0.48 lower to 0.48 higher)
EDE-Weight concerns Adults FU	68 (1 study)	⨁⨁⨀⨀ LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-weight concerns adults fu in the intervention groups was 0.2 standard deviations higher (0.27 lower to 0.68 higher)
EDE-Shape concerns Adults FU	68 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-shape concerns adults fu in the intervention groups was 0 standard deviations higher (0.48 lower to 0.48 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

[1] Unclear if allocation concealment was performed. It was also unclear if investigators, participants were blind, however, the assessors were blind.
[2] 95% CI crossed 1 MID (-0.5)
[3] For a continuous outcome, there were fewer than 400 participants
[4] 95% CI crossed 1 MID (0.5)

Table 83: Summary of findings table for CBT compared with any other intervention in adults with severe and enduring anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with Severe AN CBT (95% CI)
BMI- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias,	Not calculable for SMD values	The mean bmi- adults in the intervention groups was 0.00 standard deviations higher (0.49 lower to 0.49 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with Severe AN CBT (95% CI)
		imprecision		
Depression- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression- adults in the intervention groups was 0.24 standard deviations lower (0.74 lower to 0.25 higher)
EDE-Global- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede- global- adults in the intervention groups was 0.39 standard deviations lower (0.89 lower to 0.11 higher)
Quality of life- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean quality of life- adults in the intervention groups was 0.28 standard deviations lower (0.78 lower to 0.22 higher)
BMI FU- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean BMI fu- adults in the intervention groups was 0.11 standard deviations higher (0.38 lower to 0.61 higher)
Depression FU- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression fu- adults in the intervention groups was 0.27 standard deviations lower (0.77 lower to 0.22 higher)
EDE-Global FU- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede- global fu- adults in the intervention groups was 0.57 standard deviations lower (1.08 lower to 0.07 higher)
Quality of life FU- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean quality of life fu- adults in the intervention groups was 0.14 standard deviations lower (0.64 lower to 0.35 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with Severe AN CBT (95% CI)

[1] Unclear if allocation concealment was performed. It was unclear if the participants and investigators were blind. High dropouts were reported >20%
[2] For a continuous outcome, there were fewer than 400 participants.
[3] 95% CI crossed 1 MID (-0.5)
[4] 95% CI crossed 1 MID (0.5)

Table 84: Summary of findings table for SSCM versus any other intervention at the end of treatment and at follow up for adults with severe and enduring anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with SSCM (95% CI)
BMI- Adults	63 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean bmi- adults in the intervention groups was 0.00 standard deviations higher (0.49 lower to 0.49 higher)
EDE-Global- Adults	63 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede-global- adults in the intervention groups was 0.39 standard deviations higher (0.11 lower to 0.99 higher)
Quality of life- Adults	63 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean quality of life- adults in the intervention groups was 0.28 standard deviations higher (0.22 lower to 0.78 higher)
Depression- Adults	63 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression- adults in the intervention groups was 0.24 standard deviations higher (0.25 lower to 0.74 higher)
BMI FU- Adults	63 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	Not calculable for SMD values	The mean BMI fu- adults in the intervention groups was 0.11 standard deviations lower (0.61 lower to 0.38 higher)
EDE-Global FU- Adults	63	⊕⊕⊖⊖	Not	The mean ede-global fu- adults in the intervention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with SSCM (95% CI)
	(1 study)	LOW1,4 due to risk of bias, imprecision	calculable for SMD values	groups was 0.57 standard deviations higher (0.07 to 1.08 higher)
Quality of life FU- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean quality of life fu- adults in the intervention groups was 0.14 standard deviations higher (0.35 lower to 0.64 higher)
Depression FU- Adults	63 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression fu- adults in the intervention groups was 0.27 standard deviations higher (0.22 lower to 0.77 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear if allocation concealment was performed. It was unclear if the participants and investigators were blind. High dropouts were reported >20%
2 For a continuous outcome, there were fewer than 400 participants.
3 95% CI crossed 1 MID (0.5)
4 95% CI crossed 1 MID (-0.5)

6.3.6 Group therapy

No evidence was identified.

6.3.7 Self-help

Table 85: Summary table of findings for guided self-help with an eating disorder focus (GHS-ED) versus another intervention (other) for adults with anorexia nervosa at the end of treatment and follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with AN Internet GSH (ED) (95% CI)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with AN Internet GSH (ED) (95% CI)
EDI-Total	219 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi - total in the intervention groups was 0.27 standard deviations lower (0.53 lower to 0 higher)
EDI- Drive for thinness	219 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi- drive for thinness in the intervention groups was 0.17 standard deviations lower (0.44 lower to 0.09 higher)
EDI- Bulimia	219 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi- bulimia in the intervention groups was 0.15 standard deviations lower (0.42 lower to 0.11 higher)
EDI- Body dissatisfaction	219 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi- body dissatisfaction in the intervention groups was 0.24 standard deviations lower (0.51 lower to 0.02 higher)
Depression	219 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression in the intervention groups was 0.2 standard deviations lower (0.46 lower to 0.07 higher)
Global Clinical Score (PSR)	239 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean global clinical score (psr) in the intervention groups was 0.21 standard deviations lower (0.47 lower to 0.04 higher)
Bulimic symptoms	226 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean bulimic symptoms in the intervention groups was 0.26 standard deviations lower (0.52 lower to 0 higher)
Morgan-Russell Menstrual Function	239 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean Morgan-Russell menstrual function in the intervention groups was 0.18 standard deviations lower (0.44 lower to 0.07 higher)
General psychopathology	239	⊕⊕⊖⊖	Not	The mean general psychopathology in the intervention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Other	Risk difference with AN Internet GSH (ED) (95% CI)
	(1 study)	LOW1,3 due to risk of bias, imprecision	calculable for SMD values	groups was 0.1 standard deviations lower (0.35 lower to 0.15 higher)
General psychopathology FU	208 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean general psychopathology fu in the intervention groups was 0.07 standard deviations lower (0.34 lower to 0.21 higher)
Morgan-Russell Menstrual Function FU	208 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean Morgan-Russell menstrual function fu in the intervention groups was 0.07 standard deviations higher (0.2 lower to 0.35 higher)
Bulimic symptoms FU	208 (1 study)	See comment	Not calculable for SMD values	The mean bulimic symptoms fu in the intervention groups was 0.21 standard deviations lower (0.48 lower to 0.07 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

[1] It was unclear if allocation concealment was performed. Assessors were blind but it was unclear if investigators and participants were blind.
[2] 95% CI crossed 1 MID (-0.5)
[3] For a continuous outcome, there were fewer than 400 participants.

6.3.81 Family therapy in young people

Table 86: Summary table of findings for family therapy-ED and treatment as usual versus treatment as usual (TAU) in young inpatients with anorexia nervosa

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with TAU	Risk difference with Family Therapy-ED (95% CI)
Remission (ITT) Morgan-Russell Good or Intermediate outcome	60 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision	RR 2.4 (0.96 to 5.98)	167 per 1000	233 more per 1000 (from 7 fewer to 830 more)
BMI (raw)	60 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean BMI (raw) in the intervention groups was 0.1 standard deviations higher (0.41 lower to 0.6 higher)
#>=BMI 10th Percentile (age-sex corrected)	59 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision	RR 1.93 (0.98 to 3.81)	276 per 1000	257 more per 1000 (from 6 fewer to 775 more)
EDI Total	59 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean edi global in the intervention groups was 0.03 standard deviations higher (0.48 lower to 0.54 higher)
Global Functioning Global Outcome Assessment Scale	59 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean global functioning in the intervention groups was 0.22 standard deviations higher (0.29 lower to 0.74 higher)
Amenorrheic patients	59 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision	RR 0.56 (0.33 to 0.96)	655 per 1000	288 fewer per 1000 (from 26 fewer to 439 fewer)
Hospitalizations to EoT	59 (1 study)	⊕⊕⊖⊖ LOW2 due to imprecision	RR 0.69 (0.37 to 1.3)	483 per 1000	150 fewer per 1000 (from 304 fewer to 145 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Family Therapy-ED (95% CI)

1 CI crosses either 0.75 or 1.25 (Risk Ratio), or either -0.5 or -0.5 (SMD).
2 CI crosses both 0.75 and 1.25 (Risk Ratio).

Table 87: Summary table of findings for family therapy-ED versus any other type of family intervention in young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Any other type of family intervention	Risk difference with Family Therapy-ED (95% CI)
% of Ideal Body Weight	25 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean % of ideal body weight in the intervention groups was 0.62 standard deviations lower (1.43 lower to 0.19 higher)
EDI Bulimia	25 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi bulimia in the intervention groups was 0.54 standard deviations lower (1.34 lower to 0.26 higher)
EDI Drive for Thinness	25 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi drive for thinness in the intervention groups was 0.13 standard deviations lower (0.91 lower to 0.66 higher)
EDI Body Dissatisfaction	25 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean edi body dissatisfaction in the intervention groups was 0.2 standard deviations lower (0.99 lower to 0.59 higher)
General Psychopathology BSI GSI	25 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0 standard deviations higher (0.78 lower to 0.78 higher)
Depression CDI	25 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean depression in the intervention groups was 0.5 standard deviations lower (1.3 lower to 0.3 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Any other type of family intervention	Risk difference with Family Therapy-ED (95% CI)
Family Functioning FAM-III	25 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean family functioning in the intervention groups was 0.43 standard deviations lower (1.23 lower to 0.37 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up				
1 Geist 2000: Unclear randomization method, allocation concealment, no participant blinding, unclear assessor blinding. *2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).* *3 CI crosses both 0.74 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).*				

Table 88: Summary table of findings for family therapy-ED versus general family therapy in young people with anorexia nervosa at end of treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Family Therapy	Risk difference with Family Therapy-ED (95% CI)
Remission (ITT) % of patients achieving ≥ 95% IBW1	164 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision	RR 1.3 (0.79 to 2.14)	244 per 1000	73 more per 1000 (from 51 fewer to 278 more)
% of Ideal Body Weight	158 (1 study) 12 months	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean % of ideal body weight in the intervention groups was 0.16 standard deviations higher (0.15 lower to 0.47 higher)
EDE Global	158 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.26 standard deviations lower (0.58 lower to 0.05 higher)
Yale-Brown-Cornell Eating Disorder Scale	158 (1 study) 12 months	⊕⊕⊖⊖ LOW2,4 due to risk of bias,		Not calculable for SMD values	The mean Yale-Brown-Cornell eating disorder scale in the intervention groups was 0.18 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Family Therapy	Risk difference with Family Therapy-ED (95% CI)
		imprecision			(0.49 lower to 0.13 higher)
Depression BDI	158 (1 study) 12 months	⨁⨁⨀⨀ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.09 standard deviations higher (0.22 lower to 0.4 higher)
Quality of Life Quality of Life and Enjoyment Scale (Short-Form)	158 (1 study) 12 months	⨁⨁⨀⨀ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.15 standard deviations lower (0.46 lower to 0.16 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Combines data for 'full remission' and 'partial remission'. **2 Agras 2014: dropout rate for both arms>20% (Family Therapy 26%, Systematic Family Therapy 25%).** **3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).** **4 <400 participants.**					

Table 89: Summary table of findings for family therapy-ED versus general family therapy in young people with anorexia nervosa at follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Family Therapy	Risk difference with Family Therapy-ED (95% CI)
Remission FU (ITT) % of patients achieving ≥ 95% IBW	164 (1 study)	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 1.03 (0.7 to 1.52)	378 per 1000	11 more per 1000 (from 113 fewer to 197 more)
% of Ideal Body Weight FU	158 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean % of ideal body weight fu in the intervention groups was 0.16 standard deviations higher (0.15 lower to 0.47 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Family Therapy	Risk difference with Family Therapy-ED (95% CI)
EDE Global FU	158 (1 study)	⨁⨁⨀⨀ LOW1 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.26 standard deviations lower (0.58 lower to 0.05 higher)
Yale-Brown-Cornell Eating Disorder Scale FU	158 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean Yale-Brown-Cornell eating disorder scale fu in the intervention groups was 0.18 standard deviations lower (0.49 lower to 0.13 higher)
Depression FU BDI	158 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.09 standard deviations higher (0.22 lower to 0.4 higher)
Quality of Life FU Quality of Life and Enjoyment Scale (Short-Form)	158 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life fu in the intervention groups was 0.15 standard deviations lower (0.46 lower to 0.16 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Agras 2014: dropout rate for both arms>20% (Family Therapy 26%, Systematic Family Therapy 25%).
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 <400 participants.

Table 90: Summary table of findings for multi-family therapy-ED versus family therapy-ED at end of treatment in young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy	Risk difference with Multi-Family Therapy (95% CI)
Remission (ITT)	167 (1 study) 6 months	⨁⨁⨀⨀ LOW1,2 due to indirectness,	RR 1.31 (1.05 to 1.62)	585 per 1000	181 more per 1000 (from 29 more to 363 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy	Risk difference with Multi-Family Therapy (95% CI)
		imprecision			
BMI - Change Scores	167 (1 study) 6 months	⨁⨁⨀⨀ LOW1,2 due to indirectness, imprecision		Not calculable for SMD values	The mean BMI - change scores in the intervention groups was 0.39 standard deviations higher (0.09 to 0.7 higher)
%mBMI - Change Scores	167 (1 study) 6 weeks	⨁⨁⨀⨀ LOW1,2 due to indirectness, imprecision		Not calculable for SMD values	The mean %mbmi - change scores in the intervention groups was 0.45 standard deviations higher (0.14 to 0.75 higher)
EDE Restraint - Change scores	167 (1 study) 6 weeks	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede restraint - change scores in the intervention groups was 0.38 standard deviations higher (0.08 to 0.69 higher)
EDE Eating Concerns - Change scores	167 (1 study) 6 months	⨁⨀⨀⨀ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede eating concerns - change scores in the intervention groups was 0.12 standard deviations higher (0.18 lower to 0.43 higher)
EDE Shape Concerns - Change scores	167 (1 study) 6 weeks	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede shape concerns - change scores in the intervention groups was 0.42 standard deviations higher (0.11 to 0.72 higher)
EDE Weight Concerns - Change scores	167 (1 study) 6 months	⨁⨀⨀⨀ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concerns - change scores in the intervention groups was 0.35 standard deviations higher (0.04 to 0.65 higher)
Depression - Change scores	167 (1 study)	⨁⨀⨀⨀ VERY LOW1,2,3		Not calculable	The mean depression - change scores in the intervention groups was 0.28 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy	Risk difference with Multi-Family Therapy (95% CI)
	6 weeks	due to risk of bias, indirectness, imprecision		for SMD values	(0.02 lower to 0.59 higher)
Carer - Experience of Caregiving - Positive - Change scores	167 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean carer - experience of caregiving - positive - change scores in the intervention groups was 0.15 standard deviations higher (0.16 lower to 0.45 higher)
Carer - Experience of Caregiving - Negative - Change scores	167 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean carer - experience of caregiving - negative - change scores in the intervention groups was 0.09 standard deviations lower (0.39 lower to 0.22 higher)
Service user experience - young person Client Satisfaction Questionnaire score 27-32	79 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision	RR 0.88 (0.47 to 1.65)	351 per 1000	42 fewer per 1000 (from 186 fewer to 228 more)
Service user experience - carer Client Satisfaction Questionnaire score 27-32	96 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,5 due to risk of bias, indirectness, imprecision	RR 1.03 (0.73 to 1.45)	574 per 1000	17 more per 1000 (from 155 fewer to 259 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Sample consists of 120 AN and 40 Restricting EDNOS participants.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 Eisler 2016: no participant nor investigator blinding.
4 <400 participants (continuous outcome).
5 CI crosses both 0.75 and 1.25 (Risk Ratio).

Table 91: Summary table of findings for multi-family therapy-ED versus family therapy-ED in young people with anorexia nervosa at follow up

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with Family Therapy	Risk difference with Multi-Family Therapy (95% CI)
Remission FU (ITT)	167 (1 study)	⊕⊕⊖⊖ LOW1,2 due to indirectness, imprecision	RR 1.35 (1.09 to 1.69)	573 per 1000	201 more per 1000 (from 52 more to 395 more)
BMI FU - Change Scores	167 (1 study)	⊕⊕⊖⊖ LOW1,2 due to indirectness, imprecision		Not calculable for SMD values	The mean BMI fu - change scores in the intervention groups was 0.67 standard deviations higher (0.35 to 0.98 higher)
%mBMI FU - Change Scores	167 (1 study)	⊕⊕⊖⊖ LOW1,2 due to indirectness, imprecision		Not calculable for SMD values	The mean %mbmi fu - change scores in the intervention groups was 0.4 standard deviations higher (0.09 to 0.71 higher)
EDE Restraint FU - Change scores	167 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede restraint fu - change scores in the intervention groups was 0.37 standard deviations higher (0.06 to 0.67 higher)
EDE Eating Concerns FU - Change scores	167 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede eating concerns fu - change scores in the intervention groups was 0.17 standard deviations higher (0.13 lower to 0.48 higher)
EDE Shape Concerns FU - Change scores	167 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede shape concerns fu - change scores in the intervention groups was 0.42 standard deviations higher (0.12 to 0.73 higher)
EDE Weight Concerns FU - Change scores	167 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concerns fu - change scores in the intervention groups was 0.35 standard deviations higher (0.05 to 0.66 higher)
Depression FU - Change scores	167 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression fu - change scores in the intervention groups was 0.2 standard deviations higher (0.11 lower to 0.5 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy	Risk difference with Multi-Family Therapy (95% CI)
CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Sample consists of 120 AN and 40 Restricting EDNOS participants. *2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).* *3 Eisler 2016: no participant nor investigator blinding.* *4 <400 participants (continuous outcome).*					

Table 92: Summary table of findings for family therapy-ED versus any individual therapy in young people with anorexia nervosa at end of treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Individual Therapy	Risk difference with Family Therapy-ED (95% CI)
Remission (ITT) See footnote.1	179 (3 studies) 5 years	⊕⊖⊖⊖ VERY LOW2,3,4,5,6 due to risk of bias, inconsistency, imprecision	RR 1.45 (0.82 to 2.59)	506 per 1000	228 more per 1000 (from 91 fewer to 804 more)
BMI/Weight	160 (3 studies) 5 years	⊕⊕⊖⊖ LOW2,3,4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI/weight in the intervention groups was 0.51 standard deviations higher (0.19 to 0.82 higher)
Morgan-Russell Average Score	21 (1 study) 5 years	⊕⊕⊖⊖ LOW4,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean Morgan Russell average score in the intervention groups was 1.92 standard deviations higher (0.85 to 2.99 higher)
EDE Global	103 (1 study) 12 months	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.45 standard deviations lower (0.84 to 0.05 lower)
Depression Beck Depression Inventory	35 (1 study) 12 months	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.35 standard deviations higher (0.32 lower to 1.02 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Individual Therapy	Risk difference with Family Therapy-ED (95% CI)
Carer Family Functioning - Conflict PARQ Mother + Father	65 (1 study) 12 months	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning - conflict in the intervention groups was 0.04 standard deviations lower (0.53 lower to 0.44 higher)
Carer Family Functioning – Communication McMaster Family Assessment Device	84 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning – communication in the intervention groups was 0.48 standard deviations lower (0.92 to 0.05 lower)
Carer Family Functioning – Behaviour Control McMaster Family Assessment Device	84 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning – behaviour control in the intervention groups was 0.59 standard deviations lower (1.03 to 0.16 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 'Remission' here defined as follows: Lock 2010/Ciao 2014: All Ps who achieve weight more than 85% of expected IBW for sex, age and height (inc. full remission Ps and/or all Ps achieving 95% or greater IBW though who have elevated EDE scores (similar to Morgan-Russell intermediate outcome). Robin 1999: Morgan-Russell Good or Intermediate outcome (data from Eisler, I. (2005). The empirical and theoretical base of family therapy and multiple family day therapy for young people anorexia nervosa. Journal of Family Therapy, 27, 104-131). Russell 1987: Morgan-Russell Good or Intermediate outcomes.
2 Lock 2010/Ciao 2014: No participant blinding.
3 Robin 1999: inadequate randomization method, unclear allocation concealment, participant and assessor blinding, dropout data not provided.
4 Russell 1987/Eisler 1997: Unclear randomization method, allocation method, participant blinding, dropout rate both arms>20% (Family Therapy 40%, Individual Therapy 64%).
5 I2>=50%.
6 CI crosses 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
7 <400 participants.

Table 93: Summary table of findings for family therapy-ED versus any individual therapy in young people with anorexia nervosa at follow up

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Individual Therapy	Risk difference with Family Therapy-ED (95% CI)
Remission FU (ITT)	179 (3 studies) 5 years	⊕⊕⊖⊖ LOW1,2,3,4 due to risk of bias, imprecision	RR 1.01 (0.8 to 1.27)	618 per 1000	6 more per 1000 (from 124 fewer to 167 more)
BMI or Weight FU	150 (3 studies) 5 years	⊕⊕⊖⊖ LOW1,2,3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI or weight fu in the intervention groups was 0.24 standard deviations higher (0.08 lower to 0.56 higher)
EDE Global FU	93 (1 study) 12 months	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.23 standard deviations lower (0.63 lower to 0.18 higher)
Depression FU Beck Depression Inventory	35 (1 study) 12 months	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.87 standard deviations higher (0.17 to 1.57 higher)
Carer Family Functioning FU PARQ Mother +Father	65 (1 study) 12 months	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning fu in the intervention groups was 0.03 standard deviations higher (0.46 lower to 0.52 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Lock 2010: No participant blinding.
2 Robin 1999: inadequate randomization method, unclear allocation concealment, participant and assessor blinding, dropout data not provided.
3 Russell 1987/Eisler 1997: Unclear randomization method, allocation method, participant blinding, dropout rate both arms>20% (Family Therapy 40%, Individual Therapy 64%).
4 CI crosses 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 94: Summary table of findings for family therapy-ED-1 (conjoint family therapy) versus family therapy-ED-2 (separated family therapy) at end of treatment in young people with anorexia nervosa

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Family Therapy-ED (2)	Risk difference with Family Therapy-ED (1((95% CI)
Full Remission (ITT) Morgan-Russell Good outcome; >=95% mBMI and EDE global <= 1.59	146 (2 studies) 12 months	⊕⊕⊖⊖ LOW1,2,3 due to risk of bias, imprecision	RR 0.52 (0.32 to 0.85)	444 per 1000	213 fewer per 1000 (from 67 fewer to 302 fewer)
BMI	146 (2 studies) 12 months	⊕⊕⊖⊖ LOW1,2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.34 standard deviations lower (0.67 to 0.02 lower)
% of Average Body Weight (change scores)	40 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean % of average body weight (change scores) in the intervention groups was 0.42 standard deviations lower (1.05 lower to 0.21 higher)
Morgan-Russell Outcome-Average	40 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean Morgan-Russell outcome-average in the intervention groups was 0.29 standard deviations higher (0.34 lower to 0.91 higher)
EDE Global	106 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.23 standard deviations higher (0.16 lower to 0.61 higher)
EDE Restraint	106 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint in the intervention groups was 0.21 standard deviations higher (0.17 lower to 0.59 higher)
EDE Eating Concerns	106 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concerns in the intervention groups was 0.13 standard deviations higher (0.26 lower to 0.51 higher)
EDE Weight Concerns	106 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concerns in the intervention groups was 0.26 standard deviations higher (0.12 lower to 0.64 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy-ED (2)	Risk difference with Family Therapy-ED (1((95% CI)
EDE Shape Concerns	106 (1 study) 12 months	⨁⨁⨀⨀ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concerns in the intervention groups was 0.25 standard deviations higher (0.13 lower to 0.63 higher)
Hospitalized during treatment	106 (1 study)	⨁⨁⨀⨀ LOW2,3 due to risk of bias, imprecision	RR 2.01 (0.83 to 4.89)	118 per 1000	119 more per 1000 (from 20 fewer to 458 more)
Depression Scale analogous to Morgan-Russell; CDI	146 (2 studies)	⨁⨁⨀⨀ LOW1,2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.12 standard deviations lower (0.44 lower to 0.21 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Eisler 2000: unclear randomization method, allocation concealment, participant blinding.
2 Le Grange 2016: no participant nor investigator blinding.
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
4 <400 participants.

Table 95: Summary table of findings for family therapy-ED-1 (conjoint family therapy) versus family therapy-ED-2 (separated family therapy) at follow up in young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy-ED2	Risk difference with Family Therapy-ED1 (95% CI)
Full Remission (ITT) 12-mo FU >=95% mBMI and EDE global <= 1.59	106 (1 study)	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 0.78 (0.45 to 1.35)	373 per 1000	82 fewer per 1000 (from 205 fewer to 130 more)
%mBMI 12-mo FU	106 (1 study)	⨁⨁⨀⨀ LOW1,2		Not calculable for	The mean %mBMI 12-mo fu in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy-ED2	Risk difference with Family Therapy-ED1 (95% CI)
		due to risk of bias, imprecision		SMD values	0.23 standard deviations lower (0.61 lower to 0.15 higher)
EDE Global 12-mo FU	106 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global 12-mo fu in the intervention groups was 0.19 standard deviations higher (0.19 lower to 0.57 higher)
EDE Restraint 12-mo FU	106 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint 12-mo fu in the intervention groups was 0.2 standard deviations higher (0.18 lower to 0.58 higher)
EDE Eating Concerns 12-mo FU	106 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concerns 12-mo fu in the intervention groups was 0.12 standard deviations higher (0.26 lower to 0.5 higher)
EDE Weight Concerns 12-mo FU	106 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concerns 12-mo fu in the intervention groups was 0.13 standard deviations higher (0.25 lower to 0.51 higher)
EDE Shape Concerns 12-mo FU	106 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concerns 12-mo fu in the intervention groups was 0.2 standard deviations higher (0.18 lower to 0.58 higher)
Depression 12-mo FU	106 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression 12-mo fu in the intervention groups was 0.42 standard deviations higher (0.04 to 0.81 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Le Grange 2016: no participant nor investigator blinding.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 <400 participants.

Table 96: Summary table of findings for long-term family therapy-ED versus short-term family therapy-ED at end of treatment in young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Family Therapy-ED Short-Term	Risk difference with Family Therapy-ED Long-Term (95% CI)
BMI	86 (1 study) 3.96 years	⨁⨁⨁⨀ MODERATE1 due to imprecision	Not calculable for SMD values	The mean BMI in the intervention groups was 0.22 standard deviations higher (0.2 lower to 0.65 higher)
EDE Restraint	86 (1 study) 3.96 years	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede restraint in the intervention groups was 0.24 standard deviations lower (0.67 lower to 0.18 higher)
EDE Weight Concerns	86 (1 study) 3.96 years	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede weight concerns in the intervention groups was 0.42 standard deviations lower (0.85 lower to 0.01 higher)
EDE Eating Concerns	86 (1 study) 3.96 years	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede eating concerns in the intervention groups was 0.36 standard deviations lower (0.79 lower to 0.06 higher)
EDE Shape Concerns	86 (1 study) 3.96 years	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean ede shape concerns in the intervention groups was 0.29 standard deviations lower (0.72 lower to 0.13 higher)
Yale-Brown-Cornell Eating Disorder Scale	86 (1 study) 3.96 years	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean Yale-Brown-Cornell eating disorder scale in the intervention groups was 0.54 standard deviations lower (0.97 to 0.11 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
2 Lock 2005/2006: Participant not blind, assessor blinding unclear.

Table 97: Summary table of findings for long-term family therapy-ED versus short-term family therapy-ED at follow up in young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy-ED Short-term	Risk difference with Family Therapy-ED Long-term (95% CI)
BMI (unadjusted) FU	71 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean BMI (unadjusted) fu in the intervention groups was 0.08 standard deviations higher (0.39 lower to 0.54 higher)
BMI>20 FU	71 (1 study)	⊕⊕⊖⊖ LOW2 due to imprecision	RR 0.91 (0.63 to 1.31)	649 per 1000	58 fewer per 1000 (from 240 fewer to 201 more)
# >90% Ideal BW FU	71 (1 study)	⊕⊕⊕⊖ MODERATE3 due to imprecision	RR 1.05 (0.89 to 1.24)	865 per 1000	43 more per 1000 (from 95 fewer to 208 more)
Resumed Menstruation FU	71 (1 study)	⊕⊕⊖⊖ LOW2 due to imprecision	RR 0.98 (0.63 to 1.51)	541 per 1000	11 fewer per 1000 (from 200 fewer to 276 more)
Amenorrheic patients FU	71 (1 study)	⊕⊕⊖⊖ LOW2 due to imprecision	RR 0.36 (0.04 to 3.32)	81 per 1000	52 fewer per 1000 (from 78 fewer to 188 more)
EDE Eating Concerns FU	35 (1 study)	⊕⊖⊖⊖ VERY LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concerns fu in the intervention groups was 0.06 standard deviations lower (0.73 lower to 0.61 higher)
EDE Restraint FU	35 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint fu in the intervention groups was 0.39 standard deviations lower (1.06 lower to 0.29 higher)
EDE Weight Concerns FU	35 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concerns fu in the intervention groups was 0.32 standard deviations lower (1 lower to 0.35 higher)
EDE Shape Concerns FU	35 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias,		Not calculable for SMD values	The mean ede shape concerns fu in the intervention groups was 0.39 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy-ED Short-term	Risk difference with Family Therapy-ED Long-term (95% CI)
		imprecision			(1.07 lower to 0.28 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
2 CI crosses both 0.75 and 1.25 (Risk Ratio).
3 <300 events.
4 Lock 2005/2006: Participant not blind, assessor blinding unclear.

Table 98: Summary table of findings for family therapy-ED with family meal versus family therapy-ED without family meal at end of treatment in young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy without Family Meal	Risk difference with Family Therapy with Family Meal (95% CI)
Remission Morgan-Russell Good or Intermediate outcome	23 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 2.18 (1.09 to 4.37)	417 per 1000	492 more per 1000 (from 38 more to 1000 more)
Weight (kg)	23 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) in the intervention groups was 0.31 standard deviations lower (1.13 lower to 0.52 higher)
% EBW	23 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean % ebw in the intervention groups was 0.41 standard deviations higher (0.42 lower to 1.23 higher)
Morgan-Russell Outcome - Average score	23 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean Morgan-Russell outcome - average score in the intervention groups was 0.15 standard deviations lower (0.97 lower to 0.67 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy without Family Meal	Risk difference with Family Therapy with Family Meal (95% CI)
EDI-2	23 (1 study) 6 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 in the intervention groups was 0.6 standard deviations higher (0.24 lower to 1.44 higher)
General Psychopathology SCL90-R GSI	23 (1 study) 6 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0.92 standard deviations higher (0.05 to 1.79 higher)
Menstruation resumed	21 (1 study) 6 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 2.93 (1.06 to 8.08)	273 per 1000	526 more per 1000 (from 16 more to 1000 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Herscovici 2015: unclear allocation concealment; no participant, investigator nor assessor blinding; EDI-2 and SCL-90-R GSI score significantly lower in FT group.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).

Table 99: Summary table of findings for family therapy-ED with family meal versus family therapy-ED without family meal at follow up in young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy without Family Meal	Risk difference with Family Therapy with Family Meal (95% CI)
Remission 6-mo FU Full and partial remission	23 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.45 (0.74 to 2.85)	500 per 1000	225 more per 1000 (from 130 fewer to 925 more)
Weight 6-mo FU	21 (1 study)	⨁⊖⊖⊖ VERY LOW1,2		Not calculable for SMD values	The mean weight 6-mo fu in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Family Therapy without Family Meal	Risk difference with Family Therapy with Family Meal (95% CI)
		due to risk of bias, imprecision			0.23 standard deviations lower (1.09 lower to 0.63 higher)
% EBW 6-mo FU	21 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean % ebw 6-mo fu in the intervention groups was 0.43 standard deviations higher (0.44 lower to 1.3 higher)
Morgan-Russell Outcome - Average score 6-mo FU	21 (1 study)	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean morgan-russell outcome - average score 6-mo fu in the intervention groups was 0.05 standard deviations higher (0.81 lower to 0.9 higher)
EDI-2 6-mo FU	21 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 6-mo fu in the intervention groups was 0.54 standard deviations higher (0.34 lower to 1.41 higher)
General Psychopathology 6-mo FU	21 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology 6-mo fu in the intervention groups was 0.78 standard deviations higher (0.13 lower to 1.66 higher)
Menstruation resumed 6-mo FU	20 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	RR 2.14 (0.91 to 5.04)	364 per 1000	415 more per 1000 (from 33 fewer to 1000 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Herscovici 2015: unclear allocation concealment; no participant, investigator nor assessor blinding; EDI-2 and SCL-90-R GSI score significantly lower in FT group.
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

6.3.91 Family therapy in adults

2 **Table 100: Summary table of findings for family therapy-ED versus any other family intervention at end of treatment in adult**
3 **inpatients with anorexia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Any other type of family intervention	Risk difference with Family Therapy-ED (95% CI)
BMI	47 (1 study) 36 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean bmi in the intervention groups was 0.43 standard deviations lower (1.01 lower to 0.15 higher)
SEED Anorexia Severity Scale Scale from: 0 to 3.	25 (1 study) 36 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean seed anorexia severity scale in the intervention groups was 0.2 standard deviations higher (0.61 lower to 1 higher)
SEED Bulimia Severity Scale Scale from: 0 to 3.	25 (1 study) 36 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean seed bulimia severity scale in the intervention groups was 0.48 standard deviations higher (0.34 lower to 1.29 higher)
Carer Quality of Life GHQ-12 Short Form. Scale from: 0 to 36.	77 (1 study) 36 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer quality of life in the intervention groups was 0.08 standard deviations higher (0.37 lower to 0.53 higher)
Carer Family Functioning Level of Expressed Emotion	66 (1 study) 36 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer family functioning in the intervention groups was 0.13 standard deviations higher (0.35 lower to 0.61 higher)
Carer Experience of Caregiving Inventory (ECI) Negative	75 (1 study) 36 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer experience of caregiving inventory (eci) negative in the intervention groups was 0.43 standard deviations lower (0.89 lower to 0.03 higher)
Carer Experience of Caregiving Inventory (ECI) Positive	75 (1 study) 36 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer experience of caregiving inventory (eci) positive in the intervention groups was 0.53 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Any other type of family intervention	Risk difference with Family Therapy-ED (95% CI)
				(0.99 to 0.06 lower)

**The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Whitney 2012: Unclear whether baseline properties of two arms similar. No participant nor assessor blinding.
2 CI crosses either 0.5 or -0.5 (SMD).
3 CI crosses both 0.5 and -0.5 (SMD).

Table 101: Summary table of findings for family therapy-ED versus any other family intervention at follow up in adult inpatients with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Any other family intervention	Risk difference with Family Therapy-ED (95% CI)
BMI FU	44 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.41 standard deviations higher (0.19 lower to 1 higher)
SEED Anorexia Severity Scale FU	29 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean seed anorexia severity scale fu in the intervention groups was 0.24 standard deviations lower (0.97 lower to 0.49 higher)
SEED Bulimia Severity Scale FU	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean seed bulimia severity scale fu in the intervention groups was 0.12 standard deviations higher (0.61 lower to 0.85 higher)
Carer Quality of Life FU GHQ-12 Short Form	69 (1 study) 36 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer quality of life fu in the intervention groups was 0.16 standard deviations lower (0.63 lower to 0.32 higher)
Carer Family Functioning FU Level of Expressed Emotion	58 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,	Not calculable for SMD values	The mean carer expressed emotion fu in the intervention groups was 0.11 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with Any other family intervention	Risk difference with Family Therapy-ED (95% CI)
		imprecision		(0.62 lower to 0.41 higher)
Carer Experience of Caregiving Inventory (ECI) Negative FU	63 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer experience of caregiving inventory (eci) negative fu in the intervention groups was 0.38 standard deviations lower (0.88 lower to 0.12 higher)
Carer Experience of Caregiving Inventory (ECI) Positive FU	63 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer experience of caregiving inventory (eci) positive fu in the intervention groups was 0.23 standard deviations lower (0.73 lower to 0.26 higher)

**The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Whitney 2012: Unclear whether baseline properties of two arms similar. No participant nor assessor blinding.
2 CI crosses either 0.5 or -0.5 (SMD).
3 CI crosses both 0.5 and -0.5 (SMD).

Table 102: Summary table of findings for general family therapy and any individual therapy compared to any nutritional intervention in adults with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any individual intervention	Risk difference with General Family + any individual therapy (95% CI)
Weight (kg)	30 (1 study) 12 months	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) in the intervention groups was 0.13 standard deviations lower (0.85 lower to 0.59 higher)
Regular Menstruation	30 (1 study) 12 months	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 1 (0.24 to 4.18)	200 per 1000	0 fewer per 1000 (from 152 fewer to 636 more)
Amenorrheic patients	30 (1 study) 12 months	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias,	RR 0.8 (0.44 to 1.45)	667 per 1000	133 fewer per 1000 (from 373 fewer to 300 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any individual intervention	Risk difference with General Family + any individual therapy (95% CI)
		imprecision			
Global Clinical Score	30 (1 study) 12 months	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean global clinical score in the intervention groups was 1.95 standard deviations higher (1.06 to 2.84 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Hall 1987: Randomization method and allocation concealment unclear. Control arm dropout rate was 27%.
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).
3 <400 participants.

1 **Table 103: Summary table of findings for family therapy-ED compared to any individual therapy in adults with anorexia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Individual Therapy	Risk difference with Family Therapy-ED (95% CI)
All-cause Mortality	84 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	RR 1.01 (0.9 to 1.13)	984 per 1000	10 more per 1000 (from 98 fewer to 128 more)
Recovered	84 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	RR 0.94 (0.78 to 1.14)	903 per 1000	54 fewer per 1000 (from 199 fewer to 126 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Dare 2001: Unclear method of randomization and allocation concealment. No participant, investigator nor assessor blinding. Dropout rate>20% for all four arms.
2 <300 events.

Table 104: Summary table of findings for family therapy-ED versus any individual therapy in adults with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Individual Therapy	Risk difference with Family Therapy-ED (95% CI)
All-cause Mortality	84 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.01 (0.9 to 1.13)	984 per 1000	10 more per 1000 (from 98 fewer to 128 more)
Recovered	84 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.94 (0.78 to 1.14)	903 per 1000	54 fewer per 1000 (from 199 fewer to 126 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Dare 2001: Unclear method of randomization and allocation concealment. No participant, investigator nor assessor blinding. Dropout rate>20% for all four arms.
2 <300 events.

6.3.10 Economic Evidence

6.3.10.1 Systematic literature review

The systematic search of the literature identified:

- One study that assessed the cost effectiveness of a family-based treatment (FBT) in young people with anorexia nervosa in the US (Agras et al., 2014).

No studies assessing the cost effectiveness of other psychological interventions in people with AN were identified by the systematic search of the economic literature undertaken for this guideline.

References to all included studies and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix P. Completed methodology checklists of the studies are provided in Appendix O. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix Q.

Agras and colleagues (2014) evaluated the cost effectiveness of a family-based treatment (FBT) compared with systemic family therapy (SyFT) in young people 12 to 18 years with the anorexia nervosa in the US. The economic analysis was conducted alongside an RCT (Agras 2014) (N=158). Both therapies involved 16 one hour sessions delivered over nine months. The analysis was conducted from a health care provider perspective. The study considered treatment costs and hospital admissions. The resource use estimates were based on the RCT (N=158). The unit costs were obtained from state and local sources. The measure of outcome for the economic analysis was the proportion of people in remission at the end of treatment (36 weeks). Remission was defined as ≥95% of IBW. The time horizon of the analysis was 36 weeks.

FBT resulted in a greater proportion of people achieving remission at the end of treatment (36 weeks) compared with SyFT (33% versus 25%, respectively; a difference 8%, p = 0.22). From a health care provider perspective the mean total costs per participant over 36 weeks were $8,963 for FBT and $18,005 for SyFT, a difference of -$9,042 (p-value not reported) in 2007 US dollars. Based on the above, FBT was the dominant intervention (that is, it led to a reduction in costs and a greater proportion of people in remission at the end of treatment).

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it was conducted in the US. The authors did not attempt to estimate quality adjusted life years (QALYs) but interpretation of findings was straightforward, as FBT was found to be dominant. This study was judged by the committee to have potentially serious methodological limitations, including its short time horizon (36 weeks), the fact that significance levels for costs not reported and some unit costs being from local sources.

6.3.11 Clinical evidence statements

6.3.11.1 Individual Therapy

CBT-ED versus another intervention in adults with anorexia nervosa at the end of treatment

Very low quality evidence from two RCTs (n=298) showed that CBT-ED may be more effective at increasing body weight compared with any other treatment but there was some uncertainty.

Very low quality evidence from two RCTs (n=130) showed that CBT-ED may be more effective at improving general psychopathology compared with any other treatment but there was some uncertainty.

Low quality evidence from one RCT (n=56) showed no difference in the effect of CBT-ED on depression compared with any other treatment.

Very low quality evidence from two RCTs (n=275) showed no difference in the effect of CBT-ED on remission compared with any other treatment.

Low quality evidence from one RCT (n=56) showed no difference in the effect of CBT-ED on EDE-restraint, EDE-EDE-weight concern, EDE-shape concern, EDE-eating concern, EDI-drive for thinness, EDI- body dissatisfaction or EDI-bulimia compared with any other treatment.

Low quality evidence from one RCT (n=242) showed no difference in the effect of CBT-ED on EDI-total compared with any other treatment.

CBT-ED versus another intervention in adults with anorexia nervosa at follow up

Very low quality evidence from two RCTs (n=274) showed no difference in the effect of CBT-ED on weight compared with any other treatment.

Very low quality evidence from one RCT (n=242) showed no difference in the effect of CBT-ED on general psychopathology compared with any other treatment.

Low quality evidence from one RCT (n=43) showed no difference in the effect of CBT-ED on general function compared with any other treatment.

Low quality evidence from one RCT (n=43) showed no difference in the effect of CBT-ED on depression compared with any other treatment.

Low quality evidence from one RCT (n=142) showed no difference in the effect of CBT-ED on remission compared with any other treatment.

Low quality evidence from one RCT (n=43) showed no difference in the effect of CBT-ED on EDE-restraint, EDE-EDE-weight concern, EDE-shape concern, EDE-eating concern, EDI-drive for thinness, EDI-body dissatisfaction or EDI-bulimia compared with any other treatment.

Low quality evidence from one RCT (n=242) showed no difference in the effect of CBT-ED on EDI-total compared with any other treatment.

CBT-ED versus another intervention in young people with anorexia nervosa at follow up

Low quality evidence from one RCT (n=98) showed no difference in the effect of CBT-ED on BMI compared with any other treatment.

Low quality evidence from one RCT (n=82) showed no difference in the effect of CBT-ED on EDI-total compared with any other treatment.

Low quality evidence from one RCT (n=110) showed no difference in the effect of CBT-ED on remission compared with any other treatment.

Psychodynamic therapy versus another intervention in adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=242) showed that psychodynamic therapy may be less effective at improving BMI compared with any other treatment but there was some uncertainty.

Low quality evidence from one RCT (n=242) showed no difference in the effect of psychodynamic therapy on EDI-total compared with any other treatment.

Low quality evidence from one RCT (n=242) showed no difference in the effect of psychodynamic therapy on general psychopathology compared with any other treatment.

Very low quality evidence from two RCTs (n=84) showed no difference in the effect of psychodynamic therapy on all-cause mortality compared with any other treatment.

Low quality evidence from two RCTs (n=326) showed psychodynamic therapy is more effective at improving remission compared with any other treatment, but there was some uncertainty.

Psychodynamic therapy versus another intervention in adults with anorexia nervosa at follow up

Low quality evidence from two RCTs (n=302) showed no difference in the effect of psychodynamic therapy on body weight compared with any other treatment.

Low quality evidence from one RCT (n=242) showed no difference in the effect of psychodynamic therapy on EDI-total compared with any other treatment.

Low quality evidence from one RCT (n=30) showed no difference in the effect of psychodynamic therapy on Morgan-Russell symptoms compared with any other treatment.

Low quality evidence from one RCT (n=242) showed no difference in the effect of psychodynamic therapy on general psychopathology compared with any other treatment.

Very low quality evidence from two RCTs (n=84) showed no difference in the effect of psychodynamic therapy on bulimia compared with any other treatment.

Low quality evidence from two RCTs (n=272) showed psychodynamic therapy was more effective on remission rates compared with any other treatment.

Supportive therapy versus another intervention in young people with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=21) showed supportive therapy is less effective on body weight compared with any other treatment.

Low quality evidence from one RCT (n=21) showed supportive therapy is less effective on remission rates compared with any other treatment.

Supportive therapy versus another intervention in young people with anorexia nervosa at follow up

Low quality evidence from one RCT (n=21) showed no difference in body weight between supportive therapy and any other treatment.

Low quality evidence from one RCT (n=21) showed no difference in remission rates between supportive therapy and any other treatment.

Young people focused therapy versus another intervention in young people with anorexia nervosa at end of treatment

Low quality evidence from two RCTs (n=139) showed young people focused therapy is less effective on body weight compared with any other treatment but there was some uncertainty.

Very low quality evidence from two RCTs (n=158) showed young people focused therapy is less effective on remission rates compared with any other treatment but there was some uncertainty.

Young people focused therapy versus another intervention in young people with anorexia nervosa at follow up

Low quality evidence from two RCTs (n=139) showed no difference in body weight between young people focused therapy and any other treatment.

Very low quality evidence from two RCTs (n=158) showed no difference in remission rates between young people focused therapy and any other treatment.

Interpersonal psychotherapy versus another intervention in adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=56) showed no difference in the effect of interpersonal psychotherapy on body weight compared with any other treatment.

Low quality evidence from one RCT (n=56) showed interpersonal psychotherapy was less effective on general psychopathology compared with any other treatment.

Low quality evidence from one RCT (n=56) showed no difference in the effect of interpersonal psychotherapy on depression compared with any other treatment.

Low quality evidence from one RCT (n=56) showed interpersonal psychotherapy was less effective on EDE-restraint and EDE-eating concern compared with any other treatment.

Low quality evidence from one RCT (n=56) showed no difference in the effect of interpersonal psychotherapy on EDE-weight concern and EDE-shape concern compared with any other treatment.

Low quality evidence from one RCT (n=48) showed no difference in the effect of interpersonal psychotherapy on EDI-drive for thinness, EDI-body dissatisfaction or EDI-bulimia compared with any other treatment.

Interpersonal psychotherapy versus another intervention in adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=43) showed no difference in the effect of interpersonal psychotherapy on body weight compared with any other treatment.

Low quality evidence from one RCT (n=43) showed no difference in the effect of interpersonal psychotherapy on general function compared with any other treatment.

Low quality evidence from one RCT (n=56) showed no difference in the effect of interpersonal psychotherapy on depression compared with any other treatment.

Very low to low quality evidence from one RCT (n=43) showed no difference in the effect of interpersonal psychotherapy on EDE-restraint, EDE-weight concern, EDE-shape concern, EDE-eating concern, EDI-drive for thinness, EDI-body dissatisfaction and EDI-bulimia compared with any other treatment.

SSCM versus another intervention in adults with anorexia nervosa at end of treatment

Low quality evidence from two RCTs (n=285) showed no difference in the effect of SSCM on BMI compared with any other treatment.

Low quality evidence from two RCTs (n=213) showed no difference in the effect of SSCM on EDE-global compared with any other treatment.

Low quality evidence from three RCTs (n=269) showed SSCM is more effective on depression compared with any other treatment but there was some uncertainty.

Low quality evidence from one RCT (n=56) showed SSCM is more effective on general function compared with any other treatment.

Very low quality evidence from one RCT (n=213) showed no difference in the effect of SSCM on remission compared with any other treatment.

Very low quality evidence from two RCTs (n=198) showed no difference in the effect of SSCM on EDE-restraint compared with any other treatment.

Low quality evidence from two RCTs (n=198) showed no difference in the effect of SSCM on EDE-weight concern, EDE-shape concern and EDE-eating concern compared with any other treatment.

Low quality evidence from one RCT (n=56) showed no difference in the effect of SSCM on EDI-drive for thinness, EDI-body dissatisfaction and EDI-bulimia compared with any other treatment.

SSCM versus another intervention in adults with anorexia nervosa at follow up

Low quality evidence from three RCTs (n=286) showed no difference in the effect of SSCM on BMI compared with any other treatment.

Low quality evidence from two RCTs (n=213) showed no difference in the effect of SSCM on EDE-global compared with any other treatment.

Low quality evidence from three RCTs (n=256) showed no difference in the effect of SSCM on depression compared with any other treatment.

Very low quality evidence from one RCT (n=43) showed no difference in the effect of SSCM on general function compared with any other treatment.

Very low quality evidence from one RCT (n=43) showed no difference in the effect of SSCM on bulimia compared with any other treatment.

Very low quality evidence from one RCT (n=43) showed no difference in the effect of SSCM on remission rates compared with any other treatment.

Low quality evidence from two RCTs (n=198) showed SSCM is more effective on EDE-restraint compared with any other treatment but there was some uncertainty.

Low quality evidence from two RCTs (n=185) showed SSCM is more effective on EDE-restraint compared with any other treatment.

Low quality evidence from two RCTs (n=185) showed no difference in the effect of SSCM on EDE-weight concern, EDE-shape concern and EDE-eating concern compared with any other treatment.

Low quality evidence from one RCT (n=43) showed no difference in the effect of SSCM on EDI-drive for thinness, EDI-body dissatisfaction and EDI–bulimia compared with any other treatment.

Inpatient CBT-ED versus another inpatient CBT-ED in adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=72) showed no difference in the effect of one type of inpatient CBT-ED on BMI, general psychiatric features, EDE-restraint, EDE-weight concern, EDE-shape concern and EDE-eating concern compared with another inpatient CBT-ED therapy.

Inpatient CBT-ED versus another inpatient CBT-ED in adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=68) showed no difference in the effect of one type of inpatient CBT-ED on BMI, general psychiatric features, EDE-restraint, EDE-weight concern, EDE-shape concern and EDE-eating concern compared with another inpatient CBT-ED therapy.

SSCM versus another interventions in adults with severe and enduring anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=63) showed no difference in the effect of SSCM on BMI and depression compared with any other treatment.

Low quality evidence from two RCTs (n=213) showed no difference in the effect of SSCM on EDE-Global and quality of life compared with any other treatment.

SSCM versus another interventions in adults with severe and enduring anorexia nervosa at follow up

Low quality evidence from one RCT (n=63) showed no difference in the effect of SSCM on BMI, depression, EDE-global and quality of life compared with any other treatment.

MANTRA versus another interventions in adults with anorexia nervosa at the end of treatment

Low quality evidence from two RCTs (n=213) showed no difference in the effect of MANTRA on BMI, depression, EDE-global and quality of life compared with any other treatment.

CBT-ED versus another interventions in adults with severe and enduring anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=63) showed no difference in the effect of CBT-ED on BMI, depression, EDE-global and quality of life compared with any other treatment.

CBT-ED versus another interventions in adults with severe and enduring anorexia nervosa at follow up

Low quality evidence from one RCT (n=63) showed no difference in the effect of CBT-ED on BMI, depression and quality of life compared with any other treatment.

Low quality evidence from one RCT (n=63) showed CBT-ED is more effective on improving EDE-global compared with any other treatment.

Psychiatric counselling versus another interventions in adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=104) showed no difference in the effect of psychiatric counselling on all-cause mortality and remission compared with any other treatment.

6.3.11.2 Group Therapy

No evidence on group therapy in people with anorexia nervosa was identified.

6.3.11.3 Self-help

Internet guided self-help versus treatment as usual in adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=219 to 226) showed internet guided self-help is more effective on EDI-total and bulimic symptoms compared with treatment as usual.

Low quality evidence from one RCT (n=219 to 239) showed internet guided self-help is more effective on EDI-drive for thinness, EDI-body dissatisfaction, depression, global clinical score, Morgan-Russell menstrual function compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=219) showed no difference in the effect of internet guided self-help on EDI-bulimia and general psychopathology compared with treatment as usual, although there was some uncertainty

Internet guided self-help versus treatment as usual in adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=208) showed no difference in the effect of internet guided self-help on Morgan-Russell menstrual function and general psychopathology compared with treatment as usual.

Low quality evidence from one RCT (n=208) showed internet guided self-help is more effective on bulimic symptoms compared with treatment as usual, although there was some uncertainty.

6.3.11.4 Family Therapy

Family therapy-ED and treatment as usual versus treatment as usual in young people with anorexia nervosa at end of treatment

Moderate quality evidence from one RCT (n=60) showed family therapy-ED may be more effective on remission and the number of people in the BMI 10th percentile (age-sex corrected) compared with treatment as usual, although there was some uncertainty.

Moderate to low quality evidence from one RCT (n=59) showed no difference in the effect of family therapy-ED on BMI, EDI-total, global functioning and number of people hospitalised during treatment compared with treatment as usual.

Moderate quality evidence from one RCT (n=59) showed family therapy-ED may be more effective on reducing the number of amenorrheic patients compared with treatment as usual, although there was some uncertainty.

Family therapy-ED versus any other type of family intervention in young people with anorexia nervosa at end of treatment

Very low to low quality evidence from one RCT (n=25) showed no difference in the effect of family therapy-ED on % ideal body weight, EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction, general psychopathology, depression and family functioning compared with any other type of family intervention.

Family therapy-ED versus general family therapy in young people with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=164) showed family therapy-ED may be less effective on full remission compared with general family therapy although there was some uncertainty.

Low quality evidence from one RCT (n=158) showed family therapy-ED may be more effective on EDE-global compared with general family therapy, although there was some uncertainty.

Low quality evidence from one RCT (n=158) showed no difference in the effect of family therapy-ED on % ideal body weight, YBC-EDS scores, depression and quality of life compared with general family therapy.

Family therapy-ED versus general family therapy in young people with anorexia nervosa at follow up

Very low quality evidence from one RCT (n=164) showed no difference in the effect of family therapy-ED on remission compared with general family therapy.

Low quality evidence from one RCT (n=158) showed no difference in the effect of family therapy-ED on % ideal body weight , YBC-EDS score, depression and quality of life compared with general family therapy.

Low quality evidence from one RCT (n=158) showed family therapy-ED may be more effective on EDE-global compared with general family therapy, although there was some uncertainty.

Multi-family therapy-ED versus family therapy-ED in young people with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=167) showed multi-family therapy-ED is more effective on remission and change in BMI compared with family therapy-ED.

Very low quality evidence from one RCT (n=167) showed multi-family therapy-ED is less effective on change in EDE-restraint, EDE-shape concern and EDE-weight concern compared with family therapy-ED.

Very low quality evidence from one RCT (n=167) showed no difference in the effect of multi-family therapy-ED on change in EDE-eating concerns and both positive and negative experience of caregiving compared with family therapy-ED.

Very low quality evidence from one RCT (n=167) showed multi- family therapy-ED may be less effective on change in depression compared with family therapy-ED, although there was some uncertainty.

Very low quality evidence form one RCT (n=79) showed no difference in the effect of multi-family therapy-ED on the young person's service user compared with family therapy-ED.

Very low quality evidence form one RCT (n=96) showed no difference in the effect of multi-family therapy-ED on the carer's service user compared with family therapy-ED.

Multi-family therapy-ED versus family therapy-ED in young people with anorexia nervosa at follow up

Low quality evidence from one RCT (n=167) showed multi-family therapy-ED was more effective on remission and change in BMI compared with family therapy-ED.

Very low quality evidence from one RCT (n=167) showed multi-family therapy-ED is less effective on change in both EDE-restraint, EDE-shape concern and EDE-weight concern compared with family therapy-ED.

Very low quality evidence from one RCT (n=167) showed no difference in the effect of multi-family therapy-ED on EDE-eating concerns and depression compared with family therapy-ED.

Family therapy-ED versus individual therapy in young people with anorexia nervosa at end of treatment

Very low quality evidence from three RCTs (n=179) showed family therapy-ED may be more effective on remission compared with individual therapy, although there was some uncertainty.

Low quality evidence from three RCTs (n=160) showed family therapy-ED is more effective on BMI/Weight compared with individual therapy.

Low quality evidence from one RCT (n=21) showed family therapy-ED is more effective on Morgan-Russell Average score compared with individual therapy.

Low quality evidence from one RCT (n=103) showed family therapy-ED is more effective on EDE-global compared with individual therapy.

Low quality evidence from one RCT (n=35) showed no difference in the effect of family therapy-ED on depression compared with individual therapy.

Low quality evidence from one RCT (n=65) showed no difference in the effect of family therapy-ED on family functioning-conflict compared with individual therapy.

Low quality evidence from one RCT (n=84) showed family therapy-ED is more effective on family functioning-communication and family functioning-behaviour control compared with individual therapy.

Family therapy-ED versus individual therapy in young people with anorexia nervosa at follow up

Low quality evidence from three RCTs (n=179) showed no difference in the effect of family therapy-ED on remission compared with individual therapy.

Low quality evidence from three RCTs (n=150) showed family therapy-ED may be more effective on BMI/Weight compared with any individual therapy, although there was some uncertainty.

Low quality evidence from one RCT (n=93) showed no difference in the effect of family therapy-ED on EDE-global compared with individual therapy.

Low quality evidence from one RCT (n=35) showed family therapy-ED is less effective on depression compared with individual therapy.

Low quality evidence from one RCT (n=65) showed no difference in the effect of family therapy-ED on carer family functioning compared with individual therapy.

Family therapy-ED 1 (conjoint family therapy) versus family therapy-ED 2 (separated family therapy) at end of treatment in young people with anorexia nervosa

Low quality evidence from two RCTs (n=146) showed family therapy-ED1 is less effective on full remission and BMI compared with family therapy-ED2.

Low quality evidence from one RCT (n=40) showed no difference in the effect of family therapy-ED1 on change in % average body weight and Morgan-Russell outcome-average compared with family therapy-ED2.

Low quality evidence from one RCT (n=106) showed no difference in the effect of family therapy-ED1 on EDE-global, EDE-restraint, EDE-eating concern, EDE-weight concern and EDE-shape concern compared with family therapy-ED2.

Low quality evidence from one RCT (n=106) showed family therapy-ED1 may be less effective on the number of people hospitalized during treatment compared with family therapy-ED2, although there was some uncertainty.

Low quality evidence from two RCTs (n=146) showed no difference in the effect of family therapy-ED1 on depression compared with family therapy-ED2.

Family therapy-ED 1 (conjoint family therapy) versus family therapy-ED 2 (separated family therapy) in young people with anorexia nervosa at follow up

Very low to low quality evidence from one RCT (n=106) showed no difference in the effect of family therapy-ED1 on full remission, BMI, EDE-global, EDE-restraint, EDE-eating concern, EDE-weight concern and EDE-shape concern compared with family therapy-ED2.

Low quality evidence from one RCT (n=106) showed family therapy-ED1 is less effective on depression compared with family therapy-ED2.

Long-term family therapy-ED versus short-term family therapy-ED in young people with anorexia nervosa at end of treatment

Moderate to low quality evidence from one RCT (n=86) showed no difference in the effect of long-term family therapy-ED on BMI, EDE-restraint and EDE-shape concern compared with short-term family therapy-ED.

Low quality evidence from one RCT (n=86) showed long-term family therapy-ED may be more effective on EDE-weight concern and EDE-eating concern compared with short-term family therapy-ED, although there was some uncertainty.

Low quality evidence from one RCT (n=86) showed long-term family therapy-ED is more effective on YBC-EDS score compared with short-term family therapy-ED.

Long-term family therapy-ED versus short-term family therapy-ED in young people with anorexia nervosa at follow up

Moderate quality evidence from one RCT (n=71) showed no difference in the effect of long-term family therapy-ED on BMI compared with short-term family therapy-ED.

Low quality evidence from one RCT (n=71) showed no difference in the effect of long-term family therapy-ED on the number of people achieving a BMI greater than 20 kg/m2, nor on the number of people resuming menstruation compared with short-term family therapy-ED.

Moderate quality evidence from one RCT (n=71) long-term family therapy-ED may be more effective on increasing the number of people achieving greater than 90% Ideal Body Weight compared with short-term family therapy-ED, although there was some uncertainty.

Low quality evidence from one RCT (n=71) showed long-term family therapy-ED may be less effective on reducing the number of amenorrheic people compared with short-term family therapy-ED, although there was some uncertainty.

Very low to low quality evidence from one RCT (n=35) showed no difference in the effect of long-term family therapy-ED on EDE-eating concerns, EDE-restraint, EDE-weight concerns and EDE-shape concerns compared with short-term family therapy-ED.

Family therapy with family meal versus family therapy without family meal in young people with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=23) showed family therapy-ED with family meal is more effective on remission compared with family therapy-ED without family meal.

Very low to low quality evidence from one RCT (n=23) showed no difference in the effect of family therapy-ED with family meal on weight, % expected body weight, EDI-2-total score and Morgan-Russell outcome-average compared with family therapy-ED without family meal.

Low quality evidence from one RCT (n=21) showed family therapy-ED with family meal may be less effective on general psychopathology compared with family therapy-ED without family meal, although there was some uncertainty.

Low quality evidence from one RCT (n=21) showed family therapy-ED with family meal is more effective on the number of people who resumed menstruation compared with family therapy-ED without family meal.

Family therapy with family meal versus family therapy without family meal in young people with anorexia nervosa at follow up

Very low quality evidence from one RCT (n=23) showed no difference in the effect of family therapy-ED with family meal on remission compared with family therapy-ED without family meal.

Very low to low quality evidence from one RCT (n=21) showed no difference in the effect of family therapy-ED with family meal on weight, % expected body weight, Morgan-Russell outcome-average, EDI-2 score and general psychopathology compared with family therapy-ED without family meal.

Low quality evidence from one RCT (n=21) showed family therapy-ED with family meal may be more effective on increasing the number of people who resumed menstruation compared with family therapy-ED without family meal, although there was some uncertainty.

Family therapy-ED versus any other type of family intervention in adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=47) showed no difference in the effect of family therapy-ED on BMI compared with any other type of family intervention.

Very low to low quality evidence from one RCT (n=25) showed no difference in the effect of family therapy-ED on SEED Anorexia-severity and SEED Bulimia-severity compared with any other type of family intervention.

Low quality evidence from one RCT (n=77) showed no difference in the effect of family therapy-ED on carer quality of life compared with any other type of family intervention.

Low quality evidence from one RCT (n=66) showed no difference in the effect of Family Therapy-ED on carer family functioning compared with any other type of family intervention.

Low quality evidence from one RCT (n=75) showed family therapy-ED may be more effective on improving the negative experience of caregiving compared with any other type of family intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=75) showed family therapy-ED is less effective on positive experience of caregiving compared with any other type of family intervention.

Family therapy-ED versus any other type of family intervention in adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=44) showed no difference in the effect family therapy-ED on BMI compared with any other type of family intervention.

Very low to low quality evidence from one RCT (n=29) showed no difference in the effect of family therapy-ED on SEED Anorexia-severity and SEED Bulimia-severity compared with any other type of family intervention.

Low quality evidence from one RCT (n=69) showed no difference in the effect of family therapy-ED on carer quality of life compared with any other type of family intervention.

Low quality evidence from one RCT (n=58) showed no difference in the effect of family therapy-ED on carer family functioning compared with any other type of family intervention.

Low quality evidence from one RCT (n=63) showed no difference in the effect of family therapy-ED on the positive and negative experience of caregiving compared with any other type of family intervention.

General family and any individual therapy versus any nutritional intervention in adults with anorexia nervosa

Very low quality evidence from one RCT (n=30) showed no difference in the effect of general family therapy on weight, the number of people experiencing regular menstruation and the number of people experiencing amenorrhea compared with any nutritional intervention.

Very low quality evidence from one RCT (n=30) showed general family therapy is more effective on global clinical score compared with any nutritional intervention.

Family therapy-ED versus any individual therapy in adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=84) showed family therapy-ED may be less effective on all-cause mortality compared with individual therapy, although there was some uncertainty.

Low quality evidence from one RCT (n=84) showed no difference in the effect of family therapy-ED on the number of people recovered from anorexia nervosa compared with individual therapy.

6.3.12 Economic Evidence statements

No economic evidence on the cost effectiveness of psychological interventions for adults with anorexia nervosa was available.

The existing economic evidence on the cost effectiveness of psychological therapies for children and young people was very limited and was not directly applicable to the NICE decision-making context. According to the reviewed US study (N=158) family based treatment was dominant when compared with the systemic family therapy. The study was characterised by potentially serious limitations. No economic evidence on the cost effectiveness of other psychological interventions for children and young people with anorexia nervosa was identified.

6.3.13 Recommendations and link to evidence for the review on: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?

Psychological treatment for anorexia nervosa

	57. Be aware that a key goal of treatment for anorexia nervosa is to help people reach a healthy body weight or BMI for their age. **58. When weighing people with anorexia, consider sharing the results with them and (if appropriate) their family members or carers.** **Psychological treatment for adults** **59. Consider either individual eating-disorder-focused cognitive behavioural therapy (CBT-ED) or eating-disorder-focused focal psychodynamic therapy for adults with anorexia nervosa.** **60. Individual CBT-ED programmes for adults with anorexia nervosa should:** • use a CBT-ED manual. • consist of up to 40 sessions over 40 weeks • aim to reduce the risk to physical health and any other symptoms of the eating disorder. • encourage reaching a healthy body weight and healthy eating • cover nutrition, relapse prevention, cognitive restructuring, mood regulation, social skills, body image concern and self-esteem. • create a personalised treatment plan based on the processes that appear to be maintaining the eating problem. • explain the risks of starvation and being underweight. • enhance self-efficacy • include self-monitoring • include homework, to help the person practice what they have learned in their daily life.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating anorexia nervosa in children, young people and adults. For this population, body weight or BMI and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and

	service user experience.
Trade-off between clinical benefits and harms	**Individual CBT-ED for adults** In adults, individual CBT for eating disorders (CBT-ED) was more effective at improving body weight and remission versus any other intervention but there was some uncertainty. At 12 months follow up, the benefits of CBT-ED on body weight and remission were no longer evident. General psychopathology was also better in the CBT-ED group compared with any other intervention at the end of treatment; however, all other important outcomes showed no difference, including depression, EDE and EDI subscales. At follow up there was no difference between any of the key outcomes. No data was available on general functioning, family functioning, service user experience, all-cause mortality, adverse events, quality of life, resource use or relapse. Comparing one CBT-ED programme with another in hospital showed no difference in any outcomes including BMI, EDE-subscales, general psychiatric features at the end of treatment and 12 months follow up. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality, adverse events, quality of life, resource use or relapse. For people with severe and enduring anorexia nervosa, CBT-ED appeared to have similar effects on BMI, depression, EDE global and quality of life as any other treatment. At 12 months follow up similar results were found but there was a trend for more favourable results on EDE-global in the CBT-ED treated group. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality, adverse events, resource use or relapse. **Other treatments for adults** Psychiatric counselling compared with any other intervention showed no difference in rates of remission or all-cause mortality at the end of treatment. No data was available on body weight, general functioning, family functioning, service user experience, adverse events, quality of life, general psychopathology, resource use or relapse. Interpersonal psychotherapy showed no difference in BMI compared with any other treatment at end the end of treatment. No differences were found in EDE-weight concern, EDE-shape concern, depression, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction. IPT was less effective on EDE-restraint and showed a trend to be less effective on general function and EDE-eating concerns. At 6.7 years follow up no differences were found in any of the outcomes. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality, adverse events, quality of life, resource use or relapse. One study on self-help for anorexia nervosa was identified but it did not report data on remission or body weight. However, it did show favourable results on EDI-total at the end of treatment and a trend for favourable results on EDI-drive for thinness, EDI-body dissatisfaction, depression, global clinical score and bulimic symptoms. No difference was found in EDI-bulimia or general psychopathology. There was also a trend for less favourable results on menstrual function. At 9 months follow up, no difference was found in general psychopathology or menstrual function but a trend for favourable results on bulimic symptoms. No data was available general functioning, family functioning, service user experience, all-cause mortality, adverse events, quality of life, resource use or relapse. Refer to the following LETR for results on SSCM and MANTRA. **Family-based therapy in adults** In adults no differences were found in an eating disorder-focused family therapy compared with another family intervention (family day workshop) on BMI, severity of anorexia nervosa or bulimia nervosa, quality of life, family functioning. However, favourable results were found on negative carer experiences (but with some uncertainty) and less favourable results were found on positive carer experiences (with some uncertainty). At 36 months follow up, no differences were found in any of these reported outcomes. No evidence was found on the critical outcome of remission, nor the important outcomes of general psychopathology, general functioning, resource use, adverse events, quality of life, all-cause mortality and

	relapse. In another study, no difference was found between general family therapy and a nutritional intervention on body weight, menstruation, number who had amenorrhea, but there was a favourable outcome on global clinical score. The study did not report the critical outcome of remission, nor on the important outcomes of general psychopathology, general functioning, family functioning, service user experience, resource use, adverse events, quality of life, all-cause mortality or relapse. Finally, family therapy had a similar effect on all-cause mortality and the number of people who recovered compared with individual therapy. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of general psychopathology, general functioning, family functioning, service user experience, resource use, adverse events, quality of life or relapse. Refer to LETR on young people with anorexia nervosa for evidence on this age-group.
Trade-off between net health benefits and resource use	No economic studies assessing the cost effectiveness of psychological interventions for adults with anorexia nervosa are available. The clinical evidence indicated that individual ED-focused cognitive behavioural therapy (CBT) and ED-focussed focal psychodynamic therapy offer similar clinical benefits. According to the committee both therapies would consist of approximately 40 sessions facilitated by band 7 worker at a cost of approximately of £3,370 per person (in 2014/15 prices). The committee considered the consequences of anorexia nervosa including very high morbidity and associated high cost to the healthcare system. For example, Beat (2015) estimated the cost to health service resources of the treatment element of eating disorders was £8,850 per individual per annum. The committee also took into account the increased mortality rate; psychological and financial burden for the individual and for their families, as well as the other clinical benefits associated with the treatment. Given the benefits associated with the psychological therapies the committee expressed the view that the provision of such therapies represent good value for money and are worth the investment.
Quality of evidence	The quality of the evidence was low to very low. The evidence was downgraded for indirectness, imprecision and risk of bias for reasons such as lack of clarity on how randomisation was conducted or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded, and high dropouts were detected >20%. To account for high dropouts rates, intention-to-treat analysis was used for remission results, with the assumption that any dropouts did not recover from the eating disorder. The committee discussed how little evidence there is on people with anorexia nervosa and the difficulty in recruiting and retaining people in the study. The majority of the comparisons included 1-2 studies and a very low number of participants. A 2014 paper by Zipfel was downgraded for indirectness since a number of participants received inpatient treatment during the study. Between baseline and the end of six months' treatment, 8% of people in the focal psychodynamic group, 10% of CBT-ED and 11% of treatment as usual received additional inpatient treatment. This led to much discussion among the committee on how much any improvements in remission and body weight can be attributed to the effectiveness of the intervention and not the time in hospital. The committee also noted that inpatient hospital stay may reflect treatment failure. The adherence to the CBT-ED manual was also queried, however, the paper did report a good overall adherence of 0.74 to 0.82 (greater than 0.75 is considered good conformity). After examining other studies, it was revealed that Dare 2001, Touyz 2013 and Schmidt 2012 included a similar number of participants who required a hospital stay. Zipfel 2014 was considered a well-designed study with a relatively high number of participants. Ultimately, and for consistency, the committee included Zipfel 2014 and other studies where participants required inpatient care during the intervention. It also reflects what happens in the 'real' world. Another study included in the CBT-ED versus another analysis that was

	downgraded for indirectness was Pike 2003. This paper included a population different to other studies, wherein the participants had just been discharged from hospital. Consequently, the aim of this study is different to other studies, where the investigators aimed to maintain and consolidate on the gains achieved in the inpatient unit rather than aim to compare different therapies and/or increase remission rates. Nevertheless, the paper was still included in the analysis but downgraded for indirectness. Heterogeneity was detected for remission in response to CBT-ED versus other, however the reason could not be deciphered since Pike and Zipfel 2014 were both considered to have a high risk of bias for reasons discussed above. Also, duration or severity of illness were similar between the 2 studies (>6 years of anorexia nervosa) and comorbidities were not reported.
Other considerations	The therapies to offer adults with anorexia nervosa are not recommended in any particular order. The committee discussed the difficultly in delivering focal psychodynamic therapy, given that a number of therapists in the NHS are not trained in that type of therapy. Thus, an investment in training may be needed locally before focal psychodynamic can be offered. The other therapies included in this review but were not recommended because they were either not-effective on remission, the sample size was too small or the quality was too low included self-help, interpersonal therapy, psychiatric counselling and family therapy. SSCM and MANTRA were considered second-line options for treating adults with anorexia nervosa and their evidence is reviewed in the following LETR. The committee agreed it was important to say up to 40 sessions, since some people with anorexia nervosa may not need all 40 sessions. Some may achieve remission early or they may not respond to this type of treatment and alternatives need to be considered. The committee highlighted the importance of openly weighing the person with anorexia nervosa, so that they are made aware of the results, and if treating a child or young person that their family or carers are also aware of the results. A number of clinicians prefer not to weigh or tell the person their weight. However, because restoration of a healthy weight is a priority of treatment, the committee agreed that it was important that the person with anorexia nervosa is made aware of their progress. The committee agreed that openly weighing a person with an eating disorder should not too prescriptive. The majority of patients are willing to accept (albeit often grudgingly) the rationale for weight being a part of the discussion of treatment progress. However, if a person finds it too difficult (at least early on in treatment) it is unhelpful and often counterproductive to insist on this immediately.

1 **Focal psychodynamic programme for adults with anorexia nervosa**

	61. Eating-disorder-focused focal psychodynamic therapy programmes for adults with anorexia nervosa should: • use a focal psychodynamic manual specific to eating disorders • consist of up to 40 sessions over 40 weeks • include psychoeducation about nutrition and the effects of starvation • make a patient-centred focal hypothesis that is specific to the individual and addresses: o what the symptoms mean to the person o how the symptoms affect the person o how the symptoms influence the person's relationships with others and with the therapist.

	• in the first phase, focus on developing the therapeutic alliance between the therapist and person with anorexia nervosa, addressing pro-anorexic behaviour and ego-syntonic beliefs (beliefs, values and feelings consistent with the person's sense of self) and building self-esteem • in the second phase, focus on relevant relationships with other people and how these affect eating behaviour • in the final phase, focus on transferring the therapy experience to situations in everyday life and address any concern the person has about what will happen when treatment ends.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating anorexia nervosa in children, young people and adults. For this population, body weight or BMI and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	**Focal psychodynamic therapy** Focal psychodynamic general therapy appeared to improve remission rates but not BMI compared with any other intervention at the end of treatment. Other outcomes, including EDI total, general psychopathology and all-cause mortality were no different at the end of treatment. Similarly at 26 to 32 months follow up remission favoured focal psychodynamic therapy compared with any other intervention but no difference was found in weight, EDE-bulimia, EDI-total, Morgan–Russell symptoms, general psychopathology. No harms were identified. No data was available on general functioning, family functioning, service user experience, adverse events, quality of life, resource use or relapse. No interventions on focal psychodynamic therapy were identified in children or young people. Refer to other LETRs for outcomes from other treatments.
Trade-off between net health benefits and resource use	No economic studies assessing the cost effectiveness of different configurations of psychodynamic therapy for adults with anorexia nervosa are available. Generally, in most clinical studies psychodynamic therapy is intensive and consists of 40 sessions over 40 weeks. Also, the number of sessions and duration of treatment is in line with the recommended dose of cognitive behavioural therapy for people with anorexia nervosa. Given comparable effectiveness between psychodynamic and cognitive behavioural therapy the committee were of the view that provision of psychodynamic therapy for adults with anorexia nervosa would not incur additional healthcare resources over and above that associated with the cognitive behavioural therapy. Also, according to the expert opinion, currently people with anorexia nervosa are very likely to receive treatments of similar intensity and as such psychodynamic therapy is unlikely to incur significant additional healthcare resources. Moreover, currently provided treatments are likely to be a mixture of all available treatments (not necessarily effective); and by recommending evidence based effective treatment such as psychodynamic therapy can result in the overall cost savings to the healthcare system.
Quality of	The quality of the evidence was low to very low. The evidence was downgraded for

evidence	indirectness, imprecision and risk of bias for reasons such as it was unclear how they randomised or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded and high dropouts were detected >20%. To account for high dropouts rates, intention-to-treat analysis was used for remission results, with the assumption that any dropouts did not recover from the eating disorder. The comparisons of focal psychodynamic versus any other intervention included the paper Zipfel 2014 that was downgraded for indirectness for reasons described in the LETR on the recommendation for CBT-ED. This comparison also included the paper by Dare 2001 that was also downgraded for indirectness since it too also included participants who were hospitalised during the study. The numbers in Dare 2001 were similar to those reported in Zipfel 2014, 10% for family therapy, 9-14% for focal psychodynamic and 26% for treatment as usual (counselling). The papers were included for the same reasons discussed in the previous LETR. Like other interventions for those with anorexia nervosa, there were few studies that fed into the comparison for focal psychodynamic versus any other intervention. The highest number of participants for an outcome was for remission, n=326 and was based on only two studies. Other interventions considered for adults but not recommended included: interpersonal psychotherapy, psychodynamic counselling, nutritional counselling, Internet-guided self-help and family therapy. Only one study was identified for most of these studies and included very small numbers. For the critical outcomes, either body weight or remission was reported, or remission was excluded because the investigators did not measure symptoms over a minimum of two weeks (which was considered the minimum duration by the committee).
Other considerations	Few people in the NHS are trained in focal psychodynamic therapy and the manual is currently only available in German, however it is expected to be made available in English in 2017. The committee agreed it was important to say up to 40 sessions, since some people with anorexia nervosa may not need all 40 sessions. Some may achieve remission early or they may not respond to this type of treatment and alternatives need to be considered.

1 **Second line psychological treatments for adults with anorexia nervosa**

	62. If individual CBT-ED or focal psychodynamic-ED is ineffective, not available or not acceptable for adults with anorexia nervosa, consider specialist supportive clinical management (SSCM) or the Maudsley Anorexia Treatment for Adults (MANTRA).
Critical and important outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating anorexia nervosa in children, young people and adults. For this population, body weight or BMI and remission are of greatest concern. Other outcomes that are important but considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. Therefore, they were extracted where possible but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa – of lesser importance, but clearly still important – included general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade off benefits and harms	**SSCM and MANTRA** Comparing specialist supportive clinical management (SSCM) with any other interventions showed no difference in BMI or remission rates at the end of the intervention and at 2 months up to 6.7 years follow up. Other outcomes showed favourable results for SSCM including depression and general function at the end of treatment. All other outcomes were no different

	between the two treatment arms including EDE-global, EDI-restraint and EDE and EDI subscales. At follow up, none of the outcomes were different between SSCM and any other intervention except EDI-restraint favoured SSCM. No harms were detected for SSCM. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use or relapse. Comparing SSCM to the other therapies recommended for adults with anorexia nervosa showed SSCM is equally effective as CBT-ED on BMI at the end of the intervention and at follow up. Remission rates were not considered since the duration over which the symptoms were measured was not clear. The committee agreed that a minimum of two weeks was required for remission to be considered. Comparing SSCM with focal psychodynamic showed no difference in BMI or remission rates at follow up. No data was available at the end of treatment nor on the outcomes of general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, general psychopathology or relapse. Comparing MANTRA with any other intervention showed no difference in the critical outcomes BMI or remission at the end of treatment and at follow up. Other outcomes, depression and EDI-total, also showed no difference at the end of treatment and at follow up between the two treatment arms. No harms were detected. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use or relapse. Finally, comparing SSCM directly with MANTRA showed no difference in BMI or remission at the end of treatment or at 2 to 6 months follow up. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, general psychopathology or relapse. One study investigated the effects of SSCM on adults with severe and enduring anorexia nervosa. Compared with any other treatment (CBT-ED) it showed no difference at the end of treatment or at follow up for BMI, depression or quality of life. Only at 12 months follow up was SSCM less effective on EDE-global compared with any other treatment. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, resource use or relapse. No interventions on SSCM or MANTRA were identified in children or young people.
Trade-off between net health benefits and resource use	No economic studies assessing the cost effectiveness of SSCM or MANTRA for adults with anorexia nervosa are available. According to the committee both therapies are expected to have similar intervention costs to individual ED-focused cognitive behavioural therapy (CBT) and ED-focused focal psychodynamic therapy. However, there was a greater uncertainty pertaining to the clinical benefits. As a result, the committee expressed the view that only if CBT-ED or focal psychodynamic-ED is ineffective, not available or not acceptable for adults with anorexia nervosa SSCM or MANTRA should be considered.
Quality of the evidence	The quality of the evidence was mostly low. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded and high dropouts were detected >20%. To account for high dropouts rates, intention-to-treat analysis was used for remission results, with the assumption that any dropouts did not recover from the eating disorder. The studies were small and few studies could be meta-analysed.
Other considerations	The committee agreed that it was important to consider SSCM and MANTRA as alternatives to CBT-ED and focal psychodynamic therapy because many healthcare professionals currently deliver these therapies to adults with anorexia nervosa and both are considered effective. Although there was no evidence that compared SSCM and MANTRA with wait list controls, the data showed they were

	similarly effective as any other interventions including CBT-ED and focal psychodynamic therapy, nor did they show any harms. Also, the data shows SSCM and MANTRA have similar effects when compared with each other. For these reasons, the committee agreed they should be considered alternatives to the first-line treatment options if they are ineffective, not available or not acceptable. There was some evidence to show SSCM is equally effective as CBT-ED in adults with severe (BMI 11.8 to 18.5 with a mean of 16.2) and enduring anorexia nervosa (more than seven years of illness with a mean duration of 16.6 years), however, the committee were not confident in making a recommendation explicitly for this population given that it was only 1 study with n=63. In the end they agreed it was better to develop a research recommendation, specifically 'what is the effectiveness of adapted treatments for those with anorexia nervosa who are not responding to treatment?' The committee discussed whether people with severe anorexia nervosa (BMI <16.5) should be offered inpatient care but until more evidence becomes available, the committee agreed to treat severe and enduring anorexia nervosa the same as those with less severe anorexia nervosa, unless there is a physical health problem that needs inpatient care.

3. Research recommendation: "what is the effectiveness of adapted treatments in those with anorexia nervosa who are not responding to treatment?"

Psychological treatment for young people with anorexia nervosa

	63. Consider anorexia-nervosa-focused family therapy for young people with anorexia nervosa, delivered as single- or multi-family therapy and with sessions provided either: • separately for the young person and for their family members and carers or • for the young person and their family together. **64. Anorexia-nervosa-focused family therapy for young people with anorexia nervosa should:** • use family-based treatment for eating disorders manual • consist of 18 – 20 sessions over at most one year • review the needs of the young person four weeks after treatment begins and then every three months, to establish how regular sessions should be and how long treatment should last • emphasise the role of the family in helping the young person to recover • not blame the young person or their family members or carers • include psychoeducation about nutrition and the effects of starvation • in the first phase, aim to establish a good therapeutic alliance with the young person, their parents or carers and other family members • help the parents or carers take charge of the young person's eating and return control to the young person when they are ready

	• in the final phase: o support the young person (with help from their parents or carers) to establish a level of independence appropriate for their level of development o focus on plans for when treatment ends (including any concerns the young person and their family have) and on relapse prevention. **65. Consider support for family members who are not involved in the family therapy, to help them to cope with distress caused by the condition.** **66. Consider giving young people with anorexia nervosa additional appointments separate from their family members or carers.**
Critical and important outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating anorexia nervosa in children, young people and adults. For this population, body weight or BMI and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade off benefits and harms	**Family-based therapy in young people** In children and young people with anorexia nervosa, family-based therapy was the most effective compared with other interventions including CBT-ED and adolescent focused therapy. Compared with treatment as usual, eating disorder-focused family therapy showed a trend to be more effective on remission rates and those who achieved a target body weight (BMI greater than 10th percentile), but showed no difference in its effectiveness on BMI at the end of treatment, EDI-global, global functioning, number who were amenorrheic or required hospitalisation. No data was available at long-term follow up. No evidence was found on the important outcomes of general psychopathology, family functioning, service user experience, resource use, adverse events, quality of life, all-cause mortality or relapse. In young people with anorexia nervosa, no differences were found in an eating disorder-focused family therapy compared with another family intervention on body weight, EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction, general psychopathology, depression and family functioning. No evidence was found on the critical outcome of remission, nor on the important outcomes of service user experience, resource use, adverse events, quality of life, all-cause mortality or relapse. In young people with anorexia nervosa, no differences were found between eating disorder-focused family therapy and general family therapy on body weight, remission rates, eating disorder scale, depression or quality of life at the end of treatment or at follow up. There was a trend for EDE-global to favour the eating disorder focused family therapy at both time points, however there was some uncertainty. No evidence was found on general functioning, family functioning, service user experience, resource use, adverse events, all-cause mortality or relapse. Comparing multifamily with single family-based therapy in young people the results showed greater improvements in body weight and rates of remission in the

	multifamily therapy group at the end of treatment and at follow up. EDE-restraint, EDE-shape concern and EDE-weight concern also favoured multifamily therapy at both time points. At end of treatment, no difference was found in EDE-eating concern, positive caregiving experience, negative caregiving experience or service user experience. There was a trend for depression scores to be higher in the multifamily-based therapy. At 6 months follow up, caregiving or service user experience was not reported and depression scores showed no difference nor EDE-eating concern. Family therapy showed favourable results on remission (although there was some uncertainty), body weight, Morgan-Russell scores and EDE-global compared with individual therapy in young people with anorexia nervosa. No difference was found in depression or carer family functioning. At 5 years follow up, no difference was found in remission, EDE-global, carer family functioning. However, body weight showed a trend to still favour family therapy over individual therapy but not depression. No evidence was found on the important outcomes of general functioning, family functioning, resource use, adverse events, quality of life, all-cause mortality and relapse. One study compared conjoint family therapy with separated family therapy in young people with anorexia nervosa and showed separated family therapy may be more effective on remission, BMI and on the number of people hospitalised (although on the latter there was some uncertainty). Other results showed no difference between the two arms including EDE-subscales and depression. At follow up, no difference in remission or BMI was found, along with the EDE-subscales. Depression appeared to favour separate family therapy. No evidence was found on the important outcomes of general functioning, family functioning, service user experience, resource use, adverse events, quality of life, all-cause mortality and relapse. Different durations of family therapy, long-term versus short-term, showed no difference at the end of treatment in BMI and most EDE-subscales, although there was some uncertainty. An eating disorder scale (YBC-EDS) did favour long-term family therapy over short-term family therapy. At 12 months follow up, mostly no differences were found between the two durations of treatment. No evidence was found on the critical outcome of remission, nor the important outcomes of general psychopathology, general functioning, family functioning, service user experience, resource use, adverse events, quality of life, all-cause mortality and relapse. If family therapy is supplemented with a family meal compared with family therapy alone, there was evidence to show that the addition of a family meal may improve remission rates. However, no difference was found in weight, EDI-2-total score and Morgan-Russell scores. The family meal may also be less effective on general psychopathology compared with family therapy alone (although there was some uncertainty). At 6 months follow up, no difference was found in remission or weight, and all other outcomes showed no difference. There was some evidence that the family meal may be more effective on restoring menstruation at follow up but there was some uncertainty. No evidence was found on the important outcomes of general functioning, family functioning, service user experience, resource use, adverse events, quality of life, all-cause mortality and relapse. See following LETR for other treatments for young people including CBT-ED, adolescent focused therapy and supportive therapy.
Trade-off between net health benefits and resource use	There was very limited UK-based existing economic evidence in young people with anorexia nervosa. The evidence indicated that family therapy may potentially be cost effective. According to the committee family therapy would consist of approximately 20 sessions facilitated by band 7 worker at a cost of approximately of £2,200 per person (in 2014/15 prices). The committee took into account the physical consequences of anorexia nervosa and high costs associated with managing these; psychological and financial burden both for children and young people and for their families. For example, Beat (2015) estimated the cost to health service resources of the treatment element of eating disorders was £8,850 per individual per annum. Given the clinical benefits associated with the family

	therapy and very high costs associated with anorexia nervosa the committee expressed the view that the provision of such treatment represent good value for money and are worth the investment. Also, providing extra support as outlined in the recommendation 1.3.8-9 may result in modest additional resource implications. However, such as that extra support will potentially improve outcomes and it is expected to lead to overall savings.
Quality of the evidence	The quality of the evidence was mostly low. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded, and high dropouts were detected >20%. To account for high dropouts rates, intention-to-treat analysis was used for remission results, with the assumption that any dropouts did not recover from the eating disorder. A committee member highlighted a degree of indirectness in the study from Le Grange 2016 that was recently added to the conjoint versus separate family therapy review. It was pointed out that 23.6% participants in the conjoint-based therapy were hospitalised during treatment and 11.8% from the separate family therapy. In addition, 57% of participants continued to receive treatment during the six months' follow up. Also, 9% of the conjoined and 19.6% of the separate family therapy were hospitalised during the follow up period. The study also included people who were up to 95% of expected weight in the trial. The majority of the comparisons did not allow for a meta-analysis since only one study was available. The study numbers were also relatively low. Nevertheless, the committee agreed that family therapy provided the most convincing evidence on young people with anorexia nervosa. In a number of the studies, Eisler 2000, Herscovici 2015, Lock 2010, Robin 1999 and Russell 1989, partial and full remission results were combined since the definition of remission varied across the studies and the partial remission definition in these studies compared closely with the full remission definition in other studies. Some of this data was not available in the papers and was received by personal communication with the authors via a committee member.
Other considerations	The committee agreed that family therapy showed the most promising results in young people with anorexia nervosa compared with CBT-ED and adolescent focused therapy (as reviewed in following LETR). Although few studies were identified, three studies with 179 participants showed family therapy resulted in a greater improvements in remission rates and BMI compared with individual therapy at the end of treatment. The difference between the two groups in remission did not appear to be maintained long-term but there was a trend for BMI to still favour family therapy at follow up. It was difficult to decipher which type of family therapy was the most effective (i.e. general versus eating disorder focused or variations in the Maudesley-based therapy) given the small number of studies but it was agreed (based on committee experience) that it should be eating disorder focused. Emerging evidence showed that multifamily therapy may be more effective than single family-based therapy on body weight and rates of remission. Given that it is only one study with 167 participants, the committee agreed that the family therapy could be delivered in either a multi- or single family based format. There was also some evidence to support a recommendation that separate family therapy (parents or carers are treated separately from the young person with the eating disorder) may be more effective than conjoint family therapy (all together). Given this is only one study and most of the evidence to date has been on conjoint family therapy, the committee agreed that either could be recommended. The committee agreed that family therapy should include all family members including siblings where possible given the impact having a person with an eating disorder in the family may also have on them. If family therapy is not being delivered or if family members such as siblings are not involved in the therapy sessions, the committee agreed that it was important to offer them support to help them to cope with distress caused by the condition. The committee also acknowledged that some young people with an eating

	disorder may need to discuss concerns independent of their family, so for those who are receiving conjoint family therapy they agreed that it was important to consider the need for additional separate appointments to their family members or carers. It is important to note that this is not evidence based but based on committee experience and expertise. Therefore, it is not a strong recommendation and only for the health professional to consider. Given the specific needs of children and young people with an eating disorder, the committee agreed that it is important that only healthcare professionals with experience of working with this age group deliver the psychological treatment. **Stepped care (reviewed in chapter 5)** There was very low quality evidence on stepped care, that is, what the options are if the first line treatment does not work whether it be an alternative treatment or increase the number of sessions of first line treatment. A review on stepped care in young people with anorexia nervosa showed if family therapy is stepped up to intensive parental coaching compared with continued family based therapy, there was no difference on remission, body weight or EGE-global at the end of treatment. The committee agreed that the evidence was too weak to make a recommendation on stepped care and instead generated a research recommendation relevant for any eating disorder, including anorexia nervosa to: "evaluate the effectiveness of stepped care for psychological treatment of eating disorders for people of all-ages."

1 **Second line psychological treatments for young people with anorexia nervosa**

	67. If family therapy is unacceptable, contraindicated or ineffective for young people with anorexia nervosa, consider individual CBT-ED or adolescent focused eating disorder therapy. **68. Assess whether family members or carers (as appropriate) need support if the young person with anorexia nervosa is having therapy on their own.**
Critical and important outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating anorexia nervosa in children, young people and adults. For this population, body weight or BMI and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade off benefits and harms	When family therapy is unacceptable, contraindicated or ineffective for children or young people with anorexia nervosa, there was some evidence that showed CBT-ED and adolescent (young people) focused psychotherapy may be viable alternatives. One study was identified that compared the effectiveness of CBT-ED with treatment as usual. No data was available at the end of treatment but at follow up there was no difference in BMI or remission compared treatment as usual. Quality of life and EDI-total were also similar between the two arms. Treatment as usual was non-prescriptive and included a multidisciplinary approach with family therapy, dietetic and individual supportive therapy. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, resource use or relapse. Adolescent focused eating disorder therapy appeared to favour any other intervention for body weight and remission at the end of treatment. No differences were found between the two arms at 12 months follow up. No other relevant

	outcomes were reported. Supportive therapy in young people appears to be less effective on weight and remission compared with any other treatment at the end of treatment. No differences were found at follow up. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, general psychopathology or relapse.
Trade-off between net health benefits and resource use	There was no evidence on the cost effectiveness of CBT-ED and young people focused psychotherapy in children or young people with anorexia nervosa. According to the committee the intervention costs are likely to be similar to those of family therapy. However, there was greater uncertainty pertaining to the clinical benefits. As a result, the committee expressed the view that if family therapy is unacceptable, contraindicated or ineffective CBT-ED and young people focused psychotherapy could potentially be considered.
Quality of the evidence	The quality of the evidence was low to very low. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded, and high dropouts were detected >20%. To account for high dropouts rates, intention-to-treat analysis was used for remission results, with the assumption that any dropouts did not recover from the eating disorder. Again, only one study was available for each comparison and a relatively small sample size was used in each study. The study on CBT-ED only provided data at follow up not at the end of treatment. However, it was based in the UK so the committee considered it a very relevant study. No heterogeneity was identified. Nor were any studies identified that compared psychological treatments in young people to wait list controls, thus making it difficult to know whether remission would happen in the absence of any treatment. However, the committee said it is unethical to offer no treatment to most patients.
Other considerations	The committee agreed that it was important to consider an alternative therapy to family therapy for young people with anorexia nervosa, especially if family therapy is contraindicated because of problems within the family. A study on CBT-ED showed it was equally on remission and body weight as treatment as usual (that included family therapy). Thus suggesting it was a good alternative to family therapy. Conversely, the study on adolescent focused psychotherapy showed it may be less effective compared with any other treatment (family therapy). The other intervention where evidence was found in young people was supportive therapy. However, the committee agreed that this was only a control intervention and not one that would ever be recommended. Besides some evidence to support a recommendation on CBT-ED as an alternative to family therapy for young people, the committee agreed the remainder of the evidence was not convincing and to instead recommend what is offered to adults. As they pointed out, a lot the young people with anorexia nervosa are close in age to the adults used in the studies. In adults the 2 psychotherapies that proved the most effective were CBT-ED and focal psychodynamic therapy. However, instead of offering adult focal psychodynamic therapy for young people, the committee agreed to offer the age-appropriate version of adolescent (young people) focused therapy.

6.4 Carer interventions

6.4.1 Review question: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 105. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all psychological interventions for the parents or carers of children or young people with AN. The interventions were categorised according to their mode of delivery (e.g. group, self-help), the age of the people with the eating disorder and the type of eating disorder and were compared to wait list controls, treatment as usual or any other intervention.

Table 105: Clinical review protocol summary for the review of: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Component	Description
Review question(s)	Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?
Population	• Family or carers of people with an eating disorder (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder).
Intervention(s)	• Psychological interventions may include: • Family-based • Parent only (not necessarily focussed on eating disorder) • Parent-focused therapy (PFT) • Group Parent-Training (GPT) • Separated family therapy • Parents with person with ED (greater focus on eating disorder) • Behavioural Family Therapy (BFT) • Behavioural family systems therapy (BFST). • Family Based Treatment (FBT) • Family Day Workshops (FDW) • Family Therapy (FT) • Family therapy for anorexia nervosa (FT-AN) • Multi-Family Group Day Treatment (MFGDT) • Multi-Family Group Therapy (MFGT) • Systemic Family Therapy (SFT) • Systemic Family Therapy for AN (SFT-AN) • Multifamily therapy (MFT) is synonymous with (MFGT; MFGDT). • Uniting couples in the treatment of AN (UCAN • Conjoint family therapy
Comparison	• Wait list control • Treatment as usual • Another intervention
Critical outcomes	Parent's or carer's general psychopathology (including mood/depression/anxiety) Family functioning Quality of life

Component	Description
	Other primary outcomes commonly reported in studies that just target the family/carer • The following outcomes will be included if the family or carer intervention includes the child or person with an eating disorder: • Remission and long-term recovery (if symptoms were measured over a minimum two week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• General functioning • Resource use. • Service user experience • All-cause mortality. • Adverse events • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion)
Study design	• Systematic Reviews • RCTs

6.4.2 Clinical Evidence for: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Four RCTs (n=518) met the eligibility criteria for this review, the majority of which were in female carers of adults with anorexia nervosa (Grover 2011 (Grover et al., 2011), Hibbs 2015 (Hibbs et al., 2015), Hoyle 2013 (Hoyle et al., 2013), Magill 2016 (Magill et al., 2016), Hodsoll 2016(Hodsoll et al., 2016), Salerno 2016 (Salerno et al., 2016),). An overview of all the trials included in the meta-analysis can be found in Table 106. Further information about both included and excluded studies can be found in Appendix J.

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Four studies (n=518; Grover 2011, Hibbs 2015/Magill 2016, Hodsoll 2016, Salerno 2016, Hoyle 2013) compared self-help or guided self-help and treatment as usual with treatment as usual in carers of people with anorexia nervosa. The majority of which were for carers of adults with anorexia nervosa. Summaries of the findings regarding carer outcomes can be found in Table 107. Summaries of the findings regarding patient outcomes can be found in

One study (n=149; Hodsoll 2016) compared self-help or guided self-help and treatment as usual with treatment as usual and self-help for carers of young people and adults with anorexia nervosa.

One study (n=64; Grover 2011) compared web-based guided self-help with treatment as usual for carers of young people and adults with anorexia nervosa. Summary of findings regarding carer outcomes can be found in Table 114 and Table 115.

One study (n=37, Hoyle 2013) compared web-based guided self-hep with web-based self-help for carers of young people and adults with anorexia nervosa. Summary of findings regarding carer outcomes for interventions for carers of anorexia nervosa can be found in Table 116 and Table 117.

Table 106: Study information for trials included in review of self-help versus any other intervention for carers of people with anorexia nervosa.

Study_ID	Mean Age of Carer (SD)	Female carers (%)	Mean Age of patient (SD)	Female patients (%)	Sample and duration of illness	N Initially Randomised	Intervention	Comparison	Sessions N	Treatment Length
Grover 2011	48.2 (7.6)	Not reported	20.4 (6.2) (range: 12-44)	Not reported	Full- and sub-threshold AN. Duration=4.3 (4.5) years	64 carers	Web-based Guided Self-Help (OAO)	Ad-hoc usual support (Beat)	16 x weekly 20m email or telephone clinical guidance sessions	4 months + 6-mo FU
Hibbs 2015/ Magill 2015	52.7	60	26 (9)	95	AN. Mean duration=75 months (range: 9-480)	268 carers	Guided Self-Help (ECHO) + TAU	TAU	Variable; minimum of 4 calls (per family) or 75% of the book read to be considered as completer	12-mo from discharge + 24-mo FU from discharge
Hodsoll 2016*	48.3 (5.5)	69	16.9 (2.1)	92	Full- and sub-threshold AN. Mean duration: 13.3 months (range: 2-110)	149 patient-carer dyads	Guided Self-Help (ECHO) + TAU	Self-Help (ECHO) + TAU TAU	Variable; A: Guidance was 10 x 30-60min telephone sessions.	12 months
Hoyle 2013	Not reported	89	Not reported	Not reported	AN. Duration not	37 carers	Web-based Guided Self-Help (OAO)	Web-based Self-Help (OAO) only	Weekly email or telephone	7 weeks + 3-mo FU

Study_ID	Mean Age of Carer (SD)	Female carers (%)	Mean Age of patient (SD)	Female patients (%)	Sample and duration of illness	N Initially Randomised	Intervention	Comparison	Sessions N	Treatment Length
					reported				guidance	
Salerno 2016*	48.3 (5.5)	95	16.9 (2.1)	92	Full- and sub-threshold AN.	149 patient-carer dyads	Guided Self-Help or Self Help (ECHO) + TAU	TAU	Variable; A: Guidance was 10 x 30-60min telephone sessions.	12 months

*Note: *, Hodsoll 2016 and Salerno 2016 report data from the same trial. Abbreviations: ECHO, Experienced Carers Helping Others; OAO, Overcoming Anorexia Online. TAU, treatment as usual.*

Table 107: Summary table of findings for self-help or guided self-help and treatment as usual (TAU) versus treatment as usual at 12 months after inpatient admission for carers of people with anorexia nervosa – carer outcomes.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with TAU	Risk difference with Self-Help or Guided Self-Help + TAU (95% CI)
Carer General Psychopathology at 12 months DASS-21	149 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	Not calculable for SMD values	The mean carer general psychopathology at 12 months in the intervention groups was 0.03 standard deviations higher (0.31 lower to 0.37 higher)

**The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up; DASS: Depression anxiety stress scale

1 Salerno 2016: no participant blinding; dropout rate of TAU group >20%. Unclear whether baseline demographic and clinical features similar.
2 Salerno 2016: 50 carer-patient dyads received ECHO with guidance, 49 carer-patient dyads received ECHO without guidance.
3 <400 participants.

Table 108: Summary table of findings for self-help and treatment as usual (TAU) versus treatment as usual at 6 or 12 months after inpatient admission for carers of people with anorexia nervosa – carer outcomes.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Anticipated absolute effects	
			Risk with TAU	Risk difference with Self-Help+TAU (95% CI)
Carer Accommodation & Enabling at 6 months AESED	147 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer accommodation & enabling at 6 months in the intervention groups was 0.01 standard deviations higher (0.32 lower to 0.33 higher)
Carer Family Functioning at 6 months Family Questionnaire	147 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer family functioning at 6 months in the intervention groups was 0.25 standard deviations higher (0.07 lower to 0.57 higher)
Carer Skills at 12 months CASK	147 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	Not calculable for SMD values	The mean carer skills at 12 months in the intervention groups was 0.15 standard deviations higher (0.17 lower to 0.48 higher)

**The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Hodsoll 2016: Unclear, no details given of statistical significance for social demographic and clinical variables. Randomization method, allocation concealment and participant blinding unclear. No investigator blinding. Dropout rate of TAU group>20%.
2 <400 participants.
3 CI crosses either 0.5 or -0.5 (SMD).

Table 109: Summary table of findings for self-help and treatment as usual versus treatment as usual at 6 or 12 months after inpatient admission for carers of people with anorexia nervosa – patient outcomes.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Self-Help+TAU (95% CI)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Self-Help+TAU (95% CI)
BMI at 12 months	99 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI at 12 months in the intervention groups was 0.27 standard deviations higher (0.13 lower to 0.66 higher)
Gender Standardized Weight for Height Percentage at 12 months	99 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean gender standardized weight for height percentage at 12 months in the intervention groups was 0.2 standard deviations higher (0.2 lower to 0.59 higher)
SEED for AN at 12 months	99 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean seed for an at 12 months in the intervention groups was 0.01 standard deviations lower (0.41 lower to 0.38 higher)
General Psychopathology at 12 months DASS-21	99 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology at 12 months in the intervention groups was 0.03 standard deviations lower (0.42 lower to 0.37 higher)
Clinical Impairment due to ED at 12 months	99 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical impairment due to ed at 12 months in the intervention groups was 0.11 standard deviations higher (0.29 lower to 0.5 higher)
Strength & Difficulties Questionnaire - Peer Problems at 12 months	99 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of		Not calculabl	The mean strength & difficulties questionnaire - peer problems at 12 months in the intervention groups was 0.09 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Self-Help+TAU (95% CI)
		bias, imprecision		e for SMD values	(0.49 lower to 0.3 higher)
Strength & Difficulties Questionnaire - Prosocial Behaviour at 12 months	99 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean strength & difficulties questionnaire - prosocial behaviour at 12 months in the intervention groups was 0.32 standard deviations higher (0.07 lower to 0.72 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Hodsoll 2016: Unclear, no details given of statistical significance for social demographic and clinical variables. Randomization method, allocation concealment and participant blinding unclear. No investigator blinding. Dropout rate of treatment as usual group >20%. 2 CI crosses either 0.5 or -0.5 (SMD). 3 <400 participants.					

Table 110: Summary table of findings for guided self-help and treatment as usual (TAU) versus treatment as usual at 12 and 24 months after inpatient admission for carers of people with anorexia nervosa – carer outcomes

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Guided Self-Help+TAU (95% CI)
Carer Burden at 12 months EDSIS	182 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer burden at 12 months in the intervention groups was 0.24 standard deviations lower (0.54 lower to 0.05 higher)
Carer Family Functioning at 12 months	336 (2 studies)	⊕⊕⊖⊖ LOW1,3,4		Not calculabl	The mean carer family functioning at 12 months in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Guided Self-Help+TAU (95% CI)
Family Questionnaire		due to risk of bias, imprecision		e for SMD values	0.05 standard deviations lower (0.26 lower to 0.17 higher)
Carer Quality of Life at 12 months WHO-Quol	182 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer quality of life at 12 months in the intervention groups was 0.32 standard deviations higher (0.03 to 0.61 higher)
Carer Accommodation & Enabling at 12 months AESED	336 (2 studies)	⊕⊕⊖⊖ LOW1,3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer accommodation & enabling at 12 months in the intervention groups was 0.24 standard deviations lower (0.46 to 0.03 lower)
Carer Skills at 12 months CASK	154 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer skills at 12 months in the intervention groups was 0.13 standard deviations higher (0.19 lower to 0.44 higher)
Carer General Psychopathology after 12 months DASS-21	154 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer general psychopathology after 12 months in the intervention groups was 0.07 standard deviations lower (0.39 lower to 0.24 higher)
Carer Burden after 24 months EDSIS	185 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer burden after 24 months in the intervention groups was 0.2 standard deviations lower (0.49 lower to 0.09 higher)
Carer Family Functioning after 24 months Family Questionnaire	185 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning after 24 months in the intervention groups was 0.25 standard deviations lower (0.54 lower to 0.04 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Guided Self-Help+TAU (95% CI)
Carer Quality of Life after 24 months WHO-Quol	185 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer quality of life after 24 months in the intervention groups was 0.24 standard deviations higher (0.05 lower to 0.53 higher)
Carer Accommodation & Enabling after 24 months AESED	185 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer accommodation & enabling after 24 months in the intervention groups was 0.23 standard deviations lower (0.52 lower to 0.06 higher)
Carer General Psychopathology after 24 months DASS-21	185 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer general psychopathology after 24 months in the intervention groups was 0.23 standard deviations lower (0.52 lower to 0.06 higher)
Carer Time Spent Caring after 24 months CSRI	185 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer time spent caring after 24 months in the intervention groups was 0.2 standard deviations lower (0.49 lower to 0.09 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Hibbs 2015/Magill 2015: No participant nor assessor blinding. Dropout rate>50% 12 months after discharge.
2 CI crosses either 0.5 or -0.5 (SMD).
3 Hodsoll 2016: Unclear, no details given of statistical significance for social demographic and clinical variables. Randomization method, allocation concealment and participant blinding unclear. No investigator blinding. Dropout rate of TAU group>20%.
4 <400 participants.

1 **Table 111: Summary table of findings for guided self-help and treatment as usual versus treatment as usual at 12 and 24 months**
2 **after inpatient admission for carers of people with anorexia nervosa – patient outcomes.**

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with TAU	Risk difference with Guided Self-Help+TAU (95% CI)
Patient deaths	178 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.07 (0.07 to 16.84)	11 per 1000	1 more per 1000 (from 10 fewer to 172 more)
Readmitted to hospital for ED during course of study	178 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 0.85 (0.53 to 1.35)	11 per 1000	2 fewer per 1000 (from 5 fewer to 4 more)
Patient Relapse Readmission to hospital for ED and/or drop 2 BMI points measured monthly from discharge	178 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.82 (0.6 to 1.13)	522 per 1000	94 fewer per 1000 (from 209 fewer to 68 more)
BMI at 12 months	212 (2 studies)	⨁⨁⊖⊖ LOW1,4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI at 12 months in the intervention groups was 0.02 standard deviations lower (0.29 lower to 0.25 higher)
Gender Standardized Weight for Height Percentage at 12 months	100 (1 study)	⨁⨁⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean gender standardized weight for height percentage 12-mo at 12 months in the intervention groups was 0.12 standard deviations lower (0.51 lower to 0.27 higher)
EDE-Q Global at 12 months	112 (1 study)	⨁⨁⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD	The mean ede-q global at 12 months in the intervention groups was 0.08 standard deviations lower (0.45 lower to 0.29 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Guided Self-Help+TAU (95% CI)
				values	
SEED for AN at 12 months	100 (1 study)	⨁⨁⨀⨀ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean seed for an at 12 months in the intervention groups was 0.15 standard deviations higher (0.24 lower to 0.55 higher)
General Psychopathology at 12 months DASS-21	212 (2 studies)	⨁⨁⨀⨀ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology at 12 months in the intervention groups was 0.09 standard deviations lower (0.36 lower to 0.18 higher)
Clinical Impairment due to ED at 12 months	100 (1 study)	⨁⨁⨀⨀ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical impairment due to ed at 12 months in the intervention groups was 0.11 standard deviations lower (0.5 lower to 0.29 higher)
Strength & Difficulties Questionnaire - Peer Problems at 12 months	100 (1 study)	⨁⨁⨀⨀ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD value	The mean strength & difficulties questionnaire - peer problems at 12 months in the intervention groups was 0.5 standard deviations lower (0.9 to 0.11 lower)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Guided Self-Help+TAU (95% CI)
				s	
Strength & Difficulties Questionnaire - Prosocial Behaviour at 12 months	100 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean strength & difficulties questionnaire - prosocial behaviour at 12 months in the intervention groups was 0.38 standard deviations higher (0.02 lower to 0.77 higher)
Quality of Life at 12 months WHO-QL	112 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life at 12 months in the intervention groups was 0.1 standard deviations higher (0.27 lower to 0.47 higher)
BMI at 24 months	119 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI at 24 months in the intervention groups was 0.28 standard deviations higher (0.08 lower to 0.64 higher)
EDE-Q Global at 24 months	119 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global at 24 months in the intervention groups was 0.3 standard deviations lower (0.66 lower to 0.07 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Guided Self-Help+TAU (95% CI)
General Psychopathology at 24 months DASS-21	119 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology at 24 months in the intervention groups was 0.27 standard deviations lower (0.63 lower to 0.1 higher)
Quality of Life at 24 months WHO-QL	119 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life at 24 months in the intervention groups was 0.29 standard deviations lower (0.65 lower to 0.07 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Hibbs 2015/Magill 2015: No participant nor assessor blinding. Dropout rate >50% 12 months after discharge.

2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

3 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).

4 Hodsoll 2016: Unclear, no details given of statistical significance for social demographic and clinical variables. Randomization method, allocation concealment and participant blinding unclear. No investigator blinding. Dropout rate of TAU group >20%.

5 <300 events (Risk Ratio) or <400 participants (SMD).

Table 112: Summary table of findings for guided self-help and treatment as usual versus self-help and treatment as usual at 6 or 12 months after inpatient admission for carers of anorexia nervosa – carer outcomes.

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Self-Help+TAU	Risk difference with Guided Self-Help+TAU (95% CI)
Carer Accommodation & Enabling at 6 months AESED	151 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer accommodation & enabling at 6 months in the intervention groups was 0.3 standard deviations lower (0.63 lower to 0.02 higher)
Carer Family Functioning at 6 months Family Questionnaire	151 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning at 6 months in the intervention groups was 0.2 standard deviations lower (0.52 lower to 0.12 higher)
Carer General Psychopathology at 12 months DASS-21	151 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer general psychopathology at 12 months in the intervention groups was 0.04 standard deviations lower (0.36 lower to 0.28 higher)
Carer Skills at 12 months CASK	151 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer skills at 12 months in the intervention groups was 0.03 standard deviations lower (0.35 lower to 0.29 higher)
Time Spent Caregiving at 12 months	151 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean time spent caregiving at 12 months in the intervention groups was 0.01 standard deviations higher (0.31 lower to 0.33 higher)
Direct Spending at 12 months	151 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean direct spending at 12 months in the intervention groups was 0 standard deviations higher (0.32 lower to 0.32 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Hodsoll 2016: Unclear, no details given of statistical significance for social demographic and clinical variables. Randomization method, allocation concealment and participant blinding unclear. No investigator blinding. Dropout rate of TAU group >20%.

2 CI crosses either 0.5 or -0.5 (SMD).

3 <400 participants.

Table 113: Summary table of findings for guided self-help and treatment as usual (TAU) versus self-help and treatment as usual at 12 months after inpatient admission for carers of anorexia nervosa – patient outcomes.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Self-Help+TAU	Risk difference with Guided Self-Help+TAU (95% CI)
BMI at 12 months	99 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI at 12 months in the intervention groups was 0.45 standard deviations lower (0.85 to 0.05 lower)
Gender Standardized Weight for Height Percentage at 12 months	99 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean gender standardized weight for height percentage at 12 months in the intervention groups was 0.34 standard deviations lower (0.73 lower to 0.06 higher)
SEED for AN at 12 months	99 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean seed for an at 12 months in the intervention groups was 0.19 standard deviations higher (0.2 lower to 0.59 higher)
General Psychopathology at 12 months DASS-21	99 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology at 12 months in the intervention groups was 0.13 standard deviations lower (0.52 lower to 0.27 higher)
Clinical Impairment due to ED at 12 months	99 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical impairment due to ed at 12 months in the intervention groups was 0.21 standard deviations lower (0.6 lower to 0.19 higher)
Strength & Difficulties Questionnaire - Peer Problems at 12 months	99 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean strength & difficulties questionnaire - peer problems at 12 months in the intervention groups was 0.43 standard deviations lower (0.83 to 0.03 lower)
Strength & Difficulties Questionnaire	99	⊕⊕⊖⊖		Not	The mean strength & difficulties questionnaire -

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Self-Help+TAU	Risk difference with Guided Self-Help+TAU (95% CI)
- Prosocial Behaviour at 12 months	(1 study)	LOW1,3 due to risk of bias, imprecision		calculable for SMD values	prosocial behaviour at 12 months in the intervention groups was 0.05 standard deviations higher (0.35 lower to 0.44 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Hodsoll 2016: Unclear, no details given of statistical significance for social demographic and clinical variables. Randomization method, allocation concealment and participant blinding unclear. No investigator blinding. Dropout rate of TAU group >20%. 2 CI crosses either 0.5 or -0.5 (SMD). 3 <400 participants.					

Table 114: Summary table of findings for web-based guided self-help versus treatment as usual at end of treatment for carers of people with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Web-based Guided Self-Help (95% CI)
Carer Accommodation & Enabling AESED	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer accommodation & enabling in the intervention groups was 0.34 standard deviations lower (0.84 lower to 0.16 higher)
Carer Family Functioning Level of Expressed Emotion	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning in the intervention groups was 0.46 standard deviations lower (0.96 lower to 0.05 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Web-based Guided Self-Help (95% CI)
Carer Burden EDSIS; Experience of Caregiving Inventory (ECI) Negative	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer burden in the intervention groups was 0.32 standard deviations lower (0.67 lower to 0.04 higher)
Carer Experience of Caregiving (ECI) Positive	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer experience of caregiving (eci) positive in the intervention groups was 0.06 standard deviations higher (0.44 lower to 0.55 higher)
Carer General Psychopathology DASS-21	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer general psychopathology (distress) in the intervention groups was 0.55 standard deviations lower (1.05 to 0.05 lower)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Grover 2011: Participant not blinded. Unclear whether baseline similar. 2 CI crosses either 0.5 or -0.5 (SMD).					

1 **Table 115: Summary table of findings for web-based guided self-help versus treatment as usual at follow up for carers of people**
2 **with anorexia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Web-based Guided Self-Help (95% CI)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Web-based Guided Self-Help (95% CI)
Carer Accommodation & Enabling FU AESED	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer accommodation & enabling fu in the intervention groups was 0.02 standard deviations lower (0.52 lower to 0.47 higher)
Carer Family Functioning FU Level of Expressed Emotion	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning fu in the intervention groups was 0.18 standard deviations lower (0.67 lower to 0.32 higher)
Carer Burden FU EDSIS; Experience of Caregiving Inventory (ECI) Negative	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer burden fu in the intervention groups was 0.15 standard deviations lower (0.5 lower to 0.2 higher)
Experience of Caregiving (ECI) Positive FU	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean experience of caregiving (eci) positive fu in the intervention groups was 0.18 standard deviations higher (0.32 lower to 0.67 higher)
Carer General Psychopathology (Distress) FU HADS	63 (1 study) 3 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer general psychopathology (distress) fu in the intervention groups was 0.01 standard deviations lower (0.5 lower to 0.49 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Web-based Guided Self-Help (95% CI)
CI: Confidence interval;					
1 Grover 2011: Participant not blinded. Unclear whether baseline similar. 2 CI crosses either 0.5 or -0.5 (SMD). 3 <400 participants.					

Table 116: Summary table of findings for web-based guided self-help versus web-based self-help at end of treatment for carers of people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Web-based Self-Help	Risk difference with Web-based Guided Self-Help (95% CI)
Carer Family Functioning LEE	27 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning in the intervention groups was 0.56 standard deviations lower (1.33 lower to 0.21 higher)
Carer Burden EDSIS; ECI negative	27 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer burden in the intervention groups was 0.31 standard deviations higher (0.23 lower to 0.85 higher)
Carer Experience of Caregiving (ECI) Positive	27 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer experience of caregiving (eci) positive in the intervention groups was 0.45 standard deviations higher (0.32 lower to 1.21 higher)
Carer Quality of Life GHQ-28; SF-36	27 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer quality of life in the intervention groups was 0.15 standard deviations lower (0.69 lower to 0.39 higher)
Carer General Psychopathology (Distress) DASS-21	27 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer general psychopathology (distress) in the intervention groups was 0.48 standard deviations lower (1.25 lower to 0.28 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Web-based Self-Help	Risk difference with Web-based Guided Self-Help (95% CI)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Hoyle 2013: Unclear randomization method, allocation concealment, participant and assessor blinding. 2 CI crosses 0.5 or -0.5 (SMD).					

Table 117: Summary table of findings for web-based guided self-help versus web-based self-help at follow up for carers of people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Web-based Self-Help	Risk difference with Web-based Guided Self-Help (95% CI)
Carer Family Functioning FU LEE	29 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer family functioning fu in the intervention groups was 1.01 standard deviations lower (1.8 to 0.23 lower)
Carer Burden FU EDSIS, ECI Negative	29 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer burden fu in the intervention groups was 0.46 standard deviations higher (0.06 lower to 0.99 higher)
Carer Experience of Caregiving (ECI) Positive FU	29 (1 study) 3 months	⨁⨀⨀⨀ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer experience of caregiving (eci) positive fu in the intervention groups was 0.18 standard deviations higher (0.56 lower to 0.91 higher)
Carer Quality of Life FU	29 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer quality of life fu in the intervention groups was 0.11 standard deviations lower (0.63 lower to 0.4 higher)
Carer General Psychopathology (Distress) FU DASS-21	29 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias,		Not calculable for SMD	The mean carer general psychopathology (distress) fu in the intervention groups was 0.35 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Web-based Self-Help	Risk difference with Web-based Guided Self-Help (95% CI)
		imprecision		values	(1.09 lower to 0.39 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Hoyle 2013: Unclear randomization method, allocation concealment, participant and assessor blinding.
2 CI crosses 0.5 or -0.5 (SMD).
3 CI crosses both 0.5 and -0.5 (SMD).

1

2

6.4.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with anorexia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.4.4 Clinical evidence statements

6.4.4.1 Self-help or guided self-help and treatment as usual versus treatment as usual for carers of anorexia nervosa – carer outcomes

Very low quality evidence from one RCT (n=149) showed no difference in the effect of self-help or guided self-help and treatment as usual on general psychopathology (at 12 months) compared with treatment as usual.

6.4.4.2 Self-help and treatment as usual versus treatment as usual for carers of anorexia nervosa – carer outcomes

Low quality evidence from one RCT (n=147) showed self-help and treatment as usual may be less effective on family functioning (at 6 months) compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=147) showed no difference in the effect of self-help and treatment as usual on accommodation and enabling (at 6 months) and carer skills (at 12 months) compared with treatment as usual.

6.4.4.3 Self-help and treatment as usual versus treatment as usual for carers of anorexia nervosa – patient outcomes

Low quality of evidence from one RCT (n=99) showed no difference in the effect of self-help and treatment as usual on BMI, percentage of gender standardized weight for height, SEED, general psychopathology, clinical impairment and scores on the Strength and Difficulties Questionnaire-peer problems compared with treatment as usual.

Low quality of evidence from one RCT (n=99) showed self-help and treatment as usual may be less effective on the Strength and Difficulties Questionnaire-prosocial behaviour subscale compared with treatment as usual, although there was some uncertainty.

6.4.4.4 Guided self-help and treatment as usual versus treatment as usual for carers of anorexia nervosa – carer outcomes

Low quality evidence from one RCT (n=182) showed guided self-help and treatment as usual may be more effective on carer burden at 12 months compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=185) showed guided self-help and treatment as usual may be more effective on carer burden at 24 months compared with treatment as usual, although there was some uncertainty.

Low quality evidence from two RCTs (n=336) showed guided self-help and treatment as usual is more effective on carer accommodation and enabling at 12 months compared to treatment as usual.

Low quality evidence from one RCT (n=185) showed guided self-help and treatment as usual may be more effective on carer accommodation and enabling at 24 months compared to treatment as usual, although there was some uncertainty.

Low quality evidence from two RCTs (n=336) showed no difference in the effect of guided self-help and treatment as usual on family functioning at 12 months compared with treatment as usual.

Low quality evidence from one RCT (n=185) showed guided self-help and treatment as usual may be more effective on family functioning at 24 months compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=182) showed guided self-help and treatment as usual is more effective on quality of life at 12 months compared with treatment as usual.

Low quality evidence from one RCT (n=185) showed guided self-help and treatment as usual may be more effective on quality of life at 24 months compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=154) showed no difference in the effect of guided self-help and treatment as usual on carer skills and general psychopathology at 12 months compared with treatment as usual.

Low quality evidence from one RCT (n=185) showed guided self-help and treatment as usual may be more effective on general psychopathology and time spent caring at 24 months compared with treatment as usual, although there was some uncertainty.

6.4.4.5 Guided self-help and treatment as usual versus treatment as usual for carers of anorexia nervosa – patient outcomes

Very low to low quality evidence from one RCT (n=178) showed guided self-help and treatment as usual may be more effective on reducing the number of deaths, the number of people readmitted to hospital for treatment and the number of people suffering relapse during the course of the study compared with treatment as usual, although there was some uncertainty.

Low quality evidence from two RCTs (n=212) showed no difference in the effect of guided self-help and treatment as usual at 12 months on BMI and general psychopathology compared with treatment as usual.

Low quality evidence from one RCT (n=100) showed no difference in the effect of guided self-help and treatment as usual at 12 months on percentage gender standardized weight for height compared with treatment as usual.

Low quality evidence from one RCT (n=112) showed no difference in the effect of guided self-help and treatment as usual at 12 months on EDE-Q-global, SEED score, clinical impairment due to eating disorder and quality of life compared with treatment as usual.

Low quality evidence from one RCT (n=100) showed guided self-help and treatment as usual may be less effective at 12 months on the Strength and Difficulties Questionnaire-prosocial behaviour compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=100) showed guided self-help and treatment as usual at 12 months is more effective on the Strength and Difficulties Questionnaire-peer problems compared with treatment as usual.

Low quality of evidence from one RCT (n=119) showed guided self-help and treatment as usual may be more effective at 24 months on BMI, EDE-Q-global and quality of life compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=119) showed no difference in the effect of guided self-help and treatment as usual at 24 months on general psychopathology compared with treatment as usual.

6.4.4.6 Guided self-help and treatment as usual versus self-help and treatment as usual for carers of anorexia nervosa – carer outcomes

Low quality evidence from one RCT (n=151) showed guided self-help and treatment as usual may be more effective at 6 months on carer accommodation and enabling compared with self-help and treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=151) showed no difference in the effect of guided self-help and treatment as usual at 12 months on family functioning, general psychopathology, carer skills, time spent caring and direct spending compared with self-help and treatment as usual.

6.4.4.7 Guided self-help and treatment as usual versus self-help and treatment as usual for carers of anorexia nervosa – patient outcomes

Low quality evidence from one RCT (n=99) showed guided self-help and treatment as usual is less effective at 12 months on BMI compared with self-help and treatment as usual.

Low quality evidence from one RCT (n=99) showed guided self-help and treatment as usual may be less effective at 12 months on percentage of gender standardized weight for height compared with self-help and treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=99) showed no difference in the effect of guided self-help and treatment as usual at 12 months on SEED score, general psychopathology, clinical impairment due to the eating disorder and Strength and Difficulties Questionnaire -prosocial behaviour score compared with self-help and treatment as usual.

Low quality evidence from one RCT (n=99) showed self-help and treatment as usual is more effective at 12 months in improving scores on the Strength and Difficulties Questionnaire-peer problems compared with self-help and treatment as usual.

6.4.4.8 Web-based guided self-help versus treatment as usual for carers of anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=63) showed no difference in the effect of web-based guided self-help on accommodation and enabling and the positive experience of caregiving compared with treatment as usual.

Low quality from one RCT (n=63) showed web-based guided self-help may be more effective on family functioning and carer burden compared with treatment as usual, although there was some uncertainty.

Low quality evidence from one RCT (n=63) showed web-based guided self-help is more effective on general psychopathology compared with treatment as usual.

6.4.4.9 Web-based guided self-help versus treatment as usual for carers of anorexia nervosa at follow up

Low quality evidence from one RCT (n=63) showed no difference in the effect of web-based guided self-help on accommodation and enabling, family functioning, carer burden, positive experience of caregiving and general psychopathology compared with treatment as usual.

6.4.4.10 Web-based guided self-help versus web-based self-help at end of treatment for carers of anorexia nervosa

Low quality of evidence from one RCT (n=27) showed no difference in the effect of web-based guided self-help on family functioning compared with web-based self-help.

Low quality of evidence from one RCT (n=27) showed no difference in the effect of web-based guided self-help on carer burden, positive experience of caregiving, quality of life and general psychopathology compared with web-based self-help.

6.4.4.11 Web-based guided self-help versus web-based self-help at follow up for carers of anorexia nervosa

Low quality of evidence from one RCT (n=29) showed that web-based guided self-help is more effective on family functioning than web-based self-help.

Low quality of evidence from one RCT (n=29) showed web-based guided self-help may be less effective on carer burden compared with web-based self-help, though there was some uncertainty.

Very low to low quality of evidence from one RCT (n=29) showed no difference in the effect of web-based guided self-help on positive experience of caregiving, quality of life and general psychopathology compared with web-based self-help.

6.4.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with anorexia nervosa was available.

6.4.6 Recommendations and link to evidence for the review on: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Working with family members and carers

	69. Be aware that the family members or carers of a person with an eating disorder may experience severe distress. Offer them an assessment of their own needs, including: • what impact the eating disorder has on them • what support they need, including practical support and emergency plans for increasing medical or psychiatric risk. **70. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes, when assessing whether any interventions help the parents and carers of children and young people with an eating disorder. The critical outcomes for the parents and carers were: general psychopathology, family functioning, quality of life and other primary outcomes reported by the study. Other outcomes that are critical for the child or young person with the eating disorder include remission and bingeing or body weight, depending on the eating disorder. Other outcomes that are of lesser importance but clearly important outcomes include, general functioning, service user experience, all-cause mortality, adverse events and eating disorder psychopathology.
Trade-off between	Anorexia nervosa (reviewed in this chapter) One randomised controlled trial (RCT), aimed at carers of young people with

clinical benefits and harms	anorexia nervosa, and compared the effectiveness of guided self-help or self-help (and treatment as usual) with treatment as usual alone. After 12 months there was no difference in carer general psychopathology. No evidence was found on the critical outcomes of carer general psychopathology, carer family functioning, carer quality of life, nor the important outcomes of eating psychopathology, carer general functioning, service user experience, resource use, adverse events or all-cause mortality. Another study compared self-help (and treatment as usual) with treatment as usual and showed no difference in the carer's general psychopathology or carer skills after 6 to 12 months but a trend for poorer results for family functioning but there was some uncertainty. In the young people with anorexia nervosa whom they care for, there was no difference in BMI, weight, severity index (SEED), general psychopathology, clinical improvement, peer related problems between the two treatment arms. However, there was a trend for poorer outcomes in prosocial behaviour in the self-help group but there was some uncertainty. No evidence was found on the critical outcomes of remission, carer general psychopathology, nor on the important outcomes of service user experience, resource use, adverse events or all-cause mortality. Comparing guided self-help (and treatment as usual) with treatment as usual showed at 12 months a trend for positive outcomes in the combined treatment group on carer burden and quality of life, but no difference in family functioning, carer skills or carer psychopathology. There was a trend for poorer outcomes in carer accommodation and enabling. At 24 months, there was a trend for a positive result on carer burden, quality of life (though there was some uncertainty), carer accommodation and enabling and carer psychopathology. In addition, a trend for poorer outcomes in family functioning and time spent caring. No evidence was found on the critical outcome of carer general psychopathology, nor on the important outcomes of service user experience and resource use. In the same intervention, the guided self-help for the carers did not translate too many benefits in the young people with anorexia nervosa whom they care for. At 12 months, no differences were found in any of the outcomes including mortality, admission to hospital, patient relapse, BMI, EDE-global, severity index (SEED), general psychopathology and clinical improvement. However there was improvement in peer problems but a trend for a negative result in prosocial behaviour. At 24 months, there was a trend for positive increase in BMI and EDE-global, no difference in general psychopathology, and a trend for a negative result in quality of life. No evidence was found on the critical outcome of remission, nor on the important outcomes of adverse events and all-cause mortality. Comparing two active treatments generally showed no difference in effectiveness in the carer-related outcomes. Guided self-help compared with self-help were equally effective on all outcomes 6 to 12 months after the young people with anorexia nervosa had been discharged from inpatient care, except there was a trend for carer accommodation to favour guided self-help. No evidence was found on the critical outcomes of carer quality of life, nor on the important outcomes of carer general functioning, service user experience and resource use. In the young people with anorexia nervosa, there was a trend for poorer results on BMI and peer problems in the guided self-help group compared with self-help. No difference was found in clinical severity (SEED), general psychopathology, clinical improvement, prosocial behaviour but there was a trend for better results in peer problems. No evidence was found on the critical outcome of remission. Web-based guided self-help also failed to show convincing benefits for the carers of young people with anorexia nervosa compared with treatment as usual. At the end of treatment, a poorer outcome in distress was found but there was some uncertainty. The other outcomes, such as carer accommodation and enabling, family functioning, carer burden and caregiving experience showed no difference. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. Web-based guided self-help compared with web-based self-help showed no difference in the outcomes for carers at the end of treatment. At follow up, favourable results were found on family functioning in the guided web-based self-

	help group, but no difference in carer experience, quality of life, and general psychopathology. There was a trend for poorer results in carer burden, but there was some uncertainty. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. **Bulimia nervosa or binge eating disorder (chapter 7 and 8)** No relevant published evidence was found on parents or carers of children and young people with bulimia nervosa or binge eating disorder. **Any eating disorder (as reviewed in chapter 9)** Randomised control trials investigating interventions for the carers of young people with any eating disorder failed to show many favourable outcomes. Psychoeducation compared with waitlist control showed a positive effect on carer self-efficacy and a trend to improve carer knowledge of eating disorders at the end of treatment. Long-term follow up (unclear duration) showed favourable results in both but carer burden (only measured at follow up) was not different compared with wait list controls. No evidence was found on the critical outcomes of carer general psychopathology, family functioning, and quality of life, nor on the other important outcomes. Comparing guided self-help with self-help showed no difference in any of the carer-related outcomes at the end of treatment. No evidence was found on the other important outcomes.
Trade-off between net health benefits and resource use	The committee expressed the view that offering family members and carers an assessment of their own needs may incur additional healthcare resources (that is, time required to perform such assessment). However, the committee considered the cost of providing such assessment to be small, taking into account the potential reduction in family and carers' burden, potential depression and other health vulnerabilities that may be costly to other parts of the healthcare system, especially considering that the burden on family and carers can last for many years and increase their morbidity and stress. Consequently, the committee judged that assessment that aims to improve family and carers' experience are likely to represent good value for money.
Quality of evidence	The quality of the evidence was mostly very low. The outcomes were downgraded because of unclear randomisation method, it was unclear if allocation concealment was performed or if participants and investigators were blinded. In some, not all, assessors were blinded. High dropout rates were also detected in some groups >20%. Imprecision was detected in most outcomes due to the 95% confidence interval crossing one or two minimal important differences or because it did not meet the optimal information size. Outcomes were not always measured at the end of treatment or at follow up. It is not known if any improvements in the carer's general psychopathology also translated to benefits in the children with the eating disorder.
Other considerations	Given the very low quality of the data with very few positive findings favouring one arm over the other, the committee came to the consensus that there was not enough evidence to support a recommendation on any specific treatment for parents or carers of people with an eating disorder. Nevertheless, the committee acknowledged the stress and burden that a person with an eating disorder, in particular anorexia nervosa, can have on family members over a long period of time. Therefore, they agreed that offering family members and carers an assessment of their own needs, including: personal, social and emotional support available to them, the need for support in the caring role for example if the child should need urgent care and there are other children to take care of, and to offer advice on where they can get some practical support. The extent to which the family need to be involved in treatment depends on the age and developmental needs of the person with the eating disorder, the severity of the illness, the risk from harm and the person receiving treatment's wishes. In general, parents and other family members will want to be involved in the treatment. If a parent or carer cannot attend a meeting the healthcare professional should provide written information on the outcome of an assessment or treatment where appropriate. The committee acknowledged the importance of consent and confidentiality, and

	their discussion can be found in the LETRs relating to this. They also discussed that although the evidence found was for carers and parents of people with anorexia nervosa or any eating disorder, the recommendation is relevant for parents and carers of people with bulimia nervosa and binge eating disorder. This is mostly because no specific intervention was recommended, rather to offer an assessment of their needs and to help them find the necessary support. In absence of good evidence, the committee agreed to generate a research recommendation to address the question "What is the effectiveness of a carer-focused psychological intervention in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?"

4. **Research recommendation: "What is the effectiveness of a carer-focused psychological intervention in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?"**

6.5 Nutritional interventions

6.5.1 Review question: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 118. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all nutritional interventions that may be delivered to children, young people and adults with an eating disorder with or without a pharmacological intervention. The interventions were categorised according to the type of eating disorder being treated.

Table 118: Clinical review protocol summary for the review of: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

Topic	Interventions to treat eating disorders in children, young people and adults
Review question	Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder Strata: children (≤12), young people (13-≤17 years), adults ≥18 years eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder and atypical eating disorder)
Intervention	• Nutritional intervention • Nutritional intervention in combination with a pharmacological intervention
Control	• Wait list control • Placebo • Treatment as usual • Another intervention
Critical outcomes for decision-making	• Remission and long-term recovery (if symptoms were measured over a minimum two week period)

Topic	Interventions to treat eating disorders in children, young people and adults
	• Binge eating for BN and BED. • Body weight / BMI for AN.
Important, but not critical outcomes	• Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF). • Family functioning. • Adverse events • Resource use. • All-cause mortality. • Quality of life. • Relapse. • Service user experience (in patient vs. community).
Study design	Systematic Reviews RCTs

6.5.2 Clinical evidence: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

6.5.2.1 Nutritional intervention versus any other intervention or wait list control

Three RCTs (N =117) met the eligibility criteria for this review and were all on adults with an eating disorder (Birmingham 1994 (Birmingham et al., 2004a), Hall 1987 (Hall and Crisp, 1987), Pike 2003 (Pike et al., 2003)).

No nutritional intervention studies on any other eating disorder were identified that met the eligibility criteria for this review.

Summary of findings for those on anorexia nervosa can be found in Table 116. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 119: Study information for trials included in the meta-analysis of a nutritional intervention versus any other intervention or wait list controls in people with anorexia nervosa.

Study	Age mean (SD)	Body weight	Stage of illness for AN	Female	Randomised	Experimental arm	Control arm	Duration
Birmingham 1994	20.6 (3.8)	15.6 (1.2) BMI	Years since diagnosis: 3.6 (2.0)	100%	54	Zinc supplement (50 mg) containing 14 mg of elemental zinc	Placebo	25 days
Hall 1987	19.6 (14 to 25)	Deviation from matched populatio n mean weight 25.4%	Mean duration of illness 29.7 months	100%	30	Nutritional counselling	Family/individual psychotherapy	12-24 weeks 6 months FU
Pike 2003	26.1 (6.2)	16.0 (2.1) BMI	1 year after hospitalis ation	100%	33	Nutritional counselling	CBT-ED	12 months

Abbreviations: CBT-ED – cognitive behavioural therapy with an eating disorder focus; FU – follow up.

Table 120: Summary table of findings for nutritional counselling versus any other intervention.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with AN. Nutritional counselling (95% CI)
Did not achieve remission (ITT)	33 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.68 (1.09 to 2.59)	444 per 1000	302 more per 1000 (from 40 more to 707 more)
Relapse	33 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 2.40 (0.9 to 6.43)	222 per 1000	311 more per 1000 (from 22 fewer to 1000 more)
Weight FU	30	⊕⊕⊖⊖		Not	The mean weight fu in the intervention

	(1 study)	LOW4,5 due to risk of bias, imprecision		calculable for SMD values	groups was 0.11 standard deviations higher (0.61 lower to 0.82 higher)
Menstruation absent FU	30 (1 study)	⊕⊖⊖⊖ VERY LOW4,6 due to risk of bias, imprecision	RR 1.25 (0.69 to 2.26)	533 per 1000	133 more per 1000 (from 165 fewer to 672 more)
Menstruation regular FU	30 (1 study)	⊕⊖⊖⊖ VERY LOW4,6 due to risk of bias, imprecision	RR 1 (0.24 to 4.18)	200 per 1000	0 fewer per 1000 (from 152 fewer to 636 more)
Did not achieve remission (ITT) FU	30 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, indirectness, imprecision	RR 1.35 (0.98 to 1.85)	267 per 1000	93 more per 1000 (from 5 fewer to 227 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how randomisation was conducted, and if allocation concealment was performed. It was unclear if either the participants, investigators or assessors were blind. High dropouts were reported >20%.

2 95% CI crossed 1 MID (0.75).

3 95% CI crossed 1 MID (1.25).

4 It was unclear how randomisation was conducted, and if allocation concealment was performed. It was unclear if either the participants or investigators were blind. The assessors were blinded. High dropouts were reported >20%.

5 95% CI crossed 2 MIDs (-0.5 and 0.5).

6 95% CI crossed 2 MIDs (0.75 and 1.25).

7 No definition provided. Based on investigators decision if further treatment is required

1 .

2 **Table 121: Summary table of findings for zinc versus placebo in adults with anorexia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with AN. Zinc (95% CI)
BMI gain/day	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias,		Not calculable for SMD	The mean BMI gain/day in the intervention groups was 0.6 standard deviations higher

		imprecision		values	(0.08 lower to 1.29 higher)
Did not have side-effects	35 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 1 (0.9 to 1.11)		
				Not estimable.	Not estimable.
% body fat gain/day	35 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean % body fat gain/day in the intervention groups was 0.67 standard deviations higher (0.02 lower to 1.36 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 It was unclear how the random sequence was generated or if they performed allocation concealment. Participants and staff were blind but it was unclear if assessors were blind. High dropout rates were detected >20%. 2 95% CI crossed 2 MIDs (-0.5 and 0.5). 3 For a dichotomous outcome, there were fewer than 300 events. 4 95% CI crossed 1 MID (0.5).					

1

2

3

6.5.3 Economic evidence

No economic evidence on the cost effectiveness of nutritional interventions for people with anorexia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.5.4 Clinical evidence statements for nutritional interventions people with anorexia nervosa

6.5.4.1 Nutritional counselling versus another intervention in adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=33) showed nutritional counselling is equally effective as any other intervention on relapse rates.

Low quality evidence form one RCT (n=33) showed nutritional counselling is less effective than any other intervention on remission rates.

6.5.4.2 Nutritional counselling versus another intervention in adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=30) showed nutritional counselling is equally effective as any other intervention on body weight.

Very low quality evidence from one RCT (n=30) showed nutritional counselling is equally effective as any other intervention on menstrual function (absent and regular).

Low quality evidence from one RCT (n=30) showed nutritional counselling is less effective than any other intervention on remission rates.

6.5.4.3 Zinc supplementation versus placebo in adults with anorexia nervosa at end of treatment

Very low quality evidence from one RCT (n=35) showed zinc supplementation is more effective on BMI gain per day than compared with placebo but there was some uncertainty.

Low quality evidence from one RCT (n=35) showed no difference in side-effects reported between zine supplementation and placebo.

Low quality evidence from one RCT (n=35) showed zinc supplementation may be more effective on percent fat gain per day than compared with placebo but there was some uncertainty.

6.5.5 Economic Evidence Statements

No economic evidence on the cost effectiveness of nutritional interventions for people with anorexia nervosa was available.

6.5.6 Recommendations and link to evidence for the review on: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

Dietary counselling for anorexia nervosa

	71. Only offer dietary counselling as part of a multidisciplinary approach. **72. Encourage people with anorexia nervosa to take an age-appropriate oral multi-vitamin and multi-mineral supplement until their diet includes enough to meet their dietary reference values.** **73. Include family members or carers (as appropriate) in any dietary education or meal planning for children and young people with anorexia nervosa who are having therapy on their own.** **74. Offer individualised supplementary dietary advice to children and young people with anorexia nervosa and their parents or carers (if appropriate) to help them meet their nutritional needs for growth and development (particularly during puberty).**
Critical and important outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of nutritional interventions for treating anorexia nervosa in children, young people and adults. For this population, body weight or BMI and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade off benefits and harms	**Anorexia nervosa** Two studies were identified that investigated the effects of nutritional counselling on adults with anorexia nervosa one reported data at end of treatment, the other at 6 months follow up. At the end of treatment, remission rates were lower in the nutritional counselling group compared with any other treatment. The other outcome reported was relapse rates and this was not different between the two treatment arms. At follow up, another study showed body weight was not different between nutritional counselling and any other treatment. Menstrual function (absent and regular) was also similar. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, general psychopathology or relapse. No nutritional intervention studies were identified in children and young people One study on zinc supplementation was found. Compared with placebo, BMI gain per day and percent gain in body fat per day was higher in the zinc supplemented arm but there was some uncertainty (results just crossed the line of no effect). No harms were detected. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, general psychopathology or relapse. No other supplementation studies were found.
Trade-off between net health benefits and resource use	The committee expressed the view that dietary advice is an integral part of most eating disorder specific psychological interventions and providing such supplementary advice would not incur significant extra resource implications to the healthcare system.
Quality of the	The quality of the evidence low to very low quality. In the nutritional counselling

evidence	and zinc supplementation studies, it was unclear how randomisation was conducted and if allocation concealment was performed. It was unclear if either the participants or investigators were blind. In one nutritional study at follow up, the assessors were blinded, however, high dropouts were reported >20%. Imprecision was also detected mostly because either the 95% confidence interval crossed one or more minimal important differences or the evidence did not meet the optimal information size (300 events or 400 participants).
Other considerations	The committee agreed that the evidence was not strong enough to recommend nutritional counselling as the sole treatment for adults with anorexia nervosa. However, they highlighted that dietary advice and counselling are an integral part of CBT-ED, SSCM, MANTRA and family therapy, so it is not generally needed if the person is receiving therapy. Usually this is delivered by the therapist and sometimes in collaboration with a dietician. Whether or not it is part of the psychotherapy the children and young people with anorexia nervosa are receiving, dietary advice should be offered to ensure their food intake (including calcium and vitamin D) meets their nutritional needs for growth and development particularly during puberty. The committee considered the zinc supplementation evidence and were not confident recommending one supplement on its own or using evidence from 1 study alone. They agreed that people with anorexia nervosa may be deficient in many nutrients and decided it is best to encourage them to take age-appropriate multivitamins and multi-mineral supplements in oral form until dietary intake meets Dietary Reference Values. For children and young people with anorexia nervosa, there was no RCT evidence on how to address the nutritional needs for growth and development particularly during puberty. However, the nutritionist and paediatrician on the committee discussed the importance of offering advice to children and young people on this, in particular the need for an adequate intake of calcium and vitamin D for the growth and development of healthy bones. There is ample evidence on the effectiveness of calcium supplements in children but no studies have been published in those with anorexia nervosa. Nevertheless, these studies have shown calcium supplements are the most effective in children with a low intake of calcium. For this reason, the committee suggested that health professionals should promote adequate calcium and vitamin D through diet, but if intake is below dietary reference values then consider supplementation. If children or young people are receiving individual treatment, and not family-based therapy, the committee agreed that it was important to generate a recommendation that parents or carers (as appropriate) are included in the provision of any dietary education or meal planning. The health professional, most likely a nutritionist, are likely to discuss the nutritional value of certain foods, help devise a plan and take the personal preferences into account. The primary goal of dietary education and/or meal planning, as with all treatments for anorexia nervosa, is to achieve and maintain a healthy body weight.

6.6 Pharmacological interventions

6.6.1 Review Question: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 122. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all pharmacological interventions that may be delivered to children, young people and adults with an eating disorder. The interventions were categorised according to type of physical intervention, the age of the participants and the type of eating

disorder and were compared to wait list controls, placebo, treatment as usual or any other intervention.

Table 122: Clinical review protocol summary for the review of: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

Topic	Interventions to treat eating disorders in children, young people and adults
Review question	Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder Strata: children (≤12), young people (13-≤17 years), adults ≥18 years eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder and atypical eating disorder)
Intervention	Pharmacological intervention Pharmacological + psychological: Pharmacological interventions may include: Anti-depressants i.e. SSRIs, Fluoxetine – Prozac Anxiolytic (antianxiety) Antipsychotic Anti-emetic medication. i.e. Ondansetron Anticonvulsant topiramate/antiepileptic (Topomax) Appetite suppressant (i.e. lisdexamf(ph)etamine dimesylate)
Control	Placebo Wait list control Treatment as usual Another intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes for decision-making	Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) Binge eating for BN and BED. Body weight / BMI for AN. Adverse events
Study design	Systematic Reviews RCTs

6.6.2 Clinical evidence for: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

16 RCTs (n=555) fulfilled the criteria for the review on the effect of pharmacological interventions on people with anorexia nervosa, the majority were on adults (Andries 2014 (Andries et al., 2014), Andries 2015 (Andries et al., 2015), Attia 1998 (Attia et al., 1998), Attia 2011 (Attia et al., 2011), Bissada 2008 (Bissada et al., 2008), Brambilla 2007b (Brambilla et al., 2007), Brambilla 2007a (Brambilla et al., 2007), Court 2010 (Court et al., 2010), Fassino 2002 (Fassino et al., 2002), Hagman 2011 (Hagman et al., 2011), Halmi 1986 (Halmi et al., 1986), Kaye 2001 (Kaye et al., 2001), Kafantaris 2011 (Kafantaris et al., 2011), Misra 2013 (Misra et al., 2013), Ruggiero 2001 (Ruggiero et al., 2001), Walsh 2006 (Walsh et al., 2006)).

Summary of findings for those on anorexia nervosa can be found in Table 121. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 123: Study information for trials included in the analysis of pharmacological interventions versus any other intervention or wait list control for people with anorexia nervosa.

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N Initially Random-ised	Intervention Category	Comparison Arm Category	Treatment Length
Individual therapy								
Andries 2014 & Andries 2015	33.3 (28.0 to 38.7)	15.7 (15 to 16.4)	100%	Duration of illness 15 (10.2 to 19.9) years	25	Dronabinol (synthetic cannabinoid agonist) 5mg/day	Placebo	4 weeks
Attia 1998	26.2 (7.4)	15.0 (4.2)	100%	In patient	33	fluoxetine (SSRI, 20mg to 60 mg/day) + CBT	Placebo+ CBT	37 days
Attia 2011	27.7 (9.1)	17.1 (1.3)	99%	NR	23	Olanzapine (Antipsychotic, 7.95 mg)	Placebo	8 weeks
Bissada 2008	23.61 (6.5)	16.4 (1.1)	100%	NR	34	Olanzapine. (Antipsychotic, 2.5 to 10 mg/day.	Placebo	10 week
Brambilla 2007b	23 (4.8)	15.7 (2.1)	100%	Duration of illness 5.1 (4.0) years	20	Olanzapine (Antipsychotic, 2.5 to 5 mg/day)+CBT	CBT+Nutrition+Placebo	3 months
Brambilla 2007a	23.7 (4.8)	15.5 (1.9)	100%	Duration of illness 6.3 (5.0) years	35	Olanzapine (Antipsychotic, 2.5 to 5 mg/day)+CBT	Placebo + Psychotherapy	3 months
Court 2010	23.8 (9.4)	16.9 (1.7)	99%	NR	33	Quetiapine (antipsychotic, 50 to 400 mg/day)	Treatment as usual	12 weeks
Fassino 2002	24.3 (5.4)	16.2 (0.8)	100%	Duration of illness: 5.7 (4.9) years	52	Citalopram (SSRI)	Wait list control	3 months 52 weeks FU
Hagman 2011	16.2 (2.5)	IBW 78 (5)%	100%	NR	41	Risperidone (antipsychotic, 0.5 to 4 mg/day)	Placebo	7 weeks
Halmi 1986	20.56 (5.1)	79 (7)% target	100%	Duration of illness 2.9 (2.3)	72	Amitriptyline (TCA, max 160 mg/day)	Placebo	unclear up to 35 days

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N Initially Random-ised	Intervention Category	Comparison Arm Category	Treatment Length
		weight		years		Cyproheptadine (Antihistamine, max 32 mg/day)		
Kaye 2001	23.0 (9.0)	% Average BW: 89 (6)	100%	Duration of illness: 7 years	39	Fluoxetine (SSRI, 20 mg/day up)	Placebo	12 months
Kafantaris 2011	17.1 (2.4)	16.4 (1.2)	100%	NR	20	Olanzapine(antipsychotic, 85 mg/day)	Placebo	10 weeks
Misra 2013	16.9 (0.2)	17.2 (0.2)	100%	Duration since diagnosis: 14.5 (2.8) months	72	Transdermal 17β-estradiol	Placebo	18 months
Ruggiero 2001	24.5 (5.1)	40.9 (7.0) kg	NR	Hospitalized weight restoration treatment participated in the study.	35	Fluoxetine (SSRI, 26 mg/day) Clomipramine (TCA. 57 mg/day	Antipsychotic. D2 and D3 receptor antagonist, 50mg/day	3 months
Walsh 2006	22.4 (4.5)	20.2 (0.5)	100%	Duration of illness: 4.1 (3.1) years	93	Fluoxetine + CBT (SSRI, 20 to 60mg/day + CBT)	Placebo + CBT	12 months

1 *Abbreviations: CBT – cognitive behavioural therapy; FU – follow up; SSRI – selective serotonin reuptake inhibitor; TCA - tricyclic antidepressants.*

2 **Table 124: Summary of findings table for antidepressant (SSRI or TCA) compared with placebo in adults with anorexia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Antidepressant (95% CI)
BMI. Adults - SSRIs	52 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean bmi. adults - ssris in the intervention groups was 0.72 standard deviations higher (0.16 to 1.29 higher)

Change in % average body weight. Adults - SSRIs	23 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean change in % average body weight. adults - ssris in the intervention groups was 0.61 standard deviations lower (1.45 lower to 0.23 higher)
Depression. Adults	88 (2 studies)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression. adults in the intervention groups was 0.58 standard deviations lower (1.01 to 0.15 lower)
Depression. Adults - SSRI	52 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression. adults - ssri in the intervention groups was 0.67 standard deviations lower (1.23 to 0.11 lower)
Depression. Adults - TCA	36 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression. adults - tca in the intervention groups was 0.45 standard deviations lower (1.12 lower to 0.22 higher)
EDI-Bulimia. Adults - SSRI	52 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - bulimia. adults - ssri in the intervention groups was 0.26 standard deviations lower (0.81 lower to 0.28 higher)
Achieved target weight. Adults - TCA	48 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, imprecision, publication bias	RR 1.15 (0.7 to 1.42)	640 per 1000	96 more per 1000 (from 192 fewer to 269 more)

Relapse (LSE because of deteriorating clinical state). Adults - SSRIs	35 (1 study)	⨁⊖⊖⊖ VERY LOW3,4,8 due to risk of bias, imprecision, publication bias	RR 0.45 (0.23 to 0.86)	842 per 1000 463 fewer per 1000 (from 118 fewer to 648 fewer)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how random sequence was generated and if allocation concealment was conducted. Neither the participants, assessors nor investigators were blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (0.5).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 It was unclear how random sequence was generated and if allocation concealment was conducted. The participants and investigators were blind but it was unclear if the assessors were blind. High dropouts were reported >20%.
5 95% CI crossed 1 MID (-0.5).
6 It was unclear how random sequence was generated and if allocation concealment was conducted. In one study, neither the participants, assessors nor investigators were blind. The other study was double blind but it was unclear if assessors were blind. High dropouts were reported >20%.
7 95% CI crossed 1 MID (1.25).
8 95% CI crossed 1 MID (0.75).

Table 125: Summary of findings table for antidepressant (SSRI) compared with another antidepressant (TCA) in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with Antidepressant	**Risk difference with Antidepressant (95% CI)**
No episodes of vomiting. Adults - SSRI vs. TCA	23 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias	RR 0.61 (0.37 to 1.01)	0 per 1000	-
Bingeing. Adults - SSRI vs. TCA	23 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias	RR 1.3 (0.68 to 2.48)	538 per 1000	162 more per 1000 (from 172 fewer to 797 more)
Amenorrhea. Adults - SSRI vs. TCA	23 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias,	RR 1.3 (0.68 to 2.48)	538 per 1000	162 more per 1000 (from 172 fewer to 797 more)

		imprecision, publication bias			

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how random sequence was generated and if allocation concealment was conducted. The participants and investigators were blind but it was unclear if the assessors were blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (0.75).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 95% CI crossed 2 MIDs (0.75 and 1.25).

1 **Table 126: Summary of findings table for antipsychotic compared with placebo in adults with anorexia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Antipsychotic (95% CI)
Weight - Adults	57 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean weight - adults in the intervention groups was 0.15 standard deviations lower (0.67 lower to 0.37 higher)
Depression - Adults	26 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression - adults in the intervention groups was 0.54 standard deviations higher (0.25 lower to 1.32 higher)
No side-effects - Young people	60 (2 studies)	⊕⊖⊖⊖ VERY LOW3,6,7 due to risk of bias, imprecision, publication bias	RR 1.04 (0.91 to 1.18)	62 per 1000	2 more per 1000 (from 6 fewer to 11 more)
No side-effects - Adults	34 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,8 due to risk of bias, imprecision, publication bias	Not estimable	See comment	-
Remission - Young people_ITT	41 (1 study)	⊕⊖⊖⊖ VERY LOW3,9,10	RR 0.69 (0.31 to	455 per 1000	141 fewer per 1000 (from 314 fewer to 250 more)

		due to risk of bias, imprecision, publication bias	1.55)		

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 High dropouts were reported in one study.
2 95% CI crossed 1 MID (-0.5).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 95% CI crossed 1 MID (0.5).
5 Studies were randomised, however it was unclear if allocation concealment was conducted. Two studies were triple-blinded and one was double-blinded. High dropouts were reported >20%.
6 For a dichotomous outcome there were fewer than 300 events.
7 Studies were randomised, however it was unclear if allocation concealment was conducted. One study was triple-blinded and one was double-blinded. High dropouts were reported >20%.
8 It was unclear if allocation concealment was conducted. The study was triple-blinded. High dropouts were reported >20%
9 Studies were randomised, however it was unclear if allocation concealment was conducted. The study was double-blinded but it was unclear if assessors were blind.
10 95% CI crossed 2 MIDs (0.75 and 1.25)

Table 127: Summary of findings table for antipsychotic and psychotherapy (CBT and nutritional) compared with placebo and psychotherapy (CBT and nutritional) in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo + Therapy	Risk difference with Combined Antipsychotic + Psychotherapy (95% CI)
BMI. Adults	30 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean bmi. adults in the intervention groups was 0.18 standard deviations higher (0.54 lower to 0.89 higher)
EDI-Total. Adults	30 (1 study)	⨁⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - total. adults in the intervention groups was 0.47 standard deviations higher (0.26 lower to 1.19 higher)
EDI-Drive for thinness. Adults	30	⨁⊖⊖⊖		Not calculable	The mean edi - drive for thinness. adults in

	(1 study)	VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		for SMD values	the intervention groups was 0.36 standard deviations higher (0.37 lower to 1.08 higher)
EDI-Bulimia. Adults	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - bulimia. adults in the intervention groups was 0.18 standard deviations higher (0.54 lower to 0.9 higher)
EDI - Body dissatisfaction. Adults	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - body dissatisfaction. Adults in the intervention groups was 0.43 standard deviations higher (0.29 lower to 1.16 higher)
Yale - eating disorder rating scale. Adults	30 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean yale - eating disorder rating scale. adults in the intervention groups was 0.53 standard deviations lower (1.26 lower to 0.2 higher)
No side-effects. Adults	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias	RR: 1.00 (0.90 to 1.10)	See comment**	-

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

**Because the results included zero events an absolute risk difference could not be calculated.
CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how random sequence was generated or if allocation concealment was conducted. The study was double-blind but it was unclear if allocation concealment was conducted.
2 95% CI crossed 1 MID (0.5).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 It was unclear how random sequence was generated or if allocation concealment was conducted in both studies. The study was double-blind but it was unclear if allocation concealment was conducted.
5 95% CI crossed 1 MID (-0.5).

Table 128: Summary of findings table for antidepressant and psychotherapy compared with psychotherapy in adults with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Therapy	Risk difference with Combined Antidepressant + Psychotherapy (95% CI)
Weight % Ideal BW (final)-SSRI Adult	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean weight % ideal bw (final)-ssri adult in the intervention groups was 0.14 standard deviations lower (0.85 lower to 0.56 higher)
Weight % Ideal BW (change)-SSRI Adult	93 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean weight % ideal bw (change)-ssri adult in the intervention groups was 0.46 standard deviations lower (0.87 to 0.04 lower)
Depression (change and final) SSRI Total	124 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The difference in mean depression score (change and final) ssri total in the intervention groups was 0.32 standard deviations higher (0.03 lower to 0.68 higher)
Quality of life SSRI Adults	93 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean quality of life ssri adults in the intervention groups was 0.38 standard deviations lower (0.79 lower to 0.03 higher)
Remission SSRI Adults_ITT	93 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,6 due to risk of bias, imprecision, publication bias	RR 1.57 (0.49 to 5.01)	91 per 1000	52 more per 1000 (from 46 fewer to 365 more)
Global Improvement (CGI) SSRI Adults	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean global improvement (cgi) ssri adults in the intervention groups was 0.20 standard deviations lower (0.91 lower to 0.51 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. Studies were triple blinded. High dropouts were reported >20%.
2 95% CI crossed 1 MID (0.5).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 95% CI crossed 1 MID (-0.5).
5 For a continuous outcome there were fewer than 400 participants.
6 95% CI Crossed 2 MIDs (0.75 and 1.25).
7 95% CI crossed 2 MID (-0.5 and 0.5).

Table 129: Summary of findings table for antihistamine compared with placebo in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Other medication (not antidepressants) (95% CI)
Achieved target weight. Adults - Antihistamine	48 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias	RR 1.15 (0.79 to 1.69)	640 per 1000	96 more per 1000 (from 134 fewer to 442 more)
Depression, Adults - Antihistamine	38 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression, adults - antihistamine in the intervention groups was 0.58 standard deviations lower (1.23 lower to 0.07 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how random sequence was generated or if allocation concealment was conducted. The study was double-blind but it was unclear if assessor was blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (1.25).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 95% CI crossed 1 MID (-0.5).
.

Table 130: Summary of findings table for antipsychotics compared with antidepressant in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant	Risk difference with AN Antipsychotics (95% CI)
No bingeing	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias	RR 0.87 (0.61 to 1.24)	174 per 1000	23 fewer per 1000 (from 68 fewer to 42 more)
No vomiting	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias	RR 0.87 (0.6 to 1.25)	130 per 1000	17 fewer per 1000 (from 52 fewer to 33 more)
Amenorrhea	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias	RR 1.08 (0.65 to 1.81)	609 per 1000	49 more per 1000 (from 213 fewer to 493 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how the randomisation sequence was generated or if allocation concealment was conducted. Participants were blind, but investigators were not. It was unclear if the assessors were blind.
2 95% CI crossed 1 MID (0.75).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 95% CI crossed 2 MIDs (0.75 and 1.25).

Table 131: Summary of findings table for cannabinoid agonist compared with placebo in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Cannaboid agonist (95% CI)
Weight gain. Adults	48 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD	The mean weight gain. adults in the intervention groups was 1.6 standard deviations higher

		imprecision		values	(0.95 to 2.26 higher)
Intensity of physical activity. Adults	48 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean intensity of physical activity. adults in the intervention groups was 0.18 standard deviations higher (0.39 lower to 0.74 higher)
Change in total EDI-2. Adults	48 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in total edi-2. adults in the intervention groups was 0.78 standard deviations lower (1.36 to 0.19 lower)
Change in EDI-2 Body dissatisfaction. Adults	48 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in edi-2 body dissatisfaction. adults in the intervention groups was 0.07 standard deviations lower (0.64 lower to 0.5 higher)
Change EDI-2 Drive for thinness. Adults	48 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change edi-2 drive for thinness. adults in the intervention groups was 1.15 standard deviations higher (0.53 to 1.76 higher)
Change in EDI-2 Bulimia. Adults	48 (1 study)	⊕⊕⊕⊖ MODERATE1 due to risk of bias		Not calculable for SMD values	The mean change in bulimia. adults in the intervention groups was 0.72 standard deviations higher (0.13 to 1.3 higher)
No adverse events. Adults	25 (1 study)	⊕⊕⊕⊖ MODERATE1 due to risk of bias	Not estimable	See comment	-

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 The study was double-blind but it was unclear if investigator was blind.
2 95% CI crossed 1 MID (0.5).
3 95% CI crossed 1 MID (-0.5).

Table 132: Summary of findings table for oestrogen treatment compared with placebo in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo (AN)	Risk difference with oestrogen(95% CI)
Change in body weight – Young people	37 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in body weight - young people in the intervention groups was 0.38 standard deviations higher (0.27 lower to 1.03 higher)
Dropout for any reason - Young people	72 (1 study)	⊕⊖⊖⊖ VERY LOW2,3 due to risk of bias, imprecision	RR 0.95 (0.59 to 1.52)	500 per 1000	25 fewer per 1000 (from 205 fewer to 260 more)
Change in fat mass - Young people	37 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in fat mass - young people in the intervention groups was 0.98 standard deviations lower (1.67 to 0.29 lower)
Change in EDI-Drive for thinness - Young people	37 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in edi - drive for thinness - young people in the intervention groups was 0.38 standard deviations lower (1.03 lower to 0.28 higher)
Change in EDI-Bulimia - Young people	37 (1 study)	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in edi - bulimia - young people in the intervention groups was 1.41 standard deviations lower (2.14 to 0.68 lower)
Change EDI-Body dissatisfaction - Young people	37 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change edi - body dissatisfaction - young people in the intervention groups was 0.8 standard deviations higher (0.12 to 1.47 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 95% CI crossed 1 MID (0.5).
2 Unclear if allocation concealment was performed. It is also unclear either the participants, investigators or assessors were blind. High dropouts were reported >20%.
3 95% CI crossed 2 MIDs (0.75 and 1.25).
4 95% CI crossed 1 MID (-0.5).
5 Fewer than 400 participants were used in meta-analysis.

1

2

6.6.3 Economic evidence

No economic evidence on the cost effectiveness of pharmacological interventions for people with anorexia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.6.4 Clinical evidence statements

6.6.4.1 Antidepressant versus placebo in adults with anorexia nervosa at the end of treatment

SSRIs

Very low quality evidence from one RCT (n=35 to 52) showed SSRI's may have a positive effect on BMI, depression and relapse rates compared with placebo.

Very low quality evidence from one RCT (n=23) showed no difference on the effect of SSRI's on change in percent body weight and EDI- bulimia compared with placebo.

TCAs

Very low quality evidence from one RCT (n=36 to 48) showed no difference on the effect of TCAs on achieving target body weight and depression compared with placebo.

6.6.4.2 Antidepressant versus another antidepressant in adults with anorexia nervosa at the end of treatment

SSRIs vs. TCAs

Very low quality evidence from one RCT (n=23) showed no difference in the effect of SSRI's on bingeing and amenorrhea compared with TCAs.

Very low quality evidence from one RCT (n=23) showed SSRI's may have a positive effect on reducing vomiting rates compared with TCAs, but there was some uncertainty.

6.6.4.3 Antipsychotics versus placebo in adults with anorexia nervosa at the end of treatment

Very low quality evidence from two RCTs (n=57) showed no difference in the effect of antipsychotics on weight compared with placebo.

Very low quality evidence from one RCT (n=26 to 34) showed no difference on the effect of antipsychotics on depression and the number of side-effects compared with placebo.

6.6.4.4 Antipsychotics versus placebo in young people with anorexia nervosa at the end of treatment

Very low quality evidence from two RCTs (n=60) showed no difference in the effect of antipsychotics on the number of side-effects compared with placebo.

Very low quality evidence from one RCT (n=41) showed no difference in the effect of antipsychotics on remission compared with placebo.

6.6.4.51 Antidepressant and psychotherapy (CBT and nutritional) compared with placebo and psychotherapy (CBT and nutritional) in adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n=31) showed no difference in the effect of antidepressant and psychotherapy on body weight and global improvement compared with psychotherapy.

Very low quality evidence from one RCT (n=91) showed no difference in the effect of antidepressant and psychotherapy on remission compared with psychotherapy.

Very low quality evidence from two RCTs (n=93) showed the change in weight gain is less in the antidepressant and psychotherapy group compared with psychotherapy.

Very low quality evidence from one RCT (n=93) showed the quality of life is less in the antidepressant and psychotherapy group compared with psychotherapy, but there was some uncertainty.

6.6.4.64 Antihistamine versus placebo in adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n=48) showed no difference in the effect of antidepressant and psychotherapy group on the numbers who achieved their target body weight compared with placebo.

Very low quality evidence from one RCT (n=38) showed antidepressant and psychotherapy group may be more effective on depression compared with placebo, but there was some uncertainty.

6.6.4.71 Antipsychotics versus antidepressant in adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n=35) showed no difference in the effect of antipsychotics on vomiting, bingeing and amenorrhea compared with antidepressants.

6.6.4.85 Cannabinoid agonist versus placebo in adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=48) showed no difference in the effect of cannabinoid agonists on intense physical activity, change in EDI-2-body dissatisfaction, or the number of adverse events compared with placebo.

Low quality evidence from one RCT (n=48) showed cannabinoid agonists may be more effective on gain in body weight, change in EDI-2-drive for thinness and change in EDI-bulimia compared with placebo.

6.6.4.93 Oestrogen versus placebo in young people with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=37) showed no difference in the effect of oestrogen on change in body weight or change in EDI-drive for thinness compared with placebo.

Low quality evidence from one RCT (n=72) showed no difference in drop-out rates in the oestrogen treated group compared with placebo.

Low quality evidence from one RCT (n=37) showed oestrogen treatment is less effective on change in fat mass and change in EDI-body dissatisfaction compared with placebo.

Low quality evidence from one RCT (n=37) showed oestrogen treatment is more effective on change in EDI-bulimia compared with placebo.

6.6.51 Economic Evidence Statements

2 No economic evidence on the cost effectiveness of pharmacological interventions for people
3 with anorexia nervosa was available.

6.6.64 Recommendations and link to evidence for the review on: Does any 5 pharmacological intervention produce benefits/harms on specified outcomes 6 in people with eating disorders?

7 **Medication for anorexia nervosa**

	75. Do not offer medication as the sole treatment for anorexia nervosa.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of pharmacological for treating anorexia nervosa in children, young people and adults. For this population and review, body weight or BMI, depression and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	To investigate the effectiveness of a pharmacotherapy on treating anorexia nervosa, studies comparing the drug with a placebo are the first to consider. Trials on two antidepressants, selective serotonin reuptake inhibitor (SSRIs) and tricyclic antidepressants (TCA), were identified. SSRIs showed a benefit on body weight and depression at the end of treatment, remission was not reported. However, relapse rates were lower in the SSRI group but no difference was found in EDI-bulimia scores. Outcomes at follow up were not reported. Nor was any data available on remission, general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, or relapse. TCA treatment was no different to placebo on depression scores or at achieving target weight at the end of treatment. No long-term outcomes were reported. Nor was data available on remission, general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, or relapse. Directly comparing the two anti-depressants, SSRI and TCA, showed no difference bingeing episodes or the number of participants who had amenorrhea. There was, however, a reduced number of vomiting episodes in the SSRI group compared with TCA. No data was available at follow up nor on body weight, remission, general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, general psychopathology or relapse. Antipsychotic medication showed no difference in its effectiveness compared with placebo at the end of treatment. In young people, the following outcomes were measured and shown to be no different: body weight, depression, EDI-total, adverse events or remission. No data was available on remission, general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, or relapse. For adults, the following outcomes were no different at the end of treatment: body weight and depression. No data was available on remission, general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, eating disorder psychopathology, resource use, or relapse. A head to head trial of an antidepressant compared with antipsychotic medication

	showed they were equally effective on the number of episodes of vomiting, bingeing and the number of adults with amenorrhea. No long-term outcomes were reported nor any data on remission, general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use, general psychopathology or relapse. Comparing the combined treatment of an anti-psychotic and psychotherapy with psychotherapy alone showed no difference in the following outcomes at the end of treatment: BMI, EDI-total, EDI-drive for thinness, EDI-bulimia, EDI- body dissatisfaction, Yales eating disorder scale and the number of side-effects. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, general psychopathology or relapse. Combined treatment of an anti-depressant and psychotherapy compared with psychotherapy alone showed the change in body weight was lower in the combined treatment group, however, the mean percent ideal body weight in both groups was similar at the end of treatment. Change in depression, quality of life, and remission rates were similar at the end of treatment between the two groups. Change in global improvement was also less in the combined treatment group but there was some uncertainty. No data was available on family functioning, service user experience, adverse events, all-cause mortality, resource use or relapse. An RCT on antihistamines compared with placebo was identified and showed no difference in the number who achieved the target body weight and a trend for depression scores to be lower in the antihistamine group. No data was available on general functioning, family functioning, service user experience, adverse events, all-cause mortality, quality of life, resource use or relapse. A cannabinoid agonist showed greater improvements in body weight at the end of treatment compared with placebo. There was no difference in the other reported outcomes, including intensity of physical activity, EDI- body dissatisfaction or adverse events. However, the cannabinoid agonist was less beneficial on change in EDI-total but had favourable results on change in EDI-drive for thinness and EDI-bulimia. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use or relapse. A combined treatment of an antipsychotic and CBT and nutritional therapy compared with placebo, CBT and nutritional therapy showed no difference in their effectiveness at the end of treatment on BMI. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, resource use, general psychopathology or relapse. Finally, oestrogen treatment showed mixed results in adults with anorexia nervosa compared with placebo. No difference in change in body weight, change in EDI-drive for thinness was found between the two arms. However, change in fat mass and EDI-body dissatisfaction was less in the oestrogen treatment group, but change in EDI-bulimia favoured the oestrogen treated group compared with placebo. No data was available on remission, general functioning, family functioning, service user experience, all-cause mortality, quality of life, general psychopathology, resource use or relapse.
Trade-off between net health benefits and resource use	There was no evidence for the effectiveness of pharmacological interventions for the management of people with anorexia nervosa. As a result, such treatments are also likely to be not cost effective.
Quality of evidence	The quality of the evidence was low to very low. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded and high dropouts were detected >20%. To account for high dropouts rates, intention-to-treat analysis was used for remission results, with the assumption that dropouts did not recover from the eating disorder. The committee agreed that all outcomes should be downgraded for risk of publication bias since the studies were sponsored by pharmaceutical companies

	and in the 1980s, 1990s and early 2000s there is a risk that only positive findings were being published, cases of selective outcome reporting are occurring and outliers being excluded. The committee's confidence in the outcomes was often compromised by the small number of participants (n=16 to 124) and a small number of studies (n=1 to 2) for each comparison. Imprecision was often detected because the 95% confidence interval crossed 1 or 2 minimal important differences or the result did not meet the optimal information size (300 events or 400 participants). The committee noted that in many of studies, participants in both were receiving additional treatment such as individual psychotherapy. One of the two critical outcomes, remission was not often reported, thus making it difficult for the committee to know just how effective pharmacological agents are in treating people with anorexia nervosa. Also, very little evidence was found in children and young people with anorexia nervosa (1 study that compared antipsychotics with placebo). The setting in which the study was conducted was another factor the committee considered important if they were to recommend a drug. The following comparisons included studies that treated the participants in hospital: TCA versus placebo, SSRI versus TCA, antipsychotic versus placebo in young people and adults, antipsychotics versus antidepressants, some combined therapies, antihistamine versus placebo. The following studies were conducted in outpatient settings: SSRI versus placebo, antipsychotic versus placebo study, some combined therapies and the cannabinoid agonist was in and outpatient. Nevertheless, none of these drugs were recommended. The committee were also interested whether the participants had any mental health comorbidities to ascertain it may explain some of the findings. However, this was difficult to explore since some studies such as Attia 1998, Attia 2001 Andries 2014, Bissada 2008, Fassino 2002, Ruggiero 2001 excluded participants who had a mental health comorbidity, but it was not always explicitly stated. Interestingly, the cannabinoid agonist showed a positive effect on weight gain compared with placebo but there was some uncertainty. And no data was available on remission or depression at the end of treatment or long-term follow up. There was also only one study and 48 participants. The committee discussed the possibility of making a research recommendation on this.
Other considerations	The committee agreed that the evidence was not strong enough to recommend pharmacological treatments as the sole or primary treatment for anorexia nervosa. The combined treatment of psychotherapy and a pharmacological agent also showed no benefit compared with psychotherapy alone. The committee mentioned new pharmacological treatments may come onto the market that may be worth considering at some time in the future. The committee discussed that the pharmacological studies included in this review do not capture the potential benefits of treating people with anorexia nervosa who have a mental health comorbidity. Most of the studies excluded these people. The committee thus agreed it was best to refer them to the relevant NICE guidelines.

6.7 Physical interventions

6.7.1 Review Question: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 133. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all physical interventions that may be delivered to children, young people and adults with an eating disorder. The interventions were categorised according to

type of physical intervention, the age of the participants and the type of eating disorder and were compared to wait list controls, placebo, treatment as usual or any other intervention.

Table 133: Clinical review protocol summary for the review of: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Component	Description
Review question(s)	Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?
Population	Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). Strata: children (≤12), young people (13-≤17 years), adults ≥18 years eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	Physical interventions may include: transcranial magnetic stimulation deep brain stimulation physiotherapy yoga physical exercise acupuncture mandometer massage
Comparison	Placebo Wait list control Treatment as usual Another intervention
Critical outcomes	Remission and long-term recovery (if symptoms were measured over a minimum two week period) Binge eating for BN and BED Body weight / BMI for AN
Important outcomes	Adverse events All-cause mortality Cost effectiveness Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) Family functioning General psychopathology (including mood/depression/anxiety) General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) Quality of life Relapse Resource use Service user experience (in patient vs. community)
Study design	Systematic Reviews RCTs

6.7.21 Clinical Evidence for: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Six RCTs (n=169) and one systematic review (n=372) met the eligibility criteria for this review, all of which were conducted in an inpatient or day care setting (Birmingham 2004 (Birmingham et al., 2004b), del Valle 2010 (del Valle et al., 2010), del Valle 2014 (Del Valle et al., 2014), Janas-Kozik 2011 (Janas-Kozik et al., 2011), McClelland 2016 (McClelland et al., 2016), Smith 2014 (Smith et al., 2014), Touyz 1994 (Touyz et al., 1994), Yang 2016)). The majority of participants in these studies were young people females, whilst three of the studies concerned people diagnosed with restrictive anorexia nervosa. An overview of the trials included in the analysis can be found in Table 134. Further information about both included and excluded studies can be found in Appendix J.

One systematic review (n=372) found five studies comparing the traditional Chinese medicine version of chiropractic therapy with traditional Chinese or western medicines (e.g. supplements) or any other intervention. Summary of findings can be found in Table 143

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 134: Study information for trials included in the analysis of physical interventions versus any other intervention or wait list control for people with anorexia nervosa.

Study ID	N Random-ized	Female (%)	Mean BMI, kg/m2 (SD)	Sample	Intervention	Comparison	Duration
Birmingham 2004	21	100	Admission: 17.7 (2.8); Discharge: 18.4 (2.9)	Adult AN inpatients hospitalized for refeeding	Warming therapy + refeeding Duration of illness: 11.7 (7.1) years (n=7)	No warming + refeeding Length of illness: 15 (6.3) years (n=9)	3 weeks
del Valle 2010	22	91	18.5 (1.6)	Young people AN-R in day care program	Resistance training + TAU Time since initial diagnosis=42 days (11)	TAU Time since initial diagnosis=72 days (31)	3 months
del Valle 2014	44	100	17.7 (2.3)	Young people AN-R in day care program	Resistance training + TAU Time since admission: 50.8 (36.4) days	TAU Time since admission: 61.6 (37.3) days	8 weeks + 4 week FU
Janas-Kozik 2011	24	100	15.3 (2)	Young people AN-R inpatients with concomitant depressive symptoms	Bright light therapy + CBT	CBT	6 weeks
McClelland 2016	60	100	16.5 (1.7)	Adult AN	Repetitive transcranial magnetic stimulation Duration of illness: 9.1 (7.0) years	'Sham' repetitive transcranial magnetic stimulation Duration of illness: 11.3 (8.0) years	20 min + 1 day FU
Smith 2014	26	96	17 (2.6)	Adult AN inpatients	Acupuncture + TAU	Acupressure + massage + TAU	6 weeks
Touyz 1994	32	100	15.3 (1.5)	Young people AN inpatients	Video feedback of eating behaviour + TAU	TAU	6 weeks
Yang 2016 (5 studies)	372	not reported	Not reported	Children and young people AN	Traditional Chinese Chiropractic therapy	Traditional Chinese or Western medicine or other interventions	Variable

Abbreviations: AN-R- anorexia nervosa restricted; CBT- cognitive behavioural therapy; FU – follow up; TAU – treatment as usual

Table 135: Summary table of findings for repetitive transcranial magnetic stimulation versus 'sham' repetitive transcranial magnetic stimulation at end of treatment in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with RTMS (95% CI)
VAS Core AN symptoms	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas core an symptoms in the intervention groups was 0.57 standard deviations lower (1.14 lower to 0.01 higher)
VAS Restrict	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas restrict in the intervention groups was 0.2 standard deviations lower (0.77 lower to 0.36 higher)
VAS Feeling Full	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas feeling full in the intervention groups was 0.45 standard deviations lower (1.02 lower to 0.12 higher)
VAS Feeling Fat	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas feeling fat in the intervention groups was 0.71 standard deviations lower (1.29 to 0.13 lower)
VAS Mood	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas mood in the intervention groups was 0.17 standard deviations higher (0.4 lower to 0.73 higher)
VAS Hunger	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas hunger in the intervention groups was 0.24 standard deviations lower (0.81 lower to 0.33 higher)
VAS Urge to Eat	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas urge to eat in the intervention groups was 0.16 standard deviations lower (0.73 lower to 0.4 higher)
VAS Urge to Binge Eat	49 (1 study) 1 days	⨁⨁⨁⨀ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas urge to binge eat in the intervention groups was 0.3 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with RTMS (95% CI)
					(0.87 lower to 0.27 higher)
VAS Urge to be Sick/Purge	49 (1 study) 1 days	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas urge to be sick/purge in the intervention groups was 0.53 standard deviations lower (1.11 lower to 0.04 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 CI crosses either 0.5 or -0.5 (SMD).

Table 136: Summary table of findings for repetitive transcranial magnetic stimulation versus 'sham' repetitive transcranial magnetic stimulation at follow up for adult anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with RTMS (95% CI)
VAS Restrict 24-hr FU	49 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas restrict 24-hr fu in the intervention groups was 0.53 standard deviations lower (1.1 lower to 0.05 higher)
VAS Feeling Full 24-hr FU	49 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas feeling full 24-hr fu in the intervention groups was 0.65 standard deviations lower (1.23 to 0.06 lower)
VAS Feeling Fat 24-hr FU	49 (1 study)	⊕⊕⊕⊖ MODERATE1 due to imprecision		Not calculable for SMD values	The mean vas feeling fat 24-hr fu in the intervention groups was 0.71 standard deviations lower (1.29 to 0.13 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with RTMS (95% CI)
CI: Confidence interval; FU: follow up					
1 CI crosses either 0.5 or -0.5 (SMD).					

Table 137: Summary table of findings for bright light treatment and CBT versus any other intervention in young people and adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT only	Risk difference with Light Therapy+CBT (95% CI)
Depression	24 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 1.14 standard deviations lower (2.01 to 0.27 lower)
Remission of Depression (HAM-D<=8)	24 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision	RR 0.27 (0.1 to 0.74)	917 per 1000	669 fewer per 1000 (from 238 fewer to 825 fewer)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Janas-Kozik 2011: Unclear randomization method and allocation concealment. No participant, investigator, nor assessor blinding. 2 Sample was participants diagnosed with Anorexia Nervosa-Restricting type with concomitant depressive symptoms. 3 CI crosses -0.5. 4 <300 events.					

Table 138: Summary table of findings for warming therapy and refeeding versus any other intervention.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Refeeding	Risk difference with Warming (95% CI)
BMI - change scores	21	⨁⊖⊖⊖		Not calculable	The mean BMI - change scores in the

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Refeeding	Risk difference with Warming (95% CI)
	(1 study)	VERY LOW1,2 due to risk of bias, imprecision		for SMD values	intervention groups was 0.02 standard deviations higher (0.84 lower to 0.87 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Birmingham 2004: Unclear randomization method, unclear allocation concealment. No participant, investigator, nor assessor blinding. Dropout rate of control group>20%, reasons not stated. 2 CI crosses both 0.5 and -0.5 (SMD).					

Table 139: Summary table of findings for video feedback and treatment as usual versus treatment as usual in young people with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Video Feedback + TAU (95% CI)
BMI (change scores)	32 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI (change scores) in the intervention groups was 0.16 standard deviations higher (0.53 lower to 0.86 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Touyz 1994: Randomization method and allocation concealment unclear. Participant, investigator and assessor blinding unclear. Significant difference at baseline in EDI Body Dissatisfaction score. 2 Participants were diagnosed according to DSM-III-R. 3 CI crosses both 0.5 and -0.5.					

Table 140: Summary table of findings for acupuncture and treatment as usual versus any other intervention in adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Acupressure+Massage+ TAU	Risk difference with Acupuncture+TAU (95% CI)
BMI - change scores	20 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI - change scores in the intervention groups was 0.07 standard deviations lower (0.94 lower to 0.81 higher)
EDI-3 Bulimia - change scores	20 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-3 bulimia - change scores in the intervention groups was 0.45 standard deviations higher (0.44 lower to 1.34 higher)
EDI-3 Drive for Thinness - change scores	20 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-3 drive for thinness - change scores in the intervention groups was 0.26 standard deviations higher (0.62 lower to 1.14 higher)
EDI-3 Body Dissatisfaction - change scores	20 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-3 body dissatisfaction - change scores in the intervention groups was 0.14 standard deviations higher (0.73 lower to 1.02 higher)
EDE-Q Global - change scores	20 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global - change scores in the intervention groups was 0.47 standard deviations higher (0.42 lower to 1.36 higher)
EDE-Q Restraint - change scores	20 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q restraint - change scores in the intervention groups was 0.67 standard deviations higher (0.24 lower to 1.58 higher)
EDE-Q Eating Concerns - change scores	20 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concerns - change scores in the intervention groups was 0.44 standard deviations higher (0.45 lower to 1.33 higher)
EDE-Q Weight Concerns -	20	⨁⊖⊖⊖ VERY LOW1,2		Not calculable for SMD	The mean ede-q weight concerns - change scores in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Acupressure+Massage+ TAU	Risk difference with Acupuncture+TAU (95% CI)
change scores	(1 study)	due to risk of bias, imprecision		values	0.07 standard deviations lower (0.94 lower to 0.81 higher)
EDE-Q Shape Concerns - change scores	20 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concerns - change scores in the intervention groups was 1.38 standard deviations lower (2.38 to 0.38 lower)
General Psychopathology - DASS Total - change scores Depression, Anxiety, and Stress Scale (DASS)	20 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology - dass total - change scores in the intervention groups was 0.03 standard deviations higher (0.84 lower to 0.91 higher)
Depression - change scores	20 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - change scores in the intervention groups was 0.03 standard deviations higher (0.85 lower to 0.91 higher)
Stress - change scores	20 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean stress - change scores in the intervention groups was 0.14 standard deviations higher (0.73 lower to 1.02 higher)
Quality of Life - EDQoL - change scores	20 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life - edqol - change scores in the intervention groups was 0.05 standard deviations higher (0.83 lower to 0.92 higher)
EDQoL Psychological - change scores	20 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edqol psychological - change scores in the intervention groups was 0.11 standard deviations lower (0.99 lower to 0.76 higher)
EDQoL Physical/Cognitive - change scores	20 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edqol physical/cognitive - change scores in the intervention groups was 0 standard deviations higher (0.88 lower to 0.88 higher)
EDQoL Financial - change	20	⊕⊖⊖⊖		Not calculable for SMD	The mean edqol financial - change scores

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Acupressure+Massage+ TAU	Risk difference with Acupuncture+TAU (95% CI)
scores	(1 study)	VERY LOW1,2 due to risk of bias, imprecision		values	in the intervention groups was 0.34 standard deviations higher (0.54 lower to 1.23 higher)
EDQoL Work/School - change scores	20 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edqol work/school - change scores in the intervention groups was 0.12 standard deviations lower (1 lower to 0.75 higher)
Withdrawn due to Adverse Events	26 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1 (0.25 to 4.07)	231 per 1000	0 fewer per 1000 (from 173 fewer to 708 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Smith 2014: No participant blinding. Dropout rate of both groups>20%.
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).
3 CI crosses either 0.5 or -0.5 (SMD).

Table 141: Summary table of findings for resistance training and treatment as usual versus treatment as usual at end of treatment in young people with anorexia nervosa-restricting.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Resistance Training + TAU (95% CI)
BMI	64 (2 studies) 3 weeks	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.21 standard deviations lower (0.70 lower to 0.29 higher)
Quality of Life SF-36 Mental, SF-36 Physical	22 (1 study) 3 weeks	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness,		Not calculable for SMD	The mean quality of life in the intervention groups was 0.39 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Resistance Training + TAU (95% CI)
		imprecision		values	(0.2 lower to 0.99 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 del Valle 2010: Unclear randomization method and allocation concealment. No participant blinding, unclear investigator and assessor blinding.
2 del Valle 2014: Unclear whether baseline similar. Randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding.
3 Sample consisted of participants diagnosed with Anorexia Nervosa-Restricting type. Participants in both groups also received psychotherapy 3 days a week and were on diet.
4 CI crosses either 0.5 or -0.5 (SMD).

Table 142: Summary table of findings for resistance training and treatment as usual versus treatment as usual at follow up in young people with anorexia nervosa-restricting.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Resistance Training + TAU (95% CI)
BMI FU	36 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.53 standard deviations lower (1.19 lower to 0.14 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; FU: follow up

1 del Valle 2014: Unclear whether baseline similar. Randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding.
2 Sample consisted of participants diagnosed with Anorexia Nervosa-Restricting type. Participants in both groups also received psychotherapy 3 days a week and were on diet.
3 CI crosses either 0.5 or -0.5 (SMD).

Table 143: Summary table of findings for traditional Chinese chiropractic therapy versus traditional Chinese or Western medicine or other interventions in children and young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other intervention	Risk difference with Chiropractic therapy (95% CI)
Efficacy rate (Recovered+Significant Improvement)/Total N	371 (5 studies)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.24 (1.14 to 1.35)	772 per 1000	185 more per 1000 (from 108 more to 270 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Yang 2016: data from meta-analysis of chiropractic therapy studies published in Chinese or English. All studies were: low risk of bias for random sequence generation, unclear allocation concealment, unclear blinding of participants/assessors/investigators. Only one study reported dropout data.
2 CI crosses either 0.75 or 1.25 (Risk Ratio).

6.7.3 Economic Evidence

No economic evidence on the cost effectiveness of physical interventions for people with anorexia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.7.4 Clinical evidence statements

6.7.4.1 Repetitive transcranial magnetic stimulation (rTMS) versus 'sham' repetitive transcranial magnetic stimulation in adults with anorexia nervosa at end of treatment

Moderate quality of evidence from one RCT (n=49) showed rTMS may improve anorexia nervosa symptoms and urge to be sick/purge compared with sham, although there was some uncertainty.

Moderate quality of evidence from one RCT (n=49) showed rTMS is more effective on feeling fat compared with sham.

Moderate quality of evidence from one RCT (n=49) showed no difference in the effect of rTMS on food restriction, feeling full, mood, hunger, urge to eat and urge to binge compared with sham.

6.7.4.2 Repetitive transcranial magnetic stimulation (rTMS) versus 'sham' repetitive transcranial magnetic stimulation in adults with anorexia nervosa at follow up

Moderate quality of evidence from one RCT (n=49) showed rTMS is more effective on feeling full and feeling fat compared with sham.

Moderate quality of evidence from one RCT (n=49) showed rTMS may be more effective on restrict compared with sham, although there was some uncertainty.

6.7.4.3 Bright light treatment and CBT versus any other intervention in young people with anorexia nervosa-restricting

Very low quality evidence from one RCT (n=24) showed bright light treatment and CBT is more effective on remission from depression and depression compared to any other intervention.

6.7.4.4 Warming therapy and refeeding versus refeeding in adults with anorexia nervosa

Very low quality evidence from one RCT (n=21) showed no difference in the effect of adding warming therapy to refeeding on change in BMI compared with refeeding alone.

6.7.4.5 Video feedback and treatment as usual versus treatment as usual in young people with anorexia nervosa

Very low quality evidence from one RCT (n=32) showed no difference in the effect of video feedback and nutritional counselling on change in BMI compared with nutritional counselling.

6.7.4.6 Acupuncture and treatment as usual versus acupressure, massage and treatment as usual in adults with anorexia nervosa

Very low quality evidence from one RCT (n=20) showed acupuncture treatment as usual is more effective in reducing scores on EDE-Q-shape concerns compared with acupressure, massage and treatment as usual.

Very low quality evidence from one RCT (n=20) showed no difference in the effect of acupuncture and treatment as usual on change in BMI, EDI-3-bulimia, EDI-3-drive for thinness, EDI-3-body dissatisfaction, EDE-Q-Global, EDE-Q-restraint, EDE-Q-eating concerns, EDE-Q-weight concerns, general psychopathology, depression, stress, EDQoL Total, EDQoL Psychological, EDQoL-physical/cognitive, EDQoL-financial and EDQoL-work/school compared with acupressure, massage and treatment as usual.

6.7.4.7 Resistance training and treatment as usual versus treatment as usual at end of treatment in young people with anorexia nervosa-restricting

Very low quality evidence from two RCTs (n=64) showed no difference in the effect of resistance training and treatment as usual on BMI compared with treatment as usual.

Very low quality evidence from one RCT (n=22) showed no difference in the effect of resistance training and treatment as usual on mental and physical general functioning compared with treatment as usual.

6.7.4.8 Resistance training and treatment as usual versus treatment as usual at follow up in young people with anorexia nervosa-restricting

Very low quality evidence from one RCT (n=36) showed no difference in the effect of resistance training and treatment as usual on BMI compared with treatment as usual.

6.7.4.9 Traditional Chinese chiropractic therapy versus traditional Chinese or Western medicine or other interventions in children and young people with anorexia nervosa

Low quality evidence from five RCTs (n=371) showed traditional Chinese chiropractic therapy is more effective on the number of people recovered or significantly improved from anorexia nervosa compared with traditional Chinese or Western medicine or other intervention.

6.7.5 Economic Evidence statements

No economic evidence on the cost effectiveness of physical interventions for people with anorexia nervosa was available.

6.7.6 Recommendations and link to evidence for the review on: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Physical therapy for any eating disorder

	76. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitization, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes for the review on the effectiveness of physical interventions, such as transcranial magnetic stimulation or physiotherapy in people with eating disorders and it was agreed that for any eating disorder remission is of greatest concern. The other critical outcomes for anorexia nervosa are body weight and BMI and for binge eating disorder and bulimia nervosa it is bingeing. Other outcomes that are important but are considered rare events or rarely measured in randomised controlled trials for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse, thus they were

	extracted where possible, but did not factor strongly in the decision making. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	Young people with anorexia nervosa (chapter 6) For young people with anorexia nervosa, bright light treatment and CBT improved depression scores compared with any other intervention. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Video feedback and nutritional counselling compared with nutritional counselling alone showed no additional benefit of the video feedback on BMI. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Resistance training and treatment as usual showed no difference on BMI nor general (mental or physical) functioning in young people with anorexia nervosa compared with treatment as usual. At 4 weeks follow up, resistance training and treatment as usual appeared to have no effect on BMI compared with treatment as usual. No evidence was found on the critical outcomes of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, family functioning, resource use or service user experience. **Adults with anorexia nervosa (chapter 6)** Repetitive transcranial magnetic stimulation versus sham showed no difference in anorexia nervosa symptoms (urge to restrict, feeling full, mood, hunger, urge to eat and urge to binge), but the treatment did improve feeling fat and core anorexia nervosa symptoms and urge to be sick/purge (however there was some uncertainty for the last two outcomes). At one day follow up some benefits were detected on anorexia nervosa symptoms including feeling full and feeling flat compared with sham, and some benefit on the urge to restrict but there was some uncertainty. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Warming therapy on top of refeeding had no effect on weight compared with refeeding alone in adults with anorexia nervosa. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Acupuncture and treatment as usual compared with acupressure, massage and treatment as usual showed acupuncture is more effective on EDE-shape concerns but no other outcome was different between the two groups including EDI-subscales, EDE-subscales, depression, general psychopathology and weight. No evidence was found on the critical outcome of remission, nor on the important outcomes of all-cause mortality, relapse, general functioning, family functioning, resource use or service user experience. **Adults with bulimia nervosa (chapter 7)** Repetitive transcranial magnetic stimulation versus sham showed no difference in the effect on bingeing and food cravings within 24 hours of treatment, nor on the urge to eat or the number who withdrew due to adverse events. There was a trend for hunger and the number of those who binged to be reduced but there was some uncertainty. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience.

Aerobic exercise appeared to be less effective on EDI-drive for thinness. No difference was found on the number of people who recovered from bulimia nervosa nor who satisfied the EDNOS criteria.
Compared with wait list control, aerobic exercise was less effective on the number who had recovered (unclear definition) from bulimia nervosa but showed no difference on the number who satisfied the criteria for EDNOS. No evidence was found on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience.

Adults with binge eating disorder (chapter 8)

Yoga appears to be effective at reducing scores on the binge eating scale compared with wait list controls. However, this did not translate to a benefit in BMI. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, or service user experience.

Aerobic exercise and group CBT-ED appeared to be more effective at reducing BMI compared with group CBT-ED alone in adults with binge eating disorder. No difference was found in depression scores. Similar results were found at follow up. When a maintenance component (12 biweekly meetings over six months) was added to both arms to make this part of the intervention more comparable with the aerobic exercise group (because they continued to meet up), there was a trend for a reduced BMI and depression in the aerobic exercise, group CBT-ED and maintenance group compared with the group CBT-ED and maintenance group at the end of treatment and for the trend in the benefit on BMI to be maintained at follow up but not depression. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use or service user experience.

Any eating disorder (chapter 9)

One study compared eye movement desensitization and reprocessing therapy with treatment as usual in adults with any eating disorder. The results showed some improvement in the outcomes reported by the body image memory questionnaire, including the earliest memory and worst memory on body image, and only a trend for the most recent memory. At 12 months follow up the worst memory on body image was still better but not the earliest or most recent. No evidence was found on the critical outcomes of remission, bingeing and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience.

An RCT was identified that compared yoga and treatment as usual with treatment as usual in adults with any eating disorder. At the end the treatment, no difference was found in any of the outcomes including BMI, EDE-total or any of the EDE- sub-scales. Similar findings were found at follow up (three weeks), however there was some improvement in EDE-restraint in the yoga and treatment as usual group compared with treatment as usual. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, resource use or service user experience.

A graded body image therapy (and maintenance treatment as usual) was compared with a maintenance treatment as usual in adults with any eating disorder. No difference was found in EDE-weight concerns or EDE-shape concerns at the end of treatment or at follow up. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience.

An acceptance-based body image mirror exposure therapy was compared with a control therapy and showed an improvement in EDE-eating concerns, EDE-weight concerns, EDE-shape concerns, but not in EDE-restraint. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse,

	general functioning, family functioning, resource use or service user experience. A psychomotor therapy and support was compared with support in females with any eating disorder and showed no difference at the end of treatment on self-expression and control anger scales. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, eating disorder psychopathology, relapse, general functioning, family functioning, resource use or service user experience. The committee requested investigating the benefits of the Mandometer on eating disorders. A Mandometer is a device that measures how much weight is lost from a dinner plate after the person with eating disorder has finished eating. This weight is stored on a computer along with how satiated the person is after eating. The evidence on this is scarce and the sample sizes were too small (less than ten per group) to meet our inclusion criteria as described in the protocol.
Trade-off between net health benefits and resource use	There was no evidence for the effectiveness of physical interventions in people with eating disorders. As a result, such interventions are not likely to be cost effective.
Quality of evidence	The evidence for physical interventions was mostly very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. Most of the outcomes were the result of a single study with a very low number of participants, only binge eating disorder had more than 100 participants in total. Because of this imprecision was often detected and the outcomes were downgraded because the 95% confidence interval crossed one or two minimal important differences or the outcome did not meet the optimal information size (300 events or 400 participants). Few studies measured remission and/or compensatory behaviours relevant to that eating disorder. Some outcomes were excluded from the study because it was either unclear over what duration they measured the symptoms or it was less than the two week minimum required by the committee.
Other considerations	The committee agreed that the evidence presented was not strong enough or of sufficient quality to offer a physical intervention to people with an eating disorder. This was mostly because very few studies were identified and few participants were included in most outcomes. However, the committee decided to make a research recommendation on adding exercise to a recommended psychotherapy to determine whether it may add any benefit to those with bulimia nervosa or binge eating disorder. The committee discussed the importance of any future research exploring what the right amount of exercise is, what is the best type of exercise and what the potential harms are. The committee suggested making a research recommendation on the effects of exercise on bulimia nervosa and binge eating disorder, as opposed to any of the other physical interventions for a number of reasons. Exercise may be useful adjunct to psychotherapy to address any co-existing weight or obesity-related issues and mood disorders, such as depression and anxiety. Exercise may also be a cost effective and drug-free alternative to other therapeutic approaches such as transcranial magnetic stimulation or anti-depressants.

5. Research recommendation: Does exercise in addition to a recommended psychotherapy add any benefit to those with bulimia nervosa or binge eating disorder?

6.8 Management of long- and short-term complications

6.8.1 Review question: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 144. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all interventions that may be delivered to manage or reduce the short- or long-term physical complications of eating disorders in children, young people and adults and includes recovered as well as current service users. The interventions were categorised according to type of physical complication and intervention, the age of the participants and the type of eating disorder and were compared to the control arm as reported in the relevant studies.

Table 144: Clinical review protocol summary for the review of: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

Component	Description
Review question(s)	What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) • recovered or current service users **Strata:** • children (≤12), young people (13-≤17 years), adults ≥18 years • eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Interventions to address the following: • Low bone mineral density (risk of fracture) • Growth (physical development) • Pubertal development • Tooth wear • Low body weight • Interventions to address the long-term physical complications may include: • GH/IGF-I • Calcium with and without Vitamin D • Bisphosphonates (age dependent and exclude pregnancy) • Exercise (low impact)/Physiotherapy • Oestrogen (patches/exogenous/pills other) • Testosterone (males/females) • Weight gain vs. Weight restoration (brain size) • Interventions to address the short-term physical complications may include

Component	Description
	• Phosphates supplementation (refeeding) • Potassium • Thiamine (refeeding) • Laxatives (for when underweight patients are constipated) • Salbutamol (reduce food intake)
Comparison	• Control arm as defined by study
Critical outcomes	• Primary outcome as reported by study
Important outcomes	• Secondary outcome as reported by study
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs) (if no RCTs)

6.8.2 Clinical Evidence for: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

6.8.2.1 Low bone mineral density

12 RCTs (n=749) were identified that addressed the effects of pharmacological interventions to treat people with anorexia nervosa and low bone mineral density (Divasta 2012, Divasta 2014, Fazeli 2014 (Fazeli et al., 2014), Golden 2005 (Golden et al., 2005), Gordon 2002 (Gordon et al., 2002), Grinspoon 2002 (Grinspoon et al., 2002), Grinspoon 2003 (Grinspoon et al., 2003), Klibanski 1995 (Klibanski et al., 1995), Misra 2011 (Misra et al., 2011), Miller 2011 (Misra et al., 2011), Nakahara 2006 (Nakahara et al., 2006), Strokosch 2006 (Strokosch et al., 2006).

6.8.2.2 Low hormone levels

One RCT (n=21) was identified that addressed what are the effects of pharmacological interventions to treat people with anorexia nervosa and low background rhIGF-I levels (recombinant human insulin-like growth [IGF] factor-I) (Fazeli 2010 (Fazeli et al., 2010)).

6.8.2.3 Low body weight and malnourishment

Two RCTs (n=117; O'Connor 2016 (O'Connor et al., 2016b), Rigaud 2007 (Rigaud et al., 2007b) and four observational studies (n=803; Born 2015 (Born et al., 2015), Diamanti 2008 (Diamanti et al., 2008), Rigaud 2010 (Rigaud et al., 2010), Robb 2002 (Robb et al., 2002) looked at how to address low body weight and/or malnourishment. All of the studies were conducted in an inpatient setting and the majority of participants were female and hospitalised for low body weight and/or malnourishment. One of the studies compared parenteral and enteral nutrition with enteral nutrition alone (Diamanti 2008), whilst the remaining studies all examined various forms of enteral nutrition. An overview of the RCTs and observational studies included in the review can be found in Table 147 and Table 148. Summary of findings can be found in Table 157, Table 158, Table 159, Table 160, Table 161, Table 162, Table 163 and Table 164.

6.8.2.4 Cardiac dysfunction

One observational study (n=28) was identified that examined the effect of a 4 week oral potassium aspartate supplementation (60 mmol) on QT-dispersion in people with anorexia

nervosa and self-induced starvation (Franzoni 2002 (Franzoni et al., 2002)). An overview of the study can be found in Table 149, whilst a summary of findings can be found in Table 165.

Table 145: Study information for RCTs included in the analysis of interventions to treat low bone mineral density or growth in people with anorexia nervosa.

Study_ID	Age years	BMI	Stage of illness	Number randomised	Females	Intervention	Other arms	Duration of treatment
Divasta 2012 Divasta 2014	18.1 (2.7)	18.0 (1.5)	Duration of AN 12 (6-40)	94	100%	DHEA (50mg daily) + Ethinyl estradiol + levonorgestrel	Placebo	18 months
Fazeli 2014	47 (2.7)	17.6 (0.4)	Years since onset of AN 20.4 (3.7)	21	100%	Teriparatide (Human PTH (1-34)	Placebo	6 months
Golden 2005	16.9 (1.6)	16.3 (1.4)	Duration of illness 25.7 (14.6) months	32	100%	Alendronate Bisphosphonate (10 mg daily)	Placebo	12 months
Gordon 2002	17.8 (2.9)	NR	NR	61	100%	DHEA 50 mg/d orally.	HRT Alesse (20 g ethinyl estradiol and 0.1 mg levonorgestrel). rhIGF-I (30 g/kg sc twice daily) & OCP (Ovcon 35, once a day, containing 35 g ethinyl estradiol and norethindrone 0.4 mg) IGF-I	12 months
Grinspoon 2002	25.2 (0.7)	17.8 (0.3)	NR	60	100%	OCP (Ovcon 35, once a day) + Placebo	Placebo. Oestrogen (Ovcon 35, 35 g	9 months

Study_ID	Age years	BMI	Stage of illness	Number randomised	Females	Intervention	Other arms	Duration of treatment
							ethinyl estradiol, and 0.4 mg of norethindrone) rhIGF-I l (30 g/kg d sc twice daily (BID))	
Grinspoon 2003	25.6 (0.8)	16.6 (0.2)	NR	59	100%	rhIGF-I + oestrogen	Placebo	3 months
Klibanski 1995	24.9 (6.9)	NR	Amenorrhea 3.3 (3.1) years	48	100%	Oestrogen + Progesterone (Premarin 0.625 mg, days l-25) and Provera (5 mg)	Control	1.5 years
Misra 2011	16.5 (0.2)	17.4 (0.1)	Amenorrhea duration 0.9 (0.08) years	110	100%	Oestrogen Mature girls = 17b-estradiol (100-mg patch applied twice weekly Immature girls = oral ethinyl estradiol (3.75 mg to 11.25 mg daily)	Placebo	18 months
Miller 2011	25.3 (6.3)	17.6 (1.2)	NR	77	100%	Risendronate (35 mg/daily) Bisphosphonates	Testosterone Risendronate (Bisphosphonate) + Testorone 35 mg + 150 mg/daily Placebo	12 months
Nakahara	26.2 (8.5)	14.4 (1.7)	Duration of	41	100%	Etidronate (200 mg/day)	Placebo	3 months

Study_ID	Age years	BMI	Stage of illness	Number randomised	Females	Intervention	Other arms	Duration of treatment
2006			illness 57.3 (27.7) months			Bisphosphonates	Calcium and vitamin D 600 mg/d calcium L-aspartate and 1 mg/day alfacalcidol	
Strokosch 2006	15.2 (1.2)	17.9 (2.3)	Duration of secondary amenorrhea 9.7 (8.0) months	146	100%	Norgestimate (180-250 g) /Ethinyl Estradiol. OCP (35 g)	Placebo	13 months

Abbreviations: g-grams; mg – miligrams; NR – not reported; rhIGF-I – recombinant human insulin-like growth factor –I; OCP – oral contraceptive pill.

Table 146: Study information for RCTs included in the analysis of interventions to treat low hormones levels in people with anorexia nervosa.

Study_ID	Age	BMI	Stage of illness	Number randomised	Females	Intervention	Other arms	Duration of treatment
Fazeli 2010	28 (2.1)	17.4 (0.4)	Duration of AN 2 years	21	100%	(rhGH) Supraphysiological recombinant human GH	Placebo	12 weeks

Table 147: Study information for RCTs included in the analysis of interventions to treat low body weight and malnourishment in people with anorexia nervosa.

Study ID	N random-ised	Age years (SD)	BMI (SD), kg/m 2	Female (%)	Sample	Intervention	Comparison	Duration
O'Connor 2016	36	13.8 (1.8)	13.5 (1.1)	94	Moderately malnourished young people	High-calorie (1200 kcal/day) refeeding diet	Low-calorie (500 kcl/day) refeeding diet	Variable, >10 days
Rigaud 2007	81	23.3 (4.2)	12.4 (1.8)	98	Malnourished adults	Enteral Nutrition: Nasogastric Tube Feeding + Meals/snacks	Enteral Nutrition: Meals/snacks only	8 weeks

Table 148: Study information for observational studies in the analysis of interventions to treat low body weight and/or malnourishment in people with anorexia nervosa.

Study ID	Number of participants	Mean Age, years (SD)	BMI (SD), kg/m2	Female (%)	Sample	Intervention	Comparison	Duration
Born 2015	100	26.5 (8.5) Range: 16-61	12.3 (1.4)	96	Severely underweight adults	Parenteral Nutrition: Compulsory Percutaneous Gastric Tube Feeding + Meals	Enteral Nutrition: Meals + either Nasogastric Tube Feeding or No Tube Feeding	Until BMI ≈ 17 kg/m2; tube removed if stable body weight ≥ 2 weeks
Diamanti 2008	261	15 (1.2)	15.1 (0.9)	100	Young people with nutritional, metabolic, or psychiatric instability or cardiac dysfunction	Parenteral + Enteral Nutrition: Peripheral or intravascular infusion + Oral Feeding	Enteral Nutrition: Oral Feeding	Parenteral nutrition suspended when patients achieve intake ≥ 50 kcl/kg/day. Discharged when agreement to continue outpatient pharmacological + behavioural treatment and ≥ 60 kcal/kg/day
Rigaud 2010	284	23.1 (5.0)	13.3 (1.3)	98	Malnourished adults	Enteral Nutrition: Normal Sodium diet (10-12g NaCl/day) via Nasogastric Tube Feeding + Meals/snacks	Enteral Nutrition: Low Sodium diet (4-5g NaCl/day) via Nasogastric Tube Feeding + Meals/snacks	2 months
Robb 2002	158	14.9 (1.8)	15.7 (1.8)	100	Young people <85% ideal body weight	Enteral Nutrition: Nocturnal Nasogastric Feeding + Meals/snacks	Enteral Nutrition: Meals/snacks only	Until ideal body weight >95%

Abbreviations: EBW, Expected Body Weight; IBW, Ideal Body Weight; IP, kcal, kilocalories: OCP, oral contraceptive pill

Table 149: Study information for observational studies for treating cardiac dysfunction in female adult anorexia nervosa.

Study ID	Number of participants	Mean Age, years (SD)	BMI (SD), kg/m2	Female (%)	Sample	Intervention	Comparison	Duration
Franzoni 2002	28	20.1 (4.5)	15.9 (2.4)	100	Adult AN with self-induced starvation	Oral potassium aspartate supplementation (60 mmol)	No supplementation	4 weeks

Table 150: Summary table of findings for DHEA versus hormone replacement therapy in young people with anorexia nervosa at the end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with DHEA vs.HRT (95% CI)
Change in Total Hip BMD - Young people	61 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in total hip bmd - young people in the intervention groups was 0.11 standard deviations lower (0.61 lower to 0.39 higher)
Change in LS BMD - Young people	61 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in ls bmd - young people in the intervention groups was 0.49 standard deviations lower (1 lower to 0.02 higher)
Did not dropout due to side effects	61 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR: 1.00 (0.94 to 1.06)	See comment**	-
Change in Weight - Young people	61 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in weight - young people in the intervention groups was 0.13 standard deviations higher (0.38 lower to 0.63 higher)
Regular menses - Young people	61 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 0.73 (0.51 to 1.03)	800 per 1000	216 fewer per 1000 (from 392 fewer to 24 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%					

confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
** Absolute effect could not be calculated in GRADE because zero events were detected in the outcome.

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. Staff and participants were blind to study allocation, but it was unclear if assessors were blind. The control arm had a 20% dropout rate.
2 95% CI crossed 1 MID (-0.5)
3 For a continuous outcome, there were fewer than 400 participants.
4 95% CI crossed 1 MID (0.5).
5 95% CI crossed 1 MID (0.75)
6 For a dichotomous outcome, there were fewer than 300 events.

Table 151: Summary table of findings for DHEA and combined oral contraceptive (COC) versus placebo in adults with anorexia nervosa at the end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with DHEA+COC (95% CI)
Change in Femoral Shaft BMD - Adults	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in femoral shaft bmd - adults in the intervention groups was 12.86 standard deviations higher (10.66 to 15.05 higher)
Change in Femoral Neck BMD - Adults	76 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in femoral neck bmd - adults in the intervention groups was 14.38 standard deviations higher (11.99 to 16.77 higher)
Change in Femoral Shaft Bone Strength Index - Adults	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in femoral shaft bone strength index - adults in the intervention groups was 18.99 standard deviations higher (15.79 to 22.19 higher)
Change in FN Bone Strength Index - Adults	76 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in fn bone strength index - adults in the intervention groups was 0.95 standard deviations lower (1.43 to 0.47 lower)
Change in Weight - Adults	60 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias,		Not calculable for SMD	The mean change in weight - adults in the intervention groups was 0.99 standard deviations higher

		imprecision		values	(0.45 to 1.53 higher)
Change in BMI (% median for age) - Adults	60 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in BMI (% median for age) - adults in the intervention groups was 0.96 standard deviations higher (0.42 to 1.5 higher)
Amenorrheic - Adults	60 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 1 (0.94 to 1.07)	1000 per 1000	0 fewer per 1000 (from 60 fewer to 70 more)
Did not dropout due to side-effects	60 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 1.00 (0.94 to 1.07)	See comment**	-

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
** Absolute effects could not be calculated because zero events were included in the original analysis.

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Randomisation method was unclear and it was unclear if allocation concealment was conducted. Participants, investigators and assessors were blind. High dropout rates were detected in both arms >20%.
2 For a continuous outcome, there were fewer than 400 participants.
3 95% CI crossed 1 MID (-0.5)
4 95% CI crossed 1 MID (0.5)
5 For a dichotomous outcome, there were fewer than 300 events.

1 **Table 152: Summary table of findings for PTH versus placebo in adults with anorexia nervosa at the end of treatment.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with PTH (95% CI)
% Change in Weight - Adults	21 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean % change in weight - adults in the intervention groups was 2.45 standard deviations lower (3.63 to 1.26 lower)
Change in Lateral Spine BMD - Adults	21 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		Not calculable for SMD values	The mean change in lateral spine bmd - adults in the intervention groups was 5.09 standard deviations higher

		indirectness, imprecision			(3.18 to 7 higher)
Change in Total Hip BMD - Adults	21 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in total hip bmd - adults in the intervention groups was 0.19 standard deviations lower (1.05 lower to 0.67 higher)
Change in FN BMD - Adults	21 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in fn bmd - adults in the intervention groups was 0.86 standard deviations lower (1.77 lower to 0.04 higher)
Change in AP Spine BMD - Adults	21 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision			The mean change in ap spine bmd - adults in the intervention groups was 4.61 standard deviations higher (2.84 to 6.38 higher)
Did not dropout due to side effects	21 (1 study)	See comment	Not estimable	See comment	-
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Randomisation method was unclear and it was unclear if allocation concealment was conducted. It was unclear if either the participants, investigators or assessors were blind. No dropouts were reported. 2 Short intervention of 6 months. 3 For a continuous outcome there were fewer than 400 participants. 4 95% CI crossed 2 MIDs (-0.5 and 0.5). 5 95% CI crossed 1 MID (-0.5).					

1 **Table 153: Summary table of findings for effects of IGF in adults with anorexia nervosa at the end of treatment.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with IGF (95% CI)
Change in Total Hip BMD - IGF-I vs. placebo	31 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in total hip bmd - igf-i vs. placebo in the intervention groups was 0.37 standard deviations higher (0.36 lower to 1.11 higher)
Change in Total Hip BMD -	31	⨁⊖⊖⊖ VERY LOW1,2,3		Not calculable	The mean change in total hip bmd - igf + ocp vs. placebo in the intervention groups

IGF + OCP vs. placebo	(1 study)	due to risk of bias, indirectness, imprecision		for SMD values	was 0.49 standard deviations higher (0.23 lower to 1.2 higher)
Change in Total Hip BMD - IGF vs. OCP	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in total hip bmd - igf vs. ocp in the intervention groups was 1.08 standard deviations higher (0.29 to 1.86 higher)
Change in Total Hip BMD - IGF-I + OCP vs. OCP	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in total hip bmd - igf-i + ocp vs. ocp in the intervention groups was 1.18 standard deviations higher (0.41 to 1.95 higher)
Change in Total Hip BMD - IGF-I + OCP vs. IGF-I	32 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in total hip bmd - igf-i + ocp vs. igf-i in the intervention groups was 0.10 standard deviations higher (0.62 lower to 0.82 higher)
Change Total Body BMD - IGF-I vs. placebo	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change total body bmd - igf-i vs. placebo in the intervention groups was 0.10 standard deviations higher (0.63 lower to 0.83 higher)
Change Total Body BMD - IGF + OCP vs. placebo	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change total body bmd - igf + ocp vs. placebo in the intervention groups was 1.27 standard deviations higher (0.49 to 2.05 higher)
Change Total Body BMD - IGF vs. OCP	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change total body bmd - igf vs. ocp in the intervention groups was 1.33 standard deviations higher (0.51 to 2.15 higher)
Change Total Body BMD - IGF-I + OCP vs. OCP	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change total body bmd - igf-i + ocp vs. ocp in the intervention groups was 2.55 standard deviations higher (1.58 to 3.53 higher)
Change Total Body BMD - IGF-I + OCP vs. IGF-I	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change total body bmd - igf-i + ocp vs. igf-i in the intervention groups was 1.17 standard deviations higher (0.38 to 1.95 higher)
Change in Radial BMD - IGF-	29	⊕⊖⊖⊖		Not	The mean change in radial bmd - igf-i vs.

I vs. placebo	(1 study)	VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		calculable for SMD values	placebo in the intervention groups was 0.25 standard deviations higher (0.48 lower to 0.98 higher)
Change in Radial BMD - OCP vs. placebo	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in radial bmd - ocp vs. placebo in the intervention groups was 0.62 standard deviations higher (0.12 lower to 1.35 higher)
Change in Radial BMD - IGF + OCP vs. placebo	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in radial bmd - igf + ocp vs. placebo in the intervention groups was 1.34 standard deviations higher (0.55 to 2.13 higher)
Change in Radial BMD - IGF vs. OCP	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in radial bmd - igf vs. ocp in the intervention groups was 0.29 standard deviations lower (1.02 lower to 0.44 higher)
Change in Radial BMD - IGF-I + OCP vs. IGF-I	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in radial bmd - igf-i + ocp vs. igf-i in the intervention groups was 0.88 standard deviations higher (0.12 to 1.63 higher)
Change in AP Spine BMD - IGF-I vs. placebo	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in ap spine bmd - igf-i vs. placebo in the intervention groups was 1.17 standard deviations higher (0.37 to 1.96 higher)
Change in AP Spine BMD - IGF + OCP vs. placebo	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in ap spine bmd - igf + ocp vs. placebo in the intervention groups was 2.34 standard deviations higher (1.4 to 3.28 higher)
Change in AP Spine BMD - IGF vs. OCP	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in ap spine bmd - igf vs. ocp in the intervention groups was 0.58 standard deviations higher (0.16 lower to 1.33 higher)
Change in AP Spine BMD - IGF-I + OCP vs. OCP	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in ap spine bmd - igf-i + ocp vs. ocp in the intervention groups was 1.75 standard deviations higher (0.91 to 2.6 higher)
Change in AP Spine BMD -	30	⊕⊖⊖⊖		Not	The mean change in ap spine bmd - igf-i +

IGF-I + OCP vs. IGF-I	(1 study)	VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		calculable for SMD values	ocp vs. igf-i in the intervention groups was 1.17 standard deviations higher (0.38 to 1.95 higher)
Change in Lean Mass - IGF-I vs. placebo	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in lean mass - igf-i vs. placebo in the intervention groups was 1.59 standard deviations higher (0.74 to 2.44 higher)
Change in Lean Mass - IGF + OCP vs. placebo	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in lean mass - igf + ocp vs. placebo in the intervention groups was 2.34 standard deviations higher (1.4 to 3.28 higher)
Change in Radial BMD - IGF-I + OCP vs. OCP	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in radial bmd - igf-i + ocp vs. ocp in the intervention groups was 0.58 standard deviations higher (0.14 lower to 1.31 higher)
Change in Lean Mass - IGF vs. OCP	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in lean mass - igf vs. ocp in the intervention groups was 1.46 standard deviations higher (0.63 to 2.29 higher)
Change in Lean Mass - IGF-I + OCP vs. OCP	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in lean mass - igf-i + ocp vs. ocp in the intervention groups was 2.12 standard deviations higher (1.22 to 3.03 higher)
Change in Lean Mass - IGF-I + OCP vs. IGF-I	59 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in lean mass - igf-i + ocp vs. igf-i in the intervention groups was 0.60 standard deviations higher (0.08 to 1.13 higher)
Change in Weight - IGF-I vs.placebo	59 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in weight - igf-i vs.placebo in the intervention groups was 0.54 standard deviations higher (0.02 to 1.07 higher)
Change in Weight - IGF-I +Oestrogen vs. placebo	59 (2 studies)	⊕⊖⊖⊖ VERY LOW1,5,7,8 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean change in weight - igf-i +oestrogen vs. placebo in the intervention groups was 0.14 standard deviations lower (0.72 lower to 0.44 higher)
Change in Weight - IGF-I +	60	⊕⊖⊖⊖		Not	The mean change in weight - igf-i +

oestrogenvs. Oestrogen	(2 studies)	VERY LOW1,5,7,8 due to risk of bias, inconsistency, indirectness, imprecision		calculable for SMD values	oestrogenvs. oestrogenin the intervention groups was 0.53 standard deviations lower (1.07 lower to 0.01 higher)
Change in Weight - IGF-I + oestrogenvs. IGF-I	60 (2 studies)	⊕⊖⊖⊖ VERY LOW1,5,7,8 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in weight - igf-i + oestrogenvs. igf-i in the intervention groups was 0.48 standard deviations lower (1.06 lower to 0.09 higher)
Change in Weight - IGF-I vs. Oestrogen	60 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,7,8 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in weight - igf-i vs. oestrogenin the intervention groups was 0.35 standard deviations higher (0.18 lower to 0.89 higher)
Change in BMI - IGF-I vs. placebo	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in BMI - igf-i vs. placebo in the intervention groups was 0.76 standard deviations higher (0 to 1.52 higher)
Change in BMI - IGF-I +Oestrogen vs. placebo	29 (1 study)	⊕⊖⊖⊖ VERY LOW1,5,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in BMI - igf-i +oestrogen vs. placebo in the intervention groups was 1.46 standard deviations lower (2.29 to 0.63 lower)
Change in BMI - IGF-I + oestrogen vs. Oestrogen	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,6,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in BMI - igf-i + oestrogenvs. oestrogenin the intervention groups was 0.97 standard deviations lower (1.74 to 0.21 lower)
Change in BMI - IGF-I + oestrogenvs. IGF-I	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,5,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in BMI - igf-i + oestrogenvs. igf-i in the intervention groups was 1.91 standard deviations lower (2.79 to 1.02 lower)
Change in BMI - IGF-I vs. Oestrogen	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean change in BMI - igf-i vs. oestrogenin the intervention groups was 1.14 standard deviations higher (0.36 to 1.93 higher)
Did not drop out due to side-effects - OCP vs. placebo	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,7, 9	RR 1.00 (0.88 to	See comment**	-

		due to risk of bias, indirectness, imprecision	1.13)		
Did not drop out due to side-effects - IGF-I + OCP vs IGF-I	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,7, 9 due to risk of bias, indirectness, imprecision	RR 1.00 (0.88 to 1.13)	See comment**	-
Did not drop out due to side-effects - IGF-I vs. OCP	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,2, 9 due to risk of bias, indirectness, imprecision	RR 0.94 (0.78 to 1.12)	Not calculable for SMD values	-
Did not drop out due to side-effects. Combined vs. placebo	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,7, 9 due to risk of bias, indirectness, imprecision	RR 1.00 (0.89 to 1.13)	See comment**	-
Did not drop out due to side-effects. IGF-I + OCP vs. OCP	31 (1 study)	⊕⊖⊖⊖ VERY LOW1,7, 9 due to risk of bias, indirectness, imprecision	RR 1.00 (0.89 to 1.13)	See comment**	-

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

** Absolute effects could not be calculated because zero events were included in the original analysis.

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Randomisation method was unclear and it was unclear if allocation concealment was conducted. Participants were blind, investigators were not and it was unclear if assessors were blind. A high dropout rate was detected in control arm >20%.
2 Relatively short period, 9 months
3 95% CI Crossed 1 MID (0.5)
4 95% CI Crossed 2 MIDs (-0.5 and 0.5)
5 For a continuous outcome, there were fewer than 400 participants.
6 95% CI Crossed 1 MID (-0.5)
7 Relatively short study duration, 3 months
8 Heterogeneity detected, I2>80%
9. For a dichotomous outcome, there were fewer than 300 events

Table 154: Summary table of findings for oestrogen versus placebo in young people and adults with anorexia nervosa at the end of treatment.

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with Placebo	Risk difference with oestrogen(95% CI)
Change LS BMD - Young people	222 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change ls bmd - young people in the intervention groups was 1.05 standard deviations higher (0.74 to 1.36 higher)
Change LS BMD - Adults	74 (2 studies)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean change ls bmd - adults in the intervention groups was 1.05 standard deviations higher (0.74 to 1.36 higher)
Change in FN BMD - Young people	112 (1 study)	⊕⊕⊖⊖ LOW1 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in fn bmd - young people in the intervention groups was 0.22 standard deviations lower (0.15 lower to 0.6 higher)
Change Total Hip BMD - Young people	222 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean change total hip bmd - young people in the intervention groups was 0.61 standard deviations higher (0.33 to 0.88 higher)
Change Total Hip BMD - Adults	30 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean change total hip bmd - adults in the intervention groups was 1.02 standard deviations lower (1.79 to 0.25 lower)
Change in Weight - Young people	222 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in weight - young people in the intervention groups was 0.34 standard deviations higher (0.07 to 0.6 higher)
Change in Weight - Adults	29 (1 study)	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in weight - adults in the intervention groups was 0.39 standard deviations lower (1.13 lower to 0.35 higher)
Change in BMI - Young people	110 (1 study)	⊕⊕⊖⊖ LOW2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in BMI - young people in the intervention groups was 0.27 standard deviations higher (0.11 lower to 0.64 higher)
Change in BMI - Adults	139 (1 study)	⊕⊕⊖⊖ LOW5,8		Not calculable	The mean change in BMI - adults in the intervention groups was

		due to risk of bias, imprecision		for SMD values	0.11 standard deviations higher (0.22 lower to 0.45 higher)
Change in Lean mass - Young people	110 (1 study)	⊕⊕⊖⊖ LOW2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in lean mass - young people in the intervention groups was 0.17 standard deviations higher (0.2 lower to 0.55 higher)
Change in Lean Mass - Adults	140 (1 study)	⊕⊕⊖⊖ LOW5,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in lean mass - adults in the intervention groups was 0.13 standard deviations higher (0.2 lower to 0.47 higher)
Change in Fat Mass - Young people	110 (1 study)	⊕⊕⊖⊖ LOW2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in fat mass - young people in the intervention groups was 0.17 standard deviations higher (0.2 lower to 0.55 higher)
Change in Total Body BMD - Adults	30 (1 study)	⊕⊕⊖⊖ LOW9,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in total body bmd - adults in the intervention groups was 1.23 standard deviations lower (2.02 to 0.44 lower)
Did not achieve normal menses Young people	110 (1 study)	⊕⊕⊖⊖ LOW7,11 due to risk of bias, imprecision	RR 1.0 (1 to 1.2)	91 per 1000	0 fewer per 1000 (from 0 more to 18 more)
Did not achieve remission - Adults	44 (1 study)	⊕⊕⊖⊖ LOW11,12 due to risk of bias, imprecision	RR 1.10 (0.9 to 1.54)	240 per 1000	24 more per 1000 (from 24 fewer to 130 more)
Did not drop out due to side-effects- Young people	123 (1 study)	⊕⊕⊖⊖ LOW1,11 due to risk of bias, imprecision	RR 0.97 (0.91 to 1.03)	16 per 1000	0 fewer per 1000 (from 1 fewer to 0 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear in all studies if allocation concealment was conducted, The investigators and participants were blind, but it was unclear if the assessors were blind. High dropouts were reported >20%.
2 95% CI Crossed 1 MID (0.5).

3 It was unclear in all studies if allocation concealment was conducted. In one study the investigators were not blind and in the other it was unclear. Participants were blind in one study but it was unclear in the other study. It was also unclear for both studies if the assessors were blind. High dropouts were reported across studies >20%.
4 Heterogeneity was detected I2 >80%.
5 It was unclear in all studies if allocation concealment was conducted. In Grinspoon, the investigators were not blind but the participants were blind and it was unclear if assessors were blind. High dropouts were reported in both studies >20%.
6 Heterogeneity was detected I2 >50%.
7 It was unclear if allocation concealment was conducted, The investigators and participants were blind, but it was unclear if the assessors were blind. High dropouts were reported >20%.
8 For a continuous outcome there were fewer than 400 participants.
9 It was unclear in all studies if allocation concealment was conducted. In both studies the participants were blind. In Grinspoon, the investigators were not blind and it was unclear if assessors were blind. High dropouts were reported >20%.
10 95% CI Crossed 1 MID (-0.5).
11 For a dichotomous outcome, there were fewer than 300 events.
12 It was unclear if allocation concealment was conducted. It was unclear in Klibanski if either the participants, investigators or assessors were blind. High dropouts were reported in both studies >20%

Table 155: Summary table of findings for bisphosphonates versus placebo in adults and young people with anorexia nervosa at the end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Bisphosphonate (95% CI)
Tibia SOS - Etidronate vs. placebo	26 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean tibia sos - etidronate vs. placebo in the intervention groups was 0.33 standard deviations higher (0.45 lower to 1.1 higher)
Tibia SOS - Etidronate vs. Calcium Vit D	29 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean tibia sos - etidronate vs. calcium vit d in the intervention groups was 0.47 standard deviations lower (1.21 lower to 0.27 higher)
Tibia Z Score - Etidronate vs. placebo	26 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean tibia z score - etidronate vs. placebo in the intervention groups was 0.64 standard deviations higher (0.15 lower to 1.43 higher)
Tibia Z Score - Etidronate vs. Calcium Vit D	29 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias,		Not calculable for SMD	The mean tibia z score - etidronate vs. calcium vit d in the intervention groups was 0.24 standard deviations lower

		imprecision		values	(0.97 lower to 0.49 higher)
Difference in Lateral spine BMD	39 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in lateral spine bmd in the intervention groups was 1.35 standard deviations higher (2.05 to 0.64 lower)
Difference in hip BMD	38 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean difference in hip bmd in the intervention groups was 1.42 standard deviations higher (2.13 to 0.71 lower)
PA Spine BMD Z score	38 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean pa spine bmd z score in the intervention groups was 1.26 standard deviations higher (0.56 to 1.96 higher)
LS BMD Z score change - Young people	29 (1 study)	⊕⊕⊖⊖ LOW6,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ls bmd z score change - young people in the intervention groups was 0.05 standard deviations lower (0.78 lower to 0.68 higher)
FN BMD Z score change - Young people	29 (1 study)	⊕⊖⊖⊖ VERY LOW6,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean fn bmd z score change - - young people in the intervention groups was 0.39 standard deviations higher (0.34 lower to 1.13 higher)
Trochanter BMD Change - Young people	29 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean trochanter bmd change - young people in the intervention groups was 4.60 standard deviations higher (3.13 to 6.07 higher)
Wards Triangle Change BMD - Young people	29 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean wards triangle change bmd - young people in the intervention groups was 0.54 standard deviations higher (0.2 lower to 1.28 higher)
Total Hip BMD Change - Young people	29 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean total hip bmd change - young people in the intervention groups was 0.24 standard deviations higher (0.49 lower to 0.97 higher)
Did not drop out due to SE - Bisphosphonates vs. placebo	95 (3 studies)	⊕⊕⊖⊖ LOW1,8 due to risk of bias, imprecision	RR 1.02 (0.94 to 1.1)	21 per 1000	0 more per 1000 (from 1 fewer to 2 more)

Did not drop out due to SE - Bisphosphonates vs. Ca Vit D	29 (1 study)	⊕⊕⊖⊖ LOW1,8 due to risk of bias, imprecision	RR 1.01 (0.96 to 1.09)	See comment	-

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
** Absolute effects could not be calculated because zero events were included in the original analysis.

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. Both the participants and investigators were blind but it was unclear if assessors were blind.
2 95% CI crossed 1 MID (0.5).
3 95% CI Crossed 1 MID (-0.5).
4 Unclear how randomisation sequence was generated or if allocation concealment was performed. Double-blind study, but unclear if the assessors were blind. Not clear what groups the dropouts were in.
5 For a continuous outcome there were fewer than 400 participants.
6 Unclear how randomisation sequence was generated and unclear if allocation concealment was conducted. The participants, investigators and assessors were blind. Low dropout rates.
7 95% CI crossed 2 MIDs (-0.5 and 0.5)
8 For a dichotomous outcome, there were fewer than 300 events.

1 **Low hormone levels**

2 **Table 156: Summary table of findings for IGF-I versus placebo in adults with anorexia nervosa at the end of treatment.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo (AN)	Risk difference with Growth Hormone (rhGH) (95% CI)
Change in body weight. Adults	21 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in body weight. adults in the intervention groups was 0.7 standard deviations lower (1.59 lower to 0.19 higher)
IGF-I. Adults	21 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean igf-i. adults in the intervention groups was 1.92 standard deviations higher (0.85 to 2.99 higher)
Change in IGF-I. Adults	21 (1 study)	⊕⊖⊖⊖ VERY LOW1,4		Not calculable for SMD values	The mean change in igf-i. adults in the intervention groups was

		due to risk of bias, imprecision			0.9 standard deviations higher (0.01 lower to 1.81 higher)
Dropout for any reason. Adults	21 (1 study)	See comment	RR 0.22 (0.01 to 4.06)	182 per 1000	142 fewer per 1000 (from 180 fewer to 556 more)
Dropout due to side-effects. Adults	21 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	Not estimable	See comment	See comment

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear methods of randomisation or if allocation concealment was performed. It is also unclear either the participants, investigators or assessors were blind.
2 95% CI crossed 1 MID (-0.5)
3 Fewer than 400 participants were available for this outcome.
4 95% CI crossed 2 MIDs (-0.5 and 0.5)
5 Fewer than 300 events were available for this outcome.

1 **Low body weight and malnourishment**

2 **Table 157: Summary table of findings for parenteral and enteral refeeding diet versus enteral refeeding diet in young females with**
3 **anorexia nervosa at end of treatment.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Enteral Refeeding	Risk difference with (Obs) Parenteral+Enteral Refeeding (95% CI)
BMI	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.28 standard deviations lower (0.56 lower to 0 higher)
% Ideal Body Weight - Young people	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean % ideal body weight - young people in the intervention groups was 0.37 standard deviations lower (0.65 to 0.09 lower)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Enteral Refeeding	Risk difference with (Obs) Parenteral+Enteral Refeeding (95% CI)
Weight Gain (g/week) - Young people	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight gain (g/week) - young people in the intervention groups was 16.27 standard deviations higher (14.63 to 17.91 higher)
Length of Treatment (days) - Young people	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean length of treatment (days) - young people in the intervention groups was 8.66 standard deviations higher (7.75 to 9.56 higher)
Maximum Energy Intake (kcal/day) - Young people	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean maximum energy intake (kcal/day) - young people in the intervention groups was 3.06 standard deviations higher (2.64 to 3.47 higher)
Abdominal Pain - Young people	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.4 (0.18 to 0.88)	191 per 1000	115 fewer per 1000 (from 23 fewer to 157 fewer)
Bloating - Young people	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.45 (0.19 to 1.07)	149 per 1000	82 fewer per 1000 (from 121 fewer to 10 more)
Constipation - Young people	198 (1 study) 33.3 months	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 0.72 (0.3 to 1.76)	106 per 1000	30 fewer per 1000 (from 74 fewer to 81 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Diamanti 2008: high selection bias(significantly higher psychiatric comorbidity, weight loss at diagnosis, and resting energy expenditure in parenteral group; significantly lower % Ideal Body Weight, Weight at diagnosis and BMI in parenteral group).
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 <300 events or <400 participants.
4 CI crosses both 0.75 and 1.25 (Risk Ratio).

Table 158: Summary table of findings for parenteral and enteral refeeding diet versus enteral refeeding diet in young females with anorexia nervosa at follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Enteral Refeeding	Risk difference with (Obs) Parenteral and Enteral Refeeding (95% CI)
Recovered after nutritional rehabilitation - Young people	129 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.95 (0.73 to 1.25)	642 per 1000	32 fewer per 1000 (from 173 fewer to 160 more)
Rehospitalized - Young people	129 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 0.89 (0.48 to 1.65)	254 per 1000	28 fewer per 1000 (from 132 fewer to 165 more)
Length of 2nd rehospitalisation - Young people	129 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean length of 2nd rehospitalisation - young people in the intervention groups was 0.62 standard deviations higher (0.27 to 0.98 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Diamanti 2008: high selection bias(significantly higher psychiatric comorbidity, weight loss at diagnosis, and resting energy expenditure in parenteral group; significantly lower % Ideal Body Weight, Weight at diagnosis and BMI in parenteral group).
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 CI crosses both 0.75 and 1.25 (Risk Ratio).

Table 159: Summary table of findings for percutaneous gastric refeeding diet versus nasogastric tube refeeding diet or No Refeeding in low body weight adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Nasogastric Feeding/No Tube	Risk difference with (Obs) Percutaneous Gastric (95% CI)
Weight Gain (kg) at discharge - Adult	68 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean weight gain (kg) at discharge - adult in the intervention groups was 0.17 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Nasogastric Feeding/No Tube	Risk difference with (Obs) Percutaneous Gastric (95% CI)
		imprecision			(0.47 lower to 0.82 higher)
Length of Treatment (days) - Adult	68 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean length of treatment (days) - adult in the intervention groups was 0.87 standard deviations higher (0.21 to 1.54 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Born 2015: high selection bias (method of allocation to groups related to potential confounding factors), high performance bias (participants received various forms of therapies).
2 CI crosses 0.5 or -0.5.

1 **Table 160: Summary table of findings for nasogastric tube and oral refeeding versus oral refeeding in young underweight females**
2 **with anorexia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Oral Refeeding	Risk difference with (Obs) Nasogastric+Oral (95% CI)
BMI - Young people	100 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI - Young people in the intervention groups was 0.48 standard deviations higher (0.08 to 0.88 higher)
BMI change at discharge - Young people	100 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI change at discharge - Young people in the intervention groups was 1 standard deviations higher (0.58 to 1.42 higher)
Weight (kg) - Young people	100 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) - Young people in the intervention groups was 0.27 standard deviations higher (0.13 lower to 0.66 higher)
Weight Gain at discharge -	100	⨁⊖⊖⊖		Not calculable	The mean weight gain at discharge - Young people

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Oral Refeeding	Risk difference with (Obs) Nasogastric+Oral (95% CI)
Young people	(1 study)	VERY LOW1,3 due to risk of bias, imprecision		for SMD values	in the intervention groups was 0.95 standard deviations higher (0.54 to 1.36 higher)
Length of Stay (days) - Young people	100 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean length of stay (days) - Young people in the intervention groups was 0.02 standard deviations higher (0.38 lower to 0.41 higher)
Maximum Caloric Intake (kcal/day) - Young people	100 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean maximum caloric intake (kcal/day) - Young people in the intervention groups was 1.27 standard deviations higher (0.84 to 1.7 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Robb 2002: high selection bias (significantly higher number of hospitalisations in nocturnal NG + oral refeeding group); high performance bias (participants received various therapies during course of treatment).
2 CI crosses 0.5 or -0.5.
3 <300 events or <400 participants.

1 **Table 161: Summary table of findings for nasogastric and oral refeeding versus oral refeeding in malnourished adults with anorexia**
2 **nervosa at end of treatment.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Oral Refeeding for adult AN	Risk difference with (RCT) Nasogastric+Oral (95% CI)
BMI>18.5	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 5.2 (1.64 to 16.49)	75 per 1000	315 more per 1000 (from 48 more to 1000 more)
Weight (kg)	81 (1 study)	⊕⊕⊖⊖ LOW1,3		Not calculable for SMD values	The mean weight (kg) in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Oral Refeeding for adult AN	Risk difference with (RCT) Nasogastric+Oral (95% CI)
	1 years	due to risk of bias, imprecision			0.63 standard deviations higher (0.18 to 1.08 higher)
Weight (kg) - AN-R	56 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) - an-r in the intervention groups was 1.13 standard deviations higher (0.56 to 1.7 higher)
Weight (kg) - AN-BP	25 (1 study) 1 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) - an-bp in the intervention groups was 1.15 standard deviations higher (0.29 to 2.01 higher)
Weight Gain (g/day)	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight gain (g/day) in the intervention groups was 4.04 standard deviations higher (3.27 to 4.82 higher)
Relapse-Free Period (weeks)	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean relapse-free period (weeks) in the intervention groups was 0.94 standard deviations higher (0.48 to 1.41 higher)
Change in Extracellular fluids (kg)	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in extracellular fluids (kg) in the intervention groups was 5.03 standard deviations lower (5.94 to 4.13 lower)
Creatinine urinary output (mg/day)	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean creatinine urinary output (mg/day) in the intervention groups was 0.67 standard deviations higher (0.22 to 1.12 higher)
Fat Free Mass (kg)	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean fat free mass (kg) in the intervention groups was 1.04 standard deviations higher (0.57 to 1.5 higher)
Fat Free Mass Gain (g/day)	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean fat free mass gain (g/day) in the intervention groups was 3.06 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Oral Refeeding for adult AN	Risk difference with (RCT) Nasogastric+Oral (95% CI)
	1 years	imprecision			(2.41 to 3.71 higher)
Fat Mass Gain (g/day)	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean fat mass gain (g/day) in the intervention groups was 0.55 standard deviations higher (0.1 to 0.99 higher)
Added Sugar (sucrose) (g/day)	81 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean added sugar (sucrose) (g/day) in the intervention groups was 0.45 standard deviations lower (0.89 to 0.01 lower)
Added Fat (g/day)	81 (1 study) 1 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean added fat (g/day) in the intervention groups was 0.24 standard deviations higher (0.2 lower to 0.68 higher)
Energy Intake (kcal/day) - AN-R	56 (1 study) 1 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean energy intake (kcal/day) - an-r in the intervention groups was 0.46 standard deviations higher (0.08 lower to 0.99 higher)
Energy Intake (kcal/day) - AN-BP	25 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean energy intake (kcal/day) - an-bp in the intervention groups was 0.93 standard deviations lower (1.77 to 0.1 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Rigaud 2007: no details of randomization method provided; unclear whether participant, investigator or assessor blinded.
2 <300 events or <400 participants.
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 162: Summary table of findings for nasogastric and oral refeeding versus oral refeeding in malnourished adults with anorexia nervosa at follow up.

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Oral Refeeding	Risk difference with (RCT) Nasogastric+Oral (95% CI)
Weight (kg) - AN-R 12-mo FU	56 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) - an-r 12-mo fu in the intervention groups was 0.99 standard deviations higher (0.43 to 1.55 higher)
Weight (kg) AN-BP 12-mo FU	25 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) an-bp 12-mo fu in the intervention groups was 1.2 standard deviations higher (0.33 to 2.06 higher)
# Relapsed 12-mo FU	81 (1 study) 1 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 0.84 (0.53 to 1.32)	525 per 1000	84 fewer per 1000 (from 247 fewer to 168 more)
Energy Intake - AN-R 12-mo FU (kcal/day)	56 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean energy intake - an-r 12-mo fu (kcal/day) in the intervention groups was 0 standard deviations higher (0.52 lower to 0.53 higher)
Energy Intake AN-BP 12-mo FU (kcal/day)	25 (1 study) 1 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean energy intake an-bp 12-mo fu (kcal/day) in the intervention groups was 0.28 standard deviations lower (1.07 lower to 0.51 higher)
# BMI>18.5 + adequate energy intake 12-mo FU	81 (1 study) 1 years	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 1.33 (0.7 to 2.53)	275 per 1000	91 more per 1000 (from 83 fewer to 421 more)
EDI Total 12-mo FU	81 (1 study) 1 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total 12-mo fu in the intervention groups was 0.15 standard deviations lower (0.59 lower to 0.28 higher)
Resumed menses 12-mo FU	26 (1 study) 1 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.11 (0.88 to 1.4)	909 per 1000	100 more per 1000 (from 109 fewer to 364 more)
# taking antidepressants 12-mo FU	81 (1 study)	⊕⊖⊖⊖ VERY LOW1,3	RR 1.17 (0.39 to	125 per 1000	21 more per 1000 (from 76 fewer to 316 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Oral Refeeding	Risk difference with (RCT) Nasogastric+Oral (95% CI)
	1 years	due to risk of bias, imprecision	3.53)		
# taking antixiolytics 12-mo FU	81 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 0.76 (0.31 to 1.84)	225 per 1000	54 fewer per 1000 (from 155 fewer to 189 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Rigaud 2007: no details of randomization method provided; unclear whether participant, investigator or assessor blinded.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).

Table 163: Summary table of findings for high-calorie diet versus low-calorie diet in malnourished young people with anorexia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low-Calorie Diet	Risk difference with (RCT) High-Calorie Diet (95% CI)
QT-corrected Interval at 4 days - QT-c (ms)	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean qt-corrected interval at 4 days - qt-c (ms) in the intervention groups was 0.01 standard deviations higher (0.64 lower to 0.67 higher)
QT-corrected Interval at 4 days - Change scores	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean qt-corrected interval at 4 days - change scores in the intervention groups was 0.24 standard deviations higher (0.42 lower to 0.89 higher)
Heart Rate at 4 days - Heart Rate (bpm)	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		Not calculable for SMD values	The mean heart rate at 4 days - heart rate (bpm) in the intervention groups was 0.58 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low-Calorie Diet	Risk difference with (RCT) High-Calorie Diet (95% CI)
		indirectness, imprecision			(0.09 lower to 1.25 higher)
Heart Rate at 4 days - Change	36 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean heart rate at 4 days - change in the intervention groups was 0 standard deviations higher (0.65 lower to 0.65 higher)
Weight (kg) at 4 days - Weight (kg)	36 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight (kg) at 4 days - weight (kg) in the intervention groups was 0.21 standard deviations lower (0.86 lower to 0.45 higher)
Weight (kg) at 4 days - Change	36 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight (kg) at 4 days - change in the intervention groups was 0.64 standard deviations higher (0.03 lower to 1.31 higher)
BMI at 4 days - BMI	36 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI at 4 days - BMI in the intervention groups was 0.37 standard deviations higher (0.29 lower to 1.03 higher)
BMI at 4 days - Change	36 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI at 4 days - change in the intervention groups was 0.44 standard deviations higher (0.22 lower to 1.11 higher)
mBMI (%) at 4 days - mBMI (%)	36 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean mbmi (%) at 4 days - mbmi (%) in the intervention groups was 0.47 standard deviations higher (0.2 lower to 1.13 higher)
mBMI (%) at 4 days - Change	36	⨁⊖⊖⊖		Not	The mean mbmi (%) at 4 days - change in the

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low-Calorie Diet	Risk difference with (RCT) High-Calorie Diet (95% CI)
	(1 study)	VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		calculable for SMD values	intervention groups was 0.56 standard deviations higher (0.11 lower to 1.23 higher)
Serum Phosphate Concentration at 4 days - Nadir (mmol/L)	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean serum phosphate concentration at 4 days - nadir (mmol/l) in the intervention groups was 0.06 standard deviations higher (0.6 lower to 0.71 higher)
Serum Phosphate Concentration at 4 days - Change (mmol/L)	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean serum phosphate concentration at 4 days - change (mmol/l) in the intervention groups was 0.17 standard deviations lower (0.82 lower to 0.49 higher)
Energy Intake at 4 days - Kcal/day	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean energy intake at 4 days - kcal/day in the intervention groups was 2.16 standard deviations higher (1.32 to 3 higher)
Energy Intake at 4 days - Kcal/g	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean energy intake at 4 days - kcal/g in the intervention groups was 1.78 standard deviations higher (0.99 to 2.56 higher)
Weight (kg) at 10 days - Weight (kg)	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight (kg) at 10 days - weight (kg) in the intervention groups was 0.18 standard deviations lower (0.84 lower to 0.47 higher)
Weight (kg) at 10 days - Change	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness,		Not calculable for SMD values	The mean weight (kg) at 10 days - change in the intervention groups was 0.49 standard deviations higher (0.17 lower to 1.16 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low-Calorie Diet	Risk difference with (RCT) High-Calorie Diet (95% CI)
		imprecision			
BMI at 10 days - BMI	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI at 10 days - BMI in the intervention groups was 0.32 standard deviations higher (0.34 lower to 0.98 higher)
BMI at 10 days - Change	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI at 10 days - change in the intervention groups was 0.55 standard deviations higher (0.11 lower to 1.22 higher)
mBMI (%) at 10 days	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean mbmi (%) at 10 days in the intervention groups was 0.5 standard deviations higher (0.17 lower to 1.16 higher)
mBMI (%) at 10 days - Change	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean mbmi (%) at 10 days - change in the intervention groups was 0.64 standard deviations higher (0.04 lower to 1.31 higher)
Energy Intake at 10 days - Kcal/day	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean energy intake at 10 days - kcal/day in the intervention groups was 0.95 standard deviations higher (0.25 to 1.64 higher)
Energy Intake at 10 days - Kcal/g	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean energy intake at 10 days - kcal/g in the intervention groups was 0.91 standard deviations higher (0.22 to 1.6 higher)
Glucose (mmol/L) at 10 days	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3		Not calculable for	The mean glucose (mmol/l) at 10 days in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low-Calorie Diet	Risk difference with (RCT) High-Calorie Diet (95% CI)
		due to risk of bias, indirectness, imprecision		SMD values	0.39 standard deviations higher (0.27 lower to 1.05 higher)
Insulin (miu mol/L) at 10 days	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean insulin (miu mol/l) at 10 days in the intervention groups was 0.34 standard deviations higher (0.32 lower to 1 higher)
HOMA at 10 days Homeostatic Model Assessment Insulin Resistance	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean homa at 10 days in the intervention groups was 0.62 standard deviations higher (0.05 lower to 1.29 higher)
White Blood Cell Count (x 10 9/L)	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean white blood cell count (x 10 9/l) in the intervention groups was 0.42 standard deviations higher (0.24 lower to 1.08 higher)
No adverse Events within first 4 days of treatment	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision	RR 1.06 (0.91 to 1.23)	944 per 1000	57 more per 1000 (from 85 fewer to 217 more)
No Oral Phosphate Supplementation due to low PO	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision	RR 1 (0.85 to 1.17)	944 per 1000	0 fewer per 1000 (from 142 fewer to 161 more)
Hypophosphatemia within first 2 days	36 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, indirectness, imprecision	RR 2.5 (0.56 to 11.25)	111 per 1000	167 more per 1000 (from 49 fewer to 1000 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low-Calorie Diet	Risk difference with (RCT) High-Calorie Diet (95% CI)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 O'Connor 2016: no info regarding allocation concealment; no participant nor investigator blinding. Two participants in each group required nasogastric tube feeding due to failing to achieve >=80% expected energy intake within 48 hours of admission. 2 CI crosses either 0.75 or 1.25 (Risk Ratio), or 0.5 or -0.5 (SMD). 3 Sample was participants diagnosed with anorexia nervosa or atypical anorexia nervosa. 4 CI crosses both 0.75 and 1.25 (Risk Ratio) or 0.5 and -0.5 (SMD). 5 <300 events (dichotomous outcome) or <400 participants (continuous outcome).					

Table 164: Summary table of findings for normal-sodium nasogastric tube and oral refeeding diet versus low-sodium nasogastric tube and oral refeeding diet in malnourished adults with anorexia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low Sodium diet	Risk difference with (Obs) Nasogastric+Oral Refeeding for adult AN: Normal Sodium (95% CI)
Weight (kg)	218 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight (kg) in the intervention groups was 0.25 standard deviations higher (0.09 lower to 0.59 higher)
BMI	218 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.13 standard deviations lower (0.47 lower to 0.21 higher)
Fat Free Mass (kg; skinfold)	218 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean fat free mass (kg; skinfold) in the intervention groups was 0.41 standard deviations higher (0.07 to 0.75 higher)
Active Fat Free Mass (kg)	218 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean active fat free mass (kg) in the intervention groups was 0.32 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Low Sodium diet	Risk difference with (Obs) Nasogastric+Oral Refeeding for adult AN: Normal Sodium (95% CI)
		indirectness, imprecision			(0.66 lower to 0.02 higher)
Fat Mass (kg; skinfold and BIA) - Fat Mass skinfold	218 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean fat mass (kg; skinfold and bia) - fat mass skinfold in the intervention groups was 0.36 standard deviations lower (0.7 to 0.03 lower)
Fat Mass (kg; skinfold and BIA) - Fat Mass BIA	218 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean fat mass (kg; skinfold and bia) - fat mass bia in the intervention groups was 0.16 standard deviations lower (0.5 lower to 0.18 higher)
Energy Input (kcal/day)	218 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean energy input (kcal/day) in the intervention groups was 0.19 standard deviations higher (0.14 lower to 0.53 higher)
Energy input tube feeding (kcal/day)	218 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean energy input tube feeding (kcal/day) in the intervention groups was 0.52 standard deviations lower (0.86 to 0.18 lower)
Edema of legs	218 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision	RR 3.43 (1.52 to 7.74)	62 per 1000	152 more per 1000 (from 32 more to 421 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Rigaud 2010: Method of analysis not clear and data throughout study not reported for all participants. No restriction in sodium and water intake in normal sodium group. Sample was 98% women, duration of illness not reported.
2 CI crosses 0.5 or -0.5.
3 <300 events (dichotomous outcome) or <400 participants (continuous outcome).

Table 165: Summary table of findings for oral potassium supplementation versus no supplementation for adults with anorexia nervosa and self-induced starvation.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with (Obs) Oral Potassium Supplementation (95% CI)
Corrected QT Dispersion (ms)	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean corrected qt dispersion (ms) in the intervention groups was 1.03 standard deviations lower (1.83 to 0.23 lower)
QT Dispersion (ms)	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean qt dispersion (ms) in the intervention groups was 1.47 standard deviations lower (2.32 to 0.62 lower)
Serum potassium (mmol l-1)	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean serum potassium (mmol l-1) in the intervention groups was 0.82 standard deviations higher (0.04 to 1.59 higher)
Urinary potassium excretion (mmol 24h-1)	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean urinary potassium excretion (mmol 24h-1) in the intervention groups was 1.79 standard deviations higher (0.9 to 2.69 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Franzoni 2002: high selection bias (unclear method of allocation to groups). Demographic and baseline details of treated and untreated group not provided.
2 CI crosses 0.5 or -0.5.
3 <400 participants.

6.8.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for the management of short and long-term physical complications of anorexia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.8.4 Clinical evidence statements

6.8.4.1 DHEA versus HRT for young people with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=61) showed HRT is more effective on change in lumbar spine BMD compared with DHEA, but there was some uncertainty.

Low quality evidence from one RCT (n=61) showed no difference in the effect of DHEA change in total hip BMD, change in body weight and dropouts due to side-effects compared with HRT.

Low quality evidence from one RCT (n=61) showed DHEA is less effective on achieving a resumption in menses compared with HRT, but there was some uncertainty.

6.8.4.2 DHEA and combined oral contraceptive pill (COC) versus placebo for adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=60 to 76) showed DHEA and COC is more effective on change in femoral shaft BMD, femoral neck BMD, change in femoral shaft bone strength index, change in body weight and BMI compared with placebo.

Low quality evidence from one RCT (n=60) showed no difference in the effect DHEA and COC has on menstrual function and dropouts due to side-effects compared with placebo.

Low quality evidence from one RCT (n= 76) showed DHEA and COC is less effective on change in femoral neck bone strength index compared with placebo.

6.8.4.3 PTH versus placebo for adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n=21) showed PTH is more effective on change in weight, change in lateral spine BMD and change in anteroposterior spine BMD compared with placebo.

Low quality evidence from one RCT (n=21) showed no difference in the effect of PTH on change in total body BMD and the number of dropouts due to side-effect compared with placebo.

Low quality evidence from one RCT (n=21) showed PTH is less effective on change in femoral neck BMD compared with placebo, but there was some uncertainty.

6.8.4.4 IGF-I versus placebo for adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n=29 to 31) showed no difference in the effect of IGF-I on change in total body, total hip and radial BMD compared with placebo.

Very low quality evidence from one RCT (n=31) showed IGF-I is more effective on change in anteroposterior BMD, change in lean mass, change in BMI compared with placebo.

Very low quality evidence from two RCTs (n=59) showed IGF-I is more effective on change in weight compared with placebo.

6.8.4.5 IGF-I and oral contraceptive pill (OCP) versus placebo for adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n=31) showed no difference in the effect of IGF-I and OCP on change in total hip BMD and dropouts due to side-effects compared with placebo.

Very low quality evidence from one RCT (n=29 to 31) showed IGF-I and OCP are more effective on change in total body, radial, anteroposterior BMD, change in lean mass, compared with placebo.

Very low quality evidence from one RCT (n=31) showed IGF-I and OCP are less effective on change in BMI compared with placebo.

Very low quality evidence from two RCTs (n=59) showed no difference in the effect of IGF- and OCP on change in weight compared with placebo.

6.8.4.6 IGF-I versus oral contraceptive pill (OCP) for adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n= 31) showed IGF-I is more effective on change in total hip and total body BMD and change in lean mass and BMI compared with placebo.

Very low quality evidence from one RCT (n= 29 to 31) showed no difference in the effect of IGF-I on change in radial and anteroposterior BMD and number of dropouts due to side-effects compared with placebo.

Very low quality evidence from two RCTs (n= 60) showed no difference in the effect of IGF-I on change in weight compared with placebo.

6.8.4.7 IGF-I and oral contraceptive pill (OCP) versus OCP for adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n= 31) showed IGF-I and OCP is more effective on change in total hip, total body, anteroposterior BMD, change in lean mass compared with OCP.

Very low quality evidence from one RCT (n= 31) showed no difference in the effect of IGF-I and OCP on change in radial BMD and dropouts due to side-effects compared with OCP.

Very low quality evidence from two RCTs (n= 60) showed IGF-I and OCP is less effective on change in weight compared with OCP, but there was some uncertainty.

Very low quality evidence from one RCT (n= 30) showed IGF-I and OCP is less effective on change in BMI compared with OCP.

6.8.4.8 IGF-I and oral contraceptive pill (OCP) versus IGF-I for adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n= 31) showed no difference in the effect of IGF-I and OCP on change in total hip and radial BMD, change in BMI dropouts due to side-effects compared with IGF-I.

Very low quality evidence from one RCT (n= 31) showed IGF-I and OCP is more effective on change in total body and anteroposterior BMD and change in lean mass compared with IGF-I.

Very low quality evidence from two RCTs (n= 60) showed IGF-I and OCP is less effective on change in weight compared with IGF-I, but there was some uncertainty.

6.8.4.9 Oral contraceptive pill (OCP) versus placebo for adults with anorexia nervosa at the end of treatment

Very low quality evidence from one RCT (n= 30) showed no difference in the effect of OCP on change in radial BMD and dropouts due to side-effects compared with placebo.

6.8.4.10 Oestrogen versus placebo for young people with anorexia nervosa at the end of treatment

Low quality evidence from two RCTs (n= 222) showed oestrogen is more effective on change in lumbar spine and total hip BMD and change in weight compared with placebo.

Low quality evidence from one RCT (n= 110) showed no difference in the effect of oestrogen change in femoral neck BMD, BMI, lean mass and fat mass, the number who achieved normal menses and who dropped out due to side-effects compared with placebo.

6.8.4.11 Oestrogen versus placebo for adults with anorexia nervosa at the end of treatment

Low quality evidence from two RCTs (n= 74) showed oestrogen is more effective on change in lumbar spine BMD compared with placebo.

Low quality evidence from one RCT (n= 29) showed no difference in the effect of oestrogen on change in weight compared with placebo.

Low quality evidence from one RCT (n= 139 to 140) showed no difference in the effect of oestrogen on change in BMI or lean mass compared with placebo.

Low quality evidence from one RCT (n= 44) showed no difference in the effect of oestrogen on remission compared with placebo.

Very low quality evidence from one RCT (n= 30) showed oestrogen is less effective on change in total hip and total body BMD compared with placebo.

6.8.4.12 Biosphosphonates versus placebo for adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n= 26) showed no difference in the effect of etidronate on ultrasound tibia speed of sound (SOS) and tibia Z score compared with placebo.

Low quality evidence from one RCT (n= 39) showed risendronate may be more effective on lateral spine, hip and anteroposterior spine BMD Z scores compared with placebo.

Low quality evidence from two RCTs (n= 66) showed no difference in the effect of bisphosphonates on dropouts due to side-effects compared with placebo.

6.8.4.13 Biosphosphonates versus placebo for young people with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n= 26 to 29) showed no difference in the effect of alendronate on change in total hip, lumbar spine, wards triangle or femoral neck BMD Z score compared with placebo.

Low quality evidence from one RCT (n= 29) showed alendronate may be more effective on change trochanter BMD compared with placebo.

Low quality evidence from one RCT (n= 29) showed no difference in the effect of bisphosphonates on dropouts due to side-effects compared with placebo.

6.8.4.14 Biosphosphonates versus calcium and vitamin D for adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n= 29) showed no difference in the effect of etidronate on ultrasound tibia speed of sound (SOS), tiba Z score or dropouts due to side-effects compared with calcium and vitamin D.

6.8.4.15 Growth hormone versus placebo for adults with anorexia nervosa at the end of treatment

Low quality evidence from one RCT (n= 21) showed no difference in the effect of growth hormone (rhGH) on change in body weight or dropouts due to any reason or because of side-effects compared with placebo.

Low quality evidence from one RCT (n= 21) showed growth hormone (rhGH) is more effective on serum IGF-I levels compared with placebo.

Low quality evidence from one RCT (n= 21) showed growth hormone (rhGH) is more effective on change in serum IGF-I levels compared with placebo, but there was some uncertainty.

6.8.4.16 Parenteral and enteral nutrition versus enteral nutrition in young people with anorexia nervosa at end of treatment

Very low quality evidence from one RCT (n=198) showed parenteral and enteral nutrition is less effective on BMI, % Ideal Body Weight and length of treatment compared with enteral nutrition.

Very low quality evidence from one RCT (n=198) showed parenteral and enteral nutrition is more effective on weekly weight gain, maximum daily energy intake and number of people experiencing abdominal pain compared with enteral nutrition.

Very low quality evidence from one RCT (n=198) showed parenteral and enteral nutrition may be more effective on the number of people experiencing bloating and constipation, although there was some uncertainty.

6.8.4.17 Parenteral and enteral nutrition versus enteral nutrition in young people with anorexia nervosa at follow up

Very low quality evidence from one RCT (n=129) showed no difference in the effect of parenteral and enteral nutrition on number of people recovering from anorexia nervosa and number of people who were rehospitalised compared with enteral nutrition.

Very low quality evidence from one RCT (n=129) showed parenteral and enteral nutrition is less effective on length of second rehospitalisation compared with enteral nutrition.

6.8.4.18 Percutaneous gastric tube feeding and meals versus meals with or without nasogastric tube feeding for underweight adults with anorexia nervosa

Very low quality evidence from one RCT (n=68) showed percutaneous gastric tube feeding with meals is less effective on length of treatment compared with meals with or without nasogastric tube feeding.

Very low quality evidence from one RCT (n=68) showed no difference in the effect of percutaneous gastric tube feeding with meals on weight gain compared with meals with or without nasogastric tube feeding.

6.8.4.19 Nasogastric tube and oral refeeding diet versus oral refeeding diet in malnourished young people with anorexia nervosa

Very low quality evidence from one RCT (n=100) showed nasogastric tube and oral refeeding diet is more effective on BMI (absolute and change scores), weight gain and daily maximum caloric intake compared with oral refeeding diet.

Very low quality evidence from one RCT (n=100) showed no difference in the effect of nasogastric tube and oral refeeding diet on weight and length of inpatient stay compared with oral refeeding.

6.8.4.20 Nasogastric and oral refeeding diet versus oral refeeding diet in malnourished adults with anorexia nervosa at end of treatment

Low quality evidence from one RCT (n=81) showed that a nasogastric and oral refeeding diet is more effective on weight, daily weight gain, extracellular fluids, fat free mass, daily fat free mass gain and daily fat mass gain compared with an oral refeeding diet alone.

Low quality evidence from one RCT (n=81) showed that a nasogastric and oral refeeding diet is more effective at increasing the number of people achieving a BMI greater than 18.5 kg/m2 and the number of people having a relapse-free period compared with an oral refeeding diet alone.

Low quality evidence from one RCT (n=25) showed that a nasogastric and oral refeeding diet is more effective at increasing weight in binge-purge anorexia nervosa participants compared with an oral refeeding diet alone.

Low quality evidence from one RCT (n=56) showed that a nasogastric and oral refeeding diet is more effective at increasing weight in restricting anorexia nervosa participants compared with an oral refeeding diet alone.

Low quality evidence from one RCT (n=81) showed that a nasogastric and oral refeeding diet is less effective at increasing daily added sucrose levels compared with an oral refeeding diet alone.

Low quality evidence from one RCT (n=25) showed that a nasogastric and oral refeeding diet is less effective at increasing energy intake compared with an oral refeeding diet alone.

Low quality evidence from one RCT (n=81) showed no difference in the effect of a nasogastric and oral refeeding diet in increasing daily added fat compared with an oral refeeding diet alone.

Low quality evidence from one RCT (n=56) showed that a nasogastric and oral refeeding diet may be more effective at increasing daily energy intake in people with restricting anorexia nervosa compared with an oral refeeding diet alone, although there was some uncertainty.

Low quality evidence from one RCT (n=25) showed a nasogastric and oral refeeding diet may be less effective at increasing daily energy intake in people with binge-purge anorexia nervosa compared with an oral refeeding diet alone, although there was some uncertainty.

6.8.4.21 Nasogastric and oral refeeding diet versus oral refeeding diet in malnourished adults with anorexia nervosa at follow up

Low quality evidence from one RCT (n=56) showed nasogastric and oral refeeding diet is more effective on weight for people with the restricting subtype of anorexia nervosa compared with oral refeeding diet alone.

Low quality evidence from one RCT (n=25) showed nasogastric and oral refeeding diet is more effective on weight for people with the binge-purge subtype of anorexia nervosa compared with oral refeeding diet alone.

Very low quality evidence from one RCT (n=81) showed no difference in the effect of nasogastric and oral refeeding diet on the number of people who relapsed compared with oral refeeding alone.

Very low quality evidence from one RCT (n=81) showed no difference in the effect of nasogastric and oral refeeding diet on improving EDI-total scores and decreasing the number of people taking antixiolytics compared with oral refeeding diet alone.

Very low quality evidence from one RCT (n=81) showed nasogastric and oral refeeding diet may be more effective on increasing the number of people achieving both a BMI greater than 18.5 kg/m2 and adequate energy intake compared with oral refeeding diet alone, although there was some uncertainty.

Very low quality evidence from one RCT (n=56) showed no difference in the effect of nasogastric and oral refeeding diet on the daily energy intake of people with the restricting subtype of anorexia nervosa compared with oral refeeding diet alone.

Very low quality evidence from one RCT (n=25) showed no difference in the effect of nasogastric and oral refeeding diet on the daily energy intake of people with the binge-purge subtype of anorexia nervosa compared with oral refeeding diet alone.

Low quality evidence from one RCT (n=26) showed no difference in the effect of nasogastric and oral refeeding diet on the number of people who had resumed menses compared with oral refeeding diet alone.

Very low quality evidence from one RCT (n=81) showed nasogastric and oral refeeding diet may be less effective on the number of people taking antidepressants compared with oral refeeding diet alone, although there was some uncertainty.

6.8.4.22 High-calorie refeeding diet versus low-calorie refeeding diet for malnourished young people with anorexia nervosa

Very low quality evidence from one RCT (n=36) showed high-calorie refeeding diet may be less effective on number of people experiencing hypophosphatemia within the first 2 days of treatment compared with low-calorie refeeding diet, although there was some uncertainty.

Very low quality evidence from one RCT (n=36) showed high-calorie refeeding diet may be more effective on the number of people experiencing adverse events within first 4 days of treatment compared with low-calorie refeeding diet, although there was some uncertainty.

Very low quality evidence from one RCT (n=36) showed no difference in the effect of high-calorie refeeding diet at 4 days on QT-corrected interval (absolute and change scores), change in heart rate, weight, BMI (absolute and change scores), %mBMI (absolute and change scores) and serum phosphate concentration (absolute and change scores) compared with low-calorie refeeding diet.

Very low quality evidence from one RCT (n=36) showed high-calorie refeeding diet at 4 days may be more effective on heart rate and change in weight compared with low-calorie refeeding diet, although there was some uncertainty.

Very low quality evidence from 1 RCT (n=36) showed high-calorie diet is more effective at 4 days on kcal/day and kcl/g energy intake (i.e. had higher energy intake) compared with low-calorie refeeding diet.

Very low quality evidence from one RCT (n=36) showed no difference in the effect of high-calorie refeeding diet at 10 days on weight (absolute and change scores), BMI (absolute and change scores), %mBMI, glucose level, insulin level, white blood cell count and number of people free of oral phosphate supplementation compared with low-calorie refeeding diet.

Very low quality evidence from one RCT (n=36) showed high-calorie refeeding diet is more effective at 10 days on kcal/day and kcal/g energy intake compared with low-calorie refeeding diet.

Very low quality evidence from one RCT (n=36) showed high-calorie refeeding diet was more effective at 10 days on change in %MBMI and insulin resistance compared with low-calorie refeeding diet, although there was some uncertainty.

6.8.4.23 Normal-sodium nasogastric and oral refeeding diet versus low-sodium diet in adults with anorexia nervosa

Very low quality evidence from one RCT (n=218) showed a normal-sodium refeeding diet may be more effective on weight compared with low-sodium refeeding diet, although there was some uncertainty.

Very low quality evidence from one RCT (n=218) showed a normal-sodium refeeding diet may be less effective active fat free mass and the number of people experiencing edema of the legs compared with low-sodium refeeding diet, although there was some uncertainty.

Very low quality evidence from one RCT (n=218) showed no difference in the effect of normal-sodium refeeding diet on BMI, daily energy input, daily energy input from tube feeding and fat mass-BIA compared with low-sodium refeeding diet.

Very low quality evidence from one RCT (n=218) showed normal-sodium refeeding diet is more effective on fat free mass-skinfold compared with low-sodium refeeding diet.

Very low quality evidence from one RCT (n=218) showed normal-sodium refeeding diet is less effective on fat mass-skinfold compared with low-sodium refeeding diet.

6.8.4.24 Oral potassium supplementation versus no supplementation for cardiac dysfunction in female adult anorexia nervosa

Very low quality evidence from one RCT (n=28) showed oral potassium supplementation is more effective in reducing QT dispersion (corrected and uncorrected), serum potassium and irunary potassium excretion compared with no supplementation at all.

6.8.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the management of short and long-term physical complications of anorexia nervosa was available.

6.8.6 Recommendations and link to evidence for the review on: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

Low bone mineral density in women with anorexia nervosa

	77. Explain to women with anorexia nervosa that the primary aim of prevention and treatment of a low bone mineral density is to

	achieve and maintain a healthy body weight or BMI for their age. **78. Do not routinely offer oral or transdermal oestrogen therapy to treat low bone mineral density in children or young people with anorexia nervosa.** **79. Seek specialist paediatric or endocrinological advice before starting any hormonal treatment for a low bone mineral density. Coordinate any treatment with the eating disorders team.** **80. Consider transdermal 17-β-estradiol (with cyclic progesterone) for young women (aged 13-17 years) with anorexia nervosa who have long-term low body weight and low bone mineral density with a bone age over 15.** **81. Consider incremental physiological doses of oestrogen in young women (aged 13-17 years) with anorexia nervosa who have delayed puberty, long-term low body weight and low bone mineral density with a bone age under 15.** **82. Consider bisphosphonates for women (18 years and over) with anorexia nervosa who have long-term low body weight and low bone mineral density. Discuss the benefits and risks (including risk of teratogenic effects) with women before starting treatment.** **83. Advise people with anorexia nervosa and osteoporosis or related bone disorders to avoid high-impact physical activities and activities that significantly increase the chance of falls or fractures.** **84. Offer a bone mineral density scan:** • after 6 months of amenorrhea in young women (aged 13 to 17) and yearly after this even if the person gains weight • after 12 months of amenorrhea in adult women (18 and above) and every 2 years after this even if the person gains weight. **Continue to offer scans until either menses has resumed or bone mineral density is within healthy limits.** **85. Monitor growth and development in children and young people with anorexia nervosa who have not completed puberty (for example, not reached menarche or final height).** **86. For guidance on osteoporosis risk assessment, see the NICE guideline on assessing the risk of fragility fractures in osteoporosis.**
Relative value of different outcomes	For the review on how to treat the short and long-term physical health conditions associated with eating disorders, the committee agreed that the critical outcomes will depend on the health condition under review. For treating low bone mineral density (BMD), the committee agreed that the critical outcomes are BMD and bone

	strength. Other outcomes that are important include quality of life, weight or BMI, dropouts due to side-effects, resumption of menses, remission and service user experience.
Trade-off between clinical benefits and harms	**Girls and young women with anorexia nervosa** One study aimed to reduce bone loss in young people with anorexia nervosa by treating them for 12 months with either didehydroepiandrosterone (DHEA) or HRT. At the end of treatment, change in BMD at the lumbar spine was lower and there were fewer girls who had resumed menses in the DHEA-treated group compared with HRT. However, there was some uncertainty in the results. Other outcomes including change in total hip BMD, weight and dropouts due to side-effects were not different between the two treatment arms at the end of treatment. No data was reported on quality of life, remission or service user experience. One study compared the effects of a bisphosphonate with placebo in young girls with anorexia nervosa and a low BMD at baseline. At the end of 12 months of treatment there was no difference in the change in BMD Z score at the spine and femoral neck nor in final BMD at the wards triangle and total hip. However, favourable changes in BMD were found in response to bisphosphonates at the trochanter compared with placebo. No harms were detected. No data was reported on body weight, remission, quality of life or service user experience. Another study investigated the effects of oestrogen versus placebo in young girls with anorexia nervosa who had low BMD compared with controls at baseline. The results showed no difference in the change in lean mass or fat mass after 18 months of treatment. However, change in weight and change in BMD at the lumbar spine and total hip was greater in the oestrogen-treated group compared with placebo. The change in femoral neck BMD and dropouts due to side-effects were similar and the number who did not achieve normal menses was lower in the oestrogen-treated group. No data was reported on remission, quality of life or service user experience. **Adults with anorexia nervosa** In adults, one study compared the effects of DHEA and oral contraceptive pill with placebo in preserving bone mass in women with anorexia nervosa. The results showed after 18 months of treatment, change in weight and BMI and change in femoral shaft and femoral neck BMD were greater in the combined treatment arm compared with placebo. Change in femoral shaft bone strength index was also higher in the combined treatment arm but not change in femoral neck bone strength index. Other outcomes were no different between the combined treatment and placebo including change in the number who had amenorrhea or who dropped out due to side-effects of treatment. No data was reported on quality of life, remission or service user experience. Six months of bisphosphonate treatment in women with anorexia nervosa with a low BMD T-score resulted in greater improvements in hip, lateral spine and anteroposterior spine BMD compared with placebo, but no benefit at the tibia. No side-effects were detected. No data was reported on body weight, quality of life, remission or service user experience. One study compared bisphosphonates with calcium and vitamin D and showed no difference after three months on tibia ultrasound density score or tibia Z-score. No side-effects were detected. No data was reported on body weight, quality of life, remission or service user experience. IGF alone or combined with oral contraceptive pill was also investigated in women with anorexia nervosa and osteopenia (T-score 1.0 or less) at the spine. Two studies were conducted by the same research group, one for three months and the other for nine months. Comparing IGF-I with placebo showed change in BMD at the spine was greater in the IGF-I-treated group, along with weight, lean mass and BMI. No difference in BMD was found at the hip, total body and radius. No drop-outs due to side-effects were reported. No data was reported on quality of life, remission or service user experience. Comparing IGF-I with oral contraceptive group showed greater improvements in the IGF-I-treated group at the hip and total body, in addition to change in lean

	mass and BMI. No difference in BMD was found at the radius and anteroposterior spine or weight. No drop-outs due to side-effects were reported. No data was reported on quality of life, remission or service user experience. Combining IGF-I with oral contraceptive pill showed it was advantageous on spine, radial and total body BMD and lean mass, but the change in BMI was less compared with placebo. All other outcomes were no different between the groups including hip BMD and change in weight. No drop-outs due to side-effects were reported. No data was reported on quality of life, remission or service user experience. Again combining IGF-I with oral contraceptive pill but comparing it to an active arm of oral contraceptive pill, showed that it was still beneficial on change in hip BMD, total body BMD and lean mass. No difference in change in radial BMD or spine BMD was detected. However, there was a reduced gain in body weight in the combined treatment group. No drop-outs due to side-effects were reported. No data was reported on quality of life, remission or service user experience. Comparing the combined treatment of IGF-I and oral contraceptive pill but with IGF-I showed that combined treatment may improve total body BMD, radial BMD and spine BMD but has no significant effect on hip BMD or lean mass. Again, a reduced gain in body weight and BMI was found in the combined treated group. No drop-outs due to side-effects were reported. No data was reported on quality of life, remission or service user experience. Parathyroid treatment showed mixed results compared with placebo in women with anorexia nervosa with low BMD. After six months, change in lateral spine and anteroposterior spine BMD was greater in the parathyroid-treated group, but change in femoral neck BMD and body weight was less. No difference was detected in total hip BMD. No drop-outs due to side-effects were reported. No data was reported on quality of life, remission or service user experience. Finally, two studies compared with effects of oestrogen treatment with placebo in adult women with anorexia nervosa who had low BMD. After nine to 18 months of treatment, the results showed no difference in lumbar spine BMD, weight, BMI or lean mass. Total BMD and hip BMD showed smaller changes over time compared with the placebo group. No difference in remission was detected. No data was reported on quality of life, remission, adverse events or service user experience.
Trade-off between net health benefits and resource use	No economic studies assessing the cost effectiveness of interventions aimed at managing short and long-term physical complications of eating disorders were identified. The committee considered the low bodyweight as a clinical risk factor for low BMD. The committee also considered the consequences of low BMD including osteoporosis and the associated increased risk of incident fractures, and the high cost of managing these (including the expensive hospital care) to the healthcare system. The committee expressed the view that timely and appropriate treatment, may prevent the need of expensive secondary care, and lead to an overall cost-savings to the healthcare system.
Quality of evidence	The majority of the evidence was low to very low quality. Outcomes were downgraded because it was unclear in the studies how the randomisation sequence was generated and if allocation concealment was performed. Across studies it was unclear at times if either the participants, investigators or assessors were blind. High dropouts of >20% were also detected. There was very little evidence for most comparisons, resulting in few studies and few participants in the meta-analysis and imprecision was often detected. Despite the evidence to support the use of pharmacological treatments in people with anorexia nervosa and a low BMD, the committee agreed that primary prevention and treatment of a low BMD is to achieve and maintain a healthy body weight. For this reason, it should not be routine treatment to offer oral or transdermal oestrogen therapy to treat low BMD in children or young people with anorexia nervosa. Restoring and maintaining a healthy body weight will help return oestrogen levels to normal and (not included in review) any delayed bone growth may be partially restored and improve BMD. The committee were also concerned that oestrogen therapy may result in premature closure of growth plates Nevertheless, the committee acknowledged there may be circumstances where

	children and young people will need hormonal treatment, for example to induce puberty or if recovery is not sufficient to normalise bone accrual. In such cases it is important that specialist paediatric or endocrinologist advice is sought before starting any treatment. The best available evidence in girls and young women with long-term low body weight and low BMD supported the use of either transdermal 17ß-estradoil (with cyclic progesterone) or physiological doses of oestrogen in girls depending on their bone age. Transdermal oestrogen is said to be less likely to suppress IGF-I. No long-term follow up data was available so it was difficult to know if the benefits were maintained. Nevertheless, the committee were confident making a consider recommendation based on this study. A limitation with this study on oestrogen treatment is that the changes in BMD were not adjusted for age or weight, only in the statistical analysis did the authors take this into account. However, given that body weight did not appear to differ between the oestrogen treated and placebo groups, it may not have a significant impact on the interpretation of the data. BMD needs to be adjusted for body weight since heavier individuals have higher BMD than individuals of lower body weight. Also, BMD should also ideally be adjusted for bone size where possible. Although BMD is a measure of bone mineral content per cm2, it does not take into account the 3-dimensional, volumetric density (grams per cubic centimetre) of the bone. Thus, for girls with thicker bones, it may suggest they have high BMD but their bone mineral apparent density may be normal. This adjustment is particularly important during growth and only one bone mineral apparent density outcome for lumbar spine was provided for young women. For adult women, the committee agreed that the most convincing data was on bisphosphonates. However, like for children and young girls with anorexia nervosa, they the first line treatment should aim to restore and maintain a healthy body weight. However, for women who have long-term low body weight and low BMD, bisphosphonates could be considered. The benefits and risks would need to be discussed given possible teratogenic effects. The evidence used to generate the bisphosphonate recommendation was low quality, with only one study for each outcome and at most 39 participants. In the study by Miller 2011, there were two additional groups, one that received testosterone and the other combined testosterone with the bisphosphonate. However, the data was only presented in graph format and could not be extracted. Despite the little data available, the committee were confident recommending bisphosphonates since NICE recommend alendronate or risedronate for people without eating disorders but at high risk of osteoporotic fracture. The lack of follow up data was one of the reasons the committee were reluctant to recommend IGF-I, despite the finding that IGF-I increased bone formation compared with those who did not receive IGF-I. Additionally, IGF-I with an oestrogen-progesterone combination pill significantly increased BMD compared with placebo, a finding that was not seen with administration of oestrogen or IGF-I alone. Because of the very low number of participants in each treatment arm (n=15) the committee agreed that more data is needed before they would consider it. They also pointed out that IGF-I is not currently recommended for treating people with osteoporosis. A critical outcome that was rarely reported was bone strength. Just like volumetric bone density, the distribution of bone or bone shape is a better estimate of bone strength compared with bone density alone. There are a number of ways of calculating bone strength and one study estimated this on the femur of adults treated with DHEA and oestrogen or placebo. The results showed the femoral shaft was stronger in the DHEA and oestrogen treated group but not the femoral neck.
Other considerations	Other limitations with this review is that no long-term follow up data was available. So it is not known if the gains in BMD are sustained years after the treatment has finished. Nor it is clear what duration of treatment is needed for BMD to be restored to normal levels, or whether the changes in BMD translate to a reduction in the future risk of fracture.

	The committee noted that no studies were available in young or adult men. The committee it was important that a collaborative care approach Is used if pharmacological agents are being prescribed for people with anorexia nervosa. Given that weight gain is the primary concern and this is not likely to change in response to oestrogen or bisphosphonate treatment, it is important that the specialists treating low BMD communicate with the eating disorder's team. Given the risk of fracture for people with an eating disorder and osteoporosis, the committee agreed it is important to advise them to avoid high-impact physical activities, such as skipping, or activities that increase the risk of falls, such as contact sports. To ensure a low BMD is detected, the committee agreed based on current good practice that a bone density scan should be offered after 6 or 12 months of amenorrhea in young women or adult women respectively. And in young women yearly thereafter and every two years thereafter in adult women even if they regain weight. For guidance on osteoporosis risk assessment, the committee recommended referring to the NICE guideline on assessing the risk of fragility fractures in osteoporosis. The committee also highlighted when estimating the risk of fracture for children and young people it is important that healthcare professionals use Z-scores not T-scores. In addition, that BMD is corrected for bone size.

1 **Growth and development**

	87. Seek specialist paediatric or endocrinology advice for delayed physical development or stunted growth in children and young people with an eating disorder.
Relative value of different outcomes	For the review on how to manage, treat or reduce the short and long-term physical health conditions associated with eating disorders, the committee agreed that the critical outcomes will depend on the health condition under review. For treating delayed physical development or stunted growth, the committee agreed that the critical outcome is growth. Other outcomes that are important include quality of life, weight or BMI, compensatory behaviours, side-effects, resumption of menses, remission and service user experience.
Trade-off between clinical benefits and harms	Only one study was identified that compared the effects of growth hormone versus placebo in adults with anorexia nervosa who had low recombinant human insulin-like growth (IGF-I) factor-I. At the end of treatment, IGF-I levels were higher in the IGH-I treated group, as was the change in IGF-I levels but there was some uncertainty. Change in body weight was not significantly different between the two treatment arms. No data on quality of life, compensatory behaviours, side-effects, resumption of menses, remission or service user experience. No relevant published evidence was identified on how to managing or reduce delayed physical development or stunted growth in people with an eating disorder.
Trade-off between net health benefits and resource use	The committee considered that providing specialist paediatric or endocrinology advice for delayed physical development or stunted growth may have resource implications in terms of the extra time required to facilitate such advice. However, the committee expressed the view that this could lead to better identification of health needs and result in appropriate subsequent treatment and management of underlying health problems at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating such specialist care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	Only one study was identified that could be considered relevant but they did not report on any growth related outcomes, such as sexual development or height. The investigators only observed whether they could significantly change IGF-I levels and body weight in a population who had low background levels. The outcomes were graded low to very low because of unclear methods of

	randomisation and it was unclear if allocation concealment was performed. It was also unclear either the participants, investigators or assessors were blinded. In the absence of evidence on how to manage or reduce delayed development or stunted growth, the committee agreed by informal consensus that it was best to seek advice specialist paediatrician or endocrinologist if identified in children and young children with an eating disorder.
Other considerations	The committee discussed how delayed physical development or stunted growth is an indicator that the child or young person may have nutritional deficiencies as a result of anorexia nervosa. The effects of starvation are extensive and negatively affect the pituitary gland, thyroid gland, adrenal glands, gonads and bones. Thus, short stature, osteoporosis and infertility are indicators of an eating disorder and even if the person recovers, there may be long-lasting complications. It is thought that with weight gain, growth can 'catch-up' at least until fusion of the epiphysis occurs. Although hormonal treatments are used to treat delayed physical development or stunted growth, there is no RCT evidence on this in young people with an eating disorder and the committee were concerned about the side-effects. For example, there is a risk that oestrogen therapy may cause premature epiphyseal fusion and growth stunting. Given the complexities of hormonal treatment, the committee agreed it was best to seek specialist advice in such cases. Because of the concerns surrounding delayed physical development and stunted growth the committee agreed that growth and development should be monitored in children and young people with an eating disorder who have not completed puberty (for example, not reached menarche or final height). The committee generated a recommendation to "monitor growth and development in children and young people with anorexia nervosa who have not completed puberty [for example, not reached menarche or final height]). Details on this can be found in the LETR on low bone mineral density.

1 **Refeeding**

	88. Ensure that staff of inpatient services for people with eating disorders are trained to recognise the symptoms of refeeding syndrome and how to manage it. **89. Use a standard operating procedure for refeeding that emphasises the need to avoid under-nutrition and refeeding syndrome. Refer to existing national guidance, for example Management of Really Sick Patients with Anorexia Nervosa (MARSIPAN) and junior MARSIPAN.**
Relative value of different outcomes	For the review on how to address the short-term complications of eating disorders the committee considered the critical outcomes to be the primary outcomes reported by the study and the important outcomes to be secondary outcomes reported by the study. This was because the physical complications could cover a large number of conditions.
Trade-off between clinical benefits and harms	A randomised control trial that compared nasogastric and oral refeeding with oral refeeding in adults with anorexia nervosa showed favourable outcomes for nasogastric feeding on weight and BMI, extracellular fluid scores, fat mass and prolonging recovery. However, it was less effective on increasing energy intake (for binge-purging anorexia nervosa for but not restrictive anorexia nervosa) and sucrose intake and no difference was found in fat intake. At 12 months' follow up benefits were still found in the nasogastric group on weight, but no difference in energy intake, relapse rates, resumption of menses, use antixiolytics compared with the oral refeeding group. However, the use of nasogastric and oral refeeding favoured the use of antidepressants (though there was some uncertainty). One study compared four days of high versus low calorie refeeding diets for young people with anorexia nervosa. The results showed no difference in weight or BMI,

	heart function, adverse events, need for oral phosphate supplementation (due to low levels), serum phosphate concentrations but it did favour high calorie intake for a higher energy intake (thus the intervention was working). After 10 days of refeeding, a high calorie refeeding diet still favoured energy intake and there were no other differences on any other outcome. **Observational studies** Observational studies were also identified that compared nasogastric tube and oral refeeding with oral refeeding in young females with anorexia nervosa. The results favoured the combined treatment with a nasogastric tube. Better outcomes were also found for BMI and weight, calorie intake but no difference in hospital stay. Another observational study compared a high versus low sodium nasogastric and oral refeeding intervention. The results showed in adults with anorexia nervosa no difference in weight or BMI, some improvement in fat free mass, but not fat mass. There appeared to be some improvement in energy input, but a side-effect of oedema of the legs was detected in the high sodium group compared with the low sodium-treated group. Percutaneous gastric feeding compared with nasogastric tube refeeding for adults with anorexia nervosa showed no difference in weight gain but length of hospital stay was longer in the percutaneous gastric feeding group. Parenteral refeeding in young people showed mixed results when compared with enteral refeeding. It showed unclear trends on weight versus BMI. Some benefit was found on energy intake, reduction in some side-effects (bloating, abdominal pain) but hospital stay was longer. At 33 months' follow up, there was no difference between the two treatment groups for rehospitalisation and recovery but if they were readmitted, the parenteral refeeding group had a longer hospital stay compared with enteral refeeding group (it was unclear what subsequent treatment they had). One study investigated the effects of oral potassium supplementation for cardiac dysfunction versus no treatment in young women with anorexia nervosa and found a benefit on cardiac function, serum potassium levels and kidney function.
Trade-off between net health benefits and resource use	The committee considered that inpatient services for people with eating disorders must already be trained to recognise the symptoms of refeeding syndrome and how to manage it. So offering it in line with the principles outlined in the recommendation 86-87 would not incur significant extra resource implications.
Quality of evidence	The quality of the RCT evidence was low to very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as unclear randomisation, it was unclear if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded and high dropouts were detected >20%. Because of the little evidence found, only one study with a small sample size was available for each comparison, imprecision was found for most outcomes. The outcomes from observational studies were all very low quality. In GRADE, all evidence from observational studies begin at very low quality because of the risk of bias in selecting the people for each cohort and can only be upgraded if large effect sizes are found (after adjustment) and a dose response is identified (this could not be done in this review). Very few of the studies adjusted for any confounders.
Other considerations	The committee considered the refeeding studies and agreed that parenteral nutrition is not something that would be used in the UK for general refeeding. The RCT evidence did not show any evidence that would lead the committee to make a specific recommendation. Also, the evidence from observational studies was all very low quality. However, it did show some favourable results for nasogastric refeeding compared with oral refeeding. Given the lack of evidence, the committee agreed that it was best that healthcare professionals use existing local standard operating procedures (MARSIPAN) or local protocols for refeeding and that they do not need to be modified in the light of the (little) evidence presented in the review. But they did agree they should include initial meal plans, guidance on monitoring of physical risk, the use of electrolyte,

vitamin and mineral supplements and when and how to use nasogastric feeding. The committee said specialist eating disorder units and general hospital wards undertaking refeeding should have standard operating procedures or local protocols.

The committee all agreed that MARSIPAN and junior MARSIPAN are a well-established national guidelines that should be referred to when offering a refeeding treatment. These guidelines were mostly generated from the knowledge and expertise of health care professionals who work in the field. They identified little evidence from their search and had a different criteria to our review strategy. However, the adult MARSIPAN guideline referred to one RCT that was also included in this review (O'Connor 2016) and the others were mostly case-reports and reviews that would not meet our inclusion criteria. In the junior MARSIPAN report, no RCTs were found, only reports on the incidence of refeeding syndrome using different refeeding programmes. Again, these papers would not meet the inclusion criteria. Nevertheless, they were included in the MARSIPAN guideline and the committee were confident that they were up to date and should be referred to in the eating disorder NICE guideline.

The committee agreed that it was important that healthcare professionals are aware of the risk of refeeding syndrome, how to recognise and treat it. Refeeding syndrome usually occurs within four days of starting to re-feed and can result in people developing fluid and electrolyte disorders, especially hypophosphatemia, along with neurologic, pulmonary, cardiac, neuromuscular, and hematologic complications. This is generally the result a shift of electrolyte and fluid balance from when the person was in a state of critical health due to negligible nutrient intake to a rise in blood sugar levels during refeeding. This shift in balance can cause an increase in cardiac workload and heat rate, potentially leading to acute heart failure.

Because of this risk, close monitoring of electrolyte disturbances and blood biochemistry is needed in the early refeeding period. If an imbalance is detected, the refeeding should be modified and treatment provided. Again, the committee highlighted that healthcare professionals should refer to the local protocols and MARSIPAN and junior MARSIPAN for further information.

The committee discussed how the risk of refeeding syndrome risk may be increased if there is rapid weight loss, no calorie intake for over five days or BMI <16kg/m^2. Compensatory behaviours such as laxative misuse, vomiting, dehydration, diet pills, diuretics, water loading or excessive exercise will substantially may also increase the risk of refeeding syndrome.

The committee highlighted how more studies are needed to better understand refeeding syndrome and what prophylactic supplements should be offered to support refeeding. It is covered in the MARSIPAN guideline that thiamine and phosphate could be offered, in conjunction with close monitoring and correction of phosphate, magnesium, calcium and potassium, body weight and glucose during the first 10-14 days of refeeding. There was some evidence from this review that a low sodium diet may be useful to prevent refeeding oedema (Rigaud, 2010).

The committee discussed how energy requirements for weight restoration/growth and maintenance of body weight vary greatly between individuals, so an individualised approach to diet and weight restoration should be implemented using the least restrictive feeding option where it is safe to do so; offering oral diet and sip feeding before enteral feeding. This is also covered in the MARSIPAN guidelines.

Because of safety concerns, the committee talked about that when a nasogastric tube is used, it is important to verify its placement prior to refeeding by a member of the medical unit. Again this is mentioned to in the MARSIPAN guidelines.

For a weight gain of 0.5-1kg a week some individuals may need in excess of 1000 extra calories a day. The committee said this can be obtained through extra snacks, increased portions or the use of sip feeds/supplementary feeds where higher calorie needs are struggling to be met. Gastrointestinal function and patient preference should be considered. Weight should be recorded 1 to 2 x a week unless clinically indicated more frequently. This is also covered in the MARSIPAN guidelines.

6.9 Management of comorbidities

6.9.1 Review question: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 166. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers whether any intervention used to treat eating disorders in children, young people and adults needs to be modified in the presence of a common long-term health condition (i.e. comorbidity).The interventions were categorised according to their type, the type of eating disorder and comorbidity examined and the age of the participants. The comparison arm was the same intervention delivered to participants with the relevant eating disorder but without the relevant comorbidity.

Table 166: Clinical review protocol summary for the review of: does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Component	Description
Review question(s)	Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) and a common comorbidity (e.g. diabetes, hypothyroidism). • Mental comorbidities may include: • Depression • Anxiety • Social anxiety • Autism • Obsessive Compulsive Disorder • Personality Disorder • Learning disability • ADHD (Bulimia) • Self-harm • Substance misuse • Physical comorbidities may include: • Celiac disease • Diabetes (type II – relevant to obesity) • Irritable Bowel Disease • Cystic Fibrosis • **Strata:** • children (≤12), young people (13-≤17 years), adults ≥18 years • eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	Trials will be included that address the ED as primary or

Component	Description
	secondary aim to treating the comorbidity. Interventions may include: Psychotherapy (including psychoeducation) Pharmacological Nutritional Physical Combination of any listed above
Comparison	• The same intervention but delivered to people with an eating disorder without a comorbidity.
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs) (if no RCTs)

6.9.2 Clinical Evidence for: does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

No published studies were found for this review question in people with anorexia nervosa.

Although this review question includes people with any eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder, EDNOS), the committee wanted to firstly consider the evidence for individual eating disorders to see if specific recommendations could be made. If none was available, or it was deemed insufficient, then they agreed to make a general recommendation for treating people with any eating disorder and a common long-term health condition.

6.9.3 Economic Evidence

No economic evidence on the cost effectiveness of modified interventions for anorexia nervosa in the presence of common long-term conditions was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

6.9.4 Clinical evidence statements

No published studies were found that met the eligibility criteria for this review in people with anorexia nervosa. However, evidence was extrapolated from other eating disorders to make a general recommendation for any eating disorder where a comorbidity is present.

6.9.5 Economic Evidence statements

No economic evidence on the cost effectiveness of modified interventions for anorexia nervosa in the presence of common long-term conditions was available.

6.9.6 Recommendations and link to evidence for the review on: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Substance and medication misuse for any eating disorder

	90. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder. **91. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.** **Substance and medication misuse** **92. For people with an eating disorder who are misusing substances, or over the counter or prescribed medication, provide treatment for the eating disorder unless the substance misuse is interfering with this treatment.** **93. If substance misuse or medication is interfering with treatment, consider a multi-disciplinary approach with substance misuse services.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of treating people with an eating disorder and a comorbidity. For binge eating disorder and bulimia nervosa, it was agreed binge eating frequency and remission are of greatest concern. For anorexia nervosa, body weight/BMI and remission are critical and for EDNOS, remission and either binge eating or body weight/BMI depending on the eating disorder they most closely resemble. The other outcomes that are critical are the primary outcomes that are relevant to the physical or mental health comorbidity being treated. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser

	importance but clearly important outcomes include, general psychopathology, body weight, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	**Anorexia nervosa** No relevant published RCT or observational evidence was found on people with anorexia nervosa and a comorbid condition. Thus, data was extrapolated from those with bulimia nervosa and EDNOS to make a recommendation relevant for people with any eating disorder. **Bulimia nervosa and ENDOS (as reviewed in chapter 9)** An observational study was identified where they extracted data from a randomised control trial and compared the outcomes in those with bulimia nervosa and EDNOS who had a low or high alcohol intake. The participants were treated with either broad or focused CBT-ED. At the end of 20 weeks of treatment, there was no difference in the number who had EDE scores one standard deviation above the community norms (i.e., relatively abnormal eating psychopathology) in those with a low or high alcohol intake. However, the number who continued to have excessive alcohol intake (defined as >21 units or >14 units/week for males and females respectively) was higher in those whose alcohol intake was high compared with those whose intake was low. At 60 weeks of follow up, there continued to be no difference in EDE scores between those who had low versus high alcohol intake. No evidence was found on the critical outcomes of remission and binge eating, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. **Bulimia nervosa (as reviewed in chapter 7)** Another observational study compared the long-term outcomes (2 to 5 years) of women with bulimia nervosa who had a history of substance abuse with those who no history of substance abuse. Both groups had received outpatient group cognitive behavioural psychotherapy and showed no different in long-term remission rates or being hospitalised for substance abuse. No evidence was found on the critical outcome of binge eating, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. No evidence from a randomised control trial was found.
Trade-off between net health benefits and resource use	The committee considered that providing care for people with an eating disorder who are misusing substances or medication may have resource implications in terms of the extra time required to facilitate care for such people (in particular the use of a multi-disciplinary approach). However, the committee expressed the view that if such care leads to better identification of health needs and this results in appropriate subsequent treatment and management of health problems (including eating disorder and substance and medication misuse) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating a multi-disciplinary care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence used to generate these recommendations was very low quality. The evidence was observational, therefore in the GRADE software used to assess quality the evidence starts at very low quality and can only be upgraded if large effect sizes are found or if a dose response is identified. Neither was the case for this review. In the absence of RCTs the committee agreed that the observational evidence was useful. In one study a reasonable number of participants were included (n=119) and they underwent a currently recommended CBT-ED programme. However, there were few outcomes reported and no remission was not reported. In another study, again there was a reasonable number of participants included (n=81), but there was no data at the end of treatment (only at follow up) and again few outcomes were reported. They did however measure remission.
Other	Limited published evidence was found on individual eating disorders, so the

considerations	committee generated a recommendation incorporating the evidence from people with BN and EDNOS and made it relevant for treating people with any eating disorder and a common long-term health condition. The observational evidence suggested that people with BN or EDNOS, with a low or high alcohol intake, may be equally responsive to an eating disorder treatment. And for people with BN alone, a positive long-term response to treatment may be equally found in those with a history of substance misuse as those with no history. Thus, the committee recommended that for people with an eating disorder who are misusing substances, offer treatment for the eating disorder unless the substance misuse is demonstrably interfering with this treatment. If substance misuse is interfering with treatment, the committee recommended considering a multidisciplinary approach. It was discussed in the committee meeting that comorbid alcoholism has been associated with an increased risk of mortality.

1 **Diabetes**

94. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.

95. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Diabetes

96. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.

97. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.

98. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes.

99. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.

100. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.

101. Offer people with an eating disorder who are misusing insulin the following treatment plan:

- a low carbohydrate diet, so that insulin can be started at a low level
- gradually increasing insulin doses to reduce blood glucose levels
- adjusted total glycaemic load and carbohydrate

	distribution to meet their individual needs and prevent rapid weight gain • carbohydrate counting when adjusting their insulin dose (including via pumps) • a diabetic educational intervention such as DAFNE • education about the problems caused by misuse of diabetes medication. **102. For more guidance on managing diabetes, refer to the NICE guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem. In the case of diabetes, HbA1c levels and insulin omission days were considered critical outcomes. The other critical outcomes depended on the eating disorder included in the study. Remission is of greatest concern for any eating disorder. In addition, for those with anorexia nervosa body weight or BMI are of greatest concern. For bulimia nervosa and binge eating disorder, binge eating is a critical outcome. For any eating disorder, other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with an eating disorder that are of lesser importance but are clearly still important outcomes include general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	The ideal study design to answer the question of whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem would be to randomise people with an eating disorder and diabetes to two different treatment groups: One modified to address both the eating disorder and diabetes and one non-modified eating disorder treatment. Anorexia nervosa (as reviewed in chapter 6) No published evidence was found on people with anorexia nervosa and diabetes, however there was a sub-group analysis from a study described above on any eating disorder that showed those with anorexia nervosa and type 1 diabetes are equally responsive to treatment as those with anorexia nervosa alone. No data was available on HbA1c scores, remission, weight, all-cause mortality, adverse events, quality of life, resource use, relapse, general psychopathology, general functioning, family functioning or service user experience. Bulimia nervosa (as reviewed in chapter 7) One observational study compared the effectiveness of inpatient integrated care with treatment as usual in adults with bulimia nervosa and type I diabetes. The integrated care provided CBT-ED, family based therapy and addressed control of diabetes. Whilst treatment as usual included outpatient counselling sessions on diabetes but not inpatient care or treatment for the eating disorder. This study showed better outcomes for the integrated care including, remission, general psychopathology, depression, EDI-total, volume of the binges, few compensatory behaviours but no difference in insulin omission. No data was available on HbA1c scores, all-cause mortality, quality of life, resource use, relapse, general functioning, family functioning or service user experience. Binge eating disorder (as reviewed in chapter 8) One study randomised adults with type II diabetes and binge eating to either group CBT-ED or a non-prescriptive control therapy (NPT). The results showed no difference in remission or binge frequency at the end of treatment. BMI showed a trend to be higher in the group CBT-ED arm, however EDI-bulimia, EDI-drive for

	thinness, EDI-body dissatisfaction and quality of life were no different. At follow up, remission rates were higher in the CBT-ED arm, but again no difference in any of the other outcomes and BMI showed a trend to be higher in the group CBT-ED arm compared with controls. No data was available on HbA1c scores, all-cause mortality, resource use, relapse, general functioning, family functioning, general psychopathology or service user experience. An observational study compared the same diabetes prevention programme but in two populations, one with bulimia nervosa and a major depressive disorder and one with just any eating disorder. The results showed no difference in the degree of weight loss between the two populations. No data was available on HbA1c scores, remission, bingeing, all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience. Any eating disorder (as reviewed in chapter 9) One randomised control trial compared group psychoeducation (combined with treatment as usual) with treatment as usual (diabetes treatment only) in people with type I diabetes and disturbed eating behaviours and showed no difference at the end of treatment on bingeing, EDE-restraint, EDE-shape concern, EDE-eating concern, EDE-weight concern, EDI-drive for thinness, EDI-bulimia, insulin omission days and HbA1c (%). One outcome, EDI-body dissatisfaction, favoured group psychoeducation over treatment as usual but there was some uncertainty. At follow up some benefit was found in response to group psychoeducation on bingeing but there was some uncertainty. No data was available on remission, all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience. An observational study was identified that compared the same CBT-ED intervention but in two populations, one with any eating disorder and type I diabetes, and one with just any eating disorder. Thus, this design allowed us to see whether those with a comorbidity would respond equally well to treatment as those with just an eating disorder. The results showed adults with any eating disorder and a comorbidity are less likely to recover than those with just an eating disorder. No difference was found in dropouts. In adults with anorexia nervosa, binge eating disorder and EDNOS there was no difference in the responsiveness to treatment. No data was available on all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience.
Trade-off between net health benefits and resource use	The committee considered that providing care for eating disorders in the presence of a long-term health problems, such as diabetes, may have resource implications in terms of extra time required to provide collaborative and comprehensive care in line with the principles outlined in the recommendations 90-96 However, the committee expressed the view that if such care arrangements (that is, multidisciplinary approach, involvement of family members and carers, and the use of treatment plans) lead to better and appropriate treatment and management of health problems (including other long-term health problems such as diabetes) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating such care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence was mostly low quality from the RCT studies and very low quality from all of the observational studies. In both types of study designs the sample size was generally small and only one study was available for most outcomes, thus imprecision was often detected due to the 95% confidence interval crossing a minimal important difference or the outcome did not meet the optimal information size. **Bulimia nervosa** In the observational study where they compared inpatient integrated care with treatment as usual, the people were selected from the same recruitment site and showed no difference in their characteristics, except that binge frequency was significantly higher in the inpatient group. The duration of follow up was different for the two groups: 36 months versus 24 months in the inpatient care and treatment as usual groups, respectively. Investigators were not blind to treatment allocation and only 18 participants were included.

	Binge eating disorder In the RCT where they compared group CBT-ED with a control therapy in the same population (people with type I diabetes and binge eating disorder) inadequate randomisation was performed and it was unclear if allocation concealment was carried out. Neither the participant or investigator was blind, nor was it clear if the assessor was blind. It was unclear how many participants completed the intervention. The observational study identified was considered indirect evidence since it was a diabetes prevention programme and the participants had major depressive disorder in addition to binge eating disorder or binge eating disorder alone. The only outcome reported was weight loss. The committee did not consider this study helpful **Any eating disorder (including subgroup analysis on anorexia nervosa)** In the RCT where they compared group psychoeducation and management (and treatment as usual) with treatment as usual (diabetes only programme), it was unclear if allocation concealment was performed. Neither the participant, investigator or assessor were blind and it was unclear how many completed the intervention. The population was also indirect since it included those with disturbed eating. Also the comparison did not show whether a modified eating disorder programme is more effective at treating people with diabetes and an eating disorder compared with an eating disorder programme alone. Rather the study compared a modified diabetes programme with a regular diabetes programme. In the observational study where they compared CBT-ED in people with eating disorder alone or with a comorbidity, the authors attempted to match the groups based on age, marital status, education, catchment area and onset of diagnosis. However, it was unclear whether the two groups were followed up for the same duration and the sample size was very small. **Overall discussion** No RCT or observational study met the criteria of what would have been the ideal study design for this review (as described above). One RCT compared the effectiveness of an intervention that addressed both the eating disorder and diabetes, but the other arm addressed just the diabetes. In another RCT, one intervention was modified but it was compared with a control therapy. In the observational studies, one study compared the same intervention but in those with either an eating disorder and diabetes or just the eating disorder alone. So it only provided insight into whether one group was more responsive to treatment than the other. In the other observational study, inpatient integrated care was compared with treatment as usual, but the treatment as usual only addressed the diabetes not the eating disorder. Thus, it did not provide insight into whether a modified eating disorder treatment was needed for those with a comorbidity.
Other considerations	In conclusion, it was difficult for the committee to draw conclusions from these studies on whether treatment for an eating disorder needs to be modified in the presence of a comorbidity such as diabetes. The committee therefore agreed that it was best to instead provide guidance on how to manage the diabetes. Usually, the committee would refer to the diabetes NICE guideline, but because the diabetes guideline refers to this guideline, the committee needed to recommend what to do in the presence of both. The committee agreed on a series of recommendations based on their experience and knowledge on how to manage the diabetes in the presence of an eating disorder. A number of the recommendations are based on what would be considered good practice. For instance: i) establish who will monitor the physical health, ii) explain to the person that they need to monitor their diabetes during the treatment for the eating disorder, and iii) be aware of the problems caused by misuse of diabetes medication. The committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in the treatment of diabetes. They highlighted that the quality of the family environment has been shown to affect treatment compliance and metabolic control among young people with an eating disorder (Hauser 1990(Hauser et al., 1990)). Family members may also need to care for someone if they hyper or phyo (which is a case for medical emergency),

	so they know how to respond. There is also the possibility that eating disturbances in young girls with diabetes are associated with significantly more family dysfunction than girls with diabetes alone (i.e. 13 to 18 years of age). Specifically, they can receive less support, and have poorer communication and less trust in their relationship with their parents than diabetic girls without eating disturbances. For these reasons, the committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in their treatment. There was some indirect evidence to support the recommendation to address insulin misuse as part of any psychological treatment. One 1 RCT (n=85) showed that a modified group psychoeducation and management programme reduced bingeing episodes at follow up compared with a programme that just addressed the diabetes alone. This study was considered with the reservation that it was indirect because: 1) it did not investigate the effectiveness of a modified eating disorder psychological treatment and 2) the population had a disturbed eating behaviour, not a specific eating disorder diagnosis. Nevertheless, it showed that a psychoeducation and management programme may help reduce eating disorder psychopathology in those who also have diabetes. The committee discussed the problem of a relatively high prevalence of EDNOS In young girls with diabetes. In girls who have body dissatisfaction, diabetes provides a unique but dangerous opportunity to control weight by deliberate insulin omission, which can lead to hyperglycaemia and glycosuria. It is therefore important that insulin misuse is addressed in any psychological intervention. It can be noted that the recommendations relating to diet control were contributed to by the expert opinion of a dietician on the committee, based on their experience of treating those with an eating disorder who misuse insulin. These recommendations are based upon the treatment approach of small, attainable and incremental goals. At the outset of treatment, intensive glucose management is not an appropriate goal. The first goal must be to establish medical safety for the person with diabetes by gradually increasing the doses of insulin and food intake (as described in the recommendation). Given the fear of weight gain in this population, the committee recommended that the diet is amended to prevent rapid weight gain. They also suggested an educational programme called Dose Adjustment for Normal Eating (DAFNE) that provides people with the skills necessary to estimate the carbohydrate in each meal and to inject the right dose of insulin. There was no evidence on how to treat the eating disorder in the presence of any other long-term physical health condition, such as cystic fibrosis, celiac disease, pregnancy or irritable bowel disease. Some eating disorder specialists on the committee highlighted that they would generally refer someone with an eating disorder and diabetes to a diabetologist rather than address the points raised in the recommendations on diabetes. However, the committee agreed that it should be collaborative approach for the healthcare professionals who treat eating disorders and diabetes. Especially for young people who may need to involve family members and carers in therapy sessions to help the person with blood glucose control. Given the lack of direct evidence to address this review question the committee agreed to make a research recommendation to ask: “Do treatments need to be modified for people of all ages with an eating disorder and a comorbidity?”

6. Research recommendation: Do treatments need to be modified for people of all ages with an eating disorder and a comorbidity?

7 Treatment and management of bulimia nervosa

7.1 Introduction

Bulimia nervosa is characterised by a recurrent cycle of dietary restriction, binge eating and purging. Although most people with bulimia nervosa are in the normal weight range, the illness often begins with a period of dieting. However, the extreme nature of the dieting invariably leads to episodes of binge eating that are commonly driven by hunger. During binge episodes an objectively large quantity of food is eaten in a relatively short period of time and this is accompanied by a sense of loss of control over eating. Following a binge and driven by an extreme fear of weight gain, those with bulimia nervosa engage in a range of compensatory behaviours which may include self-induced vomiting, misuse of laxatives or diuretics, excessive exercise, dietary restriction or, in the case of those with diabetes, insulin misuse. Over time a vicious cycle of dietary restriction, bingeing and purging develops. Bulimic behaviours are commonly associated with increased levels of impairment in day-to-day functioning (Mond and Hay, 2007).

The psychological processes of people with bulimia nervosa involve them trying to adhere to a range of strict eating and food-related rules that increase the risk of binge eating. Poor body image is very common and self-worth is appraised almost exclusively on the basis of weight and shape. Bulimia nervosa is also commonly associated with borderline personality disorder traits, including emotional dysregulation and impulsivity (Diaz-Marsa et al., 2000, Wonderlich et al., 2005) and often occurs alongside self-harm (Bulik et al., 2004, Paul et al., 2002), alcohol difficulties and substance misuse (Holderness et al., 1994). It is not uncommon for people to feel extremely guilty or ashamed of their condition and many will live with it for years before seeking help.

The prevalence of bulimia nervosa is 1% in women and 0.1% in men (van Hoeken et al., 2003). Incidence studies suggest that there was an increase in diagnoses in the 1980s and mid-1990s, followed by a decrease in the late 1990s (Currin et al., 2005). Age of onset also appears to be decreasing, with the high risk group shifting from 25-29-year-old females to 15-24-year-old females (Smink et al., 2012). It is unclear whether this reflects earlier detection or earlier age of onset. A formal diagnosis using the DSM-5 classification system requires the occurrence of binge eating and compensatory behaviours for, on average, at least once a week for three months. For diagnosis, self-evaluation must also be significantly influenced by weight and shape and these symptoms should not occur exclusively during an episode of anorexia. Given the potential physical consequences of binge eating and purging, it is essential that treatment and management takes into consideration medical as well as psychiatric risk.

7.2 Psychological interventions

7.2.1 Review question: Does any psychological intervention produce benefits/harms in children, young people or adults with an eating disorder compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in table 167. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all psychological interventions that may be delivered to children, young people and adults with an eating disorder with or without a pharmacological intervention. This chapter focuses on the results relating to those with bulimia nervosa. The interventions were categorised according to their mode of delivery, i.e. individual, group or self-help, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to any other intervention or to wait list controls.

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 167: Review protocol summary

Component	Description
Review question(s)	Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?
Population	Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder.) Strata: children (<12), young people (13-17 years), adults >18 years eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder and atypical eating disorder) mode of delivery (i. individual ii. family iii. group iv. self-help)
Intervention(s)	Psychological intervention including: Dialectical behaviour therapy (DBT) Counselling (Nutritional/Other) Integrative Cognitive-Affective Therapy for Binge Eating (ICAT) Maudsley model for treatment of adults with anorexia nervosa (MANTRA) Cognitive remediation therapy (CRT) Specialist supportive clinical management for anorexia nervosa (SSCM) Behavioural therapy (BT) CBT (General or ED specific) Dynamic (IPT, Psychodynamic General or ED specific) Guided Self Help with therapist guidance Pure self help E-therapies Psychological in combination with any pharmacological intervention.
Comparison	wait list control treatment as usual another other intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) Binge eating for bulimia nervosa and binge eating disorder; and weight/body mass index (Appropriate adjustment for age) for anorexia nervosa
Important outcomes	Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) General psychopathology (including mood/depression/anxiety) General functioning, measured by return to normal activities, or

Component	Description
	by general mental health functioning measures such as Global Assessment of Functioning (GAF) Family functioning Service user experience Quality of life. All-cause mortality Relapse Adverse events Resource use
Study design	Systematic reviews RCTs

7.2.2 Individual psychotherapy versus any other intervention or wait list control

27 RCTs (n=2223) met the eligibility criteria for this review, the majority of which were on adults (Agras (Agras et al., 1989) 1989, Agras 2000 (Agras et al., 2000), Bulik 1998 (Bulik et al., 1998), Chen 2003 (Chen et al., 2003), Cooper 1995 (Cooper and Steere, 1995), Fairburn 1986 (Fairburn, 1986), Fairburn 1991 (Fairburn et al., 1991), Fairburn 1993 (Fairburn et al., 1993), Fairburn 2009 (Fairburn et al., 2009), Fairburn 2015 (Fairburn and Rothwell, 2015), Freeman 1988 (Freeman et al., 1988), Garner 1993 (Garner et al., 1993), Ghaderi 2006 (Ghaderi, 2006), Griffiths 1994 (Griffiths et al., 1994), Hsu 2001, Le Grange 2007 (le Grange et al., 2007), Le Grange 2015 (Le Grange et al., 2015), Mitchell 2008 (Mitchell et al., 2008), Nevonen 2006 (Nevonen and Broberg, 2006), Olmsted 1991 (Olmsted et al., 1991), Poulsen 2014 (Poulsen et al., 2014), Thackway 1993 (Thackwray et al., 1993), Thiels 1988 (Thiels et al., 1998), Thomson-Brenner 2016, Treasure 1994 (Treasure et al., 1994), Wilson 1991 (Wilson et al., 1991) and Wonderlich 2014 (Wonderlich et al., 2014). The two trials by LeGrange were on young people. An overview of the trials included in the meta-analysis can be found in Table 2. Further information about both included and excluded studies can be found in Appendix J.

No studies were identified that compared a combined individual psychotherapy with a pharmacological agent with any other intervention or wait list controls.

The forest plots can be found in Appendix O, full GRADE evidence profiles be found in Appendix N. See also the study selection flow chart in Appendix M, excluded studies in Appendix J.

7.2.3 Group therapy

Ten RCTs (n=645) met the eligibility criteria for this review all of which were on adults (Bailer 2004 (Bailer et al., 2004), Chen 2003 (Chen et al., 2003), Hsu 2001 (Hsu et al., 2001), Lavender 2012 (Lavender et al., 2012), Lee 1986 (Lee and Rush, 1986), Leitenberg 1988 (Leitenberg et al., 1988) 1988, Mitchell 1993 (Mitchell et al., 1993), Nauta 2001 (Nauta et al., 2001), Olmsted 1991 (Olmsted et al., 1991), Wolf 1992 (Wolf and Crowther, 1992)). An overview of the trials included in the meta-analysis can be found in Table 2. Further information about both included and excluded studies can be found in Appendix J.

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

7.2.4 Self-help

Ten RCTs (n=1559) met the eligibility criteria for this review the majority of which were in adults (Bailer 2004 (Bailer et al., 2004),Banasiask 2005 (Banasiak et al., 2005), Bauer 2012 (Bauer et al., 2012),Carter 2003 (Carter et al., 2003), Durand 2003 (Durand and King, 2003),

Ljotsson 2007 (Ljotsson et al., 2007), Mitchell 2008 (Mitchell et al., 2008), Palmer 2002 (Palmer et al., 2002), Ruwaard 2013 (Ruwaard et al., 2013), Sanchez-Ortiz 2011 (Sanchez-Ortiz et al., 2011), Schmidt 2006 (Schmidt, 2006), Schmidt 2007 (Schmidt et al., 2007b), Steele 2008 (Steele and Wade, 2008), Thiels 1998 (Thiels et al., 1998),Treasure 1994 (Treasure et al., 1994),Wagner 2013 (Wagner et al., 2013), Walsh 2004 (Walsh et al., 2004)). The study by Schmidt 2007 was on young people. An overview of the trials included in the meta-analysis can be found in Table 165.

Further information about both included and excluded studies can found in Appendix J.

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

7.2.5 Family therapy versus any individual therapy or wait list control in young people with bulimia nervosa

Three RCTs (n=295) met the eligibility criteria for this review, all of which were for young people in an outpatient setting (Le Grange 2007 (le Grange et al., 2007), Le Grange 2015(Le Grange et al., 2015) and Schmidt 2007 (Schmidt et al., 2007a)). An overview of the trials included in the meta-analysis can be found in Table 4. Further information about both included and excluded studies can found in Appendix J.

No studies were identified that compared a combined family therapy and a pharmacological agent with any other intervention or wait list controls.

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 168: Study information for trials included in the meta-analysis of individual psychotherapy versus any other intervention or wait list controls for people with bulimia nervosa.

Study ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison	Sessions N	Treatment length	Long-term FU
Bulimia Nervosa – Individual Therapy										
Agras 1989	29.2 (8.6)	NR	100%	Duration of BN: 8.8 (6.6) years	77	CBT-ED	Wait List Control Self-monitoring BT	14	4 months	None reported
Agras 2000	28.3 (7.0)	22.7 (4.2)	100%	Duration of bingeing 11.4 (7.6) years	220	CBT-ED	IPT	19	20 weeks	8 or 12 months FU
Bulik 1998/McIntosh 2011	26.1 (6.1)	22.4 (2.5)	100%	Duration of BN 6.0 (7.0) years.	111	CBT-ED 1	CBT-ED + physical therapy (breathing) CBT-ED 2	16	14 weeks	12 months FU
Chen 2003	25.8 (7.2)	22.2 (2.8)	100%	Duration of BN: 9.6 (7.3) years	71	CBT-ED	CBT-ED Group	19	4.5 months	6 months FU
Cooper 1995	23.8	NR	100%	Mean 56 months	31	CBT-ED	BT	19	18 weeks	12 months FU
Fairburn 1986	22.9 (4.4)	The weight within the normal range (mean weight = 96.9% MPMW, SD = 9.4)	100%	High EAT (mean score 48.8 [17.8])	24	CBT-ED	Psychodynamic - General	18	18 weeks	12 months FU

Study ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison	Sessions N	Treatment length	Long-term FU
Fairburn 1991 Fairburn 1993 Fairburn 2015	24.2 (22.8-25.6)	22.2 (21.5-23.0)	100%	Duration of BN 4.4 (3.4-5.3) years	75	CBT-ED	BT IPT	19	18 weeks	12 months FU 3 year FU
Fairburn 2009	26.1 (7.0)	18.7 - 26.4	96 %	Duration of BN 8.8 (6.9) years	154	CBT-ED.1	Wait List Control CBT-ED 2	21	8 weeks	60 week FU
Fairburn 2015	25.9 (7.7)	19.1 to 26.4	98%	Duration of BN 11.4 (9.6) years	130	CBT-ED	IPT	20	20 weeks	60 week FU
Freeman 1988	24.2 (5.6)	Weight as a % of matched population mean weight 108.2% (16.1)	100%	Duration of BN 6 (4.9) years	112	CBT-ED	Wait List Control BT Nutritional counselling	15	15 weeks	None reported
Garner 1993	23.7 (4.4)	Weight as a % of matched population mean weight 95.3% (9.8)	100%	Duration of BN 71.8 (47.6) months	60	CBT-ED	Dynamic psychotherapy-ED	18	18 weeks	None reported
Ghaderi 2006	27.2 (7.8)	25.0 (5.1)	NR	Duration of BN: 9.2 (6.3) years	50	CBT-ED	CBT-ED	19	19 weeks	18 months FU
Griffiths 1994	25.91 (5.7)	21.9 (2.0)	100%	Duration of BN: 6.2 (5.2)	78	CBT-ED	Wait List Control	7	8 weeks	None reported

Study ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison	Sessions N	Treatment length	Long-term FU
				years						
Hsu 2001	24. 5 (6.4)	112.2% Ideal BW	100%	Duration of BN: 5.5 (3.2) years	100	CT-ED	Nutrition counselling Nutrition and CT-ED Support group	16	14 weeks	None reported
Le Grange 2007	16.0 (1.7)	21.8 (2.5)	98%	At least 24 bulimic episodes in past 6 months	80	CBT-ED	Family therapy	20	6 months	6 months FU
Le Grange 2015	15.8 (1.5)	Mean % expected BW=109.4 (21.7)	94%	Duration of BN: 22.2 (16.2) months.	130	CBT-ED	Family therapy Non-specific psychotherapy	18	6 months	6 months FU
Mitchell 2008	28.4 (10.4)	23.5 (5.4)	100%	NR BN 56%, EDNOS 44%	128	Guided Self-help ED	CBT-ED	20	16 weeks	12 months FU
Nevonen 2006	21.1 (2.0)	21.5 (2.1)	100%	Duration of BN: 5.1 (2.9) years	86	Hybrid - mixes/sequences therapies	Other hybrid	23	23 weeks	12 months FU
Olmsted 1991	24 (4.2)	NR	100%	NR	65	CBT-ED	Group psychoeducation	5	4 weeks	None reported
Poulsen 2014	25.8 (4.9)	22.24 (2.11)	97%	Duration of BN:12.3 (6.2) years	70	Dynamic psychotherapy - ED	CBT-ED	21	14 weeks	None reported
Thackwray 1993	31.3 (10.4)	NR	100%	Duration of BN: 6.7 (7.3) years	47	CBT-ED	BT Placebo	8	8 weeks	6 months FU
Thiels	28.7 (9.1)	21.3	NR	Duration of BN 8.5 (9.2)	62	Guided Self-	CBT-ED	8	16	43 weeks

Study ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison	Sessions N	Treatment length	Long-term FU
1998		(3.1)		years		Help-ED			weeks	FU
Thompson-Brenner 2016	25.6 (8.1)	23.7 (3.5)	100%	At least 3 months	50	CBT-ED (focused)	CBT-ED (enhanced)	16	8 weeks	6 months FU
Treasure 1994	26.0 (6.6)	26.8 (7.0)	100%	Actively bulimic criteria	110	CBT-ED	Wait List Control Self-help (ED)	8	8 weeks	None reported
Wilson 1991	19.8	22.0	90%	Suffered from ED for at least 12 months	22	CBT-ED.1	CBT-ED.2	20	20 weeks	12 months FU
Wonderlich 2014	27.3 (9.6)	NR	90%	Suffered from ED for at least 3 months	80	Integrative Cognitive-Affective Therapy (ICAT)	CBT-ED	21	19 weeks	4 months FU

Abbreviations: BN – bulimia nervosa; BT – behavioural therapy; CBT-ED – cognitive behavioural therapy with an eating disorder focus; ED – eating disorder; FU – follow up; ICAT – integrative cognitive affective therapy; IPT – interpersonal psychotherapy; N – number; NR – not reported.

Table 169: Study information for trials included in the meta-analysis of group psychotherapy versus any other intervention or wait list controls for people with bulimia nervosa.

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N Initially Random-ised	Intervention Category	Comparison Arm Category	Sessions N	Treatment Length	Long-term FU
Bailer 2004	23.3 (4.1)	21.7 (3.1)	100%	Duration of BN: 6 years	81	Group CBT-ED	Guided self-help ED	20	18 weeks	12 mo FU
Chen 2003	25.8 (7.2)	22.2 (2.8)	100%	Duration of BN: 9.6 (7.3) years	71	Group CBT-ED	CBT-ED-Individual	19	4.5 months	6 mo FU
Hsu 2001	24.5 (6.6)	112.2 % ideal body weight	100%	Duration of BN: 5.5 (3.2) years	100	Group Self-help	CBT-ED individual Hybrid (nutritional and	16	14 weeks	None reported

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N Initially Random-ised	Intervention Category	Comparison Arm Category	Sessions N	Treatment Length	Long-term FU
							CBT-ED) Nutritional counselling			
Lavender 2012	27.7 (7.6)	24.4 (5.7)	92%	Patients were referred by GP. Duration of BN: 9.3 (7.6) years	74	Group CBT-ED	Emotional and Social Mind Training (ESM)	17	4 months	None reported
Lee 1986	27.7 (5.3)	NR	100%	First experienced bulimia in their late teens or early twenties.	30	Group CBT-ED	Wait list control	12	6 weeks	None reported
Leitenbergy 1988	26 years	NR	100%	Duration of BN: 6.9 years	59	Group BT (ED)	Group BT (ED) Group CBT-ED Wait list control	24	14 weeks	6 months FU
Mitchell 1993	25.8 (6.8)	NR	100%	Average duration of illness was 8.8 years	76	Group CBT-ED	Group CBT-ED Low intensity High abstinence and Low abstinence	24	12 weeks	None reported
Nauta 2001	38.3 (7.1)	33.1 (4.3)	100%	Duration of BN:12.5 (6.4) years	37	Group BT (ED)	Group nutritional counselling	15	15 weeks	6 months U
Olmsted 1991	24 (4.2)	Unclear	100%	Between 1 and 20 years	75	Group psychoeducation	CBT-ED individual	5	4 weeks	None reported
Wolf 1992	26.5 (8.1)	NR	100%	Bingeing for 8.8 years	42	Group CBT-ED	BT Group Wait list control	10	8 weeks	3 mo FU

Abbreviations: BN – bulimia nervosa; BT – behavioural therapy; CBT-ED – cognitive behavioural therapy with an eating disorder focus; ED – eating disorder; ESM - emotional and social mind training; FU – follow up; ICAT – integrative cognitive affective therapy; IPT – interpersonal psychotherapy; N – number; NR – not reported; WLC – wait list control

Table 170: Study information for trials included in the meta-analysis of self-help versus any other intervention or wait list controls for people with bulimia nervosa

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention Category	Comparison Arm Category	Sessions N	Treatment Length	Long-term FU
Bailer 2004	23.3 (4.1)	21.7 (3.1)	100%	Duration of BN: 6 years	81	Guided self-help ED	Group CBT-ED	20	18 weeks	12 months FU
Banasiask 2005	29.5 (8.7)	22.6 (3.6)	100%	Duration of BN: 9.2 (7.0)	109	Guided self-help ED	Wait list control	9	16 weeks	None reported
Bauer 2012	29.9 (7.9)	24.8 (6.8)	100%	Discharged from hospital. BN 60%, EDNOS 40%	165	Text Messaging Intervention	Wait list control	16	16 weeks	None reported
Carter 2003	27 (8) years	23 (5)	100%	Duration of BN: 7 (6) years	85	Self-help ED	Wait list control Self-help	0	8 weeks	None reported
Durand 2003	28.3 (6.5)	NR	100%	NR	68	Guided self-help ED	CBT and IPT combination	Self-help saw GP 4.9 times	Unclear if 6 or 9 months	None reported
Ljotsson 2007	35.5 (11.4)	35.5 (11.4)	94%	52% BED, 48% BN BN for at least 3 months or BED for at least 6 months	73	Guided self-help ED	Wait list control	12	12 weeks	None reported
Mitchell 2008	28.4 (10.4)	23.5 (5.4)	100%	NR BN 56%, EDNOS 44%	128	Guided self-help ED	CBT-ED	20	16 weeks	12 months FU
Palmer 2002	26.8 (9.5)	26.2 (7.9)	99%	NR BN 60% +BED 20% +EDNOS	121	Guided self-help ED 1	Wait List Control Guided Self-help ED 2	4	4 months	12 months FU

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention Category	Comparison Arm Category	Sessions N	Treatment Length	Long-term FU
				20%			Self-help (ED)			
Ruwaard 2013	30 (10)	6% Obese, 20% overweight, 74% normal.	97%	Average duration of symptoms was 11 (9) years	105	On-line self-help ED	Self-help ED Wait list control	25	20 weeks	12 months FU
Sanchez-Ortiz 2011	23.9 (5.9)	22.0 (2.8)	99%	Duration of illness 5.2 (4.0) BN (51.3%) or EDNOS (48.7%)	76	On-line guided self-help	Wait list control	8	8-12 weeks up to 24 weeks	None reported
Schmidt 2006	29.5 (9.2)	23.5 (4)	unclear	BN and EDNOS Median duration of illness 4 (1 to 6)	61	Guided self-help ED	Self-help ED	10	10 weeks	6 months FU
Schmidt 2007	17.6 (1.7)	21.1 (2.6)	98%	BN (72%) and EDNOS (28%) Time since onset 2.5 (2.1) years	85	Guided self-help ED	Family therapy	15	6 months	6 months FU
Steele 2008	24.7 (5.5)	21.4 (2.4)	99%	BN or BED (BN 59 to 100% per group) Duration of bulimic symptoms: 9.0 years	48	Guided self-help-ED	Placebo Guided self-help	8	6 weeks	None reported
Thiels	28.7 (9.1)	21.3 (3.1)	NR	Duration of	62	Guided self-help	CBT-ED	8	16 weeks	43

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention Category	Comparison Arm Category	Sessions N	Treatment Length	Long-term FU
1988				BN: 8.5 (9.2) years		ED				weeks FU
Treasure 1994	25.7 (5.8)	24.0 (5.9)	100%	BN and atypical BN. Duration of BN 7.8 years	110	Self-Help ED	CBT-ED Wait List Control	16	8 weeks	None reported
Wagner 2013	24.2 (4.5)	20.6 (2.1)	100%	BN 90%; EDNOS 10% Duration of illness: 8.2 (5.2)	155	Self-help ED (internet)	Guided Self-help ED (traditional)	Weekly emails	7 months	11 months FU
Walsh 2004	30.6 (7.8)	> 17.5	100%	Mean duration 12.0 years (7.9)	47	Guided self-help ED	Placebo	6-8	4 months	None reported

Abbreviations: BN – bulimia nervosa; BT – behavioural therapy; CBT-ED – cognitive behavioural therapy with an eating disorder focus; ED – eating disorder; ESM - emotional and social mind training; FU – follow up; ICAT – integrative cognitive affective therapy; IPT – interpersonal psychotherapy; N – number; NR – not reported; WLC – wait list control

Table 171: Study information for trials included in the meta-analysis of family therapy versus any individual therapy or wait list control for people with bulimia nervosa

Study ID	Mean age (SD)	Mean BMI, kg/m2 (SD)	Female (%)	Sample	N random-ised	Intervention	Comparison	Number of sessions	Treatment Length	Long-term FU
Le Grange 2007	16.1	22.1 (3.0)	98	DSM-IV BN or at least 24 bulimic episodes in past 6 months	80	FBT-BN Duration: 22.3 (20.4) months	SPT-BN Duration: 20.1 (24.4) months	20	6 months	6 months
Le Grange 2015	15.8 (1.5)	Not reported	94	DSM-IV BN or partial BN	130	FBT-BN Duration: 19.6 (19.9) months	CBT-A Duration: 11.3 (10.4) months SPT-BN Duration: 22.2	18 (Mean 14.7)	6 months	12 months

Study ID	Mean age (SD)	Mean BMI, kg/m2 (SD)	Female (%)	Sample	N random-ised	Intervention	Comparison	Number of sessions	Treatm ent Length	Long-term FU
							(16.2) months			
Schmidt 2007	17.6 (1.7)	21.1 (2.6)	98	DSM-IV criteria for BN or EDNOS	85*	FT-ED Age of ED onset: 15.2 (1.8) years	gSH CBT-ED Age of ED onset: 14.9 (2.1) years	Up to 13 with close others + 2 individual sessions	6 months	6 months

**sample consisted of 61 BN and 24 EDNOS. Abbreviations: BN, bulimia nervosa; CBT-A, Cognitive Behavioural Therapy for young people; ED, eating disorder; EDNOS, eating disorder not otherwise specified; FBT-BN, Family-based treatment for bulimia nervosa; FT, Family Therapy; gSH CBT-ED, Guided Self-Help Cognitive Behavioural Therapy for Eating Disorders; SPT-BN, Supportive Psychotherapy for young people bulimia nervosa.*

7.2.6 Summary of findings tables

Individual therapy

Table 172: Summary of findings table for CBT-ED versus any other intervention at the end of treatment in young people and adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BN CBT-ED (95% CI)
Purges - Young people	86 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean purges - young people in the intervention groups was 0.33 standard deviations higher (0.1 lower to 0.75 higher)
Purges - Adults	359 (4 studies)	⨁⨀⨀⨀ VERY LOW3,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean purges - adults in the intervention groups was 0.59 standard deviations lower (0.8 lower to 0.37 higher)
Binges objective Young people	157 (2 studies)	⨁⨁⨀⨀ LOW6,7 due to risk of bias,		Not calculable for SMD values	The mean binges objective young people in the intervention groups was 0.09 standard deviations higher

		imprecision			(0.23 lower to 0.4 higher)
Binges (objective) Adults	687 (10 studies)	⊕⊕⊖⊖ LOW8,9 due to risk of bias, inconsistency		Not calculable for SMD values	The mean binges (objective) adults in the intervention groups was 0.25 standard deviations lower (0.41 to 0.1 lower)
Vomiting episodes - Young people	71 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting episodes - young people in the intervention groups was 0.64 standard deviations higher (0.16 to 1.12 higher)
Vomiting episodes Adults	484 (7 studies)	⊕⊕⊖⊖ LOW4,8 due to risk of bias, inconsistency		Not calculable for SMD values	The mean vomiting episodes adults in the intervention groups was 0.34 standard deviations lower (0.82 lower to 0.14 higher)
Laxatives use/ fortnight - Adults	284 (2 studies)	⊕⊕⊖⊖ LOW2,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxatives use/ fortnight - adults in the intervention groups was 0.27 standard deviations higher (0.01 lower to 0.55 higher)
Symptom checklist (SCL-90-R)- Adults	261 (4 studies)	⊕⊖⊖⊖ VERY LOW5,9,11 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean symptom checklist (scl-90-r)- adults in the intervention groups was 0.31 standard deviations lower (0.56 to 0.06 lower)
Quality of life	80 (1 study)	⊕⊕⊖⊖ LOW2,12 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.25 standard deviations higher (0.19 lower to 0.69 higher)
Depression - Adolescents	71 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - adolescents in the intervention groups was 0.10 standard deviations higher (0.36 lower to 0.57 higher)
Depression - Adults	630 (10 studies)	⊕⊖⊖⊖ VERY LOW5,9,13 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression - adults in the intervention groups was 0.31 standard deviations lower (0.47 to 0.14 lower)
EDE-Total Young People	70 (1 study)	⊕⊕⊖⊖ LOW1,2		Not calculable for SMD values	The mean ede-total young people in the intervention groups was

		due to risk of bias, imprecision			0.49 standard deviations higher (0.02 to 0.97 higher)
EDE - Total score - Adults	419 (5 studies)	⊕⊖⊖⊖ VERY LOW4,5,14 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - total score - adults in the intervention groups was 0.20 standard deviations lower (0.67 lower to 0.27 higher)
EDE- Dietary restraint - Young people	71 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- dietary restraint - young people in the intervention groups was 0.51 standard deviations higher (0.04 to 0.98 higher)
EDE- Dietary restraint - Adults	723 (10 studies)	⊕⊖⊖⊖ VERY LOW5,9,15 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- dietary restraint - adults in the intervention groups was 0.76 standard deviations lower (1.13 to 0.39 lower)
EDE-Attitudes to shape - Young people	71 (1 study)	⊕⊕⊖⊖ LOW2,16 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-attitudes to shape - young people in the intervention groups was 0.54 standard deviations higher (0.07 to 1.01 higher)
EDE- Attitudes to shape - Adults	725 (11 studies)	⊕⊖⊖⊖ VERY LOW4,15 due to risk of bias, inconsistency		Not calculable for SMD values	The mean ede- attitudes to shape - adults in the intervention groups was 0.17 standard deviations lower (0.69 lower to 0.36 higher)
EDE-Attitudes to weight - Young people	71 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-attitudes to weight - young people in the intervention groups was 0.51 standard deviations higher (0.04 to 0.98 higher)
EDE - Eating concern	477 (5 studies)	⊕⊕⊖⊖ LOW9,17 due to risk of bias, inconsistency		Not calculable for SMD values	The mean ede - eating concern in the intervention groups was 0.10 standard deviations lower (0.46 lower to 0.27 higher)
EDE- Attitudes to weight - Adults	725 (10 studies)	⊕⊖⊖⊖ VERY LOW4,15 due to risk of bias, inconsistency		Not calculable for SMD values	The mean ede- attitudes to weight - adults in the intervention groups was 0.43 standard deviations lower (0.89 lower to 0.03 higher)
EDI- Bulimia	242	⊕⊕⊖⊖		Not calculable for	The mean edi- bulimia in the intervention

	(4 studies)	LOW5,18 due to risk of bias, imprecision		SMD values	groups was 0.30 standard deviations lower (0.57 to 0.04 lower)
EDI - Drive for thinness	243 (4 studies)	⊕⊕⊖⊖ LOW7,18 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness in the intervention groups was 0.18 standard deviations higher (0.60 lower to 0.97 higher)
EDI - Body Dissatisfaction	200 (4 studies)	⊕⊖⊖⊖ VERY LOW4,7,18 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction in the intervention groups was 0.06 standard deviations lower (0.89 lower to 0.78 higher)
Global Clinical Score	111 (3 studies)	⊕⊕⊖⊖ LOW2,19 due to risk of bias, imprecision		Not calculable for SMD values	The mean global clinical score in the intervention groups was 0.15 standard deviations lower (0.54 lower to 0.24 higher)
General psychopathology (PSE)- Adults	22 (1 study)	⊕⊕⊖⊖ LOW5,20 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology (pse)-adults in the intervention groups was 0.77 standard deviations lower (1.64 lower to 0.1 higher)
Symptom checklist (SCL-90-R)- Adults <5 years illness	62 (1 study)	⊕⊕⊖⊖ LOW5,12 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist (scl-90-r)-adults <5 years illness in the intervention groups was 1.05 standard deviations lower (1.61 higher to 0.49 lower)
Remission - Adolescents_ITT	110 (1 study)	⊕⊕⊖⊖ LOW1,21 due to risk of bias, imprecision	RR 0.52 (0.2 to 0.9)	327 per 1000	157 fewer per 1000 (from 33 fewer to 262 fewer)
Symptom checklist (SCL-90-R)- Adults >5 years illness	197 (2 studies)	⊕⊕⊖⊖ LOW7,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist (scl-90-r)-adults >5 years illness in the intervention groups was 0.13 standard deviations lower (0.40 lower to 0.15 higher)
Remission - Adults_ITT	731 (7 studies)	⊕⊖⊖⊖ VERY LOW9,22,23 due to risk of bias,	RR 1.87 (1.43 to 2.46)	174 per 1000	151 more per 1000 (from 75 more to 254 more)

		inconsistency, imprecision			
Bulimic Inventory Test Edinburgh	47 (1 study)	⊕⊕⊖⊖ LOW5,24 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic inventory test edinburgh in the intervention groups was 0.77 standard deviations lower (1.37 to 0.18 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 The participants and investigators were not blind but the assessors were.
2 95% CI crossed 1 MID (0.5)
3 Unclear if allocation concealment was performed, except Poulsen 2014. It was unclear in two studies if assessors were blind and high dropout rates were reported in two studies >20%,
4 Heterogeneity reported, I2 >80%
5 95% CI crossed 1 MID (-0.5)
6 In LeGrange 2015, the participants and investigators were not blind but the assessors were, whilst in LeGrange 2007 neither the participants, investigators or assessors were blind.
7 For a continuous outcome, there were fewer than 400 participants.
8 In half of the studies, it is unclear how the randomisation sequence was generated. In most studies it was unclear if allocation concealment was conducted. High drop outs were reported by Fairburn.
9 Heterogeneity detected I2 >50%
10 In half of the studies it is unclear how the randomisation sequence was generated. In most studies it was unclear if allocation concealment was conducted. High drop outs were reported by Fairburn and Freeman.
11 Unclear in all studies, except Poulsen 2014, if allocation concealment was conducted. It was unclear how Fairburn 1991 generated the random sequence. A high number of drop outs were reported >20% in Agras 2000.
12 Unclear if allocation concealment was performed. Unclear if assessor, investigators and patients was blind.
13 In half of the studies, it is unclear how the randomisation sequence was generated. In all studies it was unclear if allocation concealment was conducted. High drop outs were reported Theils and Agras.
14 Unclear in all studies, except Poulsen 2014, if allocation concealment was conducted. It was unclear how Fairburn 1986 generated the random sequence. In half of the studies a high number of drop outs were reported >20%
15 In a few of the studies, it is unclear how the randomisation sequence was generated. In all studies it was unclear if allocation concealment was conducted. High drop outs (>20%) were reported by Treasure, Theils and Fairburn
16 It was unclear if allocation concealment was performed. Assessors were not blinded.
17 It was unclear is one study how randomisation was conducted and in all studies if allocation concealment was conducted. Half of the studies it was unclear if assessor was blind and high dropout rates were reported in half the studies >20%.
18 In most of the studies, it is unclear how the randomisation sequence was generated. In all studies it was unclear if allocation concealment was conducted.
19 It was unclear in two of the studies how the randomisation sequence was generated and in all studies if allocation concealment was conducted. One

study reported high drop outs >20% and one study it was unclear if assessor was blind.
20 It was unclear how randomisation sequence was generated and if allocation concealment was conducted.
21 95% CI crossed 1 MID (0.75)
22 In a few studies it was unclear how randomisation was performed and in all studies it was unclear if allocation concealment was performed. High dropout rates were reported in a number of studies.
23 For a dichotomous outcome, there are fewer than 300 events.
24 Inadequate random sequence generation and unclear if allocation concealment was performed. High dropout rates were reported >20%

Table 173: Summary of findings table for CBT-ED versus any other intervention at follow up in young people and adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention Follow-up	Risk difference with BN CBT-ED (95% CI)
Bulimic episodes Follow-up - Young people	137 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic episodes follow-up - young people in the intervention groups was 0.10 standard deviations higher (0.24 lower to 0.44 higher)
Bulimic episodes Follow-up - Adults	294 (5 studies)	⊕⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean bulimic episodes follow-up - adults in the intervention groups was 0.35 standard deviations lower (0.83 lower to 0.12 higher)
Purges Follow-up - Young people	69 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean purges follow-up - young people in the intervention groups was 0 standard deviations higher (0.48 lower to 0.48 higher)
Purges Follow-up - Adults	208 (3 studies)	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean purges follow-up - adults in the intervention groups was 0.15 standard deviations lower (0.42 lower to 0.13 higher)
Laxatives Follow-up - Adults	98 (1 study)	⊕⊕⊖⊖ LOW5,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxatives follow-up - adults in the intervention groups was 0.02 standard deviations lower (0.42 lower to 0.37 higher)
Vomiting FU -Young people	68 (1 study)	⊕⊕⊖⊖ LOW9,10		Not calculable for SMD values	The mean vomiting fu -young people in the intervention groups was

		due to risk of bias, imprecision			0.17 standard deviations higher (0.3 lower to 0.65 higher)
Vomiting Follow-up - Adults	162 (3 studies)	⊕⊕⊖⊖ LOW5,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting follow-up - adults in the intervention groups was 0.31 standard deviations lower (0.84 lower to 0.22 higher)
Symptom checklist Follow-up - Adults	166 (2 studies)	⊕⊕⊖⊖ LOW5,12 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist follow-up - adults in the intervention groups was 0.02 standard deviations higher (0.29 lower to 0.32 higher)
General psychopathology - FU - Adults	49 (2 studies)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology - fu - adults in the intervention groups was 0.5 standard deviations lower (1.07 lower to 0.07 higher)
Global clinical score FU - Adults	22 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean global clinical score fu - adults in the intervention groups was 0.81 standard deviations lower (1.67 lower to 0.07 higher)
Depression - FU - Adolescents	68 (1 study)	⊕⊕⊖⊖ LOW6,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - fu - adolescents in the intervention groups was 0.09 standard deviations lower (0.56 lower to 0.29 higher)
Depression - FU - Adults	410 (8 studies)	⊕⊕⊕⊖ MODERATE13 due to risk of bias		Not calculable for SMD values	The mean depression - fu - adults in the intervention groups was 0.14 standard deviations lower (0.34 lower to 0.05 higher)
EDI - Bulimia FU - Adults	47 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia fu - adults in the intervention groups was 0.47 standard deviations lower (1.09 lower to 0.15 higher)
EDI - Drive for thinness FU - Adults	47 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness fu - adults in the intervention groups was 0.12 standard deviations higher (0.5 lower to 0.73 higher)
EDI Body Dissatisfaction FU - Adults	27 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias,		Not calculable for SMD values	The mean edi body dissatisfaction fu - adults in the intervention groups was 0.36 standard deviations lower

		imprecision			(1.12 lower to 0.4 higher)
EDE- Total score FU - Adolescents	68 (1 study)	⊕⊕⊖⊖ LOW6,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- total score fu - adolescents in the intervention groups was 0.38 standard deviations higher (0.1 lower to 0.86 higher)
EDE - Total score Follow-up - Adults	307 (3 studies)	⊕⊕⊕⊖ LOW 6,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - total score follow-up - adults in the intervention groups was 0.11 standard deviations lower (0.34 lower to 0.11 higher)
EDE- Weight concerns FU - Adolescents	68 (1 study)	⊕⊕⊖⊖ LOW6,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concerns fu - adolescents in the intervention groups was 0.46 standard deviations higher (0.02 lower to 0.94 higher)
EDE - Weight concerns FU - Adults	126 (3 studies)	⊕⊕⊖⊖ LOW5,15 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - weight concerns fu - adults in the intervention groups was 0.08 standard deviations lower (0.43 lower to 0.27 higher)
EDE- Shape concerns FU - Adolescents	68 (1 study)	⊕⊕⊖⊖ LOW6,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concerns fu - adolescents in the intervention groups was 0.58 standard deviations higher (0.09 to 1.06 higher)
EDE - Shape concerns FU - Adults	126 (3 studies)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - shape concerns fu - adults in the intervention groups was 0.10 standard deviations lower (0.36 lower to 0.34 higher)
EDE - Eating concerns FU - Adults	52 (1 study)	⊕⊕⊖⊖ LOW5,16 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - eating concerns fu - adults in the intervention groups was 0.25 standard deviations lower (0.8 lower to 0.29 higher)
EDE- Restraint FU - Adolescents	68 (1 study)	⊕⊕⊖⊖ LOW6,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint fu - adolescents in the intervention groups was 0.38 standard deviations higher (0.1 lower to 0.86 higher)
EDE - Restraint FU -	126	⊕⊕⊖⊖		Not calculable for	The mean ede - restraint fu - adults in the

Adults	(3 studies)	LOW5,15 due to risk of bias, imprecision		SMD values	intervention groups was 0.12 standard deviations lower (0.47 lower to 0.23 higher)
Bulimic Inventory Test Edinburgh - Adults FU	47 (1 study)	⊕⊕⊖⊖ LOW2,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic inventory test edinburgh - adults fu in the intervention groups was 0.21 standard deviations lower (0.78 lower to 0.37 higher)
Quality of life FU	52 (1 study)	⊕⊕⊖⊖ LOW10,16 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life fu in the intervention groups was 0.09 standard deviations lower (0.63 lower to 0.46 higher)
Remission FU - Adolescents_ITT	110 (1 study)	⊕⊖⊖⊖ VERY LOW18,19,20 due to risk of bias, imprecision	RR 0.83 (0.43 to 1.6)	269 per 1000	46 fewer per 1000 (from 153 fewer to 162 more)
Remission FU - Adult_ITT	553 (4 studies)	⊕⊕⊖⊖ LOW3,20 due to risk of bias, imprecision	RR 1.32 (1 to 1.76)	233 per 1000	75 more per 1000 (from 0 more to 177 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; FU: follow up; ITT intention to treat

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Assessors were blind in one study (LeGrange 2015) but participants, investigators and assessors were not blind in the other study.
2 95% CI crossed 1 MID (-0.5)
3 In the majority of studies it was unclear how the randomisation sequence was generated. In all studies it was unclear if allocation concealment was performed and in half the studies a high dropout was reported >20%,
4 Heterogeneity reported I2 >50%.
5 Fewer than optimal sample size was used <400 participants.
6 Unclear if allocation concealment was performed. Assessors were blind, but it was unclear if participants and investigators were blind.
7 It was unclear in a few studies how the randomisation sequence was generated and in all studies if allocation concealment was performed. In one study high drop outs were reported >20%.

8 It was unclear in one study how the randomisation sequence was performed. Unclear in all studies if allocation concealment was performed. High drop outs were reported in one study >20%.
9 Participants, assessors and investigators were not blind.
10 95% CI crossed 1 MID (0.5)
11 It was unclear in one study how the randomisation sequence was generated and in all studies, except Poulsen, if allocation concealment was performed. In two studies high drop outs were reported >20%
12 It was unclear how the random sequence was generated in one study and if allocation concealment was performed in majority of studies. In one study it was unclear if assessor was blind.
13 In half the studies it was unclear how randomisation sequence was generated. It was unclear in all of the studies if allocation concealment was performed. In few studies, high dropout rates were reported >20%,
14 Unclear if allocation concealment was performed in majority of studies. In half the studies, a high dropout was reported >20%
15 In two of three studies it was unclear how the randomisation sequence was generated and in one study it was inadequate. It was unclear in all studies if allocation concealment was performed. In one study high dropout rates were reported >20%.
16 It was unclear how random sequence was generated and allocation concealment was performed. It was unclear if assessor was blind.
17 Allocation concealment was not performed. It was unclear if either the participants, investigators or assessors were blind. High drop outs were detected >20%.
18 Assessors were blind but participants and investigators were not.
19 95% CI crossed 1 MID (0.75)
20 95% CI crossed 1 MID (1.25)

Table 174: Summary of interpersonal psychotherapy versus another intervention at the end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BN IPT (95% CI)
EDE - Total	247 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - total in the intervention groups was 0.52 standard deviations higher (0.27 to 0.77 higher)
EDE - Restraint	309 (3 studies)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - restraint in the intervention groups was 0.71 standard deviations higher (0.02 lower to 1.43 higher)
EDE - Weight concerns	309 (3 studies)	⊕⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - weight concerns in the intervention groups was 0.63 standard deviations higher (0.53 lower to 1.79 higher)

EDE - Shape concerns	309 (3 studies)	⊕⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - shape concerns in the intervention groups was 0.14 standard deviations lower (1.06 lower to 0.78 higher)
EDE - Eating concerns	247 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - eating concerns in the intervention groups was 0.47 standard deviations higher (0.22 to 0.73 higher)
Symptom checklist (SCL-90-R)	191 (2 studies)	⊕⊕⊖⊖ LOW1,5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist (scl-90-r) in the intervention groups was 0.11 standard deviations higher (0.19 lower to 0.4 higher)
Social adjustment scale	213 (3 studies)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean social adjustment scale in the intervention groups was 0.33 standard deviations higher (0.06 lower to 0.61 higher)
Purges	129 (1 study)	⊕⊕⊖⊖ LOW2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean purges in the intervention groups was 0.42 standard deviations higher (0.07 to 0.77 higher)
Self-induced vomiting	178 (2 studies)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean self-induced vomiting in the intervention groups was 0.64 standard deviations higher (0.33 to 0.96 higher)
Bulimic episodes (objective)	98 (2 studies)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic episodes (objective) in the intervention groups was 0.29 standard deviations higher (0.01 lower to 0.6 higher)
Depression	202 (3 studies)	⊕⊖⊖⊖ VERY LOW2,6,8 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.22 standard deviations higher (0.41 lower to 0.85 higher)
Laxative taking	116 (1 study)	⊕⊕⊖⊖ LOW7,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative taking in the intervention groups was 0.37 standard deviations lower (0.73 lower to 0 higher)
Remission_ITT	425	⊕⊕⊖⊖	RR 0.33	342 per 1000	229 fewer per 1000

	(3 studies)	LOW7,10 due to risk of bias, imprecision	(0.21 to 0.5)		(from 171 fewer to 270 fewer)
General clinical score	22 (1 study)	⊕⊕⊖⊖ LOW2,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean general clinical score in the intervention groups was 0.94 standard deviations higher (0.05 to 1.83 higher)
Remission_ITT < 5 years	75 (4 studies)	⊕⊕⊖⊖ LOW7,10 due to risk of bias, imprecision	RR 1.56 (0.83 to 2.93)	280 per 1000	157 more per 1000 (from 48 fewer to 540 more)
Remission_ITT > 5 years	350 (2 studies)	⊕⊕⊖⊖ LOW7,10 due to risk of bias, imprecision	RR 0.71 (0.49 to 1.03)	297 per 1000	86 fewer per 1000 (from 152 fewer to 9 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear in all studies if allocation concealment was performed. Two studies reported high dropout rates >20%
2 95% CI crossed 1 MID (0.5)
3 It was unclear if allocation concealment was performed. In Fairburn 1991 (1993) it was unclear how the randomisation sequence was generated. Two studies reported high dropout rates >20%
4 Heterogeneity detected I2 >80%
5 Optimal sample size was not met >400 participants
6 It was unclear in one study how random sequence was generated and in all studies if allocation concealment was conducted. In one study high drop outs were reported >20%.
7 It was unclear if allocation concealment was conducted. High dropout rates were reported >20%.
8 Heterogeneity detected I2 >50%
9 95% CI crossed 1 MID (-0.5)
10 Optimal event size was not met >300 events
11 It was unclear if allocation concealment was conducted. It was unclear if participants and investigators were blind, however, assessors were blind.

1 Table 175: Summary of interpersonal psychotherapy (IPT) versus another intervention at follow up in adults with bulimia nervos**a**

Outcomes	**No of Participants (studies) Follow up**	Quality of the evidence (GRADE)	**Relative effect (95% CI)**	Anticipated absolute effects	
				Risk with another intervention Follow-up	Risk difference with BN IPT (95% CI)

EDE - Total FU	227 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - total fu in the intervention groups was 0.22 standard deviations higher (0.04 lower to 0.48 higher)
EDE - Restraint FU	264 (3 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - restraint fu in the intervention groups was 0.33 standard deviations higher (0.08 to 0.57 higher)
EDE - Weight concerns FU	264 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - weight concerns fu in the intervention groups was 0.11 standard deviations higher (0.13 lower to 0.35 higher)
EDE - Shape concerns FU	264 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - shape concerns fu in the intervention groups was 0.03 standard deviations higher (0.21 lower to 0.27 higher)
EDE - Eating concerns FU	227 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - eating concerns fu in the intervention groups was 0.15 standard deviations higher (0.11 lower to 0.41 higher)
Symptom checklist (SCL-90-R) FU	166 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist (scl-90-r) fu in the intervention groups was 0.02 standard deviations lower (0.32 lower to 0.29 higher)
Social adjustment scale FU	166 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean social adjustment scale fu in the intervention groups was 0.15 standard deviations higher (0.15 lower to 0.46 higher)
Purges FU	129 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean purges fu in the intervention groups was 0.18 standard deviations higher (0.16 lower to 0.53 higher)
Bulimic episodes (objective) FU	98 (1 study)	⊕⊕⊖⊖ LOW4 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic episodes (objective) fu in the intervention groups was 0.02 standard deviations higher (0.37 lower to 0.42 higher)
Self-induced vomiting	135	⊕⊕⊖⊖		Not calculable for SMD	The mean self-induced vomiting fu in

FU	(2 studies)	LOW2,4 due to risk of bias, imprecision		values	the intervention groups was 0.05 standard deviations higher (0.28 lower to 0.39 higher)
Laxative taking FU	98 (1 study)	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative taking fu in the intervention groups was 0.02 standard deviations higher (0.37 lower to 0.42 higher)
Depression (Becks) FU	135 (3 studies)	⊕⊖⊖⊖ VERY LOW2,4,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression (becks) fu in the intervention groups was 0.10 standard deviations higher (0.22 lower to 2.05 higher)
Remission F_ITT	425 (3 studies)	⊕⊖⊖⊖ VERY LOW4,6,7 due to risk of bias, inconsistency, imprecision	RR 0.84 (0.61 to 1.15)	293 per 1000	47 fewer per 1000 (from 114 fewer to 44 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. Across studies, investigators, participants or assessors were not blind. High dropout rates were detected >20%.
2 For continuous outcome, there were fewer than <400 participants.
3 95% CI crossed 1 MID (0.5)
4 It was unclear if allocation concealment was conducted. Across studies, investigators, participants or assessors were not blind or it was unclear. High dropout rates were detected >20%.
5 It was unclear if allocation concealment was conducted. Assessors were blind but it was unclear if participants or investigators were blind. High dropout rates were detected >20%
6 Heterogeneity was detected >50%
7 95% CI crossed 1 MID (0.75)

Table 176: Summary of findings table for integrative cognitive-affective therapy (ICAT) versus any other intervention at the end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BN ICAT (95% CI)
EDE - Total	80	⊕⊕⊖⊖		Not calculable for SMD	The mean ede - total score in the

score	(1 study)	LOW1,2 due to risk of bias, imprecision		values	intervention groups was 0.11 standard deviations lower (0.55 lower to 0.33 higher)
Purges	80 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean purges in the intervention groups was 0.05 standard deviations higher (0.39 lower to 0.49 higher)
Binges (objective)	80 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean binges (objective) in the intervention groups was 0.06 standard deviations higher (0.37 lower to 0.5 higher)
Depression	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.08 standard deviations lower (0.52 lower to 0.36 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear whether the participants, investigators or the assessors were blind.
2 95% CI crossed 1 MID (-0.5).
3. Fewer than 400 participants.
4. 95% CI crossed 1 MID (-0.5).

Table 177: Summary of findings table for integrative cognitive-affective therapy (ICAT) versus any other intervention at follow up in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention FU	Risk difference with BN ICAT (95% CI)
EDE - Total score FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - total score fu in the intervention groups was 0.19 standard deviations lower (0.63 lower to 0.25 higher)
Purges FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean purges fu in the intervention groups was 0.09 standard deviations lower

		imprecision			(0.53 lower to 0.35 higher)
Binges (objective) FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binges (objective) fu in the intervention groups was 0.25 standard deviations lower (0.69 lower to 0.19 higher)
Depression FU	80 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.14 standard deviations higher (0.3 lower to 0.58 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; FU: follow up

1 It was unclear whether the participants, investigators or the assessors were blind.
2 95% CI crossed 1 MID (-0.5).
3 95% CI crossed 1 MID (0.5).

Table 178: Summary of findings table for CBT-ED (1) versus another CBT-ED (2) program at the end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT-ED (2)	Risk difference with BN CBT-ED (1) (95% CI)
Symptom check list - 90	291 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean symptom check list - 90 in the intervention groups was 0.03 standard deviations lower (0.26 lower to 0.2 higher)
Depression	306 (5 studies)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.08 standard deviations lower (0.31 lower to 0.14 higher)
Social adjustment score	142 (2 studies)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean social adjustment score in the intervention groups was 0.21 standard deviations lower (0.54 lower to 0.12 higher)
Bingeing (objective)	242 (4 studies)	⊕⊕⊖⊖ LOW3,7		Not calculable	The mean bingeing (objective) in the intervention groups was

		due to risk of bias, imprecision		for SMD values	0.20 standard deviations lower (0.43 lower to 0.03 higher)
Vomiting	122 (2 studies)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.09 standard deviations lower (0.45 lower to 0.26 higher)
Laxatives	72 (1 study)	⊕⊕⊖⊖ LOW6,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxatives in the intervention groups was 0.23 standard deviations lower (0.7 lower to 0.23 higher)
Purging (last 2 weeks)	114 (2 studies)	⊕⊕⊖⊖ LOW6,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging (last 2 weeks) in the intervention groups was 0.11 standard deviations lower (0.48 lower to 0.26 higher)
Remission_ITT	321 (4 studies)	⊕⊕⊖⊖ LOW8,10 due to risk of bias, imprecision	RR 1.13 (0.91 to 1.41)	456 per 1000	59 more per 1000 (from 41 fewer to 187 more)
EDI- Drive for thinness	72 (1 study)	⊕⊕⊖⊖ LOW9,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- drive for thinness in the intervention groups was 0.14 standard deviations higher (0.32 lower to 0.6 higher)
EDI- Bulimia	122 (2 studies)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- bulimia in the intervention groups was 0.02 standard deviations lower (0.37 lower to 0.34 higher)
EDI- Body dissatisfaction	122 (2 studies)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- body dissatisfaction in the intervention groups was 0.02 standard deviations higher (0.34 lower to 0.37 higher)
EDI- Total	319 (3 studies)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- total in the intervention groups was 0.01 standard deviations higher (0.21 lower to 0.23 higher)
EDE - Total	361 (4 studies)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - total in the intervention groups was 0.04 standard deviations lower (0.25 lower to 0.17 higher)

Global Function (GAFS)	72 (1 study)	⊕⊕⊖⊖ LOW8,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean global function (gafs) in the intervention groups was 0.36 standard deviations higher (0.1 lower to 0.83 higher)
General psychiatric features	149 (1 study)	⊕⊕⊖⊖ LOW3,12 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychiatric features in the intervention groups was 0.16 standard deviations higher (0.16 lower to 0.48 higher)
Bingeing episodes (28 d)	149 (1 study)	⊕⊕⊖⊖ LOW10,12 due to risk of bias, imprecision	RR 1.35 (0.81 to 2.24)	250 per 1000	88 more per 1000 (from 47 fewer to 310 more)
Vomiting episodes (28 d)	149 (1 study)	⊕⊕⊖⊖ LOW10,12 due to risk of bias, imprecision	RR 1.09 (0.7 to 1.69)	333 per 1000	30 more per 1000 (from 100 fewer to 230 more)
Purging (28 d)	149 (1 study)	⊕⊕⊖⊖ LOW10,12 due to risk of bias, imprecision	RR 1.12 (0.74 to 1.71)	347 per 1000	42 more per 1000 (from 90 fewer to 247 more)
Laxative misuse	149 (1 study)	⊕⊕⊖⊖ LOW10,12 due to risk of bias, imprecision	RR 1.05 (0.43 to 2.58)	111 per 1000	6 more per 1000 (from 63 fewer to 176 more)
Depression <18 binges month	50 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression <18 binges month in the intervention groups was 0.55 standard deviations higher (0.02 lower to 1.11 higher)
Depression >18 binges month	256 (4 studies)	⊕⊖⊖⊖ VERY LOW3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression >18 binges month in the intervention groups was 0.20 standard deviations lower (0.45 lower to 0.04 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

1 It was unclear if allocation concealment was conducted. Assessors were blind, but it was unclear if either participants or investigators were blind in two studies, but in Wilson 1991 it was unclear if any were blind and high drop outs were reported >20%.
2 Heterogeneity was detected I2 >50%
3 For a continuous outcome there were fewer than 400 participants.
4 It was unclear if allocation concealment was conducted in all studies. In Ghaderi and Bulike it was unclear how randomisation was conducted. Across

studies, it was either unclear whether the assessors, participants or investigators were blind, in Chen participants were not blind and Bulik assessors were blind. High dropouts were reported >20%.
5 It was unclear if allocation concealment was conducted. Only participants were not blind in study by Chen, it was not clear in investigators or assessors were blind, but it was unclear in other study/ies. High drop outs were reported >20%.
6 95% CI crossed1MID (-0.05).
7 It was unclear if allocation concealment was conducted. Across studies, it was unclear if all or only participants, investigators or assessors were blind. High drop outs were reported >20%.
8 It was unclear if allocation concealment was conducted. Across studies, it was unclear if all or only participants, investigators or assessors were blind.
9 It was unclear how randomisation was conducted or if allocation concealment was performed. Assessors were blind but it was unclear if participants or investigators were blind.
10 95% CI crossed 1 MID (1.25).
11 95% CI crossed 1 MID (0.5)
12 It was unclear if allocation concealment was performed or if participants were blind.

1 **Table 179: Summary of findings table for CBT-ED versus another CBT-ED program at follow up in adults with bulimia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT-ED (2) - Follow up	Risk difference with BN CBT-ED (1) (95% CI)
Depression Follow up	280 (4 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression follow up in the intervention groups was 0.00 standard deviations higher (0.23 lower to 0.24 higher)
Symptom check list - 90 Follow up	269 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom check list - 90 follow up in the intervention groups was 0.09 standard deviations higher (0.15 lower to 0.33 higher)
Bingeing episodes (28 d) FU	149 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision	RR 1.13 (0.68 to 1.9)	264 per 1000	34 more per 1000 (from 84 fewer to 237 more)
Vomiting (28 d) Follow up	149 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision	RR 1.18 (0.76 to 1.84)	319 per 1000	57 more per 1000 (from 77 fewer to 268 more)
Laxative misuse	149 (1 study)	⊕⊕⊖⊖ LOW4,7	RR 1.4 (0.53 to	83 per 1000	33 more per 1000 (from 39 fewer to 228 more)

		due to risk of bias, imprecision	3.74)		
Purging (28 d) Follow up	149 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision	RR 1.21 (0.79 to 1.85)	333 per 1000	70 more per 1000 (from 70 fewer to 283 more)
Bingeing Follow up	280 (4 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing follow up in the intervention groups was 0.01 standard deviations lower (0.25 lower to 0.22 higher)
Laxatives Follow up	72 (1 study)	⊕⊕⊖⊖ LOW8,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxatives follow up in the intervention groups was 0.12 standard deviations lower (0.58 lower to 0.34 higher)
Vomiting Follow up	232 (3 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting follow up in the intervention groups was 0.1 standard deviations higher (0.16 lower to 0.35 higher)
Purging (last 2 weeks) Follow up	111 (2 studies)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging (last 2 weeks) follow up in the intervention groups was 0.09 standard deviations higher (0.29 lower to 0.46 higher)
General psychiatric features - FU	149 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychiatric features - fu in the intervention groups was 0.05 standard deviations higher (0.28 lower to 0.37 higher)
Global Function (GAFS)	72 (1 study)	⊕⊕⊖⊖ LOW8,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean global function (gafs) in the intervention groups was 0.51 standard deviations higher (0.04 to 0.98 higher)
Social adjustment score Follow up	170 (2 studies)	⊕⊕⊖⊖ LOW1,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean social adjustment score follow up in the intervention groups was 0.44 standard deviations higher (0.14 to 0.75 higher)
EDI- Bulimia Follow up	122 (2 studies)	⊕⊕⊖⊖ LOW8,9 due to risk of bias,		Not calculable for SMD values	The mean edi- bulimia follow up in the intervention groups was 0.21 standard deviations lower

		imprecision			(0.57 lower to 0.15 higher)
EDI- Body dissatisfaction Follow up	122 (1 study)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- body dissatisfaction follow up in the intervention groups was 0.10 standard deviations higher (0.25 lower to 0.46 higher)
EDI- Drive for thinness Follow up	72 (1 study)	⊕⊕⊖⊖ LOW8,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- drive for thinness follow up in the intervention groups was 0.26 standard deviations higher (0.2 lower to 0.73 higher)
EDI- Total Follow up	319 (3 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- total follow up in the intervention groups was 0.02 standard deviations lower (0.24 lower to 0.2 higher)
EDE - Total - Follow up	237 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - total - follow up in the intervention groups was 0.10 standard deviations lower (0.35 lower to 0.16 higher)
Remission - FU _ ITT	144 (3 studies)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR 1.30 (0.93 to 1.83)	408 per 1000	123 more per 1000 (from 29 fewer to 339 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. Across studies, it was unclear if either or all participants, investigators or assessors were blind.
2 Heterogeneity was detected 12 >50%
3 For a continuous outcome, fewer than 400 participants were available.
[4] It was unclear if allocation concealment was conducted. Both investigators and assessors were blind but it was unclear if participants were blind.
[5] 95% CI crossed 1 MID (0.75).
[6] 95% CI crossed 1 MID (1.25)
[7] 95% CI crossed 2 MIDs (0.75 and 1.25)
[8] It was unclear if allocation concealment was conducted. Assessors were blind but it was unclear if participants or investigators were blind.
[9] 95% CI crossed 1 MID (-0.5)

[10] 95% CI crossed 1 MID (0.5)

Table 180: Summary of findings table for behavioural therapy versus any other intervention at the end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BN BT (95% CI)
Bulimic episodes	183 (3 studies)	⊕⊕⊖⊖ LOW[1,2] due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic episodes in the intervention groups was 0.10 standard deviations lower (0.41 lower to 0.21 higher)
Depression	185 (5 studies)	⊕⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.36 standard deviations higher (0.25 lower to 0.98 higher)
Laxative use (no. tablets)	92 (1 study)	⊕⊕⊖⊖ LOW[4,5] due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative use (no. tablets) in the intervention groups was 0.33 standard deviations lower (0.77 lower to 0.11 higher)
Vomiting	160 (3 studies)	⊕⊖⊖⊖ VERY LOW[1,2,6] due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.52 standard deviations lower (0.86 to 0.18 lower)
Symptom Checklist	62 (1 study)	⊕⊕⊖⊖ LOW[1,7] due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist in the intervention groups was 0.89 standard deviations lower (0.31 lower to 1.46 higher)
EDE - Dietary restraint	89 (2 studies)	⊕⊖⊖⊖ VERY LOW[1,6,7] due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - dietary restraint in the intervention groups was 0.92 standard deviations higher (0.60 lower to 2.43 higher)
EDE - Attitudes towards weight	89 (2 studies)	⊕⊖⊖⊖ VERY LOW[1,6,8] due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - attitudes towards weight in the intervention groups was 2.23 standard deviations higher (0.68 lower to 5.15 higher)

EDE - Attitudes towards shape	89 (2 studies)	⊕⊖⊖⊖ VERY LOW1,6,7 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - attitudes towards shape in the intervention groups was 1.87 standard deviations higher (0.47 lower to 4.21 higher)
EDI - Bulimia	139 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 0.42 standard deviations lower (0.78 to 0.06 lower)
EDI - Drive for thinness	139 (2 studies)	⊕⊕⊖⊖ LOW8,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness in the intervention groups was 1.64 standard deviations lower (2.05 to 1.22 lower)
EDI - Body dissatisfaction	149 (2 studies)	⊕⊖⊖⊖ VERY LOW3,5,10 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction in the intervention groups was 1.21 standard deviations lower (2.27 to 0.16 lower)
Social adjustment scale	62 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean social adjustment scale in the intervention groups was 0.48 standard deviations higher (0.47 lower to 1.44 higher)
Remission - ITT	106 (3 studies)	⊕⊖⊖⊖ VERY LOW1,8 due to risk of bias, imprecision	RR 1.01 (0.6 to 1.69)	364 per 1000	4 more per 1000 (from 145 fewer to 251 more)
Vomiting <5 years or <18 binges/mo	41 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting <5 years or <18 binges/mo in the intervention groups was 1.81 standard deviations lower (2.55 to 1.07 lower)
Vomiting >5 years or .18 binges/mo	119 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting >5 years or .18 binges/mo in the intervention groups was 0.18 standard deviations lower (0.56 lower to 0.20 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU:Follow up

1 It was unclear how randomisation was conducted or if allocation concealment was performed. Assessors were blind but it was unclear if investigators or

participants were blind. High drop outs were reported >20%.
2 For a continuous outcome, there were fewer than 400 participants.
3 Heterogeneity was detected I2 >50%
4 It was unclear allocation concealment was performed. In Freeman, it was unclear if either participants, investigators or assessors were blind. In Thackway, the assessors were blind. High drop outs were reported >20%.
5 95% CI crossed 1 MID (-0.5)
6 Heterogeneity was detected I2 >80%
7 95% CI crossed 1 MID (0.5)
8 95% CI Crossed 2 MIDs (0.75 and 1.25).
9 It was unclear how randomisation sequence was conducted or if allocation concealment was conducted. Only assessors were blind.
10 It was unclear how random sequence was generated or if allocation concealment was performed. It was unclear if participants and investigators were blind, the assessors were blind.

Table 181: Summary of findings table for behavioural therapy versus any other intervention at follow up in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention Follow up	Risk difference with BN BT (95% CI)
Vomiting or purging FU	27 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting or purging fu in the intervention groups was 1.00 standard deviations higher (0.19 to 1.80 higher)
Bulimic episodes FU	27 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic episodes fu in the intervention groups was 0.93 standard deviations higher (0.13 to 1.73 higher)
EDE - Dietary restraint FU	27 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - dietary restraint fu in the intervention groups was 0.45 standard deviations higher (0.32 lower to 1.21 higher)
EDE- Shape concerns FU	27 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concerns fu in the intervention groups was 0.35 standard deviations higher (0.42 lower to 1.11 higher)
EDE - Weight concerns FU	27 (1 study)	⊕⊕⊖⊖ LOW1,3		Not calculable for SMD values	The mean ede - weight concerns fu in the intervention groups was

		due to risk of bias, imprecision			0.07 standard deviations higher (0.69 lower to 0.82 higher)
Depression FU	74 (2 studies)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.04 standard deviations higher (0.44 lower to 0.53 higher)
EDI - Drive for thinness FU	47 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness fu in the intervention groups was 0.78 standard deviations lower (1.41 to 0.15 lower)
EDI- Body dissatisfaction FU	27 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- body dissatisfaction fu in the intervention groups was 0.36 standard deviations higher (0.40 lower to 1.12 higher)
EDI - Bulimia FU	47 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia fu in the intervention groups was 0.34 standard deviations lower (0.96 lower to 0.28 higher)
Remission FU_ITT	75 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR 0.50 (0.21 to 1.18)	40 per 100	20 fewer per 100 (from 32 fewer to 7 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 It was unclear how randomisation sequence was generated or if allocation concealment was conducted. Assessors were blind but it was unclear if investigators or participants were blind. High drop outs were reported >20%
2 95% CI crossed 1 MID (0.5)
3 For a continuous outcome, there were fewer than 400 participants.
4 It was unclear how randomisation sequence was generated or if allocation concealment was conducted. Across studies, it was unclear if either or all of the investigators, participants and assessors were blind. High drop outs were reported >20%.

5 95% CI crossed 1 MID (-0.5)
6 95% CI crossed 1 MID (0.75)

Table 182: Summary of findings table for behavioural therapy (BT) versus wait list control (WLC) at the end of treatment for adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN BT (95% CI)
Binge frequency	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 1.11 standard deviations lower (1.72 to 0.5 lower)
Self-induced vomiting	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean self-induced vomiting in the intervention groups was 0.76 standard deviations lower (1.34 to 0.17 lower)
Laxative use (no. tablets)	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative use (no. tablets) in the intervention groups was 0.75 standard deviations lower (1.33 to 0.16 lower)
Depression	34 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.04 standard deviations higher (0.64 lower to 0.71 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear how random sequence was generated and if allocation concealment was conducted. It was unclear if either participants, assessors or investigators were blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (-0.5).
3 It was unclear if allocation concealment was conducted. It was unclear if either participants, assessors or investigators were blind. High dropouts were reported >20%.
4 95% CI crossed 2 MIDs (-0.5 and 0.5).

Table 183: Summary of findings table for hybrid treatment versus other intervention at the end of treatment and at follow up in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BN other/hybrid (95% CI)
Binge Eating	86 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating in the intervention groups was 0.21 standard deviations lower (0.63 lower to 0.21 higher)
Symptom check list - 90	86 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom check list - 90 in the intervention groups was 0 standard deviations higher (0.42 lower to 0.42 higher)
Depression - Becks	86 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - becks in the intervention groups was 0.3 standard deviations lower (0.73 lower to 0.12 higher)
EDI - 1-6 ED symptoms	86 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - 1-6 ed symptoms in the intervention groups was 0.07 standard deviations lower (0.49 lower to 0.35 higher)
Binge Eating - Follow up	86 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating - follow up in the intervention groups was 0.36 standard deviations lower (0.79 lower to 0.07 higher)
Symptom check list - 90 Follow up	86 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom check list - 90 follow up in the intervention groups was 0 standard deviations higher (0.42 lower to 0.42 higher)
Depression - Becks Follow up	86 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - becks follow up in the intervention groups was 0.16 standard deviations lower (0.58 lower to 0.26 higher)
EDI -1-6 Follow up	86 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean edi -1-6 follow up in the intervention groups was 0.18 standard deviations lower

		imprecision			(0.6 lower to 0.25 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear how randomisation sequence was generated or if allocation concealment was conducted. Assessors were blind but it was unclear if investigators or participants were blind.
2 95% CI crossed 1 MID (-0.5).
3 For a continuous outcome, fewer than 400 participants were included.

Table 184: Summary of findings table for CBT-ED versus wait list control (WLC) at the end of treatment for adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN CBT-ED (95% CI)
Laxative use (no. tablets)	52 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative use (no. tablets) in the intervention groups was 0.36 standard deviations lower (0.68 to 0.05 lower)
Bingeing	113 (3 studies)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 1.35 standard deviations lower (1.79 to 0.91 lower)
Purge frequency	21 (1 study)	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean purge frequency in the intervention groups was 2.00 standard deviations lower (3.08 to 0.91 lower)
Vomiting	92 (2 studies)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 1.56 standard deviations lower (2.03 to 1.08 lower)
Overall severity	194	⊕⊕⊖⊖		Not	The mean overall severity in the intervention

	(2 studies)	LOW3,7 due to risk of bias, imprecision		calculable for SMD values	groups was 1.92 standard deviations lower (2.28 to 1.56 lower)
EDI - Body dissatisfaction	41 (1 study)	⊕⊕⊖⊖ LOW2,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction in the intervention groups was 0.37 standard deviations lower (0.99 lower to 0.25 higher)
EDI - Drive for thinness	41 (1 study)	⊕⊕⊖⊖ LOW2,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness in the intervention groups was 1.02 standard deviations lower (1.68 to 0.36 lower)
EDI - Bulimia	41 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 1.48 standard deviations higher (2.18 to 0.78 lower)
Symptom checklist - 90 items	154 (1 study)	⊕⊕⊖⊖ LOW2,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist - 90 items in the intervention groups was 0.71 standard deviations lower (1.05 to 0.36 lower)
General psychiatric features	123 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychiatric features in the intervention groups was 0.81 standard deviations lower (1.18 to 0.43 lower)
Depression	35 (1 study)	⊕⊕⊖⊖ LOW7,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 1.43 standard deviations lower (2.18 to 0.67 lower)
Vomiting episodes	153 (1 study)	⊕⊕⊖⊖ LOW9,11 due to risk of bias, imprecision	RR 0.84 (0.62 to 1.13)	600 per 1000	96 fewer per 1000 (from 228 fewer to 78 more)

Purging	154 (1 study)	⊕⊕⊖⊖ LOW11,12 due to risk of bias, imprecision	RR 0.82 (0.63 to 1.08)	647 per 1000	116 fewer per 1000 (from 239 fewer to 52 more)
Laxative misuse	154 (1 study)	⊕⊕⊖⊖ LOW11,12 due to risk of bias, imprecision	RR 0.65 (0.34 to 1.23)	255 per 1000	89 fewer per 1000 (from 168 fewer to 59 more)
EDE - Shape concern	40 (1 study)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - shape concern in the intervention groups was 2.44 standard deviations lower (3.28 to 1.6 lower)
EDE - Weight concern	40 (1 study)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - weight concern in the intervention groups was 2.44 standard deviations lower (3.28 to 1.6 lower)
Bulimic episodes	154 (1 study)	⊕⊕⊖⊖ LOW11,12 due to risk of bias, imprecision	RR 0.81 (0.57 to 1.13)	529 per 1000	101 fewer per 1000 (from 228 fewer to 69 more)
EDE - Dietary Restraint	40 (1 study)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - dietary restraint in the intervention groups was 1.52 standard deviations lower (2.24 to 0.81 lower)
Did not achieve remission ITT	81 (1 study)	⊕⊕⊖⊖ LOW3,11 due to risk of bias, imprecision	RR 0.90 (0.77 to 1.06)	74 per 1000	7 fewer per 1000 (from 17 fewer to 4 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 It was unclear if allocation concealment was performed or if participants were blind.

2 95% CI crossed 1 MID (-0.5)
3 It was unclear if allocation concealment was performed. Across studies it was unclear if either or all of the participants, investigators or assessors were blind. High drop outs were reported >20%.
4 Heterogeneity >80%
5 It was unclear if allocation concealment was performed. In Agras 1999, assessors were blind but it was unclear if either participants or investigators were blind. It was unclear in Treasure 1994 if any were blind. High drop outs were reported >20%.
6 It was unclear if allocation concealment was conducted or if either the participants, investigators or assessors were blind. High drop outs were reported >20%.
7 For a continuous outcome, there were fewer than 400 participants.
8 It was unclear how random sequence was generated or if allocation concealment was conducted. It was unclear if either participants, investigators or assessors were blind. High drop outs were reported >20%
9 It was unclear how random sequence was generated or if allocation concealment was conducted. Participants were blind but it was unclear if assessors or investigators were blind. High drop outs were reported >20%
10 It was unclear if allocation concealment was performed. Assessors were blind but it was unclear if either participants or investigators were blind. High drop outs were reported >20%.
11 95% CI Crossed 1 MID (0.75)
12 It was unclear if allocation concealment was conducted. Assessors and investigators were blind but it was unclear if participants were blind.

Table 185: Summary of findings table for dialectical behaviour therapy (DBT) versus another other intervention at the end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BN DBT (95% CI)
Negative mood regulation score	29 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean negative mood regulation score in the intervention groups was 0.33 standard deviations lower (1.07 lower to 0.4 higher)
Depression- Becks	29 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression- becks in the intervention groups was 0.91 standard deviations lower (1.68 to 0.14 lower)
Emotional eating - anger/anxiety/depression	29 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean emotional eating - anger/anxiety/depression in the intervention groups was 0.7 standard deviations lower (1.46 lower to 0.07 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%

confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear if either participants, investigators or assessors were blind.
2 95% CI crossed 1 MID (-0.5).

Table 186: Summary of findings table for psychodynamic therapy compared to another intervention at the end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BN Psychodynamic General (95% CI)
Binge eating (28/d)	116 (2 studies)	⨁⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean binge eating (28/d) in the intervention groups was 1.02 standard deviations higher (0.60 lower to 2.65 higher)
Vomiting/purging episodes (28d)	120 (2 studies)	⨁⊖⊖⊖ VERY LOW1,2,3,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean vomiting/purging episodes (28d) in the intervention groups was 1.46 standard deviations higher (0.05 lower to 2.97 higher)
EDE - Attitudes towards weight	120 (2 studies)	⨁⊖⊖⊖ VERY LOW1,2,3,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - attitudes towards weight in the intervention groups was 0.02 standard deviations higher (1.25 lower to 1.30 higher)
EDE - Dietary restraint	120 (2 studies)	⨁⨁⊖⊖ LOW1,2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - dietary restraint in the intervention groups was 0.75 standard deviations higher (0.38 to 1.12 higher)
EDE - Attitudes towards shape	120 (2 studies)	⨁⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede - attitudes towards shape in the intervention groups was 0.71 standard deviations lower (3.56 lower to 2.13 higher)
EDI - Drive for thinness	50 (1 study)	⨁⨁⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness in the intervention groups was 0.53 standard deviations higher (0.04 lower to 1.09 higher)
EDI -Bulimia	49	⨁⨁⊖⊖		Not calculable for	The mean edi -bulimia in the intervention

	(1 study)	LOW1,5 due to risk of bias, imprecision		SMD values	groups was 0.61 standard deviations higher (0.03 to 1.18 higher)
EDI - Body dissatisfaction	49 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction in the intervention groups was 0.24 standard deviations higher (0.33 lower to 0.8 higher)
Depression	70 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.78 standard deviations lower (1.27 to 0.29 lower)
General psychopathology	70 (1 study)	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0.36 standard deviations higher (0.11 lower to 0.83 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear if allocation concealment was performed. Participants or investigators were not blind and it was unclear if assessors were blind.
2 In Poulsen, it was unclear if participants or investigators were blind. Low drop outs. There was also a large difference in the duration of therapy, CBT-ED was 5 months versus psychodynamic was 19 months.
3 Heterogeneity detected >80%
4 95% CI crossed 1 MID (-0.5)
5 95% CI crossed 1 MID (0.5)
6 95% CI crossed 2 MIDs (-0.5 and 0.5)

7.2.6.1 Group therapy

Table 187: Summary of findings table for group behavioural therapy versus an alternative group behavioural therapy at the end of treatment and follow up in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with BT.2 (ED)	Risk difference with BN Group BT (ED) (95% CI)
Vomiting	23 (1 study)	⊕⊕⊖⊖ VERY LOW1,2		Not calculable for SMD	The mean vomiting in the intervention groups was

		due to risk of bias, imprecision		values	0.06 standard deviations lower (0.87 lower to 0.76 higher)
Depression	23 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.35 standard deviations higher (0.48 lower to 1.17 higher)
Remission_ITT	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 1.00 (0.31 to 3.28)	267 per 1000	0 fewer per 1000 (from 184 fewer to 608 more)
Vomiting FU	24 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting fu in the intervention groups was 0.65 standard deviations lower (1.48 lower to 0.17 higher)
Depression FU	23 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.47 standard deviations higher (0.36 lower to 1.3 higher)
Remission_ITT FU	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 2.50 (0.57 to 10.93)	133 per 1000	200 more per 1000 (from 57 fewer to 1000 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how they randomised or if they performed allocation concealment. It was unclear if either the participants, investigators or assessors were blinded. High dropout rates were detected >20% and a difference of greater than 10% in dropout rates were detected between two of the groups.
2 95% CI crossed 2 MIDs (-0.5 and 0.5).
3 95% CI crossed 1 MID (0.5).
4 95% CI Crossed 2 MIDs (0.75 and 1.25).
5 95% CI crossed 1 MID (-0.5).

Table 188: Summary of findings table for CBT-ED versus wait list controls at the end of treatment and follow up in adults with bulimia nervosa

Outcomes	No of Participants (studies)	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN Group CBT-

	Follow up				ED (95% CI)
Bingeing frequency	54 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing frequency in the intervention groups was 0.43 standard deviations lower (0.97 lower to 0.12 higher)
Purges (per week)	28 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean purges (per week) in the intervention groups was 0.33 standard deviations lower (1.08 lower to 0.42 higher)
Vomiting	24 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.9 standard deviations lower (1.74 to 0.05 lower)
Depression	24 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 1.81 standard deviations lower (2.79 to 0.84 lower)
EDI- Drive for thinness	26 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- drive for thinness in the intervention groups was 0.66 standard deviations lower (1.46 lower to 0.15 higher)
EDI - Bulimia	26 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - bulimia in the intervention groups was 0.38 standard deviations lower (1.17 lower to 0.4 higher)
EDI- Body Dissatisfaction	26 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- body dissatisfaction in the intervention groups was 0.67 standard deviations lower (1.47 lower to 0.13 higher)
No_Remission_ITT	52 (2 studies)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 0.86 (0.72 to 1.04)	38 per 1000	5 fewer per 1000 (from 11 fewer to 2 more)
No_Remission_ITT FU	59 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 0.72 (0.55 to 0.94)	69 per 1000	19 fewer per 1000 (from 4 fewer to 31 fewer)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how randomisation was performed or if allocation concealment was performed. Neither the participant's investigators nor assessors were blind. High dropout rates were detected >20% and a difference of >10% was detected between the two groups in Less 1986.
2 95% CI crossed 1 MID (-0.5).
3 For a continuous outcome, there were fewer than 400 participants.
4 For a dichotomous outcome, there were fewer than 300 events.
5 95% CI crossed 1 MID (0.75).

Table 189: Summary of findings table for group CBT-ED versus any other intervention at the end of treatment and follow up in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other Intervention	Risk difference with BN Group CBT (ED) (95% CI)
Bingeing frequency	206 (3 studies)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing frequency in the intervention groups was 0.08 standard deviations higher (0.19 lower to 0.36 higher)
EDI- Drive for thinness	206 (3 studies)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- drive for thinness in the intervention groups was 0.15 standard deviations higher (0.13 lower to 0.42 higher)
EDI - Bulimia	206 (3 studies)	⨁⊖⊖⊖ VERY LOW1,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 0.14 standard deviations higher (0.44 lower to 0.72 higher)
EDI- Body Dissatisfaction	206 (3 studies)	⨁⊖⊖⊖ VERY LOW1,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean edi- body dissatisfaction in the intervention groups was 0.16 standard deviations higher (0.33 lower to 0.66 higher)
EDI-Global	145 (2 studies)	⨁⊖⊖⊖ VERY LOW3,4,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean edi-global in the intervention groups was 0.07 standard deviations lower (0.57 lower to 0.42 higher)

EDE-Total	120 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-total in the intervention groups was 0.13 standard deviations higher (0.23 lower to 0.49 higher)
Clinical impairment	0 (1 study)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical impairment in the intervention groups was 1.02 standard deviations lower (1.54 to 0.51 lower)
Symptom checklist	120 (1 study)	⊕⊕⊖⊖ LOW2,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist in the intervention groups was 0.07 standard deviations higher (0.27 lower to 0.43 higher)
Depression	211 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.07 standard deviations higher (0.21 lower to 0.34 higher)
Anxiety	120 (1 study)	⊕⊕⊖⊖ LOW2,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean anxiety in the intervention groups was 0.11 standard deviations lower (0.47 lower to 0.25 higher)
Vomiting	91 (2 studies)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.45 standard deviations higher (0.02 to 0.87 higher)
Laxatives	56 (1 study)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxatives in the intervention groups was 0.55 standard deviations higher (0.02 to 1.09 higher)
No_Remission_ITT	81 (1 study)	⊕⊕⊖⊖ LOW9,10 due to risk of bias, imprecision	RR 0.95 (0.86 to 1.05)	25 per 1000	1 fewer per 1000 (from 3 fewer to 1 more)
Binging frequency FU	205 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binging frequency fu in the intervention groups was 0.07 standard deviations higher (0.21 lower to 0.34 higher)
EDI- Body Dissatisfaction	205	⊕⊕⊖⊖		Not calculable for	The mean edi- body dissatisfaction fu in

FU	(3 studies)	LOW1,2 due to risk of bias, imprecision		SMD values	the intervention groups was 0.25 standard deviations lower (0.53 lower to 0.02 higher)
EDI - Bulimia FU	205 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia fu in the intervention groups was 0.06 standard deviations lower (0.33 lower to 0.22 higher)
EDI-Global FU	74 (1 study)	⊕⊕⊖⊖ LOW7,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-global fu in the intervention groups was 0.1 standard deviations lower (0.15 to 0.05 lower)
EDI- Drive for thinness FU	205 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- drive for thinness fu in the intervention groups was 0.11 standard deviations lower (0.39 lower to 0.16 higher)
EDE-Total FU	120 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-total fu in the intervention groups was 0.03 standard deviations lower (0.39 lower to 0.32 higher)
Vomiting FU	91 (2 studies)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting fu in the intervention groups was 0.38 standard deviations higher (0.05 lower to 0.81 higher)
Depression FU	210 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.04 standard deviations lower (0.31 lower to 0.24 higher)
Laxatives FU	55 (1 study)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxatives fu in the intervention groups was 0.59 standard deviations higher (0.05 to 1.13 higher)
Anxiety FU	120 (1 study)	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean anxiety fu in the intervention groups was 0.41 standard deviations lower (0.78 to 0.05 lower)
Symptom checklist FU	120 (1 study)	⊕⊕⊖⊖ LOW2,8		Not calculable for SMD values	The mean symptom checklist fu in the intervention groups was

		due to risk of bias, imprecision			0.14 standard deviations lower (0.49 lower to 0.22 higher)
Remission_ITT FU	126 (2 studies)	⊕⊕⊖⊖ LOW9,11 due to risk of bias, imprecision	RR 0.70 (0.32 to 1.56)	200 per 1000	60 fewer per 1000 (from 136 fewer to 112 more)
Clinical impairment FU	74 (1 study)	⊕⊕⊖⊖ LOW2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical impairment fu in the intervention groups was 2.29 standard deviations lower (3.43 to 1.15 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 In some studies was unclear how randomisation was performed and in all studies it was unclear if allocation concealment was performed. It was either unclear or the participants, investigators or assessors were blind. High dropout rates were detected >20%.
2 For a continuous outcome, there were fewer than 400 participants.
3 95% CI crossed 1 MID (-0.5)
4 Heterogeneity was detected, I2 >50%
5 95% CI crossed 1 MID (0.5)
6 It was unclear if allocation concealment was performed. The participants were not blinded and it was unclear if the investigators and assessors were blind.
7 It was unclear if allocation concealment was performed. The participants were not blinded, however, the investigators and assessors were blinded. It was unclear what the number of completers were.
8 It was unclear if allocation concealment was performed. Participants were not blinded in Chen, and It was either unclear in Wolf. It was also unclear if the investigators or assessors were blind.
9 It was unclear if allocation concealment was performed. It was unclear if the participants, investigators and assessors were blind. High dropout rates were detected >20% and a difference in dropout rates of more than 10%.
10 For a dichotomous outcome, there were fewer than 300 events.
11 95% CI Crossed 2 MIDs (0.75 and 1.25)

Table 190: Summary of findings table for group behavioural therapy versus wait list control (WLC) at the end of treatment and follow up in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN Group BT(ED) (95% CI)

Bingeing frequency	26 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing frequency in the intervention groups was 0.15 standard deviations higher (0.63 lower to 0.93 higher)
Vomiting	35 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 1.22 standard deviations lower (1.99 to 0.45 lower)
EDI- Drive for thinness	26 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- drive for thinness in the intervention groups was 0.39 standard deviations lower (1.17 lower to 0.4 higher)
EDI - Bulimia	26 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - bulimia in the intervention groups was 0.2 standard deviations higher (0.58 lower to 0.98 higher)
EDI- Body Dissatisfaction	26 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- body dissatisfaction in the intervention groups was 0.73 standard deviations lower (1.54 lower to 0.08 higher)
Depression	35 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 1.37 standard deviations lower (2.17 to 0.58 lower)
Did not achieve remission_ITT	44 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 0.77 (0.6 to 0.99)	0 per 1000	-
Remission_ITT FU	44 (1 study)	⊕⊖⊖⊖ VERY LOW1,6 due to risk of bias, imprecision	RR 1.07 (0.73 to 1.58)	286 per 1000	20 more per 1000 (from 77 fewer to 166 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how they randomised or if they performed allocation concealment. It was unclear whether the participants, investigators or assessors were blinded. High

dropout rates were detected >20%.
2 95% CI crossed 2 MIDs (-0.5 and 0.5).
3 95% CI crossed 1 MID (-0.5).
4 For a continuous outcome, there were fewer than 400 participants.
5 95% CI crossed 1 MID (0.75).
6.95% CI crossed 2 MIDs (0.75 and 1.25)

Table 191: Summary of findings table for group behavioural therapy versus another group intervention at the end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other Group	Risk difference with BN Group BT (ED) (95% CI)
Bingeing frequency	30 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing frequency in the intervention groups was 0.33 standard deviations higher (0.39 lower to 1.06 higher)
Vomiting	36 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.27 standard deviations lower (0.97 lower to 0.43 higher)
Depression	35 (1 study)	⊕⊕⊖⊖ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.16 standard deviations higher (0.54 lower to 0.86 higher)
EDI- Drive for thinness	30 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- drive for thinness in the intervention groups was 0.25 standard deviations higher (0.47 lower to 0.97 higher)
EDI - Bulimia	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 0.51 standard deviations higher (0.22 lower to 1.24 higher)
EDI- Body Dissatisfaction	30 (1 study)	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- body dissatisfaction in the intervention groups was 0.08 standard deviations lower (0.79 lower to 0.64 higher)

Did not achieve remission	60 (1 study)	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision	RR 0.76 (0.61 to 0.96)	33 per 1000	8 fewer per 1000 (from 1 fewer to 13 fewer)
Bingeing frequency FU	58 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing frequency fu in the intervention groups was 0.01 standard deviations lower (0.53 lower to 0.52 higher)
Vomiting FU	36 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting fu in the intervention groups was 0.38 standard deviations lower (1.08 lower to 0.33 higher)
Depression FU	63 (2 studies)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.13 standard deviations higher (0.39 lower to 0.65 higher)
EDI - Drive for thinness FU	30 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness fu in the intervention groups was 0.24 standard deviations higher (0.48 lower to 0.96 higher)
EDI - Bulimia FU	30 (1 study)	⊕⊖⊖⊖ VERY LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia fu in the intervention groups was 0.02 standard deviations higher (0.69 lower to 0.74 higher)
EDI- Body Dissatisfaction FU	30 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- body dissatisfaction fu in the intervention groups was 0.35 standard deviations higher (0.37 lower to 1.07 higher)
EDE- Shape concern FU	28 (1)	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern fu in the intervention groups was 0 standard deviations higher (0.77 lower to 0.77 higher)
EDE- Weight concern FU	28 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern fu in the intervention groups was 0.34 standard deviations higher (0.44 lower to 1.12 higher)
EDE- Eating concern FU	20	⊕⊖⊖⊖		Not calculable	The mean ede- eating concern fu in the

	(1 study)	VERY LOW1,5 due to risk of bias, imprecision		for SMD values	intervention groups was 0 standard deviations higher (0.88 lower to 0.88 higher)
EDE- Restraint FU	28 (1 study)	⊕⊖⊖⊖ VERY LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint fu in the intervention groups was 0 standard deviations higher (0.77 lower to 0.77 higher)
Remission_ITT FU	73 (2 studies)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision	RR 0.85 (0.53 to 1.35)	576 per 1000	86 fewer per 1000 (from 271 fewer to 202 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear methods of randomisation and allocation concealment. Neither the participants, investigators nor assessors were blinded.
2 95% CI crossed 1 MID (0.5).
3 Unclear how randomisation was performed or if allocation concealment was conducted. It was unclear if either the participants, investigators or assessors were blind. High dropouts >20% were reported in some groups.
4 95% CI crossed 1 MID (-0.5).
5 95% CI crossed 2 MIDs (0.5 and -0.5).
6 95% CI crossed 1 MID (0.75).
7 95% CI crossed 2 MIDs (0.75 and 1.25).

Table 192: Summary of findings table for group psychoeducation versus any other intervention at the end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with BN Group psychoeducation vs.Other (95% CI)
Bingeing	54 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.2 standard deviations higher (0.33 lower to 0.74 higher)
Vomiting	54 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.44 standard deviations higher (0.11 lower to 0.98 higher)

		imprecision			
Remission_ITT	65 (1 study)	⊕⊖⊖⊖ VEY LOW1,3 due to risk of bias, imprecision	RR 0.57 (0.23 to 1.42)	300 per 1000	129 fewer per 1000 (from 231 fewer to 126 more)
EDI-Drive for thinness	54 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-drive for thinness in the intervention groups was 0.62 standard deviations higher (0.08 to 1.17 higher)
EDI-Bulimia	54 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-bulimia in the intervention groups was 0.5 standard deviations higher (0.05 lower to 1.04 higher)
EDI-Body dissatisfaction	54 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-body dissatisfaction in the intervention groups was 0.12 standard deviations higher (0.41 lower to 0.66 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Neither the participants, investigators nor assessors appear blinded. There were differences detected at baseline, however a correlations analysis suggested it had no impact on the outcomes.
2 95% CI crossed 1 MID (0.5).
3 95% CI crossed 2 MIDs (0.75 and 1.25).

Table 193: Summary of findings table for group CBT-ED (varied intensity) versus CBT (control low intensity) at the end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT (control low)	Risk difference with BN CBT (varied intensity and focus) (95% CI)
Binging episodes	143 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binging episodes in the intervention groups was 0.37 standard deviations lower (0.76 lower to 0.02 higher)
Laxative use	143	⊕⊕⊖⊖		Not calculable for	The mean laxative use in the intervention

	(1 study)	LOW1,2 due to risk of bias, imprecision		SMD values	groups was 0.10 standard deviations higher (0.29 lower to 0.49 higher)
Vomiting episodes	143 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting episodes in the intervention groups was 0.4 standard deviations lower (0.79 to 0.01 lower)
EDI - Drive for thinness	143 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - drive for thinness in the intervention groups was 0.49 standard deviations lower (0.88 to 0.1 lower)
EDI - Bulimia	143 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 0.85 standard deviations lower (1.25 to 0.45 lower)
EDI - Body dissatisfaction	143 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi - body dissatisfaction in the intervention groups was 0.03 standard deviations lower (0.41 lower to 0.36 higher)
Depression	143 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.1 standard deviations higher (0.29 lower to 0.48 higher)
Anxiety	143 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean anxiety in the intervention groups was 0.11 standard deviations higher (0.27 lower to 0.5 higher)
Did not achieve remission_ITT	143 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 0.42 (0.3 to 0.57)		106 fewer per 1000 (from 78 fewer to 127 fewer)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear method of randomisation and if allocation concealment was performed. Neither the participants, investigators nor assessors were blind.
2 95% CI crossed 1 MID (-0.5)

3 For a continuous variable, there were fewer than 400 participants.
4 For a dichotomous outcome, there were fewer than 300 participants.

Table 194: Summary of findings table for group emotional and mind training versus any other intervention at end of treatment and follow up in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BN Group Emotional and Mind Training (95% CI)
EDE-Global	74 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-global in the intervention groups was 0.1 standard deviations lower (0.59 lower to 0.39 higher)
EDE-Global FU	74 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-global fu in the intervention groups was 0.1 standard deviations higher (0.05 to 0.15 higher)
Clinical impairment	74 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical impairment in the intervention groups was 1.02 standard deviations higher (0.51 to 1.54 higher)
Clinical impairment FU	74 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical impairment fu in the intervention groups was 2.29 standard deviations higher (1.15 to 3.43 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Unclear if allocation concealment was performed. The participants were not blinded, however the investigators and assessors were blind. It was unclear how many participants dropped out of the study.
2 95% CI crossed 1 MID (-0.5).
3 For a continuous outcome, there were fewer than 400 participants.
4 95% CI crossed 1 MID (0.5).

Table 195: Summary of findings table for group support versus any other intervention at end of treatment in adults with bulimia nervosa

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with Other	Risk difference with BN Group Support (95% CI)
Change in depression scores	100 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean change in depression scores in the intervention groups was 0.06 standard deviations higher (0.4 lower to 0.52 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 It was unclear how random sequence was generated or if allocation concealment was performed. It was unclear if either the participants, investigators or assessors were blind. High dropouts were detected >20%. *2 95% CI crossed 1 MID (0.5).*					

7.2.6.2 Self-help

Table 196: Summary of findings table for guided self-help (ED) (or self-help with support) versus any other intervention at end of treatment in young people and adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BN Guided SH (ED) (95% CI)
Bingeing	388 (6 studies)	⨁⨀⨀⨀ VERY LOW1,2,3,4 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.26 standard deviations lower (0.58 lower to 0.06 higher)
Vomiting	190 (5 studies)	⨁⨀⨀⨀ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.18 standard deviations lower (0.4 lower to 0.05 higher)
Use of laxatives	243 (5 studies)	⨁⨁⨀⨀ LOW1,3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean use of laxatives in the intervention groups was 0.33 standard deviations lower (0.58 to 0.07 lower)
Depression	280 (5 studies)	⨁⨀⨀⨀ VERY LOW2,3,6,7		Not calculable for SMD values	The mean depression in the intervention groups was

		due to risk of bias, inconsistency, indirectness, imprecision			0.33 standard deviations higher (0.21 lower to 0.87 higher)
EDI Drive for thinness	56 (1 study)	⊕⊕⊖⊖ LOW5,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness in the intervention groups was 0.48 standard deviations lower (1.01 lower to 0.06 higher)
EDI Bulimia	56 (1 study)	⊕⊕⊖⊖ LOW4,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia in the intervention groups was 0.71 standard deviations lower (1.25 to 0.17 lower)
EDI Body dissatisfaction	55 (1 study)	⊕⊕⊖⊖ LOW4,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction in the intervention groups was 0.62 standard deviations lower (1.16 to 0.09 lower)
Remission - Young People_ITT	85 (1 study)	⊕⊖⊖⊖ VERY LOW9,10,11 due to risk of bias, indirectness, imprecision	RR 1.40 (0.42 to 4.6)	98 per 1000	39 more per 1000 (from 57 fewer to 351 more)
Remission - Adults_ITT	454 (4 studies)	⊕⊖⊖⊖ VERY LOW3,11,12 due to risk of bias, indirectness, imprecision	RR 1.01 (0.66 to 1.53)	162 per 1000	2 more per 1000 (from 55 fewer to 86 more)
EDE-Global	159 (3 studies)	⊕⊖⊖⊖ VERY LOW2,5,10,12 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean ede-global in the intervention groups was 0.10 standard deviations lower (0.41 lower to 0.22 higher)
EDE- Restraint	192 (3 studies)	⊕⊖⊖⊖ VERY LOW5,10,12 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.03 standard deviations higher (0.25 lower to 0.32 higher)
EDE- Weight concern	192 (3 studies)	⊕⊖⊖⊖ VERY LOW5,10,12 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.12 standard deviations lower (0.41 lower to 0.16 higher)
EDE- Shape concern	192 (3 studies)	⊕⊖⊖⊖ VERY LOW5,10,12 due to risk of bias, indirectness,		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.00 standard deviations lower

		imprecision			(0.29 lower to 0.28 higher)
EDE- Eating concern	145 (2 studies)	⨁⊖⊖⊖ VERY LOW5,10,12 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.02 standard deviations higher (0.31 lower to 0.35 higher)
Purging	80 (1 study)	⨁⊖⊖⊖ VERY LOW7,10,13 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.34 standard deviations higher (0.1 lower to 0.78 higher)
Exercising	187 (3 studies)	⨁⊖⊖⊖ VERY LOW5,10,14 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean exercising in the intervention groups was 0.02 standard deviations higher (0.27 lower to 0.31 higher)
Satisfaction with life	80 (1 study)	⨁⊖⊖⊖ VERY LOW4,10,13 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean satisfaction with life in the intervention groups was 0.25 standard deviations lower (0.69 lower to 0.19 higher)
Bulimic Inventory Index	112 (2 studies)	⨁⊖⊖⊖ VERY LOW2,7,15 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean bulimic inventory index in the intervention groups was 0.29 standard deviations higher (0.09 lower to 0.67 higher)
Bingeing FU	270 (4 studies)	⨁⊖⊖⊖ VERY LOW3,5,16 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.04 standard deviations higher (0.2 lower to 0.28 higher)
Vomiting FU	95 (3 studies)	⨁⊖⊖⊖ VERY LOW4,10,16 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean vomiting fu in the intervention groups was 0.25 standard deviations lower (0.66 lower to 0.16 higher)
Use of laxatives FU	216 (3 studies)	⨁⊖⊖⊖ VERY LOW4,10,16 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean use of laxatives fu in the intervention groups was 0.29 standard deviations lower (0.56 lower to 0.02 higher)
Depression FU	154 (3 studies)	⨁⊖⊖⊖ VERY LOW3,4,17 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.19 standard deviations lower (0.5 lower to 0.13 higher)

EDE- Restraint FU	99 (2 studies)	⊕⊖⊖⊖ VERY LOW5,10,18 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- restraint fu in the intervention groups was 0.04 standard deviations higher (0.36 lower to 0.43 higher)
EDE- Shape concern FU	99 (2 studies)	⊕⊖⊖⊖ VERY LOW5,10,18 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- shape concern fu in the intervention groups was 0.08 standard deviations lower (0.48 lower to 0.32 higher)
EDE- Weight concern FU	99 (2 studies)	⊕⊖⊖⊖ VERY LOW5,10,18 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- weight concern fu in the intervention groups was 0.09 standard deviations higher (0.31 lower to 0.48 higher)
EDE- Eating concern FU	52 (1 study)	⊕⊖⊖⊖ VERY LOW5,10,13 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- eating concern fu in the intervention groups was 0.25 standard deviations higher (0.29 lower to 0.8 higher)
Satisfaction with life FU	52 (1 study)	⊕⊖⊖⊖ VERY LOW4,10,13 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean satisfaction with life fu in the intervention groups was 0.08 standard deviations lower (0.62 lower to 0.47 higher)
Bulimic Inventory Index FU	47 (1 study)	⊕⊕⊖⊖ LOW11,19 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic inventory index fu in the intervention groups was 0.77 standard deviations higher (0.18 to 1.37 higher)
EDI Body dissatisfaction FU	55 (1 study)	⊕⊕⊖⊖ LOW7,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction fu in the intervention groups was 0.1 standard deviations higher (0.43 lower to 0.63 higher)
EDI Drive for thinness FU	55 (1 study)	⊕⊕⊖⊖ LOW7,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness fu in the intervention groups was 0.23 standard deviations higher (0.3 lower to 0.77 higher)
EDI Bulimia FU	55 (1 study)	⊕⊕⊖⊖ LOW4,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia fu in the intervention groups was 0.23 standard deviations lower (0.76 lower to 0.31 higher)
Exercising FU	159	⊕⊖⊖⊖		Not calculable	The mean exercising fu in the

	(2 studies)	VERY LOW5,10,20 due to risk of bias, indirectness, imprecision		for SMD values	intervention groups was 0.02 standard deviations lower (0.33 lower to 0.3 higher)
Purging FU	52 (1 study)	⊕⊖⊖⊖ VERY LOW7,10,13 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean purging fu in the intervention groups was 0.40 standard deviations higher (0.15 lower to 0.95 higher)
Remission FU - Young people	85 (1 study)	⊕⊖⊖⊖ VERY LOW9,10,11 due to risk of bias, indirectness, imprecision	RR 0.70 (0.33 to 1.48)	293 per 1000	88 fewer per 1000 (from 196 fewer to 140 more)
Remission FU - Adults	454 (4 studies)	⊕⊖⊖⊖ VERY LOW3,12,21 due to risk of bias, indirectness, imprecision	RR 0.85 (0.59 to 1.14)	225 per 1000	34 fewer per 1000 (from 92 fewer to 32 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; FU: Follow up

1 It was unclear in all studies except Schmidt 2006 (where it was performed) if allocation concealment was performed. Across all studies it was unclear if patients were blind to treatment allocation, and in most studies it was unclear if the assessors and investigators were blind. High dropout rates were reported across studies.
2 Heterogeneity was detected I2 >50%.
3 A mixed population of BN and EDNOS was used for a majority of the included studies, however, the BN made up the higher number so it was not downgraded.
4 95% CI crossed 1 MID (-0.5).
5 For a continuous outcome, there were fewer than 400 participants.
6 It was unclear in all studies except Theils 1998 (where it was not performed) if allocation concealment was performed. Across all studies it was unclear if patients were blind to treatment allocation, and in most studies it was unclear if the assessors and investigators were blind. High dropout rates were reported across studies >20%.
7 95% CI crossed 1 MID (0.5).
8 It was unclear in Bailer 2004 how the randomisation sequence was generated and if allocation concealment was conducted. It was also unclear if either the participant, investigator or assessor was performed. High drop outs were detected >20%.
9 Allocation concealment was performed, but it was unclear if the patients were blind to treatment allocation. The assessors and investigators were not blinded. High dropout rates were detected >20%
10 A mixed population of BN and EDNOS was used, however, the BN made up the higher number.
11 95% CI crossed 2 MIDs (0.75 and 1.25)
12 Across studies it was unclear if allocation concealment was performed and if either or all of the participants, investigators, and assessors were blind.

High dropout rates were reported >20
13 It was unclear if they performed allocation concealment. It was unclear if participants or investigators were blind, however, assessors were blind. High drop outs were reported >20%,
14 It was unclear in all studies, except Schmidt 2006 if allocation concealment was performed. It was unclear across studies if participants and investigators were blind, assessors were blind in all studies but Schmidt. High drop outs were reported >20%.
15 It was unclear in Durand 2003 if allocation concealment was performed, in Thiels it was not performed. Neither the investigators or assessors were blind in Durand 2003, but it was unclear in participants were blind. In Thiels it was unclear if any were blind. High drop outs were reported >20%,
16 It was unclear in Bailer 2004 how the randomised sequence was generated and it was unclear across all studies except Schmidt 2006 if allocation concealment was performed. In Mitchell 2008 and Wagner 2013 assessors were blind, but it was unclear if participants or investigators were blind. High drop outs were reported >20%.
17 It was unclear in Bailer and Mitchell if allocation concealment was conducted but it was no performed in Thiels 1988. It was unclear across all studies if the participants, investigators or assessors were blind, except Mtichell 2008 the assessors were blind. High drop outs were reported >20%.
18 It was unclear in Mitchell if allocation concealment was conducted but it was no performed in Thiels 1988. It was unclear if the participants, investigators or assessors were blind, except Mtichell 2008 the assessors were blind. High drop outs were reported >20%.
19 Allocation concealment was not performed and it was unclear if either the participants, investigators or assessors were blind. High dropout rates were detected >20%.
20 It was unclear if in Wagner 2013 if allocation concealment was performed, but it was in Schmidt 2006. It was unclear if participants or investigators were blind in both studies. In Schmidt the assessors were not blind at follow-up, yet in Wagner 2013 the assessors were blind. High drop outs were reported >20%.
21 95% CI crossed 1 MID (0.75).

Table 197: Summary of findings table for guided self-help (ED) (or self-help with support) versus wait list control (WLC) at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN Guided SH (ED) (95% CI)
Bingeing	111 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.46 standard deviations lower (0.84 to 0.08 lower)
Vomiting	151 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.32 standard deviations lower (0.64 lower to 0.01 higher)
Use of laxatives	151	⊕⊖⊖⊖		Not	The mean use of laxatives in the intervention

	(2 studies)	VERY LOW1,3,4 due to risk of bias, inconsistency, imprecision		calculable for SMD values	groups was 0.55 standard deviations lower (1.80 lower to 0.69 higher)
Depression	220 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.53 standard deviations lower (0.8 to 0.26 lower)
Purging	178 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.95 standard deviations lower (1.27 to 0.63 lower)
EDI Drive for thinness	178 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,5,6 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean edi drive for thinness in the intervention groups was 0.80 standard deviations lower (1.1 to 0.49 lower)
EDI Body dissatisfaction	178 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction in the intervention groups was 0.81 standard deviations lower (1.12 to 0.51 lower)
EDI - Bulimia	69 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi - bulimia in the intervention groups was 0.15 standard deviations lower (0.62 lower to 0.32 higher)
EDE- Weight concern	178 (2 studies)	⊕⊖⊖⊖ VERY LOW2,5,6,7 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.82 standard deviations lower (1.13 to 0.51 lower)
EDE-Restraint	69 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD	The mean ede-restraint in the intervention groups was 0.31 standard deviations lower (0.78 lower to 0.17 higher)

				values	
EDE - Eating concern	69 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,8 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede - eating concern in the intervention groups was 1.19 standard deviations lower (1.71 to 0.68 lower)
EDE- Shape concern	178 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,5,6 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.70 standard deviations lower (1.01 to 0.4 lower)
EDE-Global	178 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,9 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede-global in the intervention groups was 1.31 standard deviations lower (1.64 to 0.99 lower)
Quality of life	178 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.59 standard deviations higher (0.29 to 0.89 higher)
Clinical Symptom Index	151 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical symptom index in the intervention groups was 0.38 standard deviations lower (0.71 to 0.06 lower)
Did not achieve remission_ITT	198 (2 studies)	⊕⊖⊖⊖ VERY LOW10,11,12 due to risk of bias, indirectness, imprecision	RR 0.86 (0.77 to 0.96)	70 per 1000	10 fewer per 1000 (from 3 fewer to 16 fewer)
Remission FU_ITT	89 (1 study)	⊕⊖⊖⊖ VERY LOW11,13,14 due to risk of bias, indirectness, imprecision	RR 0.99 (0.44 to 2.23)	226 per 1000	2 fewer per 1000 (from 126 fewer to 278 more)
Purging <18 binges month	69 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,7 due to risk of bias, indirectness,		Not calculable for	The mean purging <18 binges month in the intervention groups was 2.07 standard deviations lower

		imprecision		SMD values	(2.66 to 1.47 lower)
Purging <18 binge month	88 (1 study)	⊕⊕⊖⊖ LOW1,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging <18 binge month in the intervention groups was 0.49 standard deviations lower (0.87 to 0.11 lower)
EDE- Shape concern <18 binges month	69 (1 study)	⊕⊖⊖⊖ VERY LOW2,7,8 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- shape concern <18 binges month in the intervention groups was 1.05 standard deviations lower (1.56 to 0.54 lower)
EDE- Shape concern >18 binges month	109 (1 study)	⊕⊕⊖⊖ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern >18 binges month in the intervention groups was 0.51 standard deviations lower (0.89 to 0.13 lower)
EDE- Weight concern <18 binges month	69 (1 study)	⊕⊖⊖⊖ VERY LOW2,7,8 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- weight concern <18 binges month in the intervention groups was 1.29 standard deviations lower (1.81 to 0.77 lower)
EDE- Weight concern >18 binges month	109 (1 study)	⊕⊕⊖⊖ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern >18 binges month in the intervention groups was 0.56 standard deviations lower (0.95 to 0.18 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

[1] It was unclear in all studies if allocation concealment was performed. How the randomisation sequence was generated in Walsh 2004 was unclear. Across the studies it was unclear if either or all the participants, investigators or assessors were blind. High dropout rates were reported >20%.
[2] Ljotsson 2007 contained a mixture of BED (52%) and BN (48%)
[3] 95% CI crossed 1 MID (-0.5).
[4] Heterogeneity was detected I2 >80%.

[5] It was unclear in all studies if allocation concealment was performed. In Banasiask 2005 the assessors were blind, but participants and investigators were not blind. In Ljotsson 2007 the participants were not blind but it was unclear if investigators and assessors were blind. High dropout rates were reported >20%.
[6] Heterogeneity was detected I2 >50%.
[7] For a continuous outcome, there were fewer than 400 participants.
[8] It was unclear in all studies if allocation concealment was performed. In Ljotsson 2007 the participants were not blind but it was unclear if investigators and assessors were blind. High dropout rates were reported >20%.
[9] It was unclear in all studies if allocation concealment was performed. Across the studies it was unclear if either or all the participants, investigators or assessors were blind. High dropout rates were reported >20%.
[10] It was unclear in all studies if allocation concealment was performed. In Banasiask 2005 the assessors were blind, but participants and investigators were not blind. In Palmer 2002, it was unclear if participants, investigators and assessors were blind. High dropout rates were reported >20%.
[11] Palmer 2002 contained a mixed population of EDNOS (20%) and BN (80%)
[12] For a dichotomous outcome, there were fewer than 300 events.
[13] It was unclear if allocation concealment was performed. It was unclear if assessors, investigators or participants were blind. High drop outs were detected >20%.
[14] 95% CI crossed 1 MID (1.25).

Table 198: Summary of findings table for self-help (ED) (or self-help without support) versus any other intervention at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BN Self-help (ED) (95% CI)
Bingeing	162 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.18 standard deviations higher (0.52 lower to 0.88 higher)
Purging	70 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.49 standard deviations higher (0.02 to 0.97 higher)
Use of laxatives	33 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,8 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean use of laxatives in the intervention groups was 0.10 standard deviations higher (0.58 lower to 0.78 higher)

Vomiting	96 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,9,10 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.85 standard deviations higher (0.41 to 1.29 higher)
Depression	56 (1 study)	⊕⊕⊖⊖ LOW10,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.52 standard deviations higher (0.01 lower to 1.05 higher)
Exercising	33 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean exercising in the intervention groups was 0.1 standard deviations higher (0.58 lower to 0.79 higher)
Remission_ITT	173 (2 studies)	⊕⊖⊖⊖ VERY LOW12,13,14 due to risk of bias, indirectness, imprecision	RR 0.74 (0.32 to 1.7)	140 per 1000	36 fewer per 1000 (from 95 fewer to 98 more)
EDE-Global	132 (2 studies)	⊕⊕⊖⊖ LOW5,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-global in the intervention groups was 0.2 standard deviations higher (0.15 lower to 0.55 higher)
EDE- Weight concern	118 (2 studies)	⊕⊕⊖⊖ LOW10,15 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.23 standard deviations higher (0.14 lower to 0.61 higher)
EDE- Eating concern	56 (1 study)	⊕⊕⊖⊖ LOW10,16 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.45 standard deviations higher (0.08 lower to 0.98 higher)
EDE- Shape concern	118 (2 studies)	⊕⊕⊖⊖ LOW10,15 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.2 standard deviations higher (0.18 lower to 0.57 higher)
EDE- Restraint	118 (2 studies)	⊕⊖⊖⊖ VERY LOW9,10,15 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.71 standard deviations higher (0.32 to 1.1 higher)
Purging FU	70	⊕⊕⊖⊖		Not	The mean purging fu in the intervention

	(1 study)	LOW8,17 due to risk of bias, imprecision		calculable for SMD values	groups was 0 standard deviations higher (0.47 lower to 0.47 higher)
Bingeing FU	111 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,10 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.23 standard deviations higher (0.14 lower to 0.61 higher)
Vomiting FU	40 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean vomiting fu in the intervention groups was 0.07 standard deviations higher (0.55 lower to 0.69 higher)
Excessive exercising FU	37 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean excessive exercising fu in the intervention groups was 0.09 standard deviations higher (0.55 lower to 0.74 higher)
Use of laxatives FU	39 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,10 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean use of laxatives fu in the intervention groups was 0.22 standard deviations higher (0.41 lower to 0.85 higher)
EDE-Global FU	70 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-global fu in the intervention groups was 0.14 standard deviations lower (0.61 lower to 0.33 higher)
Remission FU	90 (1 study)	⊕⊖⊖⊖ VERY LOW4,12,13 due to risk of bias, indirectness, imprecision	RR 0.98 (0.43 to 2.2)	224 per 1000	4 fewer per 1000 (from 128 fewer to 269 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; FU: Follow up

1 Whilst in Schmidt 2006, allocation concealment was performed it was unclear in the other studies. It was unclear in all studies if participants, investigators or assessors were blind. High drop outs were reported.>20%.
2 Heterogeneity detected I2 >50%.
3 Schmidt 2006 included a mixed population of BN and ENDOS
4 95% CI crossed 2 MIDs (-0.5 and 0.5).
5 It was unclear if allocation concealment was performed. It was also unclear if participants, investigators and assessors were blind. High drop outs were

detected >20%.
6 95% CI crossed 1 MID (-0.5).
7 In Schmidt 2006, allocation concealment was performed. It was unclear in all studies if participants, investigators were blind. Assessors were blind at baseline but not at follow-up. High drop outs were reported.>20%.
8 For a continuous outcome there were fewer than 400 participants.
9 Heterogeneity was detected I2>80%
10 95% CI crossed 1 MID (0.5).
11 Allocation concealment was performed and assessors were blind. However, participants were not blind and it was unclear if investigators were. High drop outs were detected >20%.
12 It was unclear if allocation concealment was performed. It was also unclear if either the participants, assessors or investigators were blind. High drop outs were reported >20%.
13 Palmer 2002 contained a mixed population of EDNOS (20%) and BN (80%)
14 95% CI crossed 2 MIDs (0.75 and 1.25).
15 Allocation concealment was performed in Carter 2003, however it was unclear if it was in the other study. In Carter, the participants were not blind but the assessors were. It was unclear in the other study/ies if either the participants, assessors or investigators were blind. High drop outs were reported >20%.
16 Allocation concealment was performed in Carter 2003. The participants were not blind but the assessors were. High drop outs were reported >20%.
17 it was unclear if allocation concealment was conducted. Assessors were blind but it was unclear if participants or participants were blind.

Table 199: Summary of findings table for self-help (or self-help without support) versus wait list controls (WLC) at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN Self-help (95% CI)
Depression	56 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.02 standard deviations higher (0.5 lower to 0.54 higher)
EDE- Restraint	56 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.07 standard deviations lower (0.59 lower to 0.45 higher)
EDE-Shape concern	56 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for	The mean ede-shape concern in the intervention groups was 0.08 standard deviations lower

				SMD values	(0.6 lower to 0.44 higher)
EDE-Weight concern	56 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern in the intervention groups was 0.00 standard deviations higher (0.52 lower to 0.52 higher)
EDE- Eating concern	56 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.0 standard deviations higher (0.52 lower to 0.52 higher)
Did not achieve remission_ITT	63 (1 study)	⊕⊖⊖⊖ VERY LOW4,5,6 due to risk of bias, indirectness, imprecision	RR 0.94 (0.84 to 1.04)		-
Remission_ITT_FU	63 (1 study)	⊕⊖⊖⊖ VERY LOW4,5,7 due to risk of bias, indirectness, imprecision	RR 0.97 (0.38 to 2.44)	226 per 1000	7 fewer per 1000 (from 140 fewer to 325 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 In Carter 2003, the participants were not blinded, it was unclear if investigators were blind and the assessors were blind. Again, high dropouts were reported >20%
2 95% CI crossed 2 MIDs (-0.5 and 0.5)
3 95% CI crossed 1 MID (-0.5)
4 It was unclear if allocation concealment was performed. It was unclear if participants, assessors and investigators were blinded. High dropouts were reported >20%,
5 Palmer 2002 contained a mixed population of EDNOS (20%) and BN (80%)
6 For a dichotomous outcome, there were fewer than 300 events
7 95% CI crossed 2 MIDs (0.75 and 1.25)

Table 200: Summary of findings table for self-help (ED) (or self-help without support) versus any other intervention at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BN Self-help (ED) (95% CI)

Bingeing	162 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.49 standard deviations higher (0.17 to 0.81 higher)
Purging	70 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.49 standard deviations higher (0.02 to 0.97 higher)
Use of laxatives	33 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean use of laxatives in the intervention groups was 0.10 standard deviations higher (0.58 lower to 0.78 higher)
Vomiting	96 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,8,9 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.85 standard deviations higher (0.41 to 1.29 higher)
Depression	56 (1 study)	⊕⊕⊖⊖ LOW9,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.52 standard deviations higher (0.01 lower to 1.05 higher)
Exercising	33 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,11 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean exercising in the intervention groups was 0.1 standard deviations higher (0.58 lower to 0.79 higher)
Remission_ITT	173 (2 studies)	⊕⊖⊖⊖ VERY LOW12,13,14 due to risk of bias, indirectness, imprecision	RR 0.74 (0.32 to 1.7)	140 per 1000	36 fewer per 1000 (from 95 fewer to 98 more)
EDE-Global	132 (2 studies)	⊕⊕⊖⊖ LOW5,9 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-global in the intervention groups was 0.2 standard deviations higher (0.15 lower to 0.55 higher)
EDE- Weight concern	118 (2 studies)	⊕⊕⊖⊖ LOW9,15 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.23 standard deviations higher (0.14 lower to 0.61 higher)
EDE- Eating concern	56	⊕⊕⊖⊖		Not	The mean ede- eating concern in the

	(1 study)	LOW9,16 due to risk of bias, imprecision		calculable for SMD values	intervention groups was 0.45 standard deviations higher (0.08 lower to 0.98 higher)
EDE- Shape concern	118 (2 studies)	⊕⊕⊖⊖ LOW9,15 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.2 standard deviations higher (0.18 lower to 0.57 higher)
EDE- Restraint	118 (2 studies)	⊕⊖⊖⊖ VERY LOW8,9,15 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.71 standard deviations higher (0.32 to 1.1 higher)
Purging FU	70 (1 study)	⊕⊕⊖⊖ LOW4,17 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging fu in the intervention groups was 0 standard deviations higher (0.47 lower to 0.47 higher)
Bingeing FU	111 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,9 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.23 standard deviations higher (0.14 lower to 0.61 higher)
Vomiting FU	40 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,11 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean vomiting fu in the intervention groups was 0.07 standard deviations higher (0.55 lower to 0.69 higher)
Excessive exercising FU	37 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,11 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean excessive exercising fu in the intervention groups was 0.09 standard deviations higher (0.55 lower to 0.74 higher)
Use of laxatives FU	39 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,9 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean use of laxatives fu in the intervention groups was 0.22 standard deviations higher (0.41 lower to 0.85 higher)
EDE-Global FU	70 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-global fu in the intervention groups was 0.14 standard deviations lower (0.61 lower to 0.33 higher)
Remission FU	90 (1 study)	⊕⊖⊖⊖ VERY LOW11,12,13	RR 0.98 (0.43 to	224 per 1000	4 fewer per 1000 (from 128 fewer to 269 more)

		due to risk of bias, indirectness, imprecision	2.2)		

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Whilst in Schmidt 2006, allocation concealment was performed it was unclear in the other studies. It was unclear in all studies if participants, investigators or assessors were blind. High dropouts were reported.>20%.
2 Heterogeneity detected I2 >50%.
3 Schmidt 2006 included a mixed population of BN and EDNOS
4 For a continuous outcome there were fewer than 400 participants.
5 It was unclear if allocation concealment was performed. It was also unclear if participants, investigators and assessors were blind. High dropouts were detected >20%.
6 95% CI crossed 1 MID (-0.5).
7 In Schmidt 2006, allocation concealment was performed. It was unclear in all studies if participants, investigators were blind. Assessors were blind at baseline but not at follow up. High dropouts were reported.>20%.
8 Heterogeneity was detected I2>80%
9 95% CI crossed 1 MID (0.5).
10 Allocation concealment was performed and assessors were blind. However, participants were not blind and it was unclear if investigators were. High dropouts were detected >20%.
11 95% CI crossed 2 MIDs (-0.5 and 0.5).
12 It was unclear if allocation concealment was performed. It was also unclear if either the participants, assessors or investigators were blind. High dropouts were reported >20%.
13 Palmer 2002 contained a mixed population of EDNOS (20%) and BN (80%)
14 95% CI crossed 2 MIDs (0.75 and 1.25).
15 Allocation concealment was performed in Carter 2003, however it was unclear if it was in the other study. In Carter, the participants were not blind but the assessors were. It was unclear in the other study/ies if either the participants, assessors or investigators were blind. High dropouts were reported >20%.
16 Allocation concealment was performed in Carter 2003. The participants were not blind but the assessors were. High dropouts were reported >20%.
17 it was unclear if allocation concealment was conducted. Assessors were blind but it was unclear if participants or participants were blind

.

Table 201: Summary of findings table for internet self-help versus any other intervention at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BN Internet SH (ED) (95% CI)
Bingeing	192 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.26 standard deviations lower (0.55 lower to 0.03 higher)

Purging	70 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.49 standard deviations lower (0.97 to 0.02 lower)
Vomiting	122 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.14 standard deviations higher (0.22 lower to 0.5 higher)
EDE-Q	70 (1 study)	⊕⊕⊖⊖ LOW4,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q in the intervention groups was 0.36 standard deviations lower (0.84 lower to 0.11 higher)
Laxative use	122 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative use in the intervention groups was 0.16 standard deviations higher (0.2 lower to 0.52 higher)
Excessive exercise	122 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean excessive exercise in the intervention groups was 0.08 standard deviations higher (0.28 lower to 0.44 higher)
Remission_ITT	155 (1 study)	⊕⊕⊖⊖ LOW4,8 due to risk of bias, imprecision	RR 0.95 (0.44 to 2.01)	153 per 1000	8 fewer per 1000 (from 86 fewer to 154 more)
Binging FU	192 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean binging fu in the intervention groups was 0.21 standard deviations lower (0.49 lower to 0.08 higher)
Remission FU_ITT	155 (1 study)	⊕⊖⊖⊖ VERY LOW1,8 due to risk of bias, imprecision	RR 1.66 (0.86 to 3.2)	153 per 1000	101 more per 1000 (from 21 fewer to 336 more)
EDE-Q FU	70 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q fu in the intervention groups was 0.14 standard deviations higher (0.33 lower to 0.61 higher)
Purging FU	70 (1 study)	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision		Not calculable for SMD	The mean purging fu in the intervention groups was 0 standard deviations higher

				values	(0.47 lower to 0.47 higher)
Vomiting FU	122 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting fu in the intervention groups was 0.04 standard deviations lower (0.4 lower to 0.32 higher)
Laxative use FU	122 (1 study)	⊕⊕⊖⊖ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative use fu in the intervention groups was 0.18 standard deviations higher (0.18 lower to 0.54 higher)
Excessive exercise FU	122 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean excessive exercise fu in the intervention groups was 0.01 standard deviations lower (0.37 lower to 0.35 higher)
Bingeing <18 per month	192 (1 study)	⊕⊕⊖⊖ LOW6,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing <18 per month in the intervention groups was 0.69 standard deviations higher (1.17 to 0.2 lower)
Bingeing >18 month	122 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing >18 month in the intervention groups was 0.03 standard deviations lower (0.3 lower to 0.33 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; FU: Follow up

1 It was unclear allocation concealment was conducted. In Wagner 2013 assessors were blind but it was unclear if either the participants or investigators were blind. In Ruwaard 2013 it was unclear if either the participants, investigators or assessors were blind. High dropouts were reported >20%.
2 Heterogeneity was detected >50%
3 For a continuous outcome, there were fewer than 400 participants.
4 In Wagner 2013, it was unclear allocation concealment was conducted, or if either the participants, assessors or investigators were blind. High drop outs were reported >20%.
5 95% CI crossed 1 MID (0.5)
6 In Ruwaard 2013, it was unclear allocation concealment was conducted. Assessors were blind but it was unclear if either the participants or investigators were blind. High dropouts were reported >20%.
7 95% CI crossed 1 MID (-0.5)
8 95% CI crossed 2 MIDs (0.75 and 1.25).

Table 202: Summary of findings table for internet self-help versus wait list control at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN Internet SH (ED) (95% CI)
Bingeing	137 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.41 standard deviations lower (0.75 to 0.07 lower)
Purging	137 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.37 standard deviations lower (0.71 to 0.04 lower)
Vomiting	137 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.09 standard deviations higher (0.25 lower to 0.43 higher)
Depression	67 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 1.09 standard deviations lower (1.6 to 0.57 lower)
Quality of life	67 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.7 standard deviations higher (0.19 to 1.2 higher)
Remission Not Achieved	76 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,8 due to risk of bias, indirectness, imprecision	RR 0.84 (0.71 to 0.98)	26 per 1000	4 fewer per 1000 (from 1 fewer to 8 fewer)
EDE-Q	137 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,6		Not calculabl	The mean ede-q in the intervention groups was

		due to risk of bias, indirectness, imprecision		e for SMD values	0.71 standard deviations lower (1.05 to 0.36 lower)
EDE- Restraint	67 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.88 standard deviations lower (1.38 to 0.38 lower)
EDE- Shape concern	67 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 1.18 standard deviations lower (1.7 to 0.66 lower)
EDE- Weight concern	67 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.88 standard deviations lower (1.38 to 0.38 lower)
EDE- Eating concern	67 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.94 standard deviations lower (1.45 to 0.43 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. In Sanchez-Ortiz, the assessors were blind but it was unclear if either the investigators or participants were blind. In the other study, it was unclear if any were blind. High dropouts were reported >20%.
2 Sanchez-Ortiz 2011 included a mixed population of BN (51.3%) and EDNOS (48.7%)
3 For a continuous outcome, there were fewer than 400 participants.
4 Heterogeneity was detected, I2 >80%
5 In Sanchez-Ortiz, it was unclear if allocation concealment was conducted. The assessors were blind but it was unclear if either the investigators or participants were blind. High dropouts were reported >20%.
6 95% CI crossed 1 MID (-0.5).
7 95% CI crossed 1 MID (0.5).
8 95% CI crossed 1 MID (0.75).

Table 203: Summary of findings table for self-help (ED) (or self-help without support) versus wait list control at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN Self-help (ED) (95% CI)
Bingeing	130 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 1.23 standard deviations lower (3.95 lower to 1.49 higher)
Purging	70 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.2 standard deviations higher (0.27 lower to 0.67 higher)
Vomiting	60 (1 study)	⊕⊖⊖⊖ VERY LOW1,5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.00 standard deviations higher (0.54 lower to 0.54 higher)
EDE-Q	130 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-q in the intervention groups was 1.25 standard deviations lower (3.41 lower to 0.92 higher)
Depression	57 (1 study)	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.47 standard deviations higher (0.06 lower to 1 higher)
Remission_ITT	82 (1 study)	⊕⊖⊖⊖ VERY LOW8,9 due to risk of bias, imprecision	RR 2.21 (0.51 to 9.52)	74 per 1000	90 more per 1000 (from 36 fewer to 631 more)
EDE- Restraint	117 (2 studies)	⊕⊕⊖⊖ LOW8,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.07 standard deviations higher (0.31 lower to 0.44 higher)
EDE- Shape concern	117 (2 studies)	⊕⊕⊖⊖ LOW6,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.74 standard deviations lower (1.18 to 0.29 lower)

EDE- Weight concern	117 (2 studies)	⨁⨁⊖⊖ LOW8,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.55 standard deviations lower (0.97 to 0.13 lower)
EDE- Eating concern	57 (1 study)	⨁⨁⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.42 standard deviations higher (0.1 lower to 0.95 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. It was unclear if participants, investigators or assessors were blind, except in Mitchell 2008 assessors were not blind. High drop outs were reported >20%.
2 Heterogeneity was detected, I2 >80%.
3 95% CI crossed 2 MIDs (-0.5 and 0.5).
4 It was unclear if allocation concealment was conducted. It was unclear if participants, investigators or assessors were blind. High dropouts were reported >20%.
5 95% CI crossed 1 MID (0.5).
6 95% CI crossed 1 MID (-0.5).
7 In Carter 2003, allocation concealment was conducted. Assessors were blind, but participants were not. It was unclear if investigators were blind. High dropouts were detected >20%.
8 In Carter 2003, allocation concealment was conducted, but it was unclear if it was conducted in Treasure. In Carter, assessors were blind, but participants were not. It was unclear if investigators were blind. It was unclear if any were blind in Treasure. High dropouts were detected >20%.
9 95% CI crossed 2 MIDs (0.75 and 1.25)
10 For a continuous outcome, there were fewer than 400 participants.

1 **Table 204: Summary of findings table for text messaging versus wait list control at end of treatment in adults with bulimia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BN Text messaging (95% CI)
Remission_ITT	165 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.42 (0.99 to 2.02)	361 per 1000	152 more per 1000 (from 4 fewer to 369 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%

confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 it was unclear how the randomisation sequence was generated or if allocation concealment was conducted. It was unclear if either the participants, investigators or assessors were blind.
2 Included a mixed population of BN 60% and EDNOS 40%.
3 95% CI crossed 1 MID (1.25).

7.2.6.3 Family therapy

Table 205: Summary of findings table for family therapy for eating disorders versus any individual therapy at end of treatment in young people with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other intervention	Risk difference with Family therapy-ED (95% CI)
Remission	295 (3 studies) 12 months	⊕⊖⊖⊖ VERY LOW1,2,3,4,5 due to risk of bias, indirectness, imprecision	RR 1.68 (1.11 to 2.54)	168 per 1000	114 more per 1000 (from 18 more to 258 more)
Binge Frequency	157 (2 studies) 12 months	⊕⊕⊖⊖ LOW1,2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.09 standard deviations lower (0.4 lower to 0.23 higher)
Abstinence from vomiting EATATE	63 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, indirectness, imprecision	RR 0.87 (0.41 to 1.85)	323 per 1000	42 fewer per 1000 (from 190 fewer to 274 more)
Purge Frequency	86 (1 study) 12 months	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean purge frequency in the intervention groups was 0.33 standard deviations lower (0.75 lower to 0.1 higher)
Vomit Frequency EDE	71 (1 study) 6 months	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomit frequency in the intervention groups was 0.64 standard deviations lower (1.12 to 0.16 lower)
EDE Global	155	⊕⊕⊖⊖		Not calculable for	The mean ede global in the intervention

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other intervention	Risk difference with Family therapy-ED (95% CI)
	(2 studies) 12 weeks	LOW1,2,5 due to risk of bias, imprecision		SMD values	groups was 0.38 standard deviations lower (0.69 to 0.06 lower)
EDE Restraint	71 (1 study) 6 months	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint in the intervention groups was 0.51 standard deviations lower (0.98 to 0.04 lower)
EDE Shape Concern	71 (1 study) 6 months	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.54 standard deviations lower (1.01 to 0.07 lower)
EDE Weight Concern	71 (1 study) 6 weeks	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.48 standard deviations lower (0.95 to 0.01 lower)
Yale-Brown-Cornell Eating Disorder Scale	86 (1 study) 12 weeks	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean yale-brown-cornell eating disorder scale in the intervention groups was 0.36 standard deviations lower (0.78 lower to 0.07 higher)
Depression BDI	157 (2 studies) 12 months	⊕⊕⊖⊖ LOW1,2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.28 standard deviations lower (0.6 lower to 0.03 higher)
Hospitalized during treatment phase	109 (1 study) 12 months	⊕⊕⊖⊖ LOW2,5 due to risk of bias, imprecision	RR 0.09 (0.01 to 0.7)	207 per 1000	188 fewer per 1000 (from 62 fewer to 205 fewer)
Service User Experience Helping Relationship Questionnaire	68 (1 study) 6 months	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean service user experience in the intervention groups was 0.06 standard deviations higher (0.42 lower to 0.53 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%					

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other intervention	Risk difference with Family therapy-ED (95% CI)
confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Le Grange 2007: Unclear randomization method and allocation concealment, no participant, investigator nor assessor blinding. *2 Le Grange 2016b: Unclear randomization method and allocation concealment, no participant nor investigator blinding.* *3 Schmidt 2007: Unclear randomization and allocation concealment, No participant nor investigator blinding.* *4 Schmidt 2007: Sample consists of 61 bulimia nervosa and 24 EDNOS.* *5 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).* *6 <300 events (dichotomous outcome) or <400 participants (continuous outcome).* *7 CI crosses both 0.75 and 1.25 (Risk Ratio).*					

Table 206: Summary of findings table for family therapy for eating disorders versus any individual therapy at follow up in young people with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other intervention	Risk difference with Family Therapy-ED (95% CI)
Remission FU	215 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, indirectness, imprecision	RR 1.92 (1.24 to 2.99)	230 per 1000	211 more per 1000 (from 55 more to 457 more)
Binge Frequency FU	137 (2 studies)	⊕⊕⊖⊖ LOW1,5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 0.1 standard deviations lower (0.44 lower to 0.24 higher)
Abstinence from vomiting FU EATATE	54 (1 study)	⊕⊖⊖⊖ VERY LOW4,7 due to risk of bias, imprecision	RR 0.92 (0.56 to 1.51)	560 per 1000	45 fewer per 1000 (from 246 fewer to 286 more)
Purge Frequency FU	69 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean purge frequency fu in the intervention groups was 0 standard deviations higher (0.48 lower to 0.48 higher)
Vomit Frequency FU	68	⊕⊕⊖⊖		Not calculable for	The mean vomit frequency fu in the

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other intervention	Risk difference with Family Therapy-ED (95% CI)
EDE	(1 study)	LOW3,5 due to risk of bias, imprecision		SMD values	intervention groups was 0.17 standard deviations lower (0.65 lower to 0.3 higher)
EDE Global FU	137 (2 studies)	⨁⨁⨀⨀ LOW1,3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.38 standard deviations lower (0.72 to 0.04 lower)
EDE Restraint FU	68 (1 study) 6 weeks	⨁⨁⨀⨀ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint fu in the intervention groups was 0.38 standard deviations lower (0.86 lower to 0.1 higher)
EDE Shape Concern FU	68 (1 study) 6 months	⨁⨁⨀⨀ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern fu in the intervention groups was 0.58 standard deviations lower (1.06 to 0.09 lower)
EDE Weight Concern FU	68 (1 study) 6 months	⨁⨁⨀⨀ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern fu in the intervention groups was 0.46 standard deviations lower (0.94 lower to 0.02 higher)
Yale-Brown-Cornell Eating Disorder Scale FU	69 (1 study) 12 months	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean yale-brown-cornell eating disorder scale fu in the intervention groups was 0.37 standard deviations lower (0.85 lower to 0.11 higher)
Depression FU	137 (2 studies) 12 months	⨁⨁⨀⨀ LOW1,5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.1 standard deviations lower (0.43 lower to 0.24 higher)
Service User Experience FU Helping Relationship Questionnaire	71 (1 study) 6 months	⨁⨁⨀⨀ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean service user experience fu in the intervention groups was 0.41 standard deviations lower (0.88 lower to 0.06 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%					

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other intervention	Risk difference with Family Therapy-ED (95% CI)
confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					

1 Le Grange 2016b: Unclear randomization method and allocation concealment, no participant nor investigator blinding.
2 Schmidt 2007: Sample consists of 61 bulimia nervosa and 24 EDNOS
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
4 Schmidt 2007: Unclear randomization and allocation concealment, No participant nor investigator blinding.
5 Le Grange 2007: Unclear randomization method and allocation concealment, no participant, investigator nor assessor blinding.
6 <300 events (dichotomous outcome) or <400 participants (continuous outcome).
7 CI crosses both 0.75 and 1.25 (Risk Ratio).

1

2

7.2.7 Economic Evidence

7.2.7.1 Systematic literature review

The systematic search of the economic literature undertaken for the guideline identified:

- One US study on the cost effectiveness of CBT ED in adults with BN or EDNoS subsyndromal variants of BN (Crow et al., 2009);
- One UK study on the cost effectiveness of family therapy (FT) compared with guided self-help ED in young people (13-20 years) with BN and EDNOS (Schmidt et al., 2007b).

References to included studies and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix P. Completed methodology checklists of the studies are provided in Appendix O. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix Q.

Crow and colleagues (2009) evaluated the cost effectiveness of a CBT ED compared with guided self-help ED in adults with BN or EDNOS subsyndromal variants of BN in the US. The economic analysis was conducted alongside an RCT (Mitchell 2008) (N=128). CBT comprised 20 sessions of treatment over 16 weeks. The analysis was conducted from the intervention provider (plus travel costs) perspective. The study considered a range of costs including treatment (initial evaluation, laboratory evaluation and psychotherapy visits) and travel time for therapists and participants (time and fuel). The resource use estimates were based on the RCT (N=128). The unit costs were obtained from national sources (Medicare and Medicaid reimbursement rates). The measure of outcome for the economic analysis was remission, defined as abstinence from binge eating and purging. The time horizon of the analysis was 12 months.

CBT ED resulted in a greater proportion of people achieving full remission at 12 months compared with guided self-help ED (28.8% versus 22.6%, respectively; a difference of 6.2%). The mean total costs per participant over 12 months were $2,684 for the CBT ED and $1,648 for guided self-help ED, a difference of $1,036 in 2005 US dollars. Statistical significance levels were not reported for differences in costs and outcomes. The incremental cost-effectiveness ratio (ICER) of CBT ED when compared with guided self-help ED was $16,708 per additional participant in remission. Bootstrapping indicated that in 78.9% of iterations guided self-help ED was less effective but also less costly than CBT ED while in the 21.1% iterations guided self-help ED was both more effective and less costly (that is, guided self-help ED was dominant).

Deterministic sensitivity analysis showed that the results were robust to the assumptions. Assuming full clinical prices for treatments (as opposed to the reimbursement rates) CBT ED resulted in an ICER of $16,155 per additional remitter when compared with guided self-help ED; assuming 2008 gasoline prices (as opposed to 2005 prices) the ICER increased to $17,547; and assuming built in video camera for guided self-help ED (no additional charges for the telemedicine component) the ICER increased to $19,308.

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it has been conducted in the US. The authors did not attempt to estimate quality adjusted life years (QALYs) which made it difficult for the committee to interpret the cost-effectiveness results and to compare the findings with other studies. However, overall, given the data limitations in this area, this was a well conducted study and was judged by the committee to have only minor methodological limitations.

Schmidt and colleagues (2007b) evaluated the cost effectiveness of a family therapy (FT) compared with guided self-help ED in young people (13-20 years) with BN and EDNOS alongside an RCT (Schmidt 2007) (N=85 at baseline, N=63 at 6 months, N=54 at 12 months)

conducted in the UK. The FT involved up to 13 sessions with family members and two individual sessions over a six month period and guided self-help ED involved 10 weekly sessions, three monthly follow up sessions and 2 optional sessions with a close other (such as, parent or guardian). The main analysis was conducted from a societal perspective (health care, social care, education, productivity costs and out-of-pocket expenses). The costs were stratified and could be estimated from the NHS and personal social services (PSS) perspective only. The study considered a range of educational costs (home tuition, individual help in classes, classes in a special unit, contacts with school nurse, educational psychologist and educational welfare officer, additional meetings with tutors and other educational supports), hospital services (inpatient care, A&E department visits, outpatient appointments and day hospital attendances), primary care (health visitor, GP, dentist and optician), specialist services (child development or guidance centre, dietician, family or individual therapy and contacts with a psychiatrist or psychologist), medication, social care (social work, after-school clubs and other social care supports), family member's service use (GP, outpatient appointments and psychiatrist and psychologist), lost employment and out-of-pocket expenses. The resource use estimates were based on the RCT (N=83, N=61 at 6 month follow up and N=53 at 12 month follow up). The unit costs were obtained from national sources. The measures of outcome for the economic analysis included the proportion of participants abstinent from bingeing, abstinent from vomiting and abstinent from bingeing and purging combined. The time horizon of the analysis was 12 months. The results were reported at the end of treatment (6 months) and at 12 months.

Guided self-help ED resulted in a greater proportion of participants abstinent from bingeing at the end of treatment (0.42 versus 0.25, respectively; a difference of 0.17, p=0.03). At 12 months guided self-help ED resulted in fewer participants abstinent from bingeing (0.52 versus 0.55; a difference of 0.03; p-value was not significant). Guided self-help ED also resulted in a greater proportion of participants abstinent from vomiting at six months (0.32 versus 0.28; a difference of 0.04, p-value was not significant) and at 12 months (0.56 versus 0.52; a difference of 0.04, p-value was not significant). When looking at the proportion abstinent from bingeing and vomiting combined guided self-help ED resulted in a greater proportion of participants abstinent at six months (0.19 versus 0.13, a difference of 0.07, p-value was not significant), but not at 12 months (0.36 versus 0.41; a difference of 0.05, p-value was not significant).

The mean NHS and PSS costs at the end of treatment (six months) were £319 for FT and £849 for guided self-help ED, a difference of £530 (in favour of FT) in likely 2006 prices. The mean NHS and PSS costs at 12 months were £691 for FT and £1,286 for guided self-help ED, a difference of £595 (in favour of FT). Significance levels were not reported.

When considering costs from a societal perspective the mean costs at the end of treatment were £720 for FT and £1,096 for guided self-help ED, a difference of £377 (in favour of FT) and at 12 months the mean costs per participant were £1,269 for FT and £1,657 for guided self-help ED, a difference of £388 (in favour of FT), p-values were not significant.

Using the proportion of participants abstinent from bingeing at the end of treatment (6 months) from the NHS and PSS perspective guided self-help ED results in an ICER of £3,120 per additional abstinent person. At six months from a societal perspective guided self-help ED results in an ICER of £2,216 per additional abstinent person. At 12 month follow up FT dominates self-help ED (that is, it results in lower costs and also greater proportion abstinent from bingeing and purging combined.

The analysis was judged by the committee to be partially applicable to the NICE decision-making context. The authors did not attempt to estimate quality adjusted life years (QALYs) which made it difficult to interpret the cost-effectiveness results and to compare the findings with other studies. However, overall, given the data limitations in this area, this was a well conducted study and was judged by the committee to have only minor methodological limitations.

7.2.7.2 Economic modelling

A decision-analytical model was developed to assess the relative cost effectiveness of interventions for adults with BN. The rationale for economic modelling, the methodology adopted, the results and the conclusions from this economic analysis are described in detail in Appendix S. This section provides a summary of the methods employed and the results of the economic analysis.

7.2.7.2.1 Overview of methods

A decision-analytic model in the form of a decision-tree was constructed to evaluate the relative cost effectiveness of psychological interventions over 1.4 years. The psychological interventions assessed were self-help with support (referred to as guided self-help ED in clinical evidence) and CBT-ED individual. The model also considered no treatment (wait list) as a comparator. The choice of treatments assessed in the economic analysis was determined by the availability of respective clinical data (full remission at the end of treatment) included in the guideline systematic literature review. The economic analysis considered effective treatments, as demonstrated by the systematic review of clinical evidence, that were deemed appropriate by the committee as treatment options for people with BN in the UK. The study population comprised of adults with BN.

Clinical data were derived from studies included in the guideline systematic review of clinical evidence and other published literature. Clinical data (that is, full remission at the end of treatment) were analysed using mixed treatment comparison technique. Full remission was defined as cessation of BN-related symptoms over and above two weeks. The inconsistency checks were also undertaken. Details on the methods and clinical data utilised in the NMA that was undertaken to estimate full remission for each treatment option considered in the economic analysis are presented in Appendix R. Results of inconsistency checks are presented in the Appendix N.

The measure of outcome in the economic analysis was the number of QALYs gained. The perspective of the analysis was that of NHS and PSS. Resource use was based on the published literature and the GC expert opinion. National UK unit costs were used. The cost year was 2015. Two methods were employed for the analysis of input parameter data and presentation of the results. First, a deterministic analysis was undertaken, where data were analysed as point estimates and results were presented in the form of ICERs following the principles of incremental analysis. A probabilistic analysis was subsequently performed in which most of the model input parameters were assigned probability distributions. Subsequently, 10,000 iterations were performed, each drawing random values out of the distributions fitted onto the model input parameters. Mean costs and QALYs for each treatment option were calculated by averaging across the 10,000 iterations. This approach allowed more comprehensive consideration of the uncertainty characterising the input parameters and captured the non-linearity characterising the economic model structure. Results of probabilistic analysis were also summarised in the form of cost effectiveness acceptability curves, which express the probability of each intervention being cost effective at various levels of willingness-to-pay per QALY gained (that is, at various cost-effectiveness thresholds). As part of the sensitivity analyses two different bias adjustment scenarios were tested pertaining to the estimation of treatment effects in the network meta-analysis. The scenarios tested included: in trials of active treatments versus waitlist, active treatments were favoured; all active treatments were favoured against wait list and CBT was favoured against other active treatments. Results of bias adjustment analyses are presented in the Appendix T.

The GC expressed the opinion that CBT is associated with long-term benefits, as the effect is sustained over a longer period of time than the time horizon of the analysis. As the longer-term benefits of CBT were not fully captured by the base-case analysis, a secondary analysis was undertaken, in which the time horizon of the analysis was extended to five years. The secondary analysis also included the scenario where the relapse rate associated with CBT-

ED individual was assumed to be zero (that is, all people sustained treatment effect and no-one relapse) whereas the relapse rate for self-help with support and wait list was assumed to be equivalent to the annual relapse rate of 0.27 (this rate was applied during each year of the long-term follow-up).

7.2.7.2.2 *Findings of the NMA and base-case economic analysis*

The results of the NMA indicated that wait list had the lowest probability of full remission at 16 weeks (mean 0.10), followed by self-help with support (0.32) and CBT-ED individual (0.32). Both CBT-ED individual and self-help with support showed a significant effect compared with wait list. There was no significant difference between CBT-ED individual and self-help with support. The odds ratio of CBT-ED individual versus self-help with support was 1.14 (95% CrI: 0.36 to 2.81). The inconsistency checks did not identify any significant inconsistency in the direct and indirect evidence included in the NMA. The bias adjustment sensitivity analysis suggested that bias due to small study effects may be exaggerating the treatment effects in this network. However, as the bias coefficient included 0 in all scenarios and there was no reduction in heterogeneity as a result of the bias adjustment, no strong conclusions about the presence of bias could be made. This strengthens the conclusions from the base-case analysis.

According to deterministic analysis, self-help with support was the most cost-effective option with a cost per QALY of £8,822 versus wait list that is well below lower NICE cost-effectiveness threshold of £20,000 per QALY. CBT-ED individual was not cost effective as its ICER versus self-help with support was more than £1 million per QALY. According to sensitivity analysis the results were sensitive to the probability of remission associated with the self-help with support, the utility value for remission, and intervention costs. The results were robust under all scenarios examined in bias adjustment analyses. Conclusions of probabilistic analysis were similar to those of deterministic analysis. At the lower NICE cost-effectiveness threshold of £20,000 per QALY (NICE, 2008b) the probability of self-help with support being cost effective was 0.80 and it increased to 0.89 at the upper threshold of £30,000 per QALY.

7.2.7.2.3 *Findings of the secondary economic analysis*

According to the secondary analysis, where the impact of extending the time horizon of the analysis to five years was explored, the ICER of CBT-ED individual versus self-help with support always remained above upper NICE threshold of £30,000 per QALY. However, the ICER of CBT-ED individual versus wait list at five years was reduced to £8,171 per QALY (from £55,100 per QALY at 1 year follow up). The ICER associated with CBT-ED individual was reduced to the lower NICE cost-effectiveness threshold by approximately 2.5 years follow up. At five years the conclusions of probabilistic analysis were similar to those of deterministic analysis (that is, self-help with support remained the preferred treatment option). At lower NICE cost-effectiveness threshold of £20,000 per QALY the probability of self-help with support and CBT-ED individual being cost effective was 0.60 and 0.37, respectively. When comparing only CBT-ED individual with wait list the probability of CBT-ED individual and wait list being cost effective at lower NICE cost-effectiveness threshold of £20,000 per QALY was 0.60 and 0.40, respectively. The probability of CBT-ED individual being cost effective increased to 0.65 at upper NICE cost-effectiveness threshold of £30,000 per QALY.

Similarly, in the scenario where the relapse rate associated with CBT-ED individual was assumed to be zero (that is, everyone sustains the treatment effect) and the annual relapse rate associated with self-help with support and wait list was assumed to be equivalent to the baseline rate of 0.27 (which was applied every year during the long term follow-up) the ICER of CBT-ED individual versus self-help with support was reduced to £35,578. However, it was still above upper NICE cost-effectiveness threshold (that is, self-help with support was still the preferred treatment option). The ICER of CBT-ED individual versus wait list was reduced to £3,788.

At the long-term follow up self-help with support remained the cost-effective option. Only, when CBT-ED individual was compared with wait list the ICER of CBT-ED individual was reduced below lower NICE cost-effectiveness threshold and supported the committee's view that CBT-ED individual has potentially more favourable cost effectiveness in the long run.

7.2.7.2.4 Strengths and limitations

Clinical data on remission were synthesised using network meta-analytic techniques. Such methods enabled evidence synthesis from both direct and indirect comparisons between treatments. The base-case economic analysis considered only data on remission at the end of treatment. However, the secondary analysis was undertaken where the time horizon of the analysis was extended to five years. Due to the lack of suitable data the cost estimates during the follow up were based on the committee expert opinion.

7.2.8 Clinical evidence statements for people with bulimia nervosa

7.2.8.1 Individual Therapy

CBT-ED versus another interventions in adults with bulimia nervosa at end of treatment.

Very low quality evidence from four RCTs (n=359) showed CBT-ED is no difference in the effect of CBT-ED on purges compared with any other intervention.

Low quality evidence from ten RCTs (n=387) showed CBT-ED is more effective on frequency of bingeing compared with any other intervention.

Moderate quality evidence from seven RCTs (n=560) showed no difference in the effect of CBT-ED on vomiting compared with any other intervention.

Low quality evidence from two RCTs (n=208) showed no difference in the effect of CBT-ED on laxative use compared with any other intervention.

Low quality evidence from four RCTs (n=261) showed CBT-ED is more effective on symptom checklist compared with any other intervention but there was some uncertainty.

Low quality evidence from one RCT (n=62) showed no difference in the effect of CBT-ED on symptom checklist compared with any other intervention in people who had bulimia nervosa less than five years.

Low quality evidence from three RCTs (n=197) showed no difference in the effect of CBT-ED on symptom checklist compared with any other intervention in people who had bulimia nervosa more than five years.

Low quality evidence from one RCT (n=80) showed no difference in the effect of CBT-ED on quality of life compared with any other intervention.

Very low quality evidence from five RCTs (n=419) showed no difference in the effect of CBT-ED on EDE- total compared with any other intervention.

Very low quality evidence from 10 RCTs (n=630) showed CBT-ED is more effective on depression compared with any other intervention.

Very low quality evidence from 10 RCTs (n=723) showed CBT-ED is more effective on EDE-dietary restraint compared with any other intervention.

Very low quality evidence from 10 RCTs (n= 725) showed CBT-ED is more effective on EDE-attitude to weight compared with any other intervention but there was some uncertainty.

Low quality evidence from five RCTs (n=477) showed no difference in the effect of CBT-ED is more effective on EDE- eating concern compared with any other intervention..

Very low quality evidence from 11 RCTs (n=837) showed no difference in the effect of CBT-ED on EDE-attitude to shape compared with any other intervention but there was some uncertainty

Very low to low quality evidence from four RCTs (n=242) showed CBT-ED is more effective on EDI- bulimia compared with any other intervention but there was some uncertainty.

Very low quality evidence from four RCTs (n=200 to 243) showed no difference in the effect of CBT-ED on EDI- body dissatisfaction and EDI-drive for thinness compared with any other intervention.

Low quality evidence from three RCTs (n=111) showed no difference in the effect of CBT-ED on global clinical score compared with any other intervention.

Low quality evidence from one RCT (n=22) showed no difference in the effect of CBT-ED on general psychopathology compared with any other intervention.

Low quality evidence from one RCT (n=45) showed CBT-ED is more effective on bulimic inventory test compared with any other intervention.

Low to moderate quality evidence from one RCT (n=116) showed CBT-ED is more effective on reducing bulimic and vomiting episodes compared with any other intervention.

Very low quality evidence from seven RCTs (n=731) showed CBT-ED is more effective on remission compared with any other intervention.

CBT-ED versus another intervention in adults with bulimia nervosa at follow up

Low quality evidence from three RCTs (n=208) showed no difference in the effect of CBT-frequency of purges compared with any other intervention.

Very low quality evidence from five RCTs (n=294) showed no difference in the effect of CBT-ED on frequency of bingeing compared with any other intervention

Low quality evidence from three RCTs (n=162) showed no difference in the effect of CBT-ED on frequency of vomiting compared with any other intervention.

Low quality evidence from one RCTs (n=98) showed no difference in the effect of CBT-ED on laxative use compared with any other intervention.

Low quality evidence from two RCTs (n=166) showed no difference in the effect of CBT-ED on symptom checklist compared with any other intervention.

Low quality evidence from one RCT (n=52) showed no difference in the effect of CBT-ED on quality of life compared with any other intervention.

Low quality evidence from three RCTs (n=307) showed no difference in the effect of CBT-ED on EDE- total compared with any other intervention.

Moderate quality evidence from eight RCTs (n=410) showed CBT-ED is more effective on depression compared with any other intervention, but there was some uncertainty.

Low quality evidence from three RCTs (n=126) showed no difference in the effect of CBT-ED on EDE-dietary restraint, EDE-attitude to shape and EDE-attitude to weight compared with any other intervention.

Low quality evidence from one RCT (n=52) showed no difference in the effect of CBT-ED on EDE-eating concerns compared with any other intervention.

Low quality evidence from one RCTs (n=27 to 47) showed no difference in the effect of CBT-ED on EDI-drive for thinness, EDI-body dissatisfaction and EDI- bulimia compared with any other intervention.

Low quality evidence from one RCT (n=22) showed CBT-ED is more effective on global clinical score compared with any other intervention, but there was some uncertainty.

Low quality evidence from two RCTs (n=49) showed CBT-ED is more effective on general psychopathology compared with any other intervention but there was some uncertainty.

Low quality evidence from one RCT (n=47) showed no difference in the effect of CBT-ED bulimic inventory test compared with any other intervention.

Low quality evidence from four RCTs (n=553) showed CBT-ED is more effective on improving remission compared with any other intervention, but there was some uncertainty.

CBT-ED versus another intervention in young people with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=86) showed no difference in the effect of CBT-ED on purging compared with any other intervention, but there was some uncertainty.

Low quality evidence from two RCTs (n=157) showed no difference in the effect of CBT-ED on bingeing compared with any other intervention.

Low quality evidence from one RCT (n=71) showed CBT-ED is less effective on vomiting compared with any other intervention.

Low quality evidence from one RCT (n=71) showed no difference in the effect of CBT-ED on laxative use compared with any other intervention.

Low to very low quality evidence from one RCT (n=70 to 71) showed CBT-ED is less effective on EDE-total, EDE- dietary restraint and EDE-attitude to weight compared with any other intervention.

Low quality evidence from one RCT (n=110) showed CBT-ED is less effective on remission rates compared with any other intervention.

CBT-ED versus another intervention in young people with bulimia nervosa at follow up

Low quality evidence from one RCT (n=68 to 69) showed no difference in the effect of CBT-ED on purging, vomiting and depression compared with any other intervention.

Low quality evidence from two RCTs (n=137) showed no difference in the effect of CBT-ED on bingeing compared with any other intervention.

Low quality evidence from one RCT (n=64 to 68) showed CBT-ED is less effective on EDE-dietary restraint, EDE-attitude to weight and EDE-attitude to shape compared with any other intervention.

Low quality evidence from one RCT (n=64) showed no difference in the effect of CBT-ED on EDE-total compared with any other intervention.

Very low quality evidence from one RCT (n=110) showed CBT-ED is less effective on remission rates compared with any other intervention.

IPT versus another intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from three RCTs (n=215) showed IPT is less effective on social adjustment scale compared with any other intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=131) showed IPT is less effective on purging compared with any other intervention.

Low quality evidence from two RCTs (n=178) showed IPT is less effective on vomiting compared with any other intervention.

Low quality evidence from two RCTs (n=98) showed IPT is less effective on bingeing compared with any other intervention.

Very low quality evidence from three RCTs (n=202) showed IPT is less effective on depression compared with any other intervention.

Very low quality evidence from one RCT (n=116) showed IPT is more effective on laxative use compared with any other intervention.

Low quality evidence from three RCTs (n=425) showed IPT is less effective on remission compared with any other intervention.

Low quality evidence from two RCTs (n=247) showed IPT is less effective on EDE- total compared with any other intervention.

Low quality evidence from two RCTs (n=247) showed IPT is less effective on EDE-eating concern compared with any other intervention but there was some uncertainty

Very low quality evidence from three RCTs (n=303) showed IPT is less effective on EDE-restraint compared with any other intervention.

Very low quality evidence from three RCTs (n=303) showed no difference in the effect of IPT on EDE-weight concern compared with any other intervention.

Very low quality evidence from three RCTs (n=303) showed no difference in the effect of IPT on EDE-shape concern compared with any other intervention

Low quality evidence from two RCTs (n=247) showed IPT is less effective on EDE-eating concern compared with any other intervention.

IPT versus another intervention in adults with bulimia nervosa at follow up.

Low quality evidence from two RCTs (n=166) showed no difference in the effect of IPT on the social adjustment scale compared with any other intervention.

Low quality evidence from one RCT (n=130) showed no difference in the effect of IPT on purging compared with any other intervention.

Low quality evidence from two RCTs (n=135) showed no difference in the effect of IPT on vomiting compared with any other intervention.

Low quality evidence from one RCT (n=98) showed no difference in the effect of IPT on bulimic episodes and laxative use compared with any other intervention.

Very low quality evidence from three RCTs (n=135) showed no difference in the effect of IPT on depression compared with any other intervention.

Low quality evidence from three RCTs (n=425) showed no difference in the effect of IPT on remission compared with any other intervention.

Low quality evidence from two RCTs (n=227) showed IPT is less effective on EDE-total compared with any other intervention.

Low quality evidence from three RCTs (n=264) showed IPT is less effective on EDE-restraint compared with any other intervention.

Low quality evidence from three RCTs (n=264) showed no difference in the effect of IPT on EDE-weight concern and EDE-shape concern compared with any other intervention.

Low quality evidence from two RCTs (n=227) showed no difference in the effect of IPT on EDE-eating concern compared with any other intervention.

Low quality evidence from two RCTs (n=166) showed no difference in the effect of IPT on the symptom checklist compared with any other intervention.

ICAT versus another intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=80) showed no difference in the effect of ICAT on the EDE-global, purges, bingeing and depression compared with any other intervention.

ICAT versus another intervention in adults with bulimia nervosa at follow up.

Low quality evidence from one RCT (n=80) showed no difference in the effect of ICAT on the EDE-global, purges, bingeing and depression compared with any other intervention.

CBT-ED versus CBT-ED in adults with bulimia nervosa at end of treatment.

Low quality evidence from two RCT (n=114) showed no difference in the effect of CBT-ED on purges compared with alternative type of CBT-ED.

Very low quality evidence from one RCT (n=72) showed no difference in the effect of CBT-ED laxative use compared with alternative type of CBT-ED.

Low quality evidence from two RCTs (n=123) showed no difference in the effect of CBT-ED on vomiting compared with alternative type of CBT-ED.

Very low quality evidence from five RCTs (n=306) showed no difference in the effect of CBT-ED on depression compared with alternative type of CBT-ED.

Very low quality evidence from one RCT (n=50) showed CBT-ED is less effective on depression compared with alternative type of CBT-ED but there was some uncertainty in those who binged less than 18 times per month.

Very low quality evidence from four RCTs (n=256) showed CBT-ED is more effective on depression compared with alternative type of CBT-ED but there was some uncertainty in those who binged more than 18 times per month.

Low quality evidence from four RCTs (n=321) showed no difference in the effect of CBT-ED on remission compared with alternative type of CBT-ED.

Low quality evidence from four RCTs (n=242) showed one type of CBT-ED is more effective on bulimic episodes compared with alternative type of CBT-ED but there was some uncertainty.

Low quality evidence from one RCT (n=72) showed no difference in the effect of CBT-ED on global function compared with alternative type of CBT-ED but there was some uncertainty.

Low to very low quality evidence from one RCT (n=149) showed no difference in the effect of CBT-ED on general psychiatric features, bulimic episodes, vomiting episodes, purging and laxative misuse compared with alternative type of CBT-ED.

Low quality evidence from two RCTs (n=142) showed no difference in the effect of CBT-ED on social adjustment score compared with alternative type of CBT-ED.

Very low quality evidence from three RCTs (n=291) showed no difference in the effect of CBT-ED on symptom check-list compared with alternative type of CBT-ED.

Low quality evidence from one RCT (n=72) showed no difference in the effect of CBT-ED on EDI- drive for thinness compared with alternative type of CBT-ED.

Low quality evidence from two RCTs (n=122) showed no difference in the effect of CBT-ED on EDI- bulimia and EDI-body dissatisfaction compared with alternative type of CBT-ED.

Low quality evidence from three RCTs (n=319) showed no difference in the effect of CBT-ED on EDI- total compared with alternative type of CBT-ED.

Low quality evidence from four RCTs (n=361) showed no difference in the effect of CBT-ED on EDE-total compared with alternative type of CBT-ED.

CBT-ED versus CBT-ED in adults with bulimia nervosa at follow up.

Low to very low quality evidence from two RCTs (n=111) showed no difference in the effect of CBT-ED on purges compared with alternative type of CBT-ED.

Low quality evidence from one RCT (n=72) showed no difference in the effect of CBT-ED on laxative use compared with alternative type of CBT-ED.

Low to very low quality evidence from four RCTs (n=280) showed no difference in the effect of CBT-ED on bingeing and depression compared with alternative type of CBT-ED.

Low quality evidence from three RCTs (n=232) showed no difference in the effect of CBT-ED on vomiting compared with alternative type of CBT-ED.

Low to very low quality evidence from three RCTs (n=144) showed no difference in the effect of CBT-ED on remission compared with alternative type of CBT-ED.

Low quality evidence from one RCT (n=72) showed no difference in the effect of CBT-ED on global function compared with alternative type of CBT-ED.

Low quality evidence from one RCT (n=149) showed no difference in the effect of CBT-ED on general psychiatric features, bingeing episodes, vomiting episodes, purging and laxative use compared with alternative type of CBT-ED.

Low quality evidence from two RCTs (n=170) showed no difference in the effect of CBT-ED on social adjustment score compared with alternative type of CBT-ED.

Very low quality evidence from two RCTs (n=169) showed no difference in the effect of CBT-ED on symptom check-list compared with alternative type of CBT-ED.

Low quality evidence from one RCT (n=72) showed no difference in the effect of CBT-ED on EDI- drive for thinness compared with alternative type of CBT-ED.

Low quality evidence from two RCTs (n=122) showed no difference in the effect of CBT-ED on EDI- bulimia and EDI-body dissatisfaction compared with alternative type of CBT-ED.

Low quality evidence from three RCTs (n= 319) showed no difference in the effect of CBT-ED on EDI- total compared with alternative type of CBT-ED.

Low quality evidence from two RCTs (n=199) showed no difference in the effect of CBT-ED on EDE-total compared with alternative type of CBT-ED.

Behavioural therapy versus any other intervention in adults with bulimia nervosa at the end of treatment.

Low quality evidence from three RCTs (n=183) showed no difference in the effect of behavioural therapy on bulimic episodes compared with any other intervention.

Low quality evidence from one RCT (n=92) showed no difference in the effect of behavioural therapy laxative use compared with any other intervention.

Low quality evidence from three RCTs (n=160) showed behavioural therapy is more effective on vomiting compared with any other intervention.

Low quality evidence from one RCT (n=41) showed behavioural therapy is more effective on vomiting compared with any other intervention in people who had BN less than 5 years or binged less than 18 times per month.

Low quality evidence from two RCTs (n=119) showed no difference in the effect of behavioural therapy on vomiting compared with any other intervention in people who had BN more than five years or binged more than 18 times per month.

Very low quality evidence from five RCTs (n=185) showed no difference in the effect of behavioural therapy on depression compared with any other intervention.

Low quality evidence from three RCTs (n=106) showed no difference in the effect of behavioural therapy on remission compared with any other intervention.

Low quality evidence from one RCTs (n=62) showed no difference in the effect of behavioural therapy on the symptom check list compared with any other intervention.

Low quality evidence from one RCT (n=62) showed no difference in the effect of behavioural therapy on the social adjustment scale compared with any other intervention.

Very low quality evidence from two RCTs (n=89) showed no difference in the effect of behavioural therapy on EDE-dietary restraint, EDE-attitudes to weight and EDE-attitudes to shape compared with any other intervention.

Low quality evidence from three RCTs (n=139 to 139) showed behavioural therapy is more effective on the EDI- bulimia, EDE-body dissatisfaction and EDI- drive for thinness compared with any other intervention.

Behavioural therapy versus any other intervention in adults with bulimia nervosa at follow up.

Low quality evidence from one RCT (n=27) showed behavioural therapy is less effective on purging and vomiting compared with any other intervention.

Low quality evidence from one RCT (n=27) showed no difference in the effect of behavioural therapy on bingeing compared with any other intervention.

Low quality evidence from two RCTs (n=74) showed no difference in the effect of behavioural therapy on depression compared with any other intervention.

Low quality evidence from one RCT (n=75) showed no difference in the effect of behavioural therapy on remission compared with any other intervention.

Low quality evidence from one RCT (n=27) showed no difference in the effect of behavioural therapy on EDE-dietary restraint, EDE-attitudes to shape and EDE-attitudes to weight compared with any other intervention.

Low quality evidence from one RCT (n=47) showed behavioural therapy is more effective on EDI- drive for thinness compared with any other intervention.

Low quality evidence from one RCT (n=27) showed no difference in the effect of behavioural therapy on EDI- body dissatisfaction and EDI- bulimia compared with any other intervention.

CBT-ED versus wait list controls in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=21) showed CBT-ED is more effective on purge frequency compared with wait list controls.

Very low quality evidence from three RCTs (n=113) showed CBT-ED is more effective on binge episodes compared with wait list controls.

Low quality evidence from two RCTs (n=92) showed CBT-ED is more effective on vomiting compared with wait list controls.

Low quality evidence from one RCT (n=35) showed CBT-ED is more effective on depression compared with wait list controls.

Low quality evidence from one RCT (n=52) showed CBT-ED is more effective on laxative use compared with wait list controls.

Low quality evidence from one RCT (n=71) showed CBT-ED is more effective on remission compared with wait list controls, but there was some uncertainty.

Low quality evidence from one RCT (n=154) showed CBT-ED is more effective on symptom check list compared with wait list controls.

Low quality evidence from two RCTs (n=194) showed CBT-ED is more effective on overall severity compared with wait list controls.

Low quality evidence from one RCT (n=122) showed CBT-ED is more effective on general psychiatric features compared with wait list controls.

Low quality evidence from one RCT (n=41) showed CBT-ED is more effective on EDE-dietary restraint, EDE-attitudes to weight, EDI- bulimia and EDI-drive for thinness compared with wait list controls.

Low quality evidence from two RCTs (n=89) showed CBT-ED is more effective on EDE-attitudes to shape compared with wait list controls.

Low quality evidence from one RCT (n=41) showed no difference in the effect of CBT-ED on body dissatisfaction compared with wait list controls.

Low quality evidence from one RCT (n=153 to 154) showed CBT-ED is more effective on purge frequency, bulimic episodes, vomiting and laxative use compared with wait list controls.

Behavioural therapy (BT) versus wait list controls in adults with bulimia nervosa at the end of treatment

Low quality evidence from one RCT (n=50) showed BT is more effective on binge frequency, self-induced vomiting, laxative use compared with wait list controls.

Low quality evidence from one RCT (n=50) showed no difference in the effect of BT on depression compared with wait list controls.

Hybrid therapy versus any other intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=86) showed no difference in the effect a hybrid therapy on binge eating, symptom check list, depression and EDI- symptoms compared with any other intervention.

Hybrid therapy versus any other intervention in adults with bulimia nervosa at follow up.

Low quality evidence from one RCT (n=86) showed no difference in the effect a hybrid therapy on binge eating, symptom check list, depression and EDI- symptoms compared with any other intervention.

DBT versus any other intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=29) showed no difference in the effect of DBT on negative mood regulation compared with any other intervention.

Low quality evidence from one RCT (n=29) showed DBT is more effective on depression and emotional eating compared with any other intervention.

Psychodynamic general therapy versus any other intervention in adults with bulimia nervosa at end of treatment.

Very low evidence from two RCTs (n=116) showed no difference in the effect of psychodynamic general therapy on binge eating compared with any other intervention.

Very low evidence from two RCTs (n=120) showed psychodynamic general therapy *is* less effective on EDE-dietary restraint compared with any other intervention.

Very low evidence from two RCTs (n=120) showed psychodynamic general therapy is less effective on vomiting/purging episodes compared with any other intervention but there was some uncertainty

Very low evidence from two RCTs (n=120) showed no difference in the effect of psychodynamic general therapy on EDE-attitudes of weight and EDE-attitude to shape compared with any other intervention.

Very low evidence from one RCT (n=50) showed psychodynamic general therapy is less effective on EDI-drive for thinness and EDI- bulimia compared with any other intervention.

Very low evidence from one RCT (n=50) showed no difference in the effect of psychodynamic general therapy on EDI-body dissatisfaction compared with any other intervention.

Very low evidence from one RCT (n=50) showed psychodynamic general therapy is more effective on depression compared with any other intervention.

Very low evidence from one RCT (n=50) showed no difference in the effect of psychodynamic general therapy on general psychopathology compared with any other intervention.

7.2.8.2 Group Therapy

Group behavioural therapy versus another group behavioural therapy in adults with bulimia nervosa at end of treatment.

Very low to low quality evidence from one RCT (n=23) showed no difference in the effect of group behavioural therapy on vomiting and depression compared with another group behavioural therapy.

Low quality evidence from one RCT (n=30) showed no difference in the effect of group behavioural therapy on remission compared with another group behavioural therapy.

Group behavioural therapy versus another group behavioural therapy in adults with bulimia nervosa at follow up.

Low quality evidence from one RCT (n=23 to 24) showed no difference in the effect of group behavioural therapy on vomiting and depression compared with another group behavioural therapy.

Low quality evidence from one RCT (n=30) showed no difference in the effect of group behavioural therapy on remission compared with another group behavioural therapy.

Group CBT-ED versus wait list control in adults with bulimia nervosa at end of treatment.

Low quality evidence from two RCTs (n=54) showed no difference in the effect of group CBT-ED on bingeing compared with wait list controls.

Low quality evidence from one RCTs (n=24) showed no difference in the effect of group CBT-ED on purging compared with wait list controls.

Low quality evidence from one RCT (n=24) showed a benefit of CBT-ED on vomiting and depression compared with wait list controls.

Low quality evidence from one RCT (n=26) showed no difference in the effect of group CBT-ED on EDI- body dissatisfaction, EDI-bulimia and EDI-drive for thinness compared with wait list controls.

Low quality evidence from two RCTs (n=52) showed a benefit of CBT-ED on remission, compared with wait list controls however there was some uncertainty.

Group CBT-ED versus wait list control in adults with bulimia nervosa at follow up

Low quality evidence from two RCTs (n=59) showed a benefit of CBT-ED on remission, compared with wait list control.

Group CBT-ED versus any other intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from three RCTs (n=206) showed no difference in the effect of group CBT-ED on bingeing, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction, compared with any other intervention.

Low quality evidence from two RCTs (n=145) showed no difference in the effect of group CBT-ED on EDI-global compared with any other intervention controls.

Low quality evidence from one RCT (n=120) showed no difference in the effect of group CBT-ED on EDE-total, anxiety and symptom check-list compared with any other intervention.

Low quality evidence from one RCT (n=74) showed a benefit of group CBT-ED on clinical impairment compared with any other intervention.

Low quality evidence from three RCTs (n=211) showed no difference in the effect of group CBT-ED on depression compared with any other intervention.

Low quality evidence from two RCTs (n=91) showed a potentially harmful effect of group CBT-ED on vomiting compared with any other intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=56) showed a potentially harmful effect of group CBT-ED on laxative use compared with any other intervention.

Low quality evidence from one RCT (n=81) showed no difference in the effect of group CBT-ED on remission compared with any other intervention

Group CBT-ED versus any other intervention in adults with bulimia nervosa at follow up

Low quality evidence from three RCTs (n=205) showed no difference in the effect of group CBT-ED on bingeing, EDI-drive for thinness and EDI-bulimia, compared with any other intervention.

Low quality evidence from three RCTs (n=205) showed a benefit of group CBT-ED on EDI-body dissatisfaction compared with any other intervention but there was some uncertainty.

Low quality evidence from one RCT (n=74) showed no difference in the effect of group CBT-ED on EDI-global compared with any other intervention.

Low quality evidence from one RCT (n=120) showed no difference in the effect of group CBT-ED on EDE-total compared with any other intervention.

Low quality evidence from two RCTs (n=91) showed a potentially harmful effect of group CBT-ED on vomiting compared with any other intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=56) showed a potentially harmful effect of group CBT-ED on laxative use compared with any other intervention.

Low quality evidence from three RCTs (n=205) showed a benefit of group CBT-ED on anxiety compared with any other intervention but there was some uncertainty.

Low quality evidence from one RCT (n=120) showed no difference in the effect of group CBT-ED on symptom checklist compared with any other intervention.

Low quality evidence from one RCT (n=74) showed a benefit of group CBT-ED on clinical impairment compared with any other intervention.

Low quality evidence from three RCTs (n=211) showed no difference in the effect of group CBT-ED on depression compared with any other intervention.

Low quality evidence from two RCTs (n=126) showed no difference in the effect of group CBT-ED on remission compared with any other intervention.

Group BT-ED versus wait list control in adults with bulimia nervosa at end of treatment

Low to very low quality evidence from one RCT (n=26) showed no difference in the effect of group behavioural therapy on bingeing, EDI-drive for thinness and EDI-bulimia, compared with wait list control.

Low quality evidence from one RCT (n=35) showed group behavioural therapy is more effective on vomiting and depression compared with wait list controls.

Low quality evidence from one RCT (n=26) showed group behavioural therapy is more effective on EDI-body dissatisfaction compared with wait list controls, but there was some uncertainty.

Low quality evidence from one RCT (n=44) showed group behavioural therapy is more effective on remission compared with wait list controls.

Group BT-ED versus wait list control in adults with bulimia nervosa at follow up

Very low quality evidence from one RCT (n=44) showed no difference in the effect of group behavioural therapy on remission compared with wait list control

Group BT-ED versus another group intervention in adults with bulimia nervosa at end of treatment

Low to very low quality evidence from one RCT (n=30 to 36) showed no difference in the effect of group behavioural therapy on bingeing, vomiting, depression, EDI-drive for thinness, EDI-bulimia and EDI-body dissatisfaction compared with another group intervention.

Low quality evidence from one RCT (n=60) showed group behavioural therapy is more effective on remission compared *with* another group intervention.

Group BT-ED versus another group intervention in adults with bulimia nervosa at follow up.

Low to very low quality evidence from one RCT (n=20 to 36) showed no difference in the effect of group behavioural therapy on vomiting, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction, EDE-shape concern, EDE-weight concern, EDE-eating concern and EDE-restraint compared with another group intervention.

Low quality evidence from two RCTs (n=58 to 63) showed no difference in the effect of group behavioural therapy on bingeing and depression with compared another group intervention.

Low quality evidence from two RCTs (n=73) showed no difference in the effect of group behavioural therapy on remission compared with another group intervention.

Group psychoeducation versus another intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=54 to 65) showed no difference in the effect of group psychoeducation on bingeing, vomiting, EDI-body dissatisfaction, EDE-shape concern compared another intervention.

Very low quality evidence from one RCT (n= 65) showed no difference in the effect of group psychoeducation on remission compared with another intervention.

Low quality evidence from one RCT (n=54) showed group psychoeducation is less effective on EDI-drive for thinness compared with another group intervention.

Low quality evidence from one RCT (n=54) showed group psychoeducation is less effective on EDI-bulimia compared with another intervention, but there was some uncertainty.

Group CBT-ED of varying intensity versus control CBT-ED in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=143) showed a benefit of group CBT-ED of varying intensity on vomiting, EDI-drive for thinness, EDI- bulimia and remission compared with a control CBT-ED group.

Low quality evidence from one RCT (n=143) showed a benefit of group CBT-ED of varying intensity on bingeing compared with a control CBT-ED group, but there was some uncertainty.

Low quality evidence from one RCT (n=143) showed no difference in the effect of group CBT-ED of varying intensity on laxative use, EDI-body dissatisfaction, depression and anxiety compared with a control CBT-ED group.

Group emotional and mind training versus another intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=74) showed no difference in the effect of group emotional and mind training on EDE-global compared with another intervention.

Low quality evidence from one RCT (n=74) showed group emotional and mind training is less effective on clinical impairment compared with another intervention.

Group emotional and mind training versus another intervention in adults with bulimia nervosa at follow up.

Low quality evidence from one RCT (n=74) showed no difference in the effect of group emotional and mind training on EDE-global compared with another intervention.

Low quality evidence from one RCT (n=74) showed group emotional and mind training may be less effective on clinical impairment compared with another intervention.

Group support versus another intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n=100) showed no difference in the effect of group support on depression scores compared with another intervention.

7.2.8.3 Self-help

Guided self-help (ED) versus another intervention in adults with bulimia nervosa at end of treatment

Very low quality evidence from six RCTs (n=388) showed guided self-help (ED) is more effective on bingeing compared with another intervention but there was some uncertainty.

Very low quality evidence from five RCTs (n=190) showed guided self-help (ED) is more effective on vomiting compared with another intervention but there was some uncertainty.

Low quality evidence from one RCT (n=55 to 56) showed guided self-help (ED) is more effective on EDI-body dissatisfaction and EDI-bulimia compared with another intervention.

Low quality evidence from one RCT (n= 56) showed guided self-help (ED) is more effective on ED-drive for thinness compared with another intervention but there was some uncertainty.

Very low quality evidence from five RCTs (n=280) showed no difference in the effect of guided self-help (ED) on depression compared with any other intervention.

Very low quality evidence from four RCTs (n=454) showed no difference in the effect of guided self-help (ED) on remission compared with any other intervention.

Very low quality evidence from three RCTs (n=159 to 192) showed no difference in the effect of guided self-help (ED) on EDE-global, EDE-restraint, EDE-weight concern, EDE-shape concern and excessive exercise compared with any other intervention.

Very low quality evidence from two RCTs (n=145) showed no difference in the effect of guided self-help (ED) on EDE-eating concern compared with any other intervention.

Very low quality evidence from one RCT (n=80) showed no difference in the effect of guided self-help (ED) on purging and satisfaction with life compared with any other intervention.

Very low quality evidence from two RCTs (n=112) showed guided self-help (ED) is less effective on bulimia inventory index with life compared with any other intervention, but there was some uncertainty.

Guided self-help (ED) versus another intervention in young people with bulimia nervosa at end of treatment.

Very low quality evidence from one RCT (n=85) showed no difference in the effect of guided self-help (ED) on remission compared with any other intervention.

Guided self-help (ED) versus another intervention in adults with bulimia nervosa at follow up.

Very low quality evidence from four RCTs (n=270) showed no difference in the effect of guided self-help (ED) on bingeing compared with any other intervention.

Very low quality evidence from three RCTs (n=95) showed no difference in the effect of guided self-help (ED) on vomiting compared with any other intervention.

Very low quality evidence from three RCTs (n=154) showed no difference in the effect of guided self-help (ED) on depression compared with any other intervention.

Very low quality evidence from three RCTs (n=216) showed guided self-help (ED) is less effective on the use of laxatives compared with any other intervention.

Very low quality evidence from two RCTs (n=99) showed no difference in the effect of guided self-help (ED) on EDE-restraint, EDE-weight concern and EDE-shape concern compared with any other intervention.

Very low quality evidence from one RCT (n=52 to 55) showed no difference in the effect of guided self-help (ED) on purging, EDE eating concern and satisfaction with life, EDI-body dissatisfaction and EDI-bulimia and EDI- drive for thinness compared with any other intervention.

Very low quality evidence from two RCTs (n=159) showed no difference in the effect of guided self-help (ED) on excessive exercise compared with any other intervention.

Low quality evidence from one RCT (n=47) showed guided self-help (ED) is less effective on bulimic inventory index compared with any other intervention.

Very low quality evidence from four RCTs (n=454) showed no difference in the effect of guided self-help (ED) on remission compared with any other intervention.

Guided self-help (ED) versus another intervention in young people with bulimia nervosa at follow up.

Very low quality evidence from one RCT (n=85) showed no difference in the effect of guided self-help (ED) on remission compared with any other intervention.

Guided self-help (ED) versus wait list controls in adults with bulimia nervosa at end of treatment.

Low to very low quality evidence from two RCTs (n=111 to 151) showed guided self-help (ED) is more effective on bingeing and vomiting compared with wait list controls, but there was some uncertainty.

Very low quality evidence from two RCTs (n= 151) showed guided self-help (ED) is more effective on the clinical symptom index compared with wait list controls.

Very low quality evidence from two RCTs (n= 151) showed no difference in the effect of guided self-help (ED) on laxative use compared with wait list controls.

Very low quality evidence from two RCTs (n= 178) showed guided self-help (ED) is more effective on the purging compared with wait list controls. A subgroup analysis showed the

effect is greater in those with a less severe case of BN who binge less than 18 times per month compared with those who binge more than 18 times per month.

Very low quality evidence from three RCTs (n= 220) showed guided self-help (ED) is more effective on depression compared with wait list controls.

Very low quality evidence from two RCTs (n= 178) showed guided self-help (ED) is more effective on EDI- drive for thinness, EDI-body dissatisfaction, EDE shape concern, EDE-global, quality of life and EDE weight concern compared with wait list controls.

A subgroup analysis showed the effect of guided self-help on EDE-shape concern and EDE-weight concern is greater in those with a less severe case of BN who binge less than 18 times per month compared with those who binge more than 18 times per month.

Very low quality evidence from one RCT (n= 69) showed guided self-help (ED) is more effective on EDE-restraint and EDE-eating concern compared with wait list controls.

Very low quality evidence from one RCT (n= 69) showed no difference in the effect of guided self-help (ED) on EDI-bulimia compared with wait list controls.

Very low quality evidence from two RCTs (n=198) showed guided self-help (ED) is more effective on remission compared with wait list controls.

Guided self-help (ED) versus wait list controls in adults with bulimia nervosa at follow up.

Very low quality evidence from one RCT (n=78) showed no difference in the effect of guided self-help (ED) on remission compared with wait list controls.

General self-help versus another intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n= 56) showed general self-help is more effective on depression and EDE-eating concern compared with another intervention but there was some uncertainty.

Low quality evidence from one RCT (n= 32) showed general self-help is more effective on EDE-global compared with another intervention.

Low quality evidence from one RCT (n= 56) showed no difference in the effect of general self-help on EDE-restraint, EDE-shape concern and EDE-weight concern compared with another intervention.

General self-help versus wait list controls in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n= 56) showed no difference in the effect of general self-help on depression, EDE-restraint, EDE-shape concern, EDE-eating concern and EDE-weight concern compared with wait list controls.

Low quality evidence from one RCT (n= 63) showed general self-help may be harmful on remission compared with wait list controls.

General self-help versus wait list controls in adults with bulimia nervosa at follow up.

Low quality evidence from one RCT (n= 63) showed no difference in the effect of general self-help remission compared with wait list controls.

Self-help (ED) versus another intervention in adults with bulimia nervosa at end of treatment.

Very low quality evidence from three RCTs (n= 162) showed no difference in the effect of self-help (ED) on bingeing compared with another intervention.

Low quality evidence from one RCT (n= 70) showed self-help (ED) was less effective on purging compared with another intervention.

Very low quality evidence from one RCT (n= 33) showed no difference in the effect of self-help (ED) on the use of laxatives or exercising compared with another intervention.

Very low quality evidence from two RCTs (n= 96) showed self-help (ED) is less effective on vomiting compared with another intervention.

Low quality evidence from one RCT (n= 56) showed self-help (ED) is less effective on depression compared with another intervention but there was some uncertainty.

Very low quality evidence from two RCTs (n= 173) showed no difference in the effect of self-help (ED) on remission compared with another intervention.

Low quality evidence from two RCTs (n= 118 to 132) showed no difference in the effect of self-help (ED) on EDE-global, EDE-weight concern, EDE-shape concern compared with another intervention.

Very low quality evidence from two RCTs (n= 118) showed self-help (ED) is less effective on EDE-restraint compared with another intervention.

Very low quality evidence from one RCT (n= 56) showed self-help (ED) is less effective on EDE-eating concern compared with another intervention but there was some uncertainty.

Self-help (ED) versus another intervention in adults with bulimia nervosa at follow up.

Low quality evidence from one RCT (n= 70) showed no difference in the effect of self-help (ED) on purging and EDE-global compared with another intervention.

Very low to low quality evidence from one RCT (n=37 to 40) showed no difference in the effect of self-help (ED) on vomiting, use of laxatives and excessive exercise compared with another intervention.

Very low quality evidence from two RCTs (n= 111) showed no difference in the effect of self-help (ED) on bingeing compared with another intervention.

Very low quality evidence from one RCT (n= 90) showed no difference in the effect of self-help (ED) on remission compared with another intervention.

Internet self-help (ED) versus another intervention in adults with bulimia nervosa at end of treatment.

Low quality evidence from one RCT (n= 70) showed no difference in the effect of internet self-help (ED) on EDE-global compared with another intervention.

Low quality evidence from one RCT (n= 122) showed no difference in the effect of internet self-help (ED) on vomiting, laxative use and excessive exercise compared with another intervention.

Very low quality evidence from one RCT (n= 70) internet self-help (ED) is more effective on purging compared with another intervention.

Very low quality evidence from two RCTs (n= 192) showed internet self-help (ED) is more effective on bingeing compared with another intervention, although there was some uncertainty.

Low quality evidence from one RCT (n= 192) showed internet self-help (ED) is more effective on bingeing compared with another intervention in people who binge less than 18 times per month.

Low quality evidence from one RCT (n= 122) showed internet self-help (ED) is more effective on bingeing compared with another intervention in people who binge more than 18 times per month.

Low quality evidence from one RCT (n= 155) showed no difference in the effect of internet self-help (ED) on remission compared with another intervention.

Internet self-help (ED) versus another intervention in adults with bulimia nervosa at follow up.

Low quality evidence from two RCTs (n= 192) showed there is a benefit of internet self-help (ED) on bingeing compared with another intervention.

Low quality evidence from one RCT (n= 70) showed no difference in the effect of internet self-help (ED) on purging and EDE-global compared with another intervention.

Low quality evidence from one RCT (n= 122) showed no difference in the effect of internet self-help (ED) on vomiting, laxative use and excessive exercise compared with another intervention.

Low quality evidence from one RCT (n= 155) showed no difference in the effect of internet self-help (ED) on remission compared with another intervention.

Internet self-help (ED) versus wait list controls in adults with bulimia nervosa at end of treatment.

Very low quality evidence from two RCTs (n= 137) showed internet self-help (ED) is more effective on purging and bingeing compared with wait list controls.

Very low quality evidence from two RCTs (n= 137) showed no difference in the effect of internet self-help (ED) on vomiting compared with wait list controls.

Very low quality evidence from one RCT (n=67) showed a benefit of internet self-help (ED) on depression, EDE-restraint, EDE-shape concern, EDE-weight concern, EDE-eating concern and quality of life compared with wait list controls.

Very low quality evidence from two RCTs (n=137) showed a benefit of internet self-help (ED) on EDE-global compared with wait list controls.

Very low quality evidence from one RCT (n=76) showed a benefit of internet self-help (ED) on remission compared with wait list controls.

Self-help (ED) versus wait list controls in adults with bulimia nervosa at end of treatment

Very low quality evidence from two RCTs (n=130) showed no difference in the effect of self-help (ED) on bingeing compared with wait list controls.

Very low quality evidence from two RCTs (n=130) showed there may be a benefit of self-help (ED) on EDE-global compared with wait list controls.

Very low quality evidence from two RCTs (n=117) showed there may be a benefit of self-help (ED) on EDE-shape concern and EDE-weight concern compared with wait list controls.

Very low quality evidence from two RCTs (n=55 to 117) showed no difference in the effect of self-help (ED) on purging, vomiting, depression, EDE-restraint, EDE-eating concern and remission compared with wait list controls.

Text messaging versus wait list controls in adults with bulimia nervosa at end of treatment.

Very low quality evidence from one RCT (n=165) showed no effect of text messaging on remission compared with wait list controls.

7.2.8.4 Family Therapy

Family therapy-ED versus any individual therapy in young people with bulimia nervosa at end of treatment.

Very low quality evidence from three RCTs (n=295) showed family therapy-ED is more effective on improving remission compared with any individual therapy.

Low quality evidence from two RCTs (n=157) showed no difference in the effect of family therapy-ED on binge frequency compared with any individual therapy.

Very low quality evidence from one RCT (n=63) showed no difference in the number of people abstaining from vomiting compared with any individual therapy.

Low quality evidence from one RCT (n=155) showed that family therapy-ED is more effective on reducing EDE–global scores compared with any individual therapy.

Low quality evidence from one RCT (n=71) showed that family therapy-ED is more effective on reducing scores on EDE-vomiting, EDE-all compensatory behaviours, EDE-restraint, EDE-shape concern and EDE-weight concern, and on reducing the number of participants hospitalized during treatment compared with any individual therapy.

Low quality evidence from two RCTs (n=157) showed family therapy-ED may be effective on reducing depression compared with any individual therapy, although there was some uncertainty.

Low quality evidence from one RCT (n=86) showed family therapy-ED may be more effective on YBC-EDS compared with any individual therapy, although there was some uncertainty.

Low quality evidence from one RCT (n=86) showed no difference in the effect of family therapy-ED on frequency of purging compared with any individual therapy.

Low quality evidence from one RCT (n=71) showed no difference in the effect of family therapy-ED on EDE–objective binge eating compared with any individual therapy.

Low quality evidence from one RCT (n=68) showed no difference in the effect of family therapy-ED on service user experience compared with any individual therapy.

Family therapy-ED versus any individual therapy in young people with bulimia nervosa at follow up.

Very low quality evidence from two RCTs (n=215) showed family therapy-ED is more effective on improving remission compared with any individual therapy.

Low quality evidence from one RCT (n=137) showed that family therapy-ED is more effective on reducing scores on EDE–global compared with any individual therapy.

Low quality evidence from one RCT (n=68) showed that family therapy-ED is more effective on reducing scores on EDE–shape concern compared with any individual therapy.

Low quality evidence from two RCTs (n=137) showed no difference in the effect of family therapy-ED on binge frequency and depression compared with any individual therapy.

Low quality evidence from one RCT (n=68) showed family therapy-ED may be more effective on reducing scores on EDE–weight concern compared with any individual therapy, although there was some uncertainty.

Low quality evidence from one RCT (n=69) showed no difference in the effect of family therapy-ED on frequency of purging and YBC-EDS compared with any individual therapy.

Very low quality evidence from one RCT (n=54) showed no difference in the effect of family therapy-ED on the number of people abstaining from vomiting compared with any individual therapy.

Low quality evidence from one RCT (n=68) showed no difference in the effect of family therapy-ED on frequency of vomiting and EDE-dietary restraint compared with any individual therapy.

Low quality evidence from one RCT (n=71) showed family therapy-ED may be less effective on service user experience compared with any individual therapy, although there was some uncertainty.

7.2.9 Economic Evidence statements

There was existing economic evidence from 1 US study (N=128) comparing CBT ED and guided self-help ED in people with ED. However, due to the lack of QALYs the committee could not judge the cost effectiveness of CBT ED versus guided self-help ED. This evidence was partially applicable and characterised by minor methodological limitations.

In the economic analysis conducted for this guideline, self-help with support appeared to be the most cost-effective option for adults with BN when compared with CBT-ED individual and wait list using end of treatment effectiveness data. These results were overall robust to alternative scenarios considered in sensitivity analysis. At the lower NICE cost-effectiveness threshold of £20,000 per QALY (NICE, 2008b) the probability of self-help with support being cost effective was 0.80 and it increased to 0.89 at the upper threshold of £30,000 per QALY. The secondary analysis where the time horizon of the analysis was extended to five years indicated that the ICER of CBT-ED individual when compared with wait list was below lower NICE cost-effectiveness threshold of £20,000 per QALY by approximately 2.5 years follow up. At 5 years the probability of CBT-ED individual (when compared with wait list) being cost effective was 0.60 at lower NICE cost-effectiveness threshold of £20,000 per QALY. The evidence from the guideline economic analysis was directly applicable to the UK context and was characterised by potentially serious methodological limitations.

There evidence from one UK study (N=85) suggested that family therapy may be a cost-effective treatment option when compared with guided self-help ED in young people with BN or EDNOS. This evidence was from one partially applicable study that was characterised by minor methodological limitations. There was no evidence on the cost effectiveness of other psychological therapies for the treatment of young people with BN.

7.2.10 Recommendations and link to evidence for the review on Does any psychological intervention produce benefits/harms in children, young people or adults with an eating disorder compared with any other intervention or controls?

First line treatment for bulimia nervosa

	103. Consider bulimia-nervosa-focused guided self-help for adults with bulimia nervosa. **104. Bulimia-nervosa-focused guided self-help programmes for adults with bulimia nervosa should:** • use a cognitive behavioural self-help book for eating disorders • supplement the self-help programme with brief supportive sessions (for example four to nine sessions lasting 20 minutes each over 16 weeks running weekly at first) • be delivered by a practitioner who is competent in delivering the treatment.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating bulimia nervosa in children, young people and adults. For this population, binge eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with bulimia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience. The committee discussed whether to include weight/BMI as an outcome in the reviews on people with bulimia nervosa. They acknowledged that it was not clear whether either should be ideally higher, lower or show no change at the end of treatment because it is not a priority of treatment like it is for those with anorexia nervosa. The committee considered whether weight or BMI may still be an important outcome because a number of people would want to know if they are going to gain weight in response to treatment. They also discussed whether the results on weight and BMI may lead to a research recommendation or them not recommending the treatment. In the end, because of the uncertainty surrounding the importance of weight/BMI and how to interpret the data, they decided to exclude it from the reviews.
Trade-off between clinical benefits and harms	Guided self-help compared with wait list controls improved rates of remission and reduced the frequency of bingeing in adults with bulimia nervosa. There was a reduced frequency of other compensatory behaviours including vomiting (with some uncertainty) and purging. Other outcomes also improved in response to guided self-help, including quality of life, depression and eating disorder psychopathology (EDE-global, EDE subscales [except EDE-restraint], EDI subscales and clinical symptom index). Laxative use was not significantly different. At 12 months follow up, the benefits on remission were inconclusive (depending on what the comparison arm was). No other outcomes were reported. At no time point was data reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience Comparing guided self-help with any other intervention in adults showed there was no difference in remission rates at the end of treatment. Bingeing and vomiting were reduced but there was some uncertainty, only use of laxatives, EDI-bulimia and EDI-body dissatisfaction showed a meaningful reduction compared with any other intervention at the end of treatment. Depression scores were higher in the guided self-help group. All other outcomes showed no difference between the two groups at the end of the intervention (including EDI and EDE subscales, purging,

	exercising, satisfaction with life, bulimic inventory index). Six to twelve months follow up also showed minimal difference between guided self-help and any other treatment, and for the critical outcomes of bingeing and remission there was no difference (one of the other treatments included CBT-ED and at follow up remission was not different between guided self-help and CBT-ED). At no time point was data reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning or service user experience *Refer to the following LETR for a summary of the other treatments considered by the committee when making this recommendation.* *Refer to the LETR on young people with bulimia nervosa for a summary of the finding on guided self-help in young people.*
Trade-off between net health benefits and resource use	Existing economic evidence pertaining to the psychological therapies for adults with bulimia nervosa was very limited and the committee could not draw any conclusions from it. The network meta-analysis that was undertaken to inform the guideline economic analysis demonstrated that after excluding treatments with pooled number of participants less than 150 across all RCTs, CBT-ED individual had the highest probability of full remission (using the end of treatment data) followed by self-help with support (or guided self-help) in adults with bulimia nervosa. There was high heterogeneity and uncertainty surrounding the results as indicated by very wide credible intervals and high model standard deviation. There were no significant differences between active treatments. CBT-ED individual and self-help with support were significantly better than wait list in achieving remission. The inconsistency checks did not identify any significant inconsistency in the direct and indirect evidence included in the NMA. This strengthens the conclusions from the base case analysis. The bias adjustment sensitivity analysis suggested that bias due to small study effects may be exaggerating the treatment effects in the NMA. However, as the bias coefficient included 0 in all scenarios and there was no reduction in heterogeneity as a result of the bias adjustment, no strong conclusions about the presence of bias could be made. Although individual CBT-ED was the intervention with the highest probability of full remission, it was also associated with the highest healthcare costs due to high intervention costs. The guideline economic analysis demonstrated that self-help with support is the most cost effective first-line treatment option for people with bulimia nervosa. The probability of it being cost effective ranged from 0.80 to 0.89 at a willingness-to-pay of £20,000-£30,000 per quality-adjusted life year (QALY) (NICE, 2008b). Results were sensitive to the estimate of the probability of remission associated with self-help with support and utility value associated with the remission health state. According to the secondary analysis, where the impact of extending the time horizon of the analysis to 5 years was explored, self-help with support remained the cost-effective option. Also, in the scenario where the relapse rate associated with CBT-ED individual was assumed to be zero (that is, everyone sustains the treatment effect) and the annual relapse rate associated with self-help with support and wait list was assumed to be equivalent to the baseline rate self-help with support also remained the cost-effective option.
Quality of evidence	The quality of the evidence was mostly considered very low quality. The evidence was downgraded for indirectness, imprecision and risk of bias for reasons such as it was unclear how the randomisation sequence was generated and if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded and high dropouts were detected >20%. To account for high dropouts rates, intention-to-treat analysis was used for remission results, with the assumption that any dropouts did not recover from the eating disorder. For indirectness, a number of RCTs in both the guided self-help and self-help arms included a mixture of people with EDNOS and bulimia nervosa and in one study binge eating disorder. Nevertheless, the majority of people who contributed to the overall result for each outcome were those with bulimia nervosa (ranging from an

	estimated 66 to 100%). Heterogeneity was explored for laxative use, purging, EDE-shape concern and EDE-weight concern in the comparison guided self-help (ED) versus wait list controls. For each of these outcomes, only two studies were available so it was difficult to conduct a sensitivity analysis because removing one study meant heterogeneity could no longer be measured. Also, any differences in results may be due to random effects (or study differences), not necessarily due to a risk of bias or differences in severity of illness, duration of illness, or presence of comorbidities (subgroups to explore as described in protocol). For laxative use, one study by Walsh 2004 carried a high risk of bias because very high dropouts were reported: 88% and 63% in guided self-help and waitlist control arms respectively. Participants dropped out of the intervention arm because it was not intensive enough. However, as previously discussed a sensitivity analysis could not be conducted with only two studies and duration of illness, chronicity and severity of illness could not explain the results since they were all similar. For purging, EDE-shape concern and EDE-weight concern, duration of illness, chronicity were similar between the studies but severity of illness may explain the heterogeneity detected. One study by Ljotsson 2007 included people with a low number of binges per month (<18), and compared with Banasiask 2005 who studied people with a higher number of binges per month (>18), they showed a greater response to treatment. Suggesting that severity of illness may influence how well people respond to guided self-help.
Other considerations	The committee agreed that guided self-help was a good first line treatment because of the benefit it showed on remission and bingeing compared with no treatment (wait list controls), especially for those who may have to wait a long time before they receive CBT-ED treatment or if they have a mild case of bulimia nervosa. The committee also acknowledged that the economic evidence showed guided self-help is more cost-effective than CBT-ED. This was based on the analysis that showed the additional benefit of CBT-ED compared with guided self-help was too small to justify its additional cost according to NICE criteria of cost effectiveness. Also, CBT-ED was not found to be cost-effective versus wait list control over the model's time horizon (1 year and 4 months). The committee agreed that one of the key benefits of guided self-help is that it can be easily accessed and delivered by GPs or by other trained competent practitioners. This may help reduce waiting times for treating people with bulimia nervosa since early treatment is considered important for a good response. A subgroup analysis showed that guided self-help may be less effective in people with a more severe case of bulimia nervosa compared with those with a mild case. Although this could only be explored in a few (important) outcomes and only in a few studies, it does suggest that guided self-help may be less effective in those with a severe case of bulimia nervosa. This finding was supported by the experience of the committee members, but they also acknowledged the (limited) data does not allow us to explicitly say this in a recommendation. Nevertheless, they agreed that it was important for the practitioners to assess as early as possible (e.g. four weeks) whether the person is responding to treatment and if not, they should be offered CBT-ED (see following LETR). The long-term effects of guided self-help suggest remission is not different compared with wait list controls (one study n=78), although it does show similar rates of remission compared with any other intervention or active treatment (four studies, n=454). One study in the "other intervention" arm included CBT-ED and observing this comparison more closely showed there was no significant difference in remission rates at follow up between CBT-ED and guided self-help (29% versus 23% respectively). Thus, both guided self-help and CBT may have similar benefits long-term but this was only one small study (n=128) and more evidence is needed. Details of the guided self-help (ED) intervention in the recommendation was based on manuals used in the studies that fed into the meta-analysis. The majority of the studies referred to the same published manuals and are considered the best treatment programmes available to date.

	Please refer to the previous LETR for a discussion on the other treatments considered by the committee when making this recommendation.

1 **Second line treatment for BN**

	105. If bulimia-nervosa-focused guided self-help is ineffective after four weeks or is not acceptable, consider individual eating-disorder-focused cognitive behavioural therapy (CBT-ED). **106. Individual CBT-ED for adults with bulimia nervosa should:** • follow a CBT-ED manual • consist of up to 20 sessions over 20 weeks, with sessions held twice-weekly in the first phase • in the first phase focus on: o engagement and education o establishing a pattern of regular eating, and providing encouragement, advice and support while people do this • follow by addressing the eating disorder psychopathology (that is, the extreme dietary restraint, the concerns about body shape and weight, and the tendency to binge in response to difficult thoughts and feelings) • towards the end of treatment, spread appointments further apart and focus on maintaining positive changes and minimising the risk of relapse • if appropriate, involve significant others to help with one-to-one treatment. **107. Explain to people with bulimia nervosa that psychological treatments have a limited effect on body weight.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating bulimia nervosa in children, young people and adults. For this population, binge eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern regarding people with bulimia nervosa – of lesser importance, but still clearly important – include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience. The committee agreed not to consider the effect of psychological treatment on body weight in people with bulimia nervosa because it is not an outcome they are concerned about. There is also the risk that if weight loss is a target of treatment, it may have the negative side-effect of triggering binges. Removing weight loss as a target for treatment also ensures obesity is not a barrier to treating binge eating disorder. For all these reasons, body weight was excluded from the analysis. However, the committee agreed it was important to highlight to those receiving treatment that psychological treatments would have a limited effect on body weight. In addition, that weight loss is typically a post-therapy target.
Trade-off	Individual cognitive behavioural therapy with a focus on the eating disorder (CBT-

between clinical benefits and harms	ED) improved rates of remission and bingeing frequency in adults with bulimia nervosa compared with any other intervention at the end of treatment. Benefits were also found on EDE-dietary restraint, bulimic inventory test (Edinburgh) and depression. Some benefit were found on EDE-bulimia, symptom check-list but there was some uncertainty. A number of outcomes showed no difference between the two arms including vomiting, general psychopathology, global clinical Impressions score, quality of life, laxative use, EDE-total score, EDE-eating concern, EDI-drive for thinness, EDI-body dissatisfaction and EDE-attitude to shape. CBT-ED had long-term benefits after 4 months to 3 years on remission but no significant effect on bingeing. Other outcomes continued to favour CBT-ED at follow up including depression, global clinical impressions and general psychopathology but there was some uncertainty. Other outcomes showed no difference between CBT-ED individual and any other treatment. No data was reported at any time point on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience. CBT-ED was not only more effective on bingeing compared with any other intervention but also compared with waitlist controls. A number of other outcomes were also largely improved in the CBT-ED arm compared with the wait list controls, including purging, vomiting, depression and most of the EDE and EDI sub-scales. One study reported on whether compensatory behaviours were present or absent at the end of treatment and showed no difference between CBT-ED and wait list control on bingeing, vomiting, laxative misuse and purging. No data was available for follow up nor on all-cause mortality, adverse events, quality of life, resource use, relapse, general functioning, family functioning and service user experience. Refer to the LETR for young people with bulimia nervosa for a summary of the CBT-ED evidence in young people. **Other psychological treatments for adults considered** Other individual treatments considered by the committee for this review included interpersonal psychotherapy (IPT), integrative cognitive-affective therapy (ICAT), variations in CBT-ED, behavioural therapy (BT), hybrid therapy, dialectical behaviour therapy (DBT) and psychodynamic general therapy. Group therapies were also considered, they included group behavioural therapy, group CBT-ED, group behavioural therapy (ED), group psychoeducation, group CBT-ED of varying intensity, group emotional and mind training and group support. Self-help therapies (besides guided self-help as discussed in other LETR) were also considered, they included: general self-help; self-help (ED); internet self-help (ED); and text messaging. Of all these comparisons, the therapies that showed positive results on the critical outcomes of remission and/or bingeing included: CBT-ED versus another CBT-ED (for bingeing at the end of treatment); group CBT-ED versus wait list controls (for remission at the end of treatment and follow up); group BT versus wait list controls (for remission at the end of treatment); group BT-ED versus another group intervention (for remission at the end of treatment); group CBT-ED versus another group CBT-ED of varying intensity (on remission at the end of treatment) and internet self-help (ED) versus another intervention (for binging at the end of treatment but there was some uncertainty). Conversely, the therapies that showed negative results on remission or bingeing included: IPT versus another intervention (for remission and bingeing at the end of treatment) and psychodynamic general therapy versus another intervention (for bingeing at the end of treatment; and self-help (ED) versus another intervention (for bingeing at the end of treatment and at follow up). No difference in one or both of the critical outcomes was found in the following comparisons: IPT versus another comparison (for remission and bingeing at follow up); ICAT versus another intervention (for bingeing at the end of treatment and follow up); CBT (ED) versus another CBT (ED) (for remission at the end of treatment and bingeing and remission at follow up); BT versus another intervention (for bingeing and remission at end of treatment and follow up); hybrid versus another intervention (for bingeing at the end of treatment and follow up); group BT versus another intervention (for remission at the end of treatment and follow up);

	group CBT-ED versus wait list controls (for bingeing at the end of treatment); group CBT-ED versus another intervention (for bingeing and remission at the end of treatment and follow up); group BT versus wait list controls (for bingeing at the end of treatment and remission at follow up); group BT-ED versus another group intervention (for bingeing at the end of treatment and bingeing and remission at follow up); group psychoeducation versus another intervention (for bingeing and remission at the end of treatment); group CBT-ED versus another group CBT-ED of varying intensity (on bingeing at the end of treatment but there was some uncertainty); general self-help versus wait list controls (for remission at follow up); self-help (ED) versus another intervention (for remission at end of treatment and bingeing and remission at follow up); internet self-help (ED) versus another intervention (for remission at the end of treatment and at follow up); self-help (ED) versus wait list controls (for remission and bingeing at the end of treatment); and text messaging versus wait list controls (for remission at the end of treatment). DBT, group emotional and mind training, group support, general self-help (versus another intervention), did not measure any critical outcomes. At no time point, for any comparison, was data reported on the important outcomes: all-cause mortality, adverse events, resource use, relapse, family functioning or service user experience.
Trade-off between net health benefits and resource use	Existing economic evidence pertaining to the psychological therapies for adults with bulimia nervosa was very limited and the committee could not draw any conclusions from it. The guideline economic analysis demonstrated that self-help with support is the most cost-effective first-line treatment option for people with bulimia nervosa. However, the committee expressed the view that CBT-ED individual is better at sustaining response and may have more favourable cost effectiveness when considering the long-term follow up. There was no suitable data to populate the economic model looking at the cost effectiveness of treatments for people with bulimia nervosa during the long-term follow up. Given the lack of long-term cost-effectiveness data the committee agreed that the economic analysis does not provide a complete picture and CBT-ED individual together with self-help with support should be available treatment options for people with bulimia nervosa. The secondary analysis indicated that when extending the time horizon of the analysis to five years the ICER of CBT-ED individual versus self-help with support remains above upper NICE cost-effectiveness threshold of £30,000 per QALY. However, the ICER of CBT-ED individual when compared with wait list is well below lower NICE cost-effectiveness threshold of £20,000 per QALY by approximately 2.5 years follow up. The ICER of CBT-ED individual versus wait list represents the cost effectiveness of CBT-ED individual in people where self-help is not effective or is unacceptable. Also, at five years the probability of CBT-ED individual (when compared with wait list) being cost effective was 0.60 at lower NICE cost-effectiveness threshold of £20,000 per QALY and increased to 0.65 at upper NICE cost-effectiveness threshold. In the scenario where the relapse rate associated with CBT-ED individual was assumed to be zero (that is, everyone sustains the treatment effect) and the annual relapse rate associated with self-help with support and wait list was assumed to be equivalent to the baseline rate the ICER of CBT-ED individual versus wait list was reduced to £3,788. This supports the view that CBT-ED individual may have more favourable cost effectiveness in the long run and could potentially be considered as an option where self-help with support is ineffective or unacceptable.
Quality of evidence	The quality of the evidence ranged from very low to moderate. For the review on psychotherapies, the majority of the published evidence compared individual CBT-ED versus any other intervention. Only three studies were identified for individual CBT-ED compared with wait list controls. Outcomes were downgraded for uncertainty in the methods for randomisation and if allocation concealment was performed. It was unclear if blinding of participants, investigators or assessors was carried out and high dropouts were at times detected. Heterogeneity was also detected at times, as was imprecision where the 95% confidence interval crossed one or two minimal important differences or the

	outcome did not meet the optimum information size. One study that compared CBT-ED with an active arm, Mitchell 2011, included a mixture of bulimia nervosa (56%) and EDNOS (44%). For remission, this paper contributed 18% of the overall effect size at the end of the study and 15% at follow up. In such cases the outcomes were downgrade for indirectness. Heterogeneity was detected in a number of outcomes for the comparison of CBT-ED versus any other intervention. The exploration of this heterogeneity can be found in the appendices on forest plots. Heterogeneity was detected for remission at the end of treatment (I2=53%) and may be explained by a study by Mitchell 2008 that carried a high risk of bias. Mitchell 2008 included a mixed population of binge eating and EDNOS and a sensitivity analysis showed that when it is removed from analysis, heterogeneity is reduced to 47%. Furthermore, compared with the other studies, Mitchell 2008 showed a minimal effect of CBT-ED on remission and it was unclear how long their participants had an eating disorder for. Since risk of bias appeared to explain the heterogeneity, no subgroup analysis was performed. Bingeing also showed heterogeneity and a sensitivity analysis that excluded a paper by Poulsen 2014 that carried a high risk of bias reduced it to 0%. Poulsen 2014 compared five months of CBT-ED with psychodynamic therapy that continued for 19 months, thus making it difficult to compare end of treatment effects and to account for people recovering from the eating disorder naturally over time. If Poulsen is remove from the analysis, heterogeneity is reduced to 0%. Symptom check list also showed heterogeneity and a subgroup analysis separating the groups according to their duration of illness appeared to explain this. People with a shorter duration of illness (<5 years) had a better response to CBT-ED treatment versus any other treatment compared with those who had endured a longer duration of illness (>5 years). The forest plots showing the results of this analysis can be found in the Appendices. If more outcomes showed differences in the response to treatment based on severity of illness the committee would have made a recommendation to address this. Heterogeneity was detected in a number of “important” outcomes for other types of treatment. Exploring the reasons for this can be found in the appendices showing the forest plots.
Other considerations	The committee were confident in the results for individual CBT-ED in adults given the large number of participants and the strength of the evidence that showed CBT-ED was effective across a number of outcomes including bingeing and remission at the end of treatment and importantly on remission at follow up. For this reason, they felt it was important that if people are not responding early (approximately four weeks) to guided self-help then they should be offered CBT-ED. **Other interventions for adults considered** There were many different comparisons included in this review, so the committee focused on the outcome of remission at end of treatment and more importantly at follow up to help inform their decision. Besides individual CBT-ED, only a few treatments showed a benefit on remission at the end of treatment they included: group CBT-ED versus wait list controls; group BT versus wait list controls; group BT-ED versus another group intervention; group CBT-ED versus another group; CBT-ED of varying intensity versus another CBT-ED. None of these treatments showed a positive effect on remission at follow up compared with another treatment or wait list controls. Because of the uncertainty surrounding the results of these comparisons, due to either the small sample size, the small effect size, the economic evidence, or the lack of long-term benefits on remission the committee decided not to recommend any of these treatments. The committee considered the alternative of delivering individual CBT-ED in a group format. The evidence, however, was limited and based on very small numbers (critical outcomes n=52 to 126). Group CBT-ED showed no difference in bingeing or remission compared with any other intervention at the end of treatment or at follow up. Compared with wait list controls, benefits from group CBT-ED were found for remission at the end of treatment (14%) and at follow up (28%). However, the results were from only two studies with 59 participants. Thus, the committee did not have confidence in recommending group CBT-ED over individual CBT-ED.

Economic analysis

The economic analysis showed that self-help with support is more cost-effective when compared with CBT-ED individual using full remission as an outcome measure at end of treatment. The secondary analysis indicated that CBT-ED individual may have more favourable long-term cost effectiveness, in particular when compared with wait list.

The committee considered the existing clinical and economic evidence. It also considered the following when recommending it as a second-line treatment (that is, in people where self-help with support is ineffective or is unacceptable):

The committee agreed that remission at follow up is the most important outcome and ideally should have been modelled in the economic analysis. However, because of the lack of follow up data it was not possible to model long term cost effectiveness of interventions for people with BN (some data was removed from the analysis because it did not meet the committee's criteria of measuring symptoms for a minimum of two weeks). Nevertheless, individual CBT-ED compared with any other intervention (interpersonal psychotherapy, behavioural therapy and guided self-help) appeared to show the most convincing results on remission at follow up. For this reason the committee agreed it should be recommended as a second-line treatment.

Based on the committee's experience of treating people with bulimia nervosa, they agreed that people with greater chronicity and severity of illness are more likely to respond to CBT-ED than guided self-help. However, because of the cost effectiveness of treatment and the lack of evidence to support such a recommendation, the committee agreed that guided self-help should be offered first. If people don't respond to treatment, possibly because they have a more severe case of bulimia nervosa, then they should be offered CBT-ED as a second line treatment.

the impact of severity of illness (average number of binges per month) and duration of illness on responsiveness to CBT-ED was difficult to decipher because of the limited number of published studies, there was a lot of overlap in the populations severity of illness, variation in reporting binge frequency, data was not always reported and no clear definition in the field on what is mild, moderate or severe bulimia nervosa. Nevertheless, there was a subgroup analysis (see previous LETR) that showed people with a more severe case of bulimia nervosa may be less responsive to guided self-help. This provided further support for the committee to recommend CBT-ED to people who do not respond early to guided self-help. However, this was based on a few outcomes and only two studies.

a secondary analysis (economic) examining the cost effectiveness of CBT-ED individual over five years showed that CBT-ED individual is potentially more cost effective when compared with wait list. Thus, this provided economic justification for the committee to recommend CBT-ED individual as an appropriate second-line treatment.

The committee agreed that individual CBT-ED has the most convincing body of evidence when you compare it to other active treatments. Remission rates were 87% higher in the CBT-ED versus any other treatment at end of treatment (7 studies, n=731) and 32% higher at follow up (four studies n=95). CBT-ED showed a trend to improve remission compared wait list controls but this was based on only one study (n=71) and no data at follow up. However, the NMA of interventions for people with BN that used a common comparator found that CBT-ED and self-help with support had a very similar effectiveness at the end of treatment and the economic analysis (that was informed by the NMA) found that self-help is the most cost-effective option when compared with CBT-ED and wait list. As such it was difficult to justify CBT-ED as a first line treatment.

Guided self-help compared with another intervention showed no difference in remission rates at end of treatment and at follow up (both time points four studies, n=454). However, compared with wait list controls, the results showed guided self-help improved remission rates by 14%. The committee commented that it was based on only two studies (n= 198) and no data at follow up. Also, the NMA of treatments for people with BN showed that self-help with support and CBT-ED individual had comparable effectiveness at the end of treatment (using full

remission as an outcome measure). Again, this provided the evidence needed to recommend guided self-help as a first line treatment.
There was 1 study (n=128) that compared guided self-help with CBT-ED at follow up and showed no difference in remission rates. Thus, supporting guided self-help as a first line treatment given that it may be equally effective as CBT-ED on remission in the long-term.

Personal correspondence from an author said that cost of individual CBT-ED may be reduced in the future based on the preliminary evidence that 20 sessions of CBT-ED may not be needed for benefits to be seen. People may only need as few as 10 sessions. To further explore this in a large scale RCT the committee agreed to generate a research recommendation "are shorter psychological treatment lengths equally effective compared with the treatment lengths recommended in this guideline for children, young people and adults with an eating disorder?"

Because of the high cost of individual CBT-ED the committee agreed it was important that those delivering treatment are aware of non-responders and do not persist with treatment in such cases.

There is a body of observational evidence that shows individual CBT-ED translates effectively from well-designed RCTs to 'real-world' clinical settings. (This evidence was only considered as part of the discussion since it did not answer the review question).

The committee were very uncertain about the quality of the clinical evidence for the alternative treatments, including behavioural therapy, dialectical behaviour therapy, psychodynamic therapy, interpersonal psychotherapy.

Details of the CBT-ED intervention in the recommendation was based on the manuals used in the studies that fed into the meta-analysis. The majority of the studies referred to the same published manuals and are considered the best treatment programmes available to date.

Stepped care (see chapter 5)

There was very low quality evidence on stepped care, that is, what the options are if the first line treatment does not work whether it be an alternative treatment or an increase in the number of sessions offered in the first line treatment. A review on stepped care in adults with bulimia nervosa showed if a stepped care approach is used, there was no difference compared with continued treatment at the end of treatment on remission, binge frequency and EGE-global. There was some evidence from a pair-wise comparison that group psychoeducation stepped-up to CBT-ED may lead to better remission rates compared with group psychoeducation and wait list control but no other outcomes favoured the stepped care approach, they all showed no difference between the two arms (binge frequency and EDE-global). Given the lack of evidence, the committee decided to generate a research recommendation relevant for any eating disorder, including bulimia nervosa to: "evaluate the effectiveness of stepped care for psychological treatment of eating disorders for people of all-ages." See chapter 5.

Assessment after 4 weeks

The GC discussed how there is RCT evidence that shows an assessment during the early stages of a psychological treatment can predict the likelihood of a full-response at the end of treatment. For this reason they felt it was important to have an early assessment after 4 weeks and if the person is not showing signs of responding then they should be offered CBT-ED.

There is no justification for ending treatment early, after only 4 weeks, because the magnitude of the response at this stage is not the same as that seen at the end of treatment, i.e. 16 weeks. For this reason it is important that if the person is showing early signs of responding to treatment that they continue-on to the end of the programme. Furthermore, there is no evidence on people with an eating disorder that 4 weeks of treatment is sufficient to show long-term benefits. Also, the programme is delivered in a way that different elements are explored over the full 16 weeks and this can not be shortened. The GC agreed that it is important to have this recommendation because there is a concern that people with a severe case of bulimia nervosa may be better suited to CBT-ED rather than guided self-help. An early assessment should detect these patients and ensure that they receive a treatment that may be more cost-effective.

7. Research recommendation: are shorter psychological treatment lengths equally effective compared with the treatment lengths recommended in this guideline for children, young people and adults with an eating disorder?

Treatment for young people with bulimia nervosa

	108. Offer bulimia-nervosa-focused family therapy to young people with bulimia nervosa **109. Bulimia-nervosa-focused family therapy for young people with bulimia nervosa should:** • use a bulimia-nervosa-focused family therapy manual • consist of 18-20 sessions over six months • support and encourage the family to help the young person recover • not blame the young person or their family members or carers • include information about regulating body weight, dieting and the adverse effects of controlling weight with self-induced vomiting or laxatives • establish a good therapeutic relationship with the young person and their family members or carers • use a collaborative approach between the parents and the young person to establish regular eating patterns and minimize compensatory behaviours • include regular meetings with the young person on their own throughout the treatment • include self-monitoring of bulimic behaviours and discussions with family members or carers • in later phases of treatment, support the young person and their family members or carers to establish a level of independence appropriate for their level of development • in the final phase of treatment, focus on plans for when treatment ends (including any concerns the young person and their family have), and on relapse prevention. **110. If family therapy is ineffective, or is not acceptable, consider bulimia-nervosa-focused guided self-help for young people with bulimia nervosa.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating bulimia nervosa in children and young people. For this population, binge eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with bulimia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and

	service user experience.
Trade-off between clinical benefits and harms	**Family therapy** Three RCTs were identified on the effects of family-based therapy on young people with bulimia nervosa. Overall, the studies showed family based therapy has a positive effect on remission but has a similar effect on binge frequency compared with any other individual intervention. Other outcomes generally showed positive results for family based therapy compared with any other individual intervention. The following outcomes favoured family-based therapy: EDE-global, EDE-restraint, EDE-shape concern, EDE-weight concern and reduced hospitalisation rates. The Yale-Brown Cornell eating disorder scale and depression showed a trend to favour family based therapy. One outcome, mean objective bingeing episodes, favoured the other intervention. The following outcomes were not different between the two arms: purging episodes, vomiting episodes, service user experience. At follow up, remission rates in young people with bulimia nervosa still favoured family-based therapy compared with any other intervention. Most other outcomes were similar between the two treatment arms, including: binge episodes, purging episodes, vomiting episodes, EDE-restraint, Yale-Brown Cornell eating disorder scale and depression. Two outcomes, EDE-weight concern and service user experience, showed a trend to favour family based therapy although there was some uncertainty. Family-based therapy did however show a benefit on EDE-shape concern and EDE-global. No evidence was found on the important outcomes of general functioning, family functioning, adverse events, quality of life, all-cause mortality or relapse. **Guided self-help** One RCT was found in young people comparing guided self-help with an active intervention (family therapy). No differences in remission were found at the end of treatment or at 6 months follow up. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning, general psychopathology or service user experience. **CBT-ED** The evidence on individual CBT-ED in young people with bulimia nervosa was less convincing and only one or two small studies were available per outcome. Remission favoured the other active arm however bingeing showed no difference between the two arms. Other compensatory behaviours and depression showed no difference and less favourable results were found for eating disorder psychopathology measured using EDE in the CBT-ED arm. These same trends were identified at 12 months follow up. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning or service user experience.
Trade-off between net health benefits and resource use	The existing economic evidence in young people with bulimia nervosa is very limited. Existing evidence from UK-study is only partially applicable and is characterised by minor methodological limitations. The committee considered the existing economic evidence which indicated that family therapy may potentially be cost effective when compared with CBT based guided self-help. The committee also took into account the physical consequences of eating disorders and high costs associated with managing these; psychological and financial burden associated with eating disorders both for children and young people and for their families, as well as the clinical benefits associated with treatment and the benefits of involving parents or carers in the recovery process. The committee also considered the intervention costs and concluded that offering treatment, such as family treatment, would represent good value for money. The committee noted that there is a lack of evidence on the effectiveness and cost effectiveness of interventions for young people with BN. The committee also noted that because treatment options are very limited guided self-help should be an option where family therapy is ineffective or unacceptable. Also, the committee considered that guided self-help was found to be cost effective for adults with BN and given the lack of data for young people the committee extrapolated the cost effectiveness of interventions (to young people with BN) from the economic

	analysis conducted for this guideline for adults with BN.
Quality of evidence	The evidence for family-based therapy was all low or very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. Imprecision was often detected because the 95% confidence interval crossed one or two minimal important differences or the outcome did not meet the optimal information size of 300 events or 400 participants. In one of the family therapy studies, i.e. Grange 2007, partial and full remission results were combined since the definition of remission varied across the studies and the partial remission definition compared closely with the full remission definition in other studies. Thus, total remission results included those who no longer fit the definition of bulimia nervosa in DSM-IV (partial) in addition to those who had no binge eating behaviours in previous four weeks as determined by EDE (full remission).
Other considerations	Few studies on psychotherapies for young people with bulimia nervosa were identified and none on children. The most convincing evidence was on family therapy where three studies were found with 295 participants. The 68% greater increase in the number who achieved remission compared with any other treatment convinced the committee that family therapy should be recommended as a first line treatment. The committee also discussed the importance of including siblings and other family members in the treatment because of the effects bulimia nervosa may have on other family members. The other evidence identified was on CBT-ED and guided self-help, however, only one study on CBT-ED was found with 110 participants and one study for guided self-help with 85 participants. CBT-ED did not show favourable results on remission and based on the costings for CBT-ED in adults, the committee agreed to not recommend it as a first or second line treatment. Conversely, guided self-help showed it was equally effective as any other treatment (family therapy) on remission rates at the end of treatment and at follow up. Moreover, given the evidence to support its recommendation as a first line treatment in adults, the committee agreed it could be offered as a second line treatment for young people with bulimia nervosa if family therapy is unacceptable, contraindicated or ineffective. Given the lack of evidence on CBT-ED and guided self-help in young people with bulimia nervosa, the committee decided to develop a research recommendation to investigate 'the effectiveness of young people focused CBT-ED and guided self-help in children and young people with bulimia nervosa'.

8. Research recommendation: What is the effectiveness of young people focused CBT-ED and guided self-help in children and young people with bulimia nervosa?

7.3 Carer interventions

7.3.1 Review question: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 207. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix J.

This review considers all psychological interventions for the parents or carers of children or young people with bulimia nervosa. The interventions were categorised according to their mode of delivery (e.g. group, self-help), the age of the people with the eating disorder, and

the type of eating disorder and were compared to wait list controls, TAU or any other intervention.

Table 207: Clinical review protocol summary for the review of does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Component	Description
Review question(s)	Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?
Population	• Family or carers of people with an eating disorder (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder)
Intervention(s)	• Psychological interventions may include: • Family-based • Parent only (not necessarily focussed on eating disorder) • Parent-focused therapy (PFT) • Group Parent-Training (GPT) • Separated family therapy • Parents with person with ED (greater focus on eating disorder) Behavioural Family Therapy (BFT) Behavioural family systems therapy (BFST). Family Based Treatment (FBT) Family Day Workshops (FDW) Family Therapy (FT) Family therapy for anorexia nervosa (FT-AN) Multi-Family Group Day Treatment (MFGDT) Multi-Family Group Therapy (MFGT) Systemic Family Therapy (SFT) Systemic Family Therapy for AN (SFT-AN) Multifamily therapy (MFT) is synonymous with (MFGT; MFGDT). Uniting couples in the treatment of AN (UCAN Conjoint family therapy
Comparison	• Wait list control • Treatment as usual • Another intervention
Critical outcomes	• Parent's or carer's general psychopathology (including mood/depression/anxiety) • Family functioning • Quality of life

Component	Description
	• Other primary outcomes commonly reported in studies that just target the family/carer • The following outcomes will be included if the family or carer intervention includes the child or person with an eating disorder: • Remission and long-term recovery (if symptoms were measured over a minimum two week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• General functioning • Resource use. • Service user experience • All-cause mortality. • Adverse events • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion)
Study design	• Systematic Reviews • RCTs

7.3.2 Clinical Evidence for: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

No studies in parents or cares of children or young people with bulimia nervosa were found that met the eligibility criteria for this review. Further information about excluded studies can found in Appendix J.

7.3.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with bulimia nervosa was identified by the systematic search of the economic literature undertaken for this guideline.

7.3.4 Clinical evidence statements

No studies in parents or carers of children or young people with bulimia nervosa were found that met the eligibility criteria for this review

7.3.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with bulimia nervosa was available.

7.3.6 Recommendations and link to evidence for the review on Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Working with family members and carers

	111. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing whether any interventions help the parents and carers of children and young people with an eating disorder. The critical outcomes for the parents and carers were: general psychopathology, family functioning, quality of life, and other primary outcomes reported by the study. Other outcomes that are critical for the child or young person with the eating disorder include remission and bingeing or body weight, depending on the eating disorder. Other outcomes that are of lesser importance but clearly important outcomes include, general functioning, service user experience, all-cause mortality, adverse events and eating disorder psychopathology.
Trade-off between clinical benefits and harms	**Bulimia nervosa or binge eating disorder (chapter 7 and 8)** No relevant published evidence was found on parents or carers of children and young people with bulimia nervosa or binge eating disorder. Anorexia nervosa (evidence in chapter 6) One randomised controlled trial (RCT), aimed at carers of young people with anorexia nervosa and compared the effectiveness of guided self-help or self-help (and treatment as usual) with treatment as usual alone. After 12 months there was no difference in carer general psychopathology. No evidence was found on the critical outcomes of carer general psychopathology, carer family functioning, carer quality of life, nor on the important outcomes of eating psychopathology, carer general functioning, service user experience, resource use, adverse events or all-cause mortality. Another study compared self-help (with treatment as usual) with treatment as usual and showed no difference in the carer's general psychopathology or carer skills after 6 to 12 months but a trend for poorer results for family functioning. However, there was some uncertainty. In the young people with anorexia nervosa whom they care for, there was no difference in BMI, weight, severity index (SEED), general psychopathology, clinical improvement, peer related problems between the two treatment arms. However, there was a trend for poorer outcomes in prosocial behaviour in the self-help group but there was some uncertainty. No evidence was found on the critical outcomes of remission, carer general psychopathology, nor on the important outcomes of service user experience, resource use, adverse events or all-cause mortality. Comparing guided self-help (and treatment as usual) with treatment as usual showed at 12 months a trend for positive outcomes in the combined treatment group on carer burden and quality of life, but no difference in family functioning, carer skills or carer psychopathology. There was a trend for poorer outcomes in carer accommodation and enabling. At 24 months, there was a trend for a positive result on carer burden, quality of life, carer accommodation and enabling and carer psychopathology. In addition, a trend for poorer outcomes in family functioning and time spent caring. No evidence was found on the critical outcome of carer general psychopathology, nor on the important outcomes of service user experience and resource use. In the same intervention, the guided self-help for the carers did not translate too many benefits in the young people with anorexia nervosa whom they care for. At 12 months, no differences were found in any of the outcomes including mortality, admission to hospital, patient relapse, BMI, EDE-global, severity index (SEED), general psychopathology, clinical improvement. However, there was improvement in peer problems but a trend for a negative result in prosocial behaviour. At 24 months, there was a trend for positive increase in BMI and EDE-global, no difference in general psychopathology and a trend for a negative result in quality of life. No evidence was found on the critical outcome of remission, nor on the important outcomes of adverse events and all-cause mortality.

	Comparing two active treatments generally showed no difference in effectiveness in the carer-related outcomes. Guided self-help compared with self-help were equally effective on all outcomes 6 to 12 months after the young people with anorexia nervosa had been discharged from inpatient care, except there was a trend for carer accommodation to favour guided self-help. No evidence was found on the critical outcomes of carer quality of life, nor on the important outcomes of carer general functioning, service user experience or resource use. In the young people with anorexia nervosa, there was a trend for poorer results on BMI and peer problems in the guided self-help group compared with self-help. No difference was found in clinical severity (SEED), general psychopathology, clinical improvement, prosocial behaviour but there was a trend for better results in peer problems. No evidence was found on the critical outcome of remission. Web-based guided self-help also failed to show convincing benefits for the carers of young people with anorexia nervosa compared with treatment as usual. At the end of treatment, a poorer outcome in distress was found but there was some uncertainty. The other outcomes, such as carer accommodation and enabling, family functioning, carer burden and caregiving experience showed no difference. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. Web-based guided self-help compared with web-based self-help showed no difference in the outcomes for carers at the end of treatment. At follow up, favourable results were found on family functioning in the guided web-based self-help group, but no difference in carer experience, quality of life and general psychopathology. There was a trend for poorer results in carer burden, but there was some uncertainty. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. Any eating disorder (as reviewed in chapter 9) Randomised control trials investigating interventions for the carers of young people with any eating disorder failed to show many favourable outcomes. Psychoeducation compared with waitlist control showed a positive effect on carer self-efficacy and a trend to improve carer knowledge of eating disorders at the end of treatment. Long-term follow up (unclear duration) showed favourable results in both but carer burden (only measured at follow up) was not different compared with wait list controls. No evidence was found on the critical outcomes of carer general psychopathology, family functioning, and quality of life, nor on the other important outcomes. Comparing guided self-help with self-help showed no difference in any of the carer-related outcomes at the end of treatment. No evidence was found on the other important outcomes.
Trade-off between net health benefits and resource use	The committee expressed the view that offering family members and carers an assessment of their own needs may incur additional healthcare resources (that is, time required to perform such assessment). However, the committee considered the cost of providing such assessment to be small, taking into account the potential reduction in family and carers' burden, potential depression and other health vulnerabilities which may be costly to other parts of the healthcare system, especially considering that the burden on family and carers can last for many years and increase their morbidity and stress. Consequently, the committee judged that assessment that aims to improve family and carers' experience are likely to represent good value for money.
Quality of evidence	The quality of the evidence was mostly very low. The outcomes were downgraded because it was unclear how they randomised, if allocation concealment was performed or if participants and investigators were blinded. In some, not all, assessors were blinded. High dropout rates were also detected in some groups >20%. Imprecision was detected in most outcomes due to the 95% confidence interval crossing one or two minimal important differences or because it did not meet the optimal information size. Outcomes were not always measured at the end of treatment or at follow up. It is not known if any improvements in the carer's general psychopathology also translated to benefits in the children with the eating disorder.

Other considerations	Given the very low quality of the data with very few positive findings favouring one arm over the other, the committee came to the consensus that there was not enough evidence to support a recommendation on any specific treatment for parents or carers of people with an eating disorder. Nevertheless, the committee acknowledged the stress and burden that a person with an eating disorder, in particular anorexia nervosa, can have on family members over a long period of time. Therefore, they agreed that offering family members and carers an assessment of their own needs, including: personal, social and emotional support available to them, the need for support in the caring role for example if the child should need urgent care and there are other children to take care of, and to offer advice on where they can get some practical support. The extent to which the family need to be involved in treatment depends on the age and developmental needs of the person with the eating disorder, the severity of the illness, the risk from harm and the person receiving treatment's wishes. In general, parents and other family members will want to be involved in the treatment. If a parent or carer cannot attend a meeting the healthcare professional should provide written information on the outcome of an assessment or treatment where appropriate. The committee acknowledged the importance of consent and confidentiality and their discussion can be found in the LETRs relating to this. They also discussed that although the evidence found was for carers and parents of people with anorexia nervosa or any eating disorder, the recommendation is relevant for parents and carers of people with bulimia nervosa and binge eating disorder. This is mostly because no specific intervention was recommended, rather to offer an assessment of their needs and to help them find the necessary support. In absence of good evidence, the committee agreed to generate a research recommendation to address the question "What is the effectiveness of a carer-focused psychological intervention in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?" See chapter 6.2.

7.4 Nutritional interventions

7.4.1 Review question: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 208. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all nutritional interventions that may be delivered to children, young people and adults with an eating disorder with or without a pharmacological intervention. The interventions were categorised according to type of nutritional intervention, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to wait list controls, placebo, TAU or any other intervention.

Table 208: Clinical review protocol summary for the review of Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

Component	Description
Review question(s)	Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical

Component	Description
	eating disorder). • Strata: • Children (≤12), young people (13-≤17 years), adults ≥18 years. • Eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder).
Intervention(s)	• Nutritional intervention • Method of feeding • Nutritional in combination with any pharmacological intervention • Examples of nutritional interventions are nutritional counselling (with or without educational and supportive groups) and supplements (e.g. zinc)
Comparison	• Placebo • Wait list control • Treatment as usual • Another intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs

7.4.2 Clinical Evidence for: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

7.4.2.1 Nutritional interventions versus any other intervention or wait list control

Six RCTs (n=295) met the eligibility criteria for this review for people with bulimia nervosa, all of which were for adult females (Beumont 1997 (Beumont et al., 1997), Burton 2006 (Burton

and Stice, 2006), Hsu 2001 (Hsu et al., 2001), Laessle 1991 (Laessle et al., 1991), Sundgot-Borgen 2002 (Sundgot-Borgen et al., 2002), Ventura 1999 (Ventura and Bauer, 1999)). An overview of the trials included in the meta-analysis can be found in Table 209. Further information about both included and excluded studies can found in Appendix J.

Summary of findings for those on bulimia nervosa can be found in Table 210, Table 211, Table 212, Table 213, Table 214 and Table 215. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 209: Study information for trials included in the meta-analysis of nutritional interventions versus any other intervention or wait list control for adult female bulimia nervosa.

Study ID	N Random-ized	Sample	Type of nutritional intervention	Intervention Age at onset and/or duration of illness (years)	Comparison(s) Age at onset and/or duration of illness (years)	Duration
Beumont 1997	67	Adult BN	Nutritional counselling	Nutritional counselling + Fluoxetine	Nutritional Counselling + Placebo	8 weeks + 3-mo FU
Burton 2006	85	Adult Full- and Sub-threshold BN	Healthy Weight Program	Healthy Weight Program	WLC	8 weeks + 3-mo FU
Hsu 2001	100	Adult BN	Nutritional counselling	Nutritional therapy + Cognitive therapy Duration of BN: 5.9 (3.7)	1. Nutritional therapy Duration of BN: 5 (4.4) 2. Cognitive therapy Duration of BN: 5.5 (3.2)	14 weeks
Laessle 1991	55	Adult BN	Nutritional counselling	Nutritional Management Duration of bulimic symptoms for sample=7.5 (3.8)	Stress Management Duration of bulimic symptoms for sample=7.5 years (3.8)	3 months + 12-mo FU
Sundgot-Borgen 2002	64	Adult BN	Nutritional counselling	Nutritional counselling Duration of BN=5 (2.3)	1. Exercise Duration of BN=7 (3.7) 2. CBT Duration of BN=5 (1.6) 3. WLC Duration of BN=6 (3.8)	16 weeks + 18-mo FU
Ventura 1999	40	Adult BN-P	Nutritional counselling	Psychobiological Nutritional Rehabilitation + CBT Duration of BN: 8.6 (4.9)	Traditional Nutritional Rehabilitation + CBT Duration of BN 6.5 (4.6)	24 weeks

Abbreviations: BN, Bulimia Nervosa; BN-P, Bulimia Nervosa – Purging subtype; CBT, Cognitive Behavioural Therapy; FU, follow up; WLC, wait list control.

Table 210: Summary table of findings for nutritional counselling versus any other intervention at the end of treatment in adult females with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Nutritional Counselling (95% CI)
Meal Frequency meals/week	100 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean meal frequency in the intervention groups was 0.34 standard deviations higher (0.11 lower to 0.78 higher)
Calories/day (kcal)	48 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean calories/day (kcal) in the intervention groups was 0.21 standard deviations higher (0.36 lower to 0.78 higher)
EDI Bulimia	48 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia in the intervention groups was 0.21 standard deviations lower (0.78 lower to 0.36 higher)
EDI Body Dissatisfaction	79 (2 studies) 12 months	⊕⊕⊖⊖ LOW2,3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction in the intervention groups was 0.54 standard deviations higher (0.09 to 0.99 higher)
EDI Drive for Thinness	48 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness in the intervention groups was 0.19 standard deviations higher (0.38 lower to 0.76 higher)
Depression - raw scores BDI	48 (1 study) 12 months	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - raw scores in the intervention groups was 0.22 standard deviations lower (0.79 lower to 0.35 higher)
Depression - Change scores HDRS	100 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - change scores in the intervention groups was 0.4 standard deviations lower (0.85 lower to 0.04 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Nutritional Counselling (95% CI)
CI: Confidence interval;					
1 Hsu 2001: Allocation concealment unclear. No participant nor investigator blinding. Dropout rate of Nutritional therapy group=46%; dropout rate of Cognitive therapy group 39%. Difference between Nutritional and Cognitive Therapy group, Nutritional Therapy group and Cognitive Therapy group>20%. 2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD). 3 Laessle 1991: No details provided regarding randomization method nor allocation concealment. Participant, investigator and assessor blinding unclear. 4 Sundgot-Borgen 2002: Unclear randomization and allocation concealment. No participant blinding, unclear investigator blinding. Physical exercise group dropout rate=20%.					

Table 211: Summary table of findings for nutritional counselling versus any other intervention at follow up in adult females with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Nutritional Counselling (95% CI)
Recovered from Bulimia Nervosa FU	43 (1 study) 18 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision	RR 0.1 (0.02 to 0.71)	500 per 1000	450 fewer per 1000 (from 145 fewer to 490 fewer)
Satisfying EDNOS criteria FU	43 (1 study) 18 months	⨁⨀⨀⨀ VERY LOW1,3 due to risk of bias, imprecision	RR 0.53 (0.2 to 1.36)	423 per 1000	199 fewer per 1000 (from 338 fewer to 152 more)
Calories/day (kcal) FU	42 (1 study) 12 months	⨁⨀⨀⨀ VERY LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean calories/day (kcal) fu in the intervention groups was 0.1 standard deviations higher (0.51 lower to 0.71 higher)
EDI Bulimia FU	73 (2 studies) 18 months	⨁⨀⨀⨀ VERY LOW1,3,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean edi bulimia fu in the intervention groups was 1.28 standard deviations higher (2.15 lower to 4.72 higher)
EDI Body Dissatisfaction FU	73 (2 studies)	⨁⨁⨀⨀ LOW1,4,6		Not calculabl	The mean edi body dissatisfaction fu in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Nutritional Counselling (95% CI)
	18 months	due to risk of bias, imprecision		e for SMD values	0.25 standard deviations higher (0.22 lower to 0.71 higher)
EDI Drive for Thinness FU	42 (1 study) 12 months	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness fu in the intervention groups was 0.16 standard deviations lower (0.77 lower to 0.46 higher)
Depression FU Beck Depression Inventory. Scale from: 0 to 63.	42 (1 study) 12 months	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.35 standard deviations lower (0.96 lower to 0.27 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio;					
1 Sundgot-Borgen 2002: Unclear randomization and allocation concealment. No participant blinding, unclear investigator blinding. Physical exercise group dropout rate=20%. 2 <300 events. 3 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD). 4 Laessle 1991: No details provided regarding randomization method nor allocation concealment. Participant, investigator and assessor blinding unclear. 5 I2>80%. 6 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).					

Table 212: Summary table of findings for nutritional counselling versus wait list control at follow up in adult females with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Nutritional Counselling (95% CI)
Does not satisfy EDNOS	32	⊕⊕⊖⊖	RR 0.77	1000 per	230 fewer per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Nutritional Counselling (95% CI)
criteria FU	(1 study) 18 months	LOW1,2 due to risk of bias, imprecision	(0.58 to 1.03)	1000	(from 420 fewer to 30 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Sundgot-Borgen 2002: Unclear randomization and allocation concealment. No participant blinding, unclear investigator blinding. Physical exercise group dropout rate=20%.
2 CI crosses either 0.75 or 1.25 (Risk Ratio).

1 **Table 213: Summary table of findings for nutritional therapy versus any other intervention in adult females with bulimia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Nutritional Therapy (95% CI)
Meal Frequency meals/week	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean meal frequency in the intervention groups was 0.021 standard deviations higher (0.47 lower to 0.52 higher)
Depression – change scores HDRS	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.17 standard deviations lower (0.66 lower to 0.33 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Hsu 2001: Allocation concealment unclear. No participant nor investigator blinding. Dropout rate of Nutritional therapy group=46%; dropout rate of Cognitive therapy group 39%.
2 CI crosses either 0.5 or -0.5 (SMD).

Table 214: Summary table of findings for healthy weight program versus wait list control at end of treatment in adult females with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Healthy Weight Program (95% CI)
Remission	85 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 6.84 (0.88 to 53.2)	24 per 1000	139 more per 1000 (from 3 fewer to 1000 more)
Binge Frequency binge episodes/month	85 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.95 standard deviations lower (1.4 to 0.5 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Burton 2006: No details of randomization method nor allocation concealment provided. No participant blinding, unclear investigator blinding. Dropout rate of 3 of 4 groups>25%. Reasons for dropout not stated.
2 Sample is participants with Full- and Sub-Threshold Bulimia Nervosa. Participants classified as Full Threshold BN if they have (i) >=8 binge eating episodes or compensatory behaviour episodes in month prior to study and (ii) overvalue weight and shape. Participants classified as Sub Threshold BN if they are not classified as Full Threshold (minimum of 4 binge eating and 4 compensatory episodes in past month).
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
4 <300 events (dichotomous outcome) or <400 participants (continuous outcome).

Table 215: Summary table of findings for healthy weight program versus wait list control follow up for adult females with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Healthy Weight Program (95% CI)
Remission FU	85 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 3.66 (1.32 to 10.13)	95 per 1000	253 more per 1000 (from 30 more to 870 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Healthy Weight Program (95% CI)
Binge Frequency FU binge episodes/month	85 (1 study) 3 weeks	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 0.86 standard deviations lower (1.3 to 0.41 lower)
General functioning FU Social Adjustment Scale (adapted)	85 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean social adjustment scale (adapted) fu in the intervention groups was 0.31 standard deviations lower (0.74 lower to 0.12 higher)
Resource use FU Health Survey Utilization Scale	85 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean health survey utilization scale fu in the intervention groups was 0.16 standard deviations lower (0.58 lower to 0.27 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Burton 2006: No details of randomization method nor allocation concealment provided. No participant blinding, unclear investigator blinding. Dropout rate of 3 of 4 groups>25%. Reasons for dropout not stated.
2 Sample is participants with Full- and Sub-Threshold Bulimia Nervosa. Participants classified as Full Threshold BN if they have (i) >=8 binge eating episodes or compensatory behaviour episodes in month prior to study and (ii) overvalue weight and shape. Participants classified as Sub Threshold BN if they are not classified as Full Threshold (minimum of 4 binge eating and 4 compensatory episodes in past month).
3 <300 events (dichotomous outcome) or <400 participants (continuous outcome).
4 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

1

2

7.4.3 Economic Evidence

No economic evidence on the cost effectiveness of nutritional interventions for people with bulimia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

7.4.4 Clinical evidence statements

7.4.4.1 Nutritional counselling versus any other intervention at end of treatment in adults with bulimia nervosa

Low quality evidence from one RCT (n=79) showed nutritional counselling may be less effective on improving scores on EDI-body dissatisfaction compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=48) showed no difference in the effect of nutritional counselling on increasing daily calorie intake, improving scores on EDI-bulimia, EDI-drive for thinness and reducing depression compared with any other intervention.

Low quality evidence from one RCT (n=100) showed no difference in the effect of nutritional counselling on improving weekly meal frequency compared with any other intervention.

Low quality evidence form one RCT (n=100) showed nutritional counselling may be more effective on change in depression compared with any other intervention, although there was some uncertainty.

7.4.4.2 Nutritional counselling versus any other intervention at follow up in adults with bulimia nervosa

Low quality evidence from one RCT (n=43) showed that nutritional counselling was less effective on the number of people who recovered from bulimia nervosa compared to any other intervention.

Very low quality evidence from one RCT (n=43) showed no difference in the effect of nutritional counselling on the number of people who satisfied EDNOS criteria compared with any other intervention.

Low to very low quality evidence from 2 RCTs (n=73) showed no difference in the effect of nutritional counselling on improving scores on EDI-body dissatisfaction and EDI-bulimia compared with any other intervention.

Very low quality evidence from one RCT (n=42) showed no difference in the effect of nutritional counselling on daily calorie intake, improving scores on EDI-drive for thinness and reducing depression compared with any other intervention.

7.4.4.3 Nutritional counselling versus wait list control at follow up in adults with bulimia nervosa

Low quality evidence from one RCT (n=32) showed nutritional counselling may be more effective on the number of people not satisfying EDNOS criteria compared with any other intervention, although there was some uncertainty.

7.4.4.4 Nutritional therapy versus any other intervention in adults with bulimia nervosa

Low quality evidence from one RCT (n=73) showed no difference in the effect of nutritional therapy on meal frequency and change in depression compared with any other intervention.

7.4.4.5 Healthy Weight Program versus wait list control at end of treatment in adults with bulimia nervosa

Very low quality evidence from one RCT (n=85) showed Healthy Weight Program is more effective on binge frequency compared with wait list control.

Very low quality evidence form one RCT (n=85) showed Healthy Weight Program may be more effective on in remission compared with wait list control, although there was some uncertainty.

7.4.4.6 Healthy Weight Program versus wait list control at follow up in adults with bulimia nervosa

Very low quality evidence from one RCT (n=85) showed Healthy Weight Program is more effective on remission and reducing binge frequency than wait list control.

Very low quality evidence from one RCT (n=85) showed no difference in the effect of Healthy Weight Program on general functioning and resource use than wait list control.

7.4.5 Economic Evidence statements

No economic evidence on the cost effectiveness of nutritional interventions for people with bulimia nervosa was available.

7.4.6 Recommendations and link to evidence for the review: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

Nutritional counselling for people with bulimia nervosa

	The committee expressed the view that nutritional counselling is an integral part of most eating disorder specific psychological interventions so they did not make a recommendation about this for people with bulimia nervosa.
Critical and important outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of nutritional interventions for treating bulimia nervosa in children, young people and adults. For this population, bingeing and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with bulimia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, eating disorder psychopathology, general functioning, family functioning and service user experience.
Trade off benefits and harms	Bulimia nervosa (chapter 7) Compared with any other intervention, nutritional counselling showed no difference on meal frequency, calories consumed per day, EDI-bulimia, EDI-drive for thinness, depression, and a trend to have a negative effect on EDI-body dissatisfaction but a more positive effect on depression. At 12 to 18 months follow up, fewer people recovered from bulimia nervosa in the nutritional counselling group compared with any other intervention. No differences were found in the number who satisfied the EDNOS criteria, EDI scores, depression or calorie intake. No evidence was found on the important outcomes of adverse events, quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, or service user experience. Compared with wait list controls, nutritional counselling showed some long-term

	benefits at 12 months follow up on the number who satisfied the EDNOS criteria but there was some uncertainty. No evidence was found on remission, bingeing, adverse events, quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, or service user experience. Nutritional therapy had a similar effect on meal frequency and change in depression scores at the end of treatment compared with any other intervention. No other outcomes were reported. A healthy weight programme showed a benefit on bingeing and a trend to improve remission compared with wait list controls at the end of the treatment. At 3 months follow up, benefits were found on bingeing and remission, but no difference on general functioning or resource use. No evidence was found on the important outcomes of adverse events, quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, family functioning, cost effectiveness or service user experience.
Trade-off between net health benefits and resource use	The committee expressed the view that dietary advice is an integral part of most eating disorder specific psychological interventions and providing such supplementary advice would not incur significant extra resource implications to the healthcare system.
Quality of the evidence	The quality of the evidence very low quality. It was unclear how randomisation was conducted and if allocation concealment was performed. It was unclear if either the participants or investigators were blind. In one nutritional study the assessors were blinded. High dropouts were also reported >20%. Imprecision was also detected because the 95% confidence interval crossed one or two minimal important differences or the numbers did not meet the optimum information size (300 events or 400 participants). Heterogeneity was not detected.
Other considerations	The committee agreed that the evidence was not strong enough to recommend nutritional counselling or a healthy weight programme as the sole treatment for adults with bulimia nervosa. They highlighted that dietary advice and counselling are an integral part of CBT-ED, SSCM, MANTRA and family therapy, so it is not generally needed if the person is receiving therapy. Usually this is delivered by the therapist and sometimes in collaboration with a dietician. No other considerations were made by the committee.

7.5 Physical interventions

7.5.1 Review Question: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 216. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all physical interventions that may be delivered to children, young people and adults with an eating disorder. The interventions were categorised according to type of physical intervention, the age of the participants and the type of eating disorder and were compared to wait list controls, placebo, TAU or any other intervention.

Table 216: Clinical review protocol summary for the review of Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Component	Description

Component	Description
Review question(s)	Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). • Strata: • Children (≤12), young people (13-≤17 years), adults ≥18 years • Eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Physical interventions may include: transcranial magnetic stimulation deep brain stimulation physiotherapy yoga physical exercise acupuncture mandometer massage
Comparison	• Placebo • Wait list control • Treatment as usual • Another intervention
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)

Component	Description
Study design	• Systematic Reviews • RCTs

7.5.2 Clinical Evidence for: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Three RCTs (n=213) met the eligibility criteria for this review, which were all for adults (Bulik 1998, Sundgot-Borgen 2002 (Sundgot-Borgen et al., 2002), Van den Eynde 2010 (Van den Eynde et al., 2010)). An overview of the trials included in the analysis can be found in Table 217. Further information about both included and excluded studies can found in Appendix J.

No studies were identified that compared a combined physical intervention and a pharmacological agent with any other intervention or wait list controls.

Table 217: Study information for trials included in the analysis of physical interventions versus any other intervention or wait list control for people with bulimia nervosa.

Study ID	N Random-ized	Female (%)	Mean BMI, kg/m2 (SD)	Intervention Age at onset and/or duration of illness (years)	Comparison(s) Age at onset and/or duration of illness (years)	Duration
Bulik 1998	111	100	22.4 (2.5)	Relaxation training	Exposure with response prevention to pre-binge cues Exposure with response prevention to pre-purge cues	6 weeks
Sundgot-Borgen 2002	64	100	21 (2.21)	Physical Exercise Duration of illness=7 (3.7)	1. Nutritional Counselling Duration of illness=5 (2.3) 2. CBT-ED Duration of illness=5 (1.6) 3. WLC Duration of illness=6 (3.8)	16 weeks
van den Eynde 2010a	38	86	25.4 (9.9)	Real rTMS Duration of illness (years): 0-5=6, 5-10=5, 10-15=3, >15=3	Sham rTMS Duration of illness (years): 0-5=12, 5-10=2, 10-15=4, >15=2	20 trains of 5 sec with 55 sec intertrain intervals.

Note: a All participants received 8 sessions of CBT-ED over 8 weeks before randomisation to groups. Intervention group includes 10 participants diagnosed with bulimia nervosa and 7 participants diagnosed with an eating disorder not otherwise specified-bulimic type; comparison group includes 10 participants with bulimia nervosa and 10 participants diagnosed with an eating disorder not otherwise specified
Abbreviations: CBT-ED, Cognitive Behavioural Therapy for eating disorders; rTMS, Repetitive Transcranial Magnetic Stimulation; WLC, wait list control.

Table 218: Summary of findings table for real repetitive transcranial magnetic stimulation (rTMS) versus 'sham' rTMS (placebo) at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with (Sham) rTMS	Risk difference with (Real) rTMS (95% CI)
Food Craving Questionnaire-State (raw scores)	37 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		Not calculable for SMD values	The mean food craving questionnaire-state (raw scores) in the intervention groups was 0.33 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with (Sham) rTMS	Risk difference with (Real) rTMS (95% CI)
		indirectness, imprecision			(0.98 lower to 0.32 higher)
Food Craving Questionnaire-State (change scores)	37 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean food craving questionnaire-state (change scores) in the intervention groups was 0.41 standard deviations lower (1.06 lower to 0.25 higher)
Not Withdrawn due to Adverse Events	38 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision	RR 0.94 (0.81 to 1.09)	1000 per 1000	60 fewer per 1000 (from 190 fewer to 90 more)
Urge To Eat (Visual Analogue Scale)	37 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean urge to eat (visual analogue scale) in the intervention groups was 0.44 standard deviations lower (1.09 lower to 0.22 higher)
Mood (Visual Analogue Scale)	37 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean mood (visual analogue scale) in the intervention groups was 0.38 standard deviations higher (0.27 lower to 1.03 higher)
Tension (Visual Analogue Scale)	37 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean tension (visual analogue scale) in the intervention groups was 0.04 standard deviations higher (0.6 lower to 0.69 higher)
Hunger (Visual Analogue Scale)	37 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean hunger (visual analogue scale) in the intervention groups was 0.58 standard deviations lower (1.25 lower to 0.08 higher)
Urge To Binge Eat (Visual Analogue Scale)	37 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean urge to binge eat (visual analogue scale) in the intervention groups was 0.03 standard deviations lower (0.68 lower to 0.61 higher)
# patients NOT binged in 24 hours after treatment	34 (1 study) 1 days	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,	RR 1.27 (0.98 to 1.66)	778 per 1000	210 more per 1000 (from 16 fewer to 513 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with (Sham) rTMS	Risk difference with (Real) rTMS (95% CI)
		indirectness, imprecision			

1 van den Eynde 2010: unclear randomization method and allocation concealment. No investigator blinding. Blinding only partially successful with 15/18 participants in real rTMS group correctly guessed treatment group; 11/20 participants in sham rTMS incorrectly guessed treatment group.
2 Sample consists of 20 BN participants and 17 EDNOS participants. EDNOS subgroup includes participants diagnosed with Binge Eating Disorder.
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
4 <300 events.
5 CI crosses both 0.5 and -0.5 (SMD).

1 **Table 219: Summary table of findings for aerobic exercise versus any other intervention in adults with bulimia nervosa at follow up**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Exercise (95% CI)
Recovery from Bulimia Nervosa FU	43 (1 study) 18 months	⨁⨁⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 5.04 (0.3 to 83.76)	161 per 1000	652 more per 1000 (from 113 fewer to 1000 more)
Satisfied EDNOS criteria FU	43 (1 study) 18 months	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 0.57 (0.11 to 3.06)	194 per 1000	83 fewer per 1000 (from 172 fewer to 399 more)
EDI Drive for Thinness FU	26 (1 study) 18 months	⨁⨁⨀⨀ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness fu in the intervention groups was 1.36 standard deviations higher (0.47 to 2.25 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Sundgot-Borgen 2002: Unclear randomization and allocation concealment. No participant blinding, unclear investigator blinding. Physical exercise group dropout rate=20%.
2 CI crosses both 0.75 and 1.25 (Risk Ratio).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with Exercise (95% CI)
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).					

1 **Table 220: Summary table of findings for aerobic exercise versus wait list control at follow up in adults with bulimia nervosa**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Exercise (95% CI)
Not recovered from Bulimia Nervosa FU	27 (1 study) 18 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.36 (0.17 to 0.76)	1000 per 1000	640 fewer per 1000 (from 240 fewer to 830 fewer)
Does not satisfy EDNOS criteria FU	27 (1 study) 18 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.91 (0.74 to 1.13)	1000 per 1000	90 fewer per 1000 (from 260 fewer to 130 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Sundgot-Borgen 2002: Unclear randomization and allocation concealment. No participant blinding, unclear investigator blinding. Physical exercise group dropout rate=20%. 2 CI crosses either 0.75 or 1.25 (Risk Ratio).					

2 **Table 221: Summary of findings for relaxation training versus any other intervention in adults with bulimia nervosa at end of**
3 **treatment**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with other intervention for adult BN	Risk difference with Relaxation training (95% CI)
Binge frequency	111 (1 study) 12 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.09 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with other intervention for adult BN	Risk difference with Relaxation training (95% CI)
		indirectness, imprecision			(0.3 lower to 0.48 higher)
Vomiting frequency	111 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean vomiting frequency in the intervention groups was 0.33 standard deviations higher (0.07 lower to 0.72 higher)
Laxative use frequency	111 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean laxative use frequency in the intervention groups was 0.37 standard deviations higher (0.03 lower to 0.76 higher)
Purge frequency	111 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean purge frequency in the intervention groups was 0.42 standard deviations higher (0.03 to 0.82 higher)
No binge or purge episodes/2 weeks	111 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision	RR 0.85 (0.57 to 1.27)	542 per 1000	81 fewer per 1000 (from 233 fewer to 146 more)
EDI Drive for Thinness	111 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi drive for thinness in the intervention groups was 0.09 standard deviations higher (0.3 lower to 0.48 higher)
EDI Bulimia	111 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi bulimia in the intervention groups was 0.55 standard deviations higher (0.15 to 0.94 higher)
EDI Body dissatisfaction	111 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction in the intervention groups was 0.1 standard deviations higher (0.29 lower to 0.49 higher)
Depression	111 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,4, 6 due to risk of bias, inconsistency, indirectness,		Not calculable for SMD values	The mean depression in the intervention groups was 0.61 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with other intervention for adult BN	Risk difference with Relaxation training (95% CI)
		imprecision			(0.21 to 1.01 higher)
Global Functioning GAFS	111 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean global functioning in the intervention groups was 0.3 standard deviations lower (0.69 lower to 0.09 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Bulik 1998: unclear randomisation method and allocation concealment. Unclear participant and investigator blinding. Seventeen participants discontinued treatment during prior CBT-ED, whilst 2 were withdrawn by investigators. Five participants discontinued treatment prior to randomization.
2 All participants received 8 sessions of CBT-ED over 8 week period prior to randomisation to intervention groups.
3 <400 participants.
4 CI crosses either 0.5 or -0.5 (SMD).
5 CI crosses both 0.75 and 1.25 (Risk Ratio).
6 I2>50%.

1 **Table 222: Summary of findings for relaxation training versus any other intervention in adults with bulimia nervosa at follow up**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with other intervention for adult BN 12-mo FU	Risk difference with Relaxation training (95% CI)
Binge frequency	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.08 standard deviations lower (0.47 lower to 0.31 higher)
Vomiting frequency	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean vomiting frequency in the intervention groups was 0.16 standard deviations higher (0.23 lower to 0.56 higher)
Laxative use frequency	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4		Not calculable for SMD values	The mean laxative use frequency in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with other intervention for adult BN 12-mo FU	Risk difference with Relaxation training (95% CI)
		due to risk of bias, indirectness, imprecision			0.4 standard deviations higher (0.01 to 0.79 higher)
Purge frequency	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean purge frequency in the intervention groups was 0.27 standard deviations higher (0.13 lower to 0.66 higher)
No binge or purge episodes/2 weeks	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4,5 due to risk of bias, inconsistency, indirectness, imprecision	RR 0.78 (0.52 to 1.19)	556 per 1000	122 fewer per 1000 (from 267 fewer to 106 more)
EDI Drive for Thinness	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi drive for thinness in the intervention groups was 0.05 standard deviations higher (0.34 lower to 0.44 higher)
EDI Bulimia	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi bulimia in the intervention groups was 0.05 standard deviations higher (0.34 lower to 0.44 higher)
EDI Body dissatisfaction	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction in the intervention groups was 0.17 standard deviations higher (0.22 lower to 0.56 higher)
Depression	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.47 standard deviations higher (0.08 to 0.87 higher)
Global Functioning GAFS	111 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean global functioning in the intervention groups was 0.44 standard deviations lower (0.84 to 0.05 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with other intervention for adult BN 12-mo FU	Risk difference with Relaxation training (95% CI)
CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Bulik 1998: unclear randomisation method and allocation concealment. Unclear participant and investigator blinding. Seventeen participants discontinued treatment during prior CBT-ED, whilst 2 were withdrawn by investigators. Five participants discontinued treatment prior to randomization. 2 All participants received 8 sessions of CBT-ED over 8 week period prior to randomisation to intervention groups. 3 <400 participants. 4 CI crosses either 0.75 or 1.25 (SMD), or either 0.5 or -0.5 (SMD). 5 I2>50%.					

1

2

7.5.3 Economic Evidence

No economic evidence on the cost effectiveness of physical interventions for people with bulimia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

7.5.4 Clinical evidence statements

7.5.4.1 Repetitive transcranial magnetic stimulation (rTMS) versus placebo in adults with bulimia nervosa

Very low quality evidence from one RCT (n=37) showed rTMS may be more effective on hunger and bingeing within 24 hours of treatment compared with placebo (sham rTMS), although there was some uncertainty.

Very low quality evidence from one RCT (n=37) showed no difference in the effect of rTMS on food craving, the urge to eat, urge to binge eat, mood, and tension compared with placebo (sham rTMS).

Very low quality evidence from one RCT (n=37) showed rTMS may be less effective on the number of people who withdrew due to adverse events compared with placebo (sham rTMS), although there was some uncertainty.

7.5.4.2 Aerobic exercise versus any other intervention at follow up in adults with bulimia nervosa

Very low quality evidence from one RCT (n=43) showed no difference in the effect of aerobic exercise on the number of people who recovered from bulimia nervosa nor on the number of people who satisfied the EDNOS criteria compared with any other intervention.

Low quality evidence from one RCT (n=26) showed exercise is less effective on scores for EDI-drive for thinness compared with any other intervention.

7.5.4.3 Aerobic exercise versus wait list control at follow up in adults with bulimia nervosa

Low quality evidence from one RCT (n=27) showed no difference in the effect of exercise on the number of people not recovered from bulimia nervosa nor the number of people not satisfying EDNOS criteria compared with wait list control.

7.5.4.4 Relaxation training versus any other intervention in adults with bulimia nervosa at end of treatment

Very low quality evidence from one RCT (n=111) showed no difference in the effect of relaxation training on binge frequency, the number of people not bingeing nor purging for 2 weeks, EDI-drive for thinness and EDI-body dissatisfaction compared with any other intervention.

Very low quality evidence from one RCT (n=111) showed relaxation training may be less effective on vomiting frequency, laxative use, and global functioning compared with any other intervention, although there was some uncertainty.

Very low quality evidence from one RCT (n=111) showed relaxation training is less effective on purge frequency, EDI-bulimia and depression compared with any other intervention.

7.5.4.5 Relaxation training versus any other intervention in adults with bulimia nervosa at follow up

Very low quality evidence from one RCT (n=111) showed no difference in the effect of relaxation training on binge frequency, vomiting frequency, purge frequency, the number of people not bingeing nor purging for two weeks, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction compared with any other intervention.

Very low quality evidence from one RCT (n=111) showed relaxation training is less effective on laxative use, depression and global functioning compared with any other intervention.

7.5.5 Economic Evidence statements

No economic evidence on the cost effectiveness of physical interventions for people with bulimia nervosa was available.

7.5.6 Recommendations and link to evidence for the review on: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Physical therapy

	112. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitisation, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes for the review on the effectiveness of physical interventions, such as transcranial magnetic stimulation or physiotherapy in people with eating disorders and it was agreed that for any eating disorder remission is of greatest concern. The other critical outcomes for anorexia nervosa are body weight and BMI and for binge eating disorder and bulimia nervosa it is bingeing. Other outcomes that are important but are considered rare events or rarely measured in randomised controlled trials for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse, thus they were extracted where possible, but did not factor strongly in the decision making. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	**Adults with bulimia nervosa (chapter 7)** Repetitive transcranial magnetic stimulation (rTMS) versus sham rTMS ('placebo') showed no difference in the effect on bingeing and food cravings within 24 hours of treatment, nor on the urge to eat. There was a trend for hunger and the number of people who binged to be reduced and the number of people who withdrew due to adverse events to be increased. However, there was some uncertainty. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Aerobic exercise appeared to be less effective on EDI-drive for thinness. No difference was found on the number of people who recovered from bulimia nervosa nor who satisfied the EDNOS criteria. Compared with wait list control, aerobic exercise was less effective on the number who had recovered (unclear definition) from bulimia nervosa but showed no difference on the number who satisfied the criteria for EDNOS. No evidence was

	112. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitisation, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
	found on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. **Young people with anorexia nervosa (chapter 6)** For young people with anorexia nervosa, bright light treatment and CBT showed benefits on depression compared with any other intervention. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Video feedback and nutritional counselling compared with nutritional counselling alone showed no additional benefit of the video feedback on BMI. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, or service user experience. Resistance training and treatment as usual showed no difference on BMI and quality of life in young people with anorexia nervosa compared with treatment as usual. At follow up, resistance training and treatment as usual appeared to be less effective on BMI compared with treatment as usual. No evidence was found on the critical outcomes of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, family functioning, resource use or service user experience. **Adults with anorexia nervosa (chapter 6)** Repetitive transcranial magnetic stimulation versus sham showed no difference in anorexia nervosa symptoms (urge to restrict, feeling full and feeling fat), urge to binge or side-effects from treatment. However, at follow up some benefits were detected on anorexia nervosa symptoms including feeling full and feeling flat compared with sham, but no difference in the symptom of urge to restrict. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Warming therapy on top of refeeding had no effect on change in BMI compared with refeeding alone in adults with anorexia nervosa. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. Acupuncture and treatment as usual compared with acupressure, massage and treatment as usual showed acupuncture is more effective on EDE-shape concerns but no other outcome was different between the two groups including EDI-subscales, EDE-subscales, depression, general psychopathology and BMI. No evidence was found on the critical outcome of remission, nor on the important outcomes of all-cause mortality, relapse, general functioning, family functioning, resource use or service user experience. **Adults with binge eating disorder (chapter 8)** Yoga appears to be effective at reducing scores on the binge eating scale compared with wait list controls. However, this did not translate to a benefit in BMI. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience.

	112. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitisation, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
	Aerobic exercise and group CBT-ED appeared to be more effective at reducing BMI compared with group CBT-ED alone in adults with binge eating disorder. No difference was found in depression scores. Similar results were found at follow up. When a maintenance component (12 biweekly meetings over six months) was added to both arms to make this part of the intervention more comparable with the aerobic exercise group (because they continued to meet up), there was a trend for a reduced BMI and depression in the aerobic exercise, group CBT-ED and maintenance group compared with the group CBT-ED and maintenance group at the end of treatment and for the trend in the benefit on BMI to be maintained at follow up but not depression. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use or service user experience. **Any eating disorder (chapter 9)** One study compared eye movement desensitization and reprocessing therapy with treatment as usual in adults with any eating disorder. The results showed some improvement in the outcomes reported by the body image memory questionnaire, including the earliest memory and worst memory on body image and only a trend for the most recent memory. At follow up the worst memory on body image was still better but not the earliest or most recent. No evidence was found on the critical outcomes of remission, bingeing and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. An RCT was identified that compared yoga and treatment as usual with treatment as usual in adults with any eating disorder. At the end the treatment, no difference was found in any of the outcomes including BMI, EDE-total or any of the EDE- sub-scales. Similar findings were found at follow up (three weeks), however there was some improvement in EDE-restraint in the yoga and treatment as usual group compared with treatment as usual. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, resource use or service user experience. A graded body image therapy (and maintenance treatment as usual) was compared with a maintenance treatment as usual in adults with any eating disorder. No difference was found in EDE-weight concerns or EDE-shape concerns at the end of treatment or at follow up. No difference was found in EDE-weight concerns or EDE-shape concerns at the end of treatment or at follow up. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. An acceptance-based body image mirror exposure therapy was compared with a control therapy and showed an improvement in EDE-eating concerns, EDE-weight concerns, EDE-shape concerns, but not in EDE-restraint. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use or service user experience. A psychomotor therapy and support was compared with support in females with any eating disorder and showed no difference at the end of treatment on self-expression and control anger scales. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, eating disorder psychopathology, relapse, general functioning, family functioning, resource use or service user experience.

	112. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitisation, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
	The committee requested investigating the benefits of the Mandometer on eating disorders. A Mandometer is a device that measures how much weight is lost from a dinner plate after the person with eating disorder has finished eating. This weight is stored on a computer along with how satiated the person is after eating. The evidence on this is scarce and the sample sizes were too small (less than 10 per group) to meet our inclusion criteria as described in the protocol.
Trade-off between net health benefits and resource use	There was no evidence for the effectiveness of physical interventions in people with eating disorders. As a result, such interventions are likely to be not cost effective.
Quality of evidence	The evidence for physical interventions was mostly very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as unclear methods of randomisation or if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. Most of the outcomes were the result of a single study with a very low number of participants, only binge eating disorder had more than 100 participants in total. Imprecision was detected in most outcomes because the 95% confidence interval crossed one or two minimal important differences or it did not meet the optimal information size. Also, few studies measured remission and/or compensatory behaviours relevant to that eating disorder. Some outcomes were excluded from the study because it was either unclear over what duration they measured the symptoms or it was less than the two week minimum required by the committee.
Other considerations	The committee agreed that the evidence presented was not strong enough or of sufficient quality to offer a physical intervention to people with an eating disorder. This was mostly because very few studies were identified and few participants were included in most outcomes. However, the committee decided to make a research recommendation on adding exercise to a recommended psychotherapy to determine whether it may add any benefit to those with bulimia nervosa or binge eating disorder. The committee discussed the importance of exploring what the right amount of exercise is, what is the best type of exercise and what the potential harms are. The committee suggested making a research recommendation on the effects of exercise on bulimia nervosa and binge eating disorder, as opposed to any of the other physical interventions for a number of reasons. Exercise may be useful adjunct to psychotherapy to address any co-existing weight or obesity-related issues and mood disorders, such as depression and anxiety. Exercise may also be a cost-effective and drug-free alternative to other therapeutic approaches such as transcranial magnetic stimulation or anti-depressants. See chapter 6.6

7.6 Pharmacological interventions

7.6.1 Review question: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

In total 28 RCTs were identified that met the criteria for this review. Some of these studies included the follow up data. Of these studies, 17 (n=1556) compared a single pharmacological intervention with another active arm or placebo (Carruba 2001 (Carruba et al., 2001), Faris 2000 (Faris et al., 2000), Fluoxetine Bulimia Nervosa Collaborative Study

Group 1992 (FBNCSG, 1992), Hedges 2003 (Hedges et al., 2003), Hoopes 2003 (Hoopes et al., 2003), Kanerva 1995 (Kanerva et al., 1995), Leombruni 2006 (Leombruni et al., 2006), Milano 2004 (Milano et al., 2004), McCann 1990 (McCann and Agras, 1990), Mitchell 1990 (Mitchell et al., 1990b), Nickel 2005 (Nickel et al., 2005), Pope 1989 (Pope et al., 1989), Romano 2002 (Romano et al., 2002), Schmidt 2004 (Schmidt et al., 2004), Walsh 1991 (Walsh et al., 1991), Walsh 1987 (Walsh et al., 1987), Walsh 2004 (Walsh et al., 2004)). All of the studies were conducted in adults.

Of all the relevant studies identified, 10 RCTs (n=732) compared a combined pharmacological therapy with another treatment (mostly psychological) and compared it with another active arm or placebo (Beumont 1997 (Beumont et al., 1997), Agras 1992 (Agras et al., 1992), Agras 1994 (Agras et al., 1994a), Goldbloom 1997 (Goldbloom et al., 1997), Jacobi 2002 (Jacobi et al., 2002), Mitchell 1990 (Mitchell et al., 1990b), Keel 2002 (Keel et al., 2002), Mitchell 2001 (Mitchell et al., 2001), (Walsh et al., 2004), (Walsh et al., 1997)).

Table 223: Clinical review protocol summary for the review: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

Topic	Interventions to treat eating disorders in children, young people and adults
Review question	Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder Strata: Children (≤12), young people (13-≤17 years), adults ≥18 years Eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder and atypical eating disorder)
Intervention	Pharmacological intervention Pharmacological + psychological: Pharmacological interventions may include: Anti-depressants i.e. SSRIs, Fluoxetine – Prozac Anxiolytic (antianxiety) Antipsychotic Anti-emetic medication. i.e. Ondansetron Anticonvulsant topiramate/antiepileptic (Topomax) Appetite suppressant (i.e. lisdexamf(ph)etamine dimesylate)
Control	Placebo Wait list control Treatment as usual Another intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes for decision making	Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) Binge eating for BN and BED. Body weight / BMI for AN. Adverse events
Study design	Systematic Reviews RCTs
Minimum sample size	N=10 per arm
Note	Consider the prescription of medications that may be misused or inappropriately prescribed by those with ED

Table 224: Study information for trials included in the meta-analysis of pharmacological versus any other intervention, placebo or wait list controls for people with bulimia nervosa.

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N	Intervention Category	Comparison Arm Category	Treatment Length
Bulimia Nervosa Single Pharmacotherapy								
Carruba 2001	25.7 (0.8)	20.4 (0.4)	100%	Not reported	77	Moclobemide - 400-600 mg/day. MAOI	Placebo	42 days
Faris 2000	29.1 (6)	21.6 (2.5)	100%	Duration of BN 11·8 years	26	Ondansetron -4 mg/day. Antiemetics	Placebo	4 weeks
FBNCSG 1992	27.4 (7.2)	22.7 (4.2)	100%	At least 6 months	387	Fluoxetine hydrochloride -20 mg/d. SSRI Fluoxetine hydrochloride 60 mg/d. SSRI	Placebo	8 weeks
Hoopes 2003/ Hedges 2003	29.0 (9.7)	61.3 (10.3) kg	99%	At least 2 binge eating episodes per week for a min of 1 year	69	Topiramate 25 to 400 mg/day. Anticonvulsant	Placebo	50 weeks
Kanerva 1995	25.2 (16-22)	62.2 (15.4) (kg)	100%	Duration of BN 5.7 (0.5-20) years	50	Fluoxetine 60 mg/day. SSRI	Placebo	8 weeks
Leombruni 2006	28.7 (8.2)	20.7 (5.0)	100%	Duration of ED 7 years	37	Citalopram 20 mg to 40 mg. SSRI	Fluoextine SSRI	3 months
Milano 2004	24 to 36 years	NR	100%	Not reported	20	Sertraline 100 mg/day. SSRI	Placebo	12 weeks
McCann 1990	20 (14-36)	31.7 (4.7) kg	100%	Not reported	30	Desipramine 50 mg to 300 mg. TCA	Placebo	84 days

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N	Intervention Category	Comparison Arm Category	Treatment Length
Mitchell 1990	24.1 (4.4)	106.4 (12.8)% ideal body weight	100%	Duration of BN: 6.5 (2.9) years	171	Imipramine Hydrochloride 50 to 300 mg/day. TCA TCA and Group therapy Group therapy + Placebo	Placebo	12 weeks
Nickel 2005	21.5 (3.1)	64.9 (5.8) kg	100%	Duration of illness at least 6 months	60	Topiramate 25 to 200 mg/day. anticonvulsant	Placebo	10 weeks
Pope 1989	26 (18-55)	Not reported	100%	Not reported	46	Trazodone -355 mg (max 400 mg). Other	Placebo	26 days
Romano 2002* (prevent relapse)	29.5 (7.0)	22.5 (3.9)	93%	Responders to 8 weeks of treatment were randomised to treatment or placebo	150	Fluoxetine 60 mg/day. SSRI	Placebo	12 months
Schmidt 2004	18-50 years	85% and 115% of ideal body weight	100%	Not reported	200	Fluvoxamine 50 to 300 mg/day. SSRI	Placebo	52 weeks.
Walsh 1991	24.8 (4.5)	136.2 (16.1)	100%	6.7 (3.6) years	80	Desipramine.200 to 300 mg/day. (TCA).	Placebo	42 days
Walsh 1987	26. 9 (4.3)	NR	100%	Duration of BN: 9.0 (4.4) years	62	Phenelzine 30-90 mg/day. (MAOI)	Placebo	56 days
Walsh 2004	30.6 (7.8)	BMI > 17.5 kg/m2	100%	Duration of BN: 12.0 (7.9) years	91	Fluoxetine (60 mg/day). SSRI Self-help + SSRI. Combination Guided Self-help	Placebo	4 months

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N	Intervention Category	Comparison Arm Category	Treatment Length
Bulimia Nervosa Combined Therapy								
Beumont 1997	24.2 (4.5)	22.0 (2.0)	100%	Not reported	67	Nutrition + Fluoxetine (SSRI, 60 mg/day)	Nutritional counselling	8 weeks + 12 weeks FU
Agras 1992 (Agras 1994 FU)	29.6 (8.9)	59.9 (9.1) ideal weight: 53.7 (5.8) kg	100%	Duration of BN: 10 years	71	Desipramine hydrochloride (TCA, 25-350mg) Combination Medication and CBT	CBT	16 or 24 weeks. 8, 12 weeks FU
Goldbloom 1997	25.8 (5.5)	23.0 (2.5)	100%	Minimum 6-month duration of illness;	76	Fluoxetine (SSRI, 60 mg/day) Combination Medication and CBT	CBT-ED	12 weeks (4 weeks FU)
Jacobi 2002	26.0 (5.8)	20.6 (2.0	100%	Binge eating had begun 7.9 (5.2) years and vomiting 7.4 (4.8) years before the start of the study	53	Fluoxetine (SSRI, 20 to 60 mg/day) Combination Medication and CBT	CBT-ED	4 months (6 mo and 1 year FU)
Mitchell 1990 (Keel 2002 FU)	24.1 (4.4)	% Ideal BW 106.5 (12.8)	100%	Duration of BN: 6.5 (2.9) years	171	Impramine hydrochloride (TCA, 50 to 300 mg/day) Group-CBT-ED (placebo) Combined TCA + Group therapy	Placebo	10 weeks (10 year FU)
Mitchell 2001	26.6 (7.1)	59.5 (13.9)	100%	Duration of	91	Fluoxetine (SSRI,	Placebo	16 weeks

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N	Intervention Category	Comparison Arm Category	Treatment Length
		kg		illness at least 6 months		60 mg/day) Fluoxetine + Self-help manual Placebo + Self-help		
Walsh 2004	24.3 (5.5)	21.9 (3.4)	100%	Duration of BN: 12.0 (7.9) years	91	Fluoxetine (SSRI, 60mg/day) Guided self-help Guided self-help+fluorextine (SSRI)	Placebo	4 months
Walsh 1997	26.1 (4.9)	21.9 (2.2)	100%	Duration of BN: 8.0 (4.0)	112	CBT + Placebo CBT+ Desipramine TCA Supportive Psychotherapy + Medication	Supportive Psychotherapy + Placebo	112 days

Abbreviations: CBT, cognitive behavioural therapy; FU, follow up; TCA, tricylic antidepressant; SSRI, selective serotonin reuptake inhibitor.

7.6.2 Clinical Evidence for: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

Table 225: Summary of findings table for an antidepressant versus placebo at end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Antidepressant (95% CI)
Binge frequency, Adults - SSRIs	42 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency, adults - ssris in the intervention groups was 0.13 standard deviations lower (0.73 lower to 0.48 higher)
Purge frequency. Adults - TCAs	78 (1 study)	⨁⊖⊖⊖ VERY LOW2,3,4		Not calculable	The mean purge frequency. adults - tcas in the intervention groups was

		due to risk of bias, imprecision, publication bias		for SMD values	0.34 standard deviations lower (0.79 lower to 0.11 higher)
Vomiting frequency. Adults - SSRI	42 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting frequency. adults - ssri in the intervention groups was 0.20 standard deviations lower (0.8 lower to 0.41 higher)
EDI Adults	123 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi adults in the intervention groups was 1.19 standard deviations higher (0.74 to 1.64 higher)
EDI Adults - SSRI	46 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi adults - ssri in the intervention groups was 0.29 standard deviations lower (0.87 lower to 0.29 higher)
EDI Adults - MAOI	77 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi adults - maoi in the intervention groups was 3.34 standard deviations higher (2.64 to 4.04 higher)
EDI - Drive for thinness. Adults - SSRI	46 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - drive for thinness. adults - ssri in the intervention groups was 0.44 standard deviations lower (1.02 lower to 0.15 higher)
EDI- Body dissatisfaction. Adults - SSRI	46 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi- body dissatisfaction. adults - ssri in the intervention groups was 0.48 standard deviations lower (1.07 lower to 0.1 higher)
EDI- Bulimia. Adults - SSRI	46 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi- bulimia. adults - ssri in the intervention groups was 0.15 standard deviations lower (0.73 lower to 0.43 higher)
Depression TCA	101 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression tca in the intervention groups was 0.35 standard deviations lower (0.74 lower to 0.04 higher)
Depression scores. Adults - SSRIs	88 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,9 due to risk of bias, imprecision,		Not calculable for SMD	The mean depression scores. adults - ssris in the intervention groups was 0.39 standard deviations lower

		publication bias		values	(0.81 to 0.03 lower)
Depression scores. Adults - MAOIs	127 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression scores. adults - maois in the intervention groups was 0.06 standard deviations lower (0.4 lower to 0.29 higher)
Depression change score - SSRI	146 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression change score - ssri in the intervention groups was 0.19 standard deviations lower (0.52 lower to 0.13 higher)
Global clinical score. Adults	312 (4 studies)	⊕⊖⊖⊖ VERY LOW3,5,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean global clinical score. adults in the intervention groups was 0.33 standard deviations lower (0.55 to 0.1 lower)
Global clinical score. Adults - TCA	78 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean global clinical score. adults - tca in the intervention groups was 0.33 standard deviations lower (0.77 lower to 0.12 higher)
Global clinical score. Adults - SSRI	234 (3 studies)	⊕⊖⊖⊖ VERY LOW3,5,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean global clinical score. adults - ssri in the intervention groups was 0.32 standard deviations lower (0.58 to 0.07 lower)
Did not have adverse event. Adults	960 (11 studies)	⊕⊖⊖⊖ VERY LOW2,3,12 due to risk of bias, imprecision, publication bias	RR 0.95 (0.92 to 0.99)	49 per 1000	2 fewer per 1000 (from 0 fewer to 4 fewer)
Did not have adverse event. Adults - TCAs	165 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,9 due to risk of bias, imprecision, publication bias	RR 0.94 (0.87 to 1.01)	14 per 1000	1 fewer per 1000 (from 2 fewer to 0 more)
Did not have adverse event. Adults- SSRIs	610 (5 studies)	⊕⊖⊖⊖ VERY LOW2,3,9,13 due to risk of bias, imprecision, publication bias	RR 0.97 (0.93 to 1.01)	49 per 1000	1 fewer per 1000 (from 3 fewer to 0 more)
Did not have adverse event. Adults - MAOIs	139 (2 studies)	⊕⊖⊖⊖ VERY LOW3,6,14,15 due to risk of bias, inconsistency, imprecision, publication bias	RR 0.87 (0.75 to 1)	86 per 1000	11 fewer per 1000 (from 21 fewer to 0 more)

Drop out due to adverse events. Adults - Other	46 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,15 due to risk of bias, imprecision, publication bias	RR 1 (0.88 to 1.13)	43 per 1000	0 fewer per 1000 (from 5 fewer to 6 more)
Did not achieve remission Adults Other_ITT	46 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,15 due to risk of bias, imprecision, publication bias	RR 0.91 (0.79 to 1.06)	0 per 1000	-
Binge frequency Adults TCA FU	38 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency adults tca fu in the intervention groups was 0.39 standard deviations lower (1.04 lower to 0.25 higher)
Laxative use Adults TCA FU	38 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,16 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean laxative use adults tca fu in the intervention groups was 0.08 standard deviations higher (0.56 lower to 0.72 higher)
Vomit frequency Adults TCA FU	38 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomit frequency adults tca fu in the intervention groups was 0.46 standard deviations lower (1.1 lower to 0.19 higher)
Depression Adults TCA FU	38 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression adults tca fu in the intervention groups was 0.27 standard deviations higher (0.37 lower to 0.91 higher)
EDI - Body dissatisfaction Adults TCA FU	38 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - body dissatisfaction adults tca fu in the intervention groups was 0.24 standard deviations lower (0.88 lower to 0.4 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how the random sequence was generated or if allocation concealment was conducted. It was unclear if either the participants, investigators or assessors were blind. High dropouts were reported in one arm >20%.
2 For continuous outcome, there were fewer than 400 participants.
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.

4 It was unclear how patients were randomised and if allocation concealment was performed. Studies were double-blind but unclear if assessors were blind.
5 95% Crossed 1 MID (-0.5).
6 It was unclear how randomised sequence was generated and if allocation concealment was conducted. Study was double-blind but it was unclear if assessors were blind. High dropouts were reported.
7 It was unclear in one study how randomised sequence was generated and if allocation concealment was conducted in both studies. Studies were double-blind but it was unclear if investigators were blind. High dropouts were reported in Romano.
8 It was unclear how randomised sequence was generated and if allocation concealment was conducted in both studies. Studies were double-blind but it was unclear if assessors were blind.
9 It was unclear in one study how random sequence was generated and in all studies if allocation concealment was performed. In was unclear if assessors were blind. High dropouts were reported >20%.
10 It was unclear if allocation concealment was performed. It was a double-blind study but it was unclear if assessors were blind. High dropouts were reported >20%.
11 It was unclear in all but one study how the randomised sequence was generated and if allocation concealment was conducted. It was unclear in one study if investigator was blind and in all studies if assessors were blind. High dropout rates were reported >20%.
12 In most studies it was unclear how patients were randomised and if allocation concealment was performed. Most studies were double-blind but unclear if assessors were blind. High dropouts were reported >20%.
13 It was unclear how the random sequence was generated or if allocation concealment was conducted. It was unclear if either the participants, investigators or assessors were blind.
14 Heterogeneity was detected I2 >80%.
15 For a dichotomous outcome, there were fewer than 300 events.
16 95% CI crossed 2 MIDs (-0.5 to 0.5).

Table 226: Summary of findings table for an antidepressant versus another antidepressant at end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another Antidepressant	Risk difference with Antidepressant (95% CI)
Depression - SSRI (Citalopram) vs. SSRI (Fluoxetine). Adults	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression - ssri (citalopram) vs. ssri (fluoxetine). adults in the intervention groups was 0.22 standard deviations lower (0.97 lower to 0.52 higher)
EDI - Drive for thinness - SSRI (Citalopram) vs. SSRI (Fluoxetine). Adults	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision,		Not calculable for SMD values	The mean edi - drive for thinness - ssri (citalopram) vs. ssri (fluoxetine). adults in the intervention groups was 0.34 standard deviations higher

		publication bias			(0.4 lower to 1.09 higher)
EDI- Body dissatisfaction - SSRI (Citalopram) vs. SSRI (Fluoxetine). Adults	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi- body dissatisfaction - ssri (citalopram) vs. ssri (fluoxetine). adults in the intervention groups was 0 standard deviations higher (0.74 lower to 0.74 higher)
EDI - Bulimia - SSRI (Citalopram) vs. SSRI (Fluoxetine). Adults	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - bulimia - ssri (citalopram) vs. ssri (fluoxetine). adults in the intervention groups was 0.04 standard deviations lower (0.78 lower to 0.7 higher)
Exercise - SSRI (Citalopram) vs. SSRI (Fluoxetine). Adults	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean exercise - ssri (citalopram) vs. ssri (fluoxetine). adults in the intervention groups was 1.23 standard deviations higher (0.41 to 2.05 higher)
Clinical Global Impression - Adverse effect - SSRI (Citalopram) vs. SSRI (Fluoxetine). Adults	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean clinical global impression - adverse effect - ssri (citalopram) vs. ssri (fluoxetine). adults in the intervention groups was 0.27 standard deviations lower (1.02 lower to 0.47 higher)
Drop outs due to any reason - SSRI (Citalopram) vs. SSRI (Fluoxetine). Adults	37 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,6 due to risk of bias, imprecision, publication bias	RR 1.18 (0.38 to 3.72)	222 per 1000	40 more per 1000 (from 138 fewer to 604 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear how random sequence was generated and if allocation concealment was conducted. Single-blind study but patients were not blinded. High dropouts were reported >20%.
2 95% CI crossed 2 MIDs (-0.5 and 0.5).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 95% CI crossed 1 MID (0.5).
5 95% CI crossed 1 MID (-0.5).
6 95% CI crossed 2 MIDs (0.75 and 1.25).

Table 227: Summary of findings table for an antidepressant versus a combined antidepressant and psychotherapy at end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Antidepressant + Psychotherapy (BN)	Risk difference with Antidepressant (95% CI)
Laxative use. Adults - Self-help	44 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean laxative use. adults - self-help in the intervention groups was 0.04 standard deviations lower (0.64 lower to 0.55 higher)
Vomiting frequency. Adults	102 (3 studies)	⊕⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting frequency. adults in the intervention groups was 0.19 standard deviations higher (0.21 lower to 0.58 higher)
Vomiting frequency. Adults - Self-help	44 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting frequency. adults - self-help in the intervention groups was 0.02 standard deviations lower (0.62 lower to 0.57 higher)
Vomiting frequency. Adults - CBT	58 (2 studies)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting frequency. adults - cbt in the intervention groups was 0.35 standard deviations higher (0.17 lower to 0.87 higher)
Binge frequency- Adults	203 (5 studies)	⊕⊖⊖⊖ VERY LOW3,5,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency- adults in the intervention groups was 0.26 standard deviations higher (0.02 lower to 0.547 higher)
Binge frequency. Adults - CBT	109 (3 studies)	⊕⊖⊖⊖ VERY LOW3,5,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency. adults - cbt in the intervention groups was 0.63 standard deviations higher (0.24 to 1.02 higher)
Binge frequency. Adults - Self-help	44 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,9 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency. adults - self-help in the intervention groups was 0.02 standard deviations higher (0.58 lower to 0.61 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Antidepressant + Psychotherapy (BN)	Risk difference with Antidepressant (95% CI)
Binge frequency. Adults - Supportive Psychotherapy	50 (1 study)	⨁⊖⊖⊖ VERY LOW3,10,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency. adults - focal/ supportive psychotherapy in the intervention groups was 0.29 standard deviations lower (0.85 lower to 0.27 higher)
Purge frequency Total Adults	159 (4 studies)	⨁⊖⊖⊖ VERY LOW3,5,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean purge frequency total adults in the intervention groups was 0.22 standard deviations higher (0.1 lower to 0.54 higher)
Purge frequency, Adults - CBT	109 (3 studies)	⨁⊖⊖⊖ VERY LOW3,5,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean purge frequency, adults - cbt in the intervention groups was 0.49 standard deviations higher (0.1 to 0.87 higher)
Purge frequency, Adults - Supportive Psychotherapy	50 (1 study)	⨁⨁⊖⊖ LOW3,10,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean purge frequency, adults - focal/ supportive psychotherapy in the intervention groups was 0.35 standard deviations lower (0.92 lower to 0.21 higher)
General psychiatric features - Total Adults	179 (4 studies)	⨁⊖⊖⊖ VERY LOW3,7,12 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general psychiatric features - total adults in the intervention groups was 0.04 standard deviations lower (0.33 lower to 0.26 higher)
General psychiatric symptoms, Adults - CBT	85 (2 studies)	⨁⊖⊖⊖ VERY LOW3,5,13 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general psychiatric symptoms, adults - cbt in the intervention groups was 0.1 standard deviations higher (0.33 lower to 0.53 higher)
General psychiatric symptoms, Adults - Self-help	44 (1 study)	⨁⊖⊖⊖ VERY LOW3,11,14 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general psychiatric symptoms, adults - self-help in the intervention groups was 0.09 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Antidepressant + Psychotherapy (BN)	Risk difference with Antidepressant (95% CI)
					(0.69 lower to 0.5 higher)
General psychiatric symptoms, Adults - Supportive Psychotherapy	50 (1 study)	⨁⊖⊖⊖ VERY LOW3,10,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general psychiatric symptoms, adults - focal/ supportive psychotherapy in the intervention groups was 0.22 standard deviations lower (0.78 lower to 0.34 higher)
Depression. Adults - CBT	125 (4 studies)	⨁⊖⊖⊖ VERY LOW3,5,15 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression. adults - cbt in the intervention groups was 0.29 standard deviations higher (0.06 lower to 0.65 higher)
Depression. Adults - Self-help	44 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,14 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression. adults - self-help in the intervention groups was 0.02 standard deviations lower (0.62 lower to 0.57 higher)
Depression. Adults - supportive psychotherapy	50 (1 study)	⨁⊖⊖⊖ VERY LOW3,10,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression. adults - focal/ supportive psychotherapy in the intervention groups was 0.26 standard deviations higher (0.3 lower to 0.83 higher)
EDE-Shape concern. Adults - CBT	24 (1 study)	⨁⊖⊖⊖ VERY LOW2,3,16 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-shape concern. adults - cbt in the intervention groups was 0.26 standard deviations higher (0.54 lower to 1.07 higher)
EDE-Weight concern. Adults - CBT	24 (1 study)	⨁⊖⊖⊖ VERY LOW2,3,16 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-weight concern. adults - cbt in the intervention groups was 0.19 standard deviations higher (0.62 lower to 0.99 higher)
EDE-Global score, Adults - CBT	51 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,16 due to risk of bias,		Not calculable for SMD values	The mean ede-global score, adults - cbt in the intervention groups was 0.54 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Antidepressant + Psychotherapy (BN)	Risk difference with Antidepressant (95% CI)
		imprecision, publication bias			(0.03 lower to 1.1 higher)
EDI-Drive for thinness. Adults - CBT	34 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi-drive for thinness. adults - cbt in the intervention groups was 0.24 standard deviations higher (0.44 lower to 0.92 higher)
EDI-Bulimia. Adults - CBT	34 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi-bulimia. adults - cbt in the intervention groups was 0.6 standard deviations higher (0.09 lower to 1.29 higher)
EDI-Body dissatisfaction. Adults - CBT	34 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi-body dissatisfaction. adults - cbt in the intervention groups was 0.34 standard deviations higher (0.34 lower to 1.02 higher)
Drop out due to adverse events. Adults - CBT	140 (2 studies)	⨁⊖⊖⊖ VERY LOW3,15,18 due to risk of bias, imprecision, publication bias	RR 0.8 (0.31 to 2.07)	114 per 1000	23 fewer per 1000 (from 79 fewer to 122 more)
Remission (100% binge free). Adults - Supportive psychotherapy	45 (1 study)	⨁⊖⊖⊖ VERY LOW3,10,18 due to risk of bias, imprecision, publication bias	RR 1.10 (0.4 to 3)	294 per 1000	29 more per 1000 (from 176 fewer to 588 more)
Remission (100% binge free). Adults - CBT ITT	155 (3 studies)	⨁⊖⊖⊖ VERY LOW3,19,20 due to risk of bias, imprecision, publication bias	RR 0.56 (0.3 to 1.06)	222 per 1000	98 fewer per 1000 (from 156 fewer to 13 more)
Did not achieve Remission (100% binge free) FU Adults - CBT ITT	52 (1 study)	⨁⊖⊖⊖ VERY LOW3,17,20 due to risk of bias, imprecision, publication bias	RR 0.99 (0.89 to 1.11)	34 per 1000	0 fewer per 1000 (from 4 fewer to 4 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Antidepressant + Psychotherapy (BN)	Risk difference with Antidepressant (95% CI)
Remission (100% purge free). Adults - CBT ITT	155 (3 studies)	⨁⊖⊖⊖ VERY LOW3,19,21 due to risk of bias, imprecision, publication bias	RR 1.15 (0.44 to 3.06)	86 per 1000	13 more per 1000 (from 48 fewer to 178 more)
Remission (100% purge free). Adults - Supportive Psychotherapy ITT	50 (1 study)	⨁⊖⊖⊖ VERY LOW3,10,18 due to risk of bias, imprecision, publication bias	RR 1.31 (0.35 to 4.89)	136 per 1000	42 more per 1000 (from 89 fewer to 530 more)
Did not achieve Remission (100% purge free) FU Adults - CBT ITT (Copy)	52 (1 study)	⨁⊖⊖⊖ VERY LOW3,17,20 due to risk of bias, imprecision, publication bias	RR 0.90 (0.76 to 1.07)	34 per 1000	3 fewer per 1000 (from 8 fewer to 2 more)
Quality of life. Adults - CBT	34 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean quality of life. adults - cbt in the intervention groups was 0.17 standard deviations higher (0.5 lower to 0.85 higher)
EDI Body dissatisfaction FU. Adults - CBT	53 (1 study)	⨁⊖⊖⊖ VERY LOW3,5, 24 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi body dissatisfaction fu. adults - cbt in the intervention groups was 0.11 standard deviations higher (0.44 lower to 0.67 higher)
Vomit frequency FU. Adults - CBT	53 (1 study)	⨁⊖⊖⊖ VERY LOW3,11, 24 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomit frequency fu. adults - cbt in the intervention groups was 0.09 standard deviations lower (0.65 lower to 0.46 higher)
Depression FU. Adults - CBT	92 (2 studies)	⨁⊖⊖⊖ VERY LOW3,20,23 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression fu. adults - cbt in the intervention groups was 0.07 standard deviations higher (0.35 lower to 0.48 higher)
Laxative FU abuse - CBT	53 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,17,24		Not calculable for SMD values	The mean laxative fu abuse - cbt in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Antidepressant + Psychotherapy (BN)	Risk difference with Antidepressant (95% CI)
		due to risk of bias, imprecision, publication bias			0.18 standard deviations higher (0.38 lower to 0.73 higher)
Binge frequency FU. Adults - CBT	53 (1 study)	⨁⊖⊖⊖ VERY LOW3,12,24 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency fu. adults - cbt in the intervention groups was 0.00 standard deviations higher (0.55 lower to 0.55 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear how random sequence was generated and if allocation concealment was conducted. Unclear if it were blinded, although placebo pills were used. High dropouts were reported >20%.
2 95% CI crossed 2 MIDs (-0.5 and 0.5).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 It was unclear how randomised sequence was generated and if allocation concealment was conducted. It was unclear if patients, investigators or assessors were blind. High dropouts were reported.
5 95% CI crossed 1 MID (0.5).
6 Unclear how random sequence was generated or if allocation concealment was performed. In one study patients were not blinded. Unclear in either study if assessors were blind. High dropouts were reported >20%.
7 In most studies it is unclear how random sequence was generated and if allocation concealment were conducted. It is unclear if assessors were blind in all studies, High dropouts were reported.
8 Unclear how random sequence was generated or if allocation concealment was performed. Unclear in most studies if participants, investigators or assessors were blind. High dropouts were reported >20%.
9 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were blind but unclear if assessors were blind, one study investigators were not blind. High dropouts were reported.
10 Unclear how random sequence was generated and if allocation concealment was conducted. It is unclear if assessors were blind, High dropouts were reported.
11 95% CI crossed 1 MID (-0.5).
12 For continuous variable, there were fewer than 400 participants.
13 It was unclear how randomised sequence was generated and if allocation concealment was conducted. Participants were not blind in one study and it

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Antidepressant + Psychotherapy (BN)	Risk difference with Antidepressant (95% CI)

was unclear if assessors were blind in all studies. High dropouts were reported.
14 It was unclear how randomised sequence was generated and if allocation concealment was conducted. Participants were blind, but it was unclear if investigators or assessors were blind. High dropouts were reported.
15 It was unclear how randomised sequence was generated and if allocation concealment was conducted. Participants were not blind in one study and it was unclear if investigators were blind or assessors were blind in all studies. High dropouts were reported.
16 Unclear how random sequence was generated and if allocation concealment was conducted. It is unclear if participants, investigator or assessors were blind, High dropouts were reported.
17 It was unclear how randomised sequence was generated and if allocation concealment was conducted. Participants were not blind and it was unclear if investigators or assessors were blind. High dropouts were reported.
18 95% CI crossed 2 MIDs (0.75 and 1.25).
19 Unclear how random sequence was generated and if allocation concealment was conducted. It is unclear if participants, investigators or assessors were blind across different studies, High dropouts were reported.
20 For a dichotomous outcome, there were fewer than 300 events.
21 95% CI crossed 1 MID (0.75).
22 Unclear how random sequence was generated and if allocation concealment was conducted. Investigators were not blind and it was unclear if either participants or assessors were blind. High dropouts were reported >20%.
23 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were not blind in one study but not the investigators and it was unclear if the assessors were blind. In the other it was unclear if they were blind, along with the investigators and assessors. High dropouts were reported >20%.
24 It was unclear how random sequence was generated and if allocation concealment was performed. Participants were blind to drug treatment, assessors were blind but investigators were not blind. High dropouts were reported >20%.

Table 228: Summary of findings table for an antidepressant and nutritional therapy versus placebo and nutritional therapy at end of treatment and follow up in adults with bulimia nervosa

.Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+Nutrition	Risk difference with Antidepressant+Nutrition (95% CI)
EDE- Weight concern FU. Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias,		Not calculable for SMD values	The mean ede- weight concern fu. adults - ssri in the intervention groups was 0.12 standard deviations lower

		imprecision, publication bias			(0.6 lower to 0.36 higher)
EDE- Weight . Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede- weight . adults - ssri in the intervention groups was 0.94 standard deviations lower (1.45 to 0.44 lower)
EDE-Eating concern. Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-eating concern. adults - ssri in the intervention groups was 0.04 standard deviations lower (0.51 lower to 0.44 higher)
EDE-Eating concern FU. Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-eating concern fu. adults - ssri in the intervention groups was 0.12 standard deviations higher (0.36 lower to 0.6 higher)
EDE-Shape concern. Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-shape concern. adults - ssri in the intervention groups was 0.63 standard deviations lower (1.13 to 0.14 lower)
EDE-Shape concern FU. Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-shape concern fu. adults - ssri in the intervention groups was 0.26 standard deviations higher (0.23 lower to 0.74 higher)
Drop out due to any reason. Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, imprecision, publication bias	RR 1.53 (0.67 to 3.45)	212 per 1000	112 more per 1000 (from 70 fewer to 520 more)
Drop out due to adverse events. Adults - SSRI	67 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,6 due to risk of bias, imprecision, publication bias	RR 0.88 (0.77 to 1.01)	0 per 1000	-

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how the randomised sequence was generated and if allocation concealment was performed. It was unclear if either the participants or investigators were blinded. Assessors were blind. High dropouts were reported >20%
2 95% CI crossed 1 MID (-0.5)
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 95% CI crossed 1 MID (0.5)
5 95% CI crossed 2 MIDs (0.75 and 1.25)
6 95% CI crossed 1 MID (0.75)

Table 229: Summary of findings table for psychotherapy versus an antidepressant at the end of treatment in adults and at follow up with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	**Quality of the evidence (GRADE)**	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant	Risk difference with Psychotherapy (95% CI)
Laxative use. Adults - Self-help (Guided)	45 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean laxative use. adults - self-help (guided) in the intervention groups was 0.56 standard deviations higher (0.04 lower to 1.16 higher)
Vomiting. Adults	183 (3 studies)	⨁⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting. adults in the intervention groups was 0.51 standard deviations higher (0.21 to 0.8 higher)
Vomiting. Adults - Self-help (Guided)	45 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting. adults - self-help (guided) in the intervention groups was 0.82 standard deviations higher (0.21 to 1.44 higher)
Vomiting. Adults - CBT	88 (2 studies)	⨁⊖⊖⊖ VERY LOW2,3,5 due to risk of bias,		Not calculable for SMD values	The mean vomiting. adults - cbt in the intervention groups was 0.36 standard deviations higher

		imprecision, publication bias			(0.06 lower to 0.78 higher)
Vomiting. Adults - Focal psychoeducation	50 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting. adults - focal psychoeducation in the intervention groups was 0.49 standard deviations higher (0.08 lower to 1.06 higher)
Binge frequency Total Adult	183 (4 studies)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency total adult in the intervention groups was 0.09 standard deviations higher (0.2 lower to 0.38 higher)
Binge frequency. Adults - CBT	88 (2 studies)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency. adults - cbt in the intervention groups was 0.10 standard deviations lower (0.52 lower to 0.32 higher)
Binge frequency. Adults - Focal/ Supportive Psychotherapy	50 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency. adults - focal/ supportive psychotherapy in the intervention groups was 0.19 standard deviations higher (0.37 lower to 0.75 higher)
Binge frequency. Adults - Self-help (Guided)	45 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency. adults - self-help (guided) in the intervention groups was 0.37 standard deviations higher (0.22 lower to 0.97 higher)
Binge frequency (follow up). Adults - CBT	106 (2 studies)	⊕⊖⊖⊖ VERY LOW3,4,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency (follow up). adults - cbt in the intervention groups was 0.13 standard deviations lower (0.51 lower to 0.26 higher)
Purge frequency Total Adults	138 (3 studies)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean purge frequency total adults in the intervention groups was 0.28 standard deviations higher (0.05 lower to 0.62 higher)

Purge frequency. Adults - CBT	88 (2 studies)	⊕⊖⊖⊖ VERY LOW3,4,7 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean purge frequency. adults - cbt in the intervention groups was 0.17 standard deviations higher (0.25 lower to 0.59 higher)
Purge frequency. Adults - Supportive Psychotherapy	50 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean purge frequency. adults - focal/ supportive psychotherapy in the intervention groups was 0.49 standard deviations higher (0.08 lower to 1.06 higher)
Purge frequency (follow up). Adults - CBT	26 (1 study)	⊕⊖⊖⊖ VERY LOW3,8,9 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean purge frequency (follow up). adults - cbt in the intervention groups was 0.36 standard deviations lower (1.14 lower to 0.42 higher)
General psychiatric symptoms. Adults - CBT	88 (2 studies)	⊕⊖⊖⊖ VERY LOW3,5,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general psychiatric symptoms. adults - cbt in the intervention groups was 0.11 standard deviations lower (0.53 lower to 0.31 higher)
General psychiatric symptoms. Adults - Self-help (Guided)	45 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general psychiatric symptoms. adults - self-help (guided) in the intervention groups was 0.48 standard deviations higher (0.11 lower to 1.08 higher)
General psychiatric symptoms. Adults - Supportive Psychotherapy	50 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general psychiatric symptoms. adults - focal/ supportive psychotherapy in the intervention groups was 0.22 standard deviations higher (0.34 lower to 0.78 higher)
EDI-Drive for thinness. Adults - CBT	35 (1 study)	⊕⊖⊖⊖ VERY LOW3,8,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi-drive for thinness. adults - cbt in the intervention groups was 0.39 standard deviations lower (1.06 lower to 0.28 higher)
EDI-Weight concern. Adults - CBT	26 (1 study)	⊕⊖⊖⊖ VERY LOW 3,10,11		Not calculable for SMD values	The mean edi-weight concern. adults - cbt in the intervention groups was

		due to risk of bias, imprecision, publication bias			0.15 standard deviations lower (0.93 lower to 0.62 higher)
EDI-Shape concern. Adults - CBT	26 (1 study)	⊕⊖⊖⊖ VERY LOW3,11,12 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi-shape concern. adults - cbt in the intervention groups was 0.25 standard deviations lower (1.03 lower to 0.52 higher)
Depression scores. Adults - CBT	141 (4 studies)	⊕⊖⊖⊖ VERY LOW3,4,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression scores. adults - cbt in the intervention groups was 0.14 standard deviations lower (0.48 lower to 0.2 higher)
Depression scores. Adults - Self-help (guided)	45 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression scores. adults - self-help (guided) in the intervention groups was 0.45 standard deviations higher (0.14 lower to 1.05 higher)
Depression scores. Adults - Supportive Psychotherapy	50 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression scores. adults - focal/ supportive psychotherapy in the intervention groups was 0.2 standard deviations higher (0.36 lower to 0.76 higher)
Depression scores (follow up). Adults - CBT	76 (2 studies)	⊕⊖⊖⊖ VERY LOW3,7,13 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression scores (follow up). adults - cbt in the intervention groups was 0 standard deviations higher (47 lower to 0.47 higher)
EDE-Global Adults - CBT	53 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,8 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-global adults - cbt in the intervention groups was 0.39 standard deviations lower (0.94 lower to 0.15 higher)
EDE-Bulimia. Adults - CBT	35 (1 study)	⊕⊖⊖⊖ VERY LOW3,8,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-bulimia. adults - cbt in the intervention groups was 0.51 standard deviations lower (1.19 lower to 0.17 higher)

EDE-Body dissatisfaction. Adults - CBT	35 (1 study)	⊕⊖⊖⊖ VERY LOW3,8,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-body dissatisfaction. adults - cbt in the intervention groups was 0.44 standard deviations lower (1.11 lower to 0.24 higher)
Did not achieve remission (100% purge free). Adults - CBT ITT	145 (3 studies)	⊕⊖⊖⊖ VERY LOW3,4,14 due to risk of bias, imprecision, publication bias	RR 0.84 (0.71 to 0.98)	108 per 1000	17 fewer per 1000 (from 2 fewer to 31 fewer)
Did not achieve remission (100% purge free). Adults Supportive Psychotherapy ITT	50 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,15 due to risk of bias, imprecision, publication bias	RR 1.11 (0.89 to 1.38)	179 per 1000	20 more per 1000 (from 20 fewer to 68 more)
Did not achieve remission (100% purge free) FU Adults - CBT ITT	47 (1 study)	⊕⊖⊖⊖ VERY LOW3,10,15 due to risk of bias, imprecision, publication bias	RR 1.05 (0.86 to 1.29)	130 per 1000	7 more per 1000 (from 18 fewer to 38 more)
Did not achieve remission (100% binge free). Adults - CBT ITT	233 (4 studies)	⊕⊖⊖⊖ VERY LOW3,4,14 due to risk of bias, imprecision, publication bias	RR 0.78 (0.67 to 0.92)	156 per 1000	34 fewer per 1000 (from 12 fewer to 52 fewer)
Did not achieve remission (100% binge free). Adults - Focal/ Supportive Psychotherapy ITT	50 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,16 due to risk of bias, imprecision, publication bias	RR 1.03 (0.75 to 1.41)	250 per 1000	7 more per 1000 (from 62 fewer to 102 more)
Did not achieve remission (100% binge free) FU. Adults - CBT ITT	47 (1 study)	⊕⊖⊖⊖ VERY LOW3,15,17 due to risk of bias, imprecision, publication bias	RR 0.87 (0.71 to 1.06)	43 per 1000	6 fewer per 1000 (from 13 fewer to 3 more)
No adverse events. Adults - CBT	123 (2 studies)	⊕⊖⊖⊖ VERY LOW3,18,19	RR 1.09 (0.99 to	86 per 1000	8 more per 1000 (from 1 fewer to 17 more)

		due to risk of bias, imprecision, publication bias	1.2)		
Quality of life - CBT	35 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean quality of life - cbt in the intervention groups was 0.49 standard deviations lower (1.17 lower to 0.19 higher)
Laxative FU abuse - CBT	45 (1 study)	⊕⊖⊖⊖ VERY LOW3,7,20 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean laxative fu abuse - cbt in the intervention groups was 0.41 standard deviations lower (1 lower to 0.18 higher)
Vomit frequency FU. Adults - CBT	45 (1 study)	⊕⊖⊖⊖ VERY LOW3,12,20 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomit frequency fu. adults - cbt in the intervention groups was 0.05 standard deviations lower (0.64 lower to 0.54 higher)
EDI Body dissatisfaction FU. Adults - CBT	45 (1 study)	⊕⊖⊖⊖ VERY LOW3,8,20 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi body dissatisfaction fu. adults - cbt in the intervention groups was 0.49 standard deviations lower (1.09 lower to 0.1 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were not blind, but it was unclear if either investigators or assessors were blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (0.5).
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 Unclear how random sequence was generated and if allocation concealment was conducted. It was unclear if either participants, investigators or assessors were blind. High dropouts were reported >20%.
5 Unclear how random sequence was generated and if allocation concealment was conducted. In one study it was unclear if participants, investigators or assessors were blind. The other study was double bind but it was unclear if assessors were blind. High dropouts were reported >20%.
6 Unclear how random sequence was generated and if allocation concealment was conducted. Study was double-blind but it was unclear if assessors were blind. High dropouts were reported >20%.

7 For a continuous outcome there were fewer than 400 participants.
8 95% CI crossed 1 MID (-0.5)
9 Unclear how random sequence was generated and if allocation concealment was conducted. It was unclear if investigators, investigators or assessors were blind. High dropouts were reported >20%,
10 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were not blind and it was unclear if investigators or assessors were blind. High dropouts were reported >20%.
11 Unclear how random sequence was generated and if allocation concealment was conducted. It was unclear if participants, investigators or assessors were blind. High dropouts were reported >20%.
12 95% CI crossed 2 MIDs (-0.5 and 0.5).
13 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were blind but not investigators in one study and it was unclear if assessors were blind. In the other study it was unclear if any were blind. High dropouts were reported >20%.
14 95% CI crossed 1 MID (0.75).
15 95% CI crossed 1 MID (1.25).
16 95% CI crossed 2 MIDs (0.75 and 1.25).
17 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were not blind in one study and it was unclear if they were in the other study. It was unclear in both studies if either investigators or assessors were blind. High dropouts were reported >20%.
18 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were blind but not investigators in one study and it was unclear if assessors were blind. In the other study participants were not blind and it was unclear if investigators or assessors were blind. High dropouts were reported >20%.
19 For a dichotomous outcome there were fewer than 300 participants.
20 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were blind, investigators were not. It was unclear if assessors were blind.

Table 230: Summary of findings table for psychotherapy versus psychotherapy and antidepressant at end of treatment in adults with bulimia nervosa.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Combined Psychotherapy+Antidepressant	Risk difference with Psychotherapy (95% CI)
Binges. Adults - CBT	37 (1 study)	⨁⊖⊖⊖ VERY LOW3,4,5 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binges. adults - cbt in the intervention groups was 0.46 standard deviations higher (0.19 lower to 1.12 higher)
Binges. Adults - Guided SH	49 (1 study)	⨁⊖⊖⊖ VERY LOW3,6,7		Not calculable for SMD values	The mean binges. adults - guided sh in the intervention

		due to risk of bias, imprecision, publication bias			groups was 0.39 standard deviations higher (0.18 lower to 0.95 higher)
Vomiting. Total Adults	204 (4 studies)	⊕⊖⊖⊖ VERY LOW1,3,8,9 due to risk of bias, inconsistency, imprecision, publication bias		Not calculable for SMD values	The mean vomiting. total adults in the intervention groups was 0.74 standard deviations higher (0.45 to 1.04 higher)
Vomiting. Adults - CBT	111 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,8 due to risk of bias, inconsistency, imprecision, publication bias		Not calculable for SMD values	The mean vomiting. adults - cbt in the intervention groups was 0.98 standard deviations higher (0.56 to 1.4 higher)
Vomiting. Adults - Guided SH	49 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting. adults - guided sh in the intervention groups was 0.75 standard deviations higher (0.16 to 1.33 higher)
Vomiting. Adults - Focal psychoeducation	44 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomiting. adults - focal psychoeducation in the intervention groups was 0.25 standard deviations higher (0.35 lower to 0.84 higher)
Objective purgers. Adults - CBT	26 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean objective purgers. adults - cbt in the intervention groups was 0.44 standard deviations higher (0.35 lower to 1.22 higher)
Laxative use - Adults - CBT	49 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,12 due to risk of bias, imprecision,		Not calculable for SMD values	The mean laxative use - adults - cbt in the intervention groups was 0.55 standard deviations

		publication bias			higher (0.02 lower to 1.12 higher)
EDE-Global score. Adults - CBT	48 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-global score. adults - cbt in the intervention groups was 0.14 standard deviations higher (0.42 lower to 0.71 higher)
EDE - Shape concern. Adults - CBT	26 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede - shape concern. adults - cbt in the intervention groups was 0 standard deviations higher (0.77 lower to 0.77 higher)
EDE-Body dissatisfaction, Adults - CBT	37 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-body dissatisfaction, adults - cbt in the intervention groups was 0.04 standard deviations lower (0.68 lower to 0.61 higher)
EDE-Weight concern, Adults - CBT	26 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,11 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean ede-weight concern, adults - cbt in the intervention groups was 0 standard deviations higher (0.77 lower to 0.77 higher)
EDI-Drive for thinness. Adults - CBT	37 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,9 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi-drive for thinness. adults - cbt in the intervention groups was 0.16 standard deviations lower (0.8 lower to 0.49 higher)
EDI-Bulimia. Adults - CBT	37 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi-bulimia. adults - cbt in the intervention groups was 0.01 standard deviations higher (0.63 lower to 0.66 higher)
Depression, Adults - CBT	108 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3		Not calculable for SMD values	The mean depression, adults - cbt in the intervention

		due to risk of bias, imprecision, publication bias			groups was 0.18 standard deviations higher (0.2 lower to 0.56 higher)
Depression, Adults - Focal psychoeducation	44 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression, adults - focal psychoeducation in the intervention groups was 0.37 standard deviations higher (0.22 lower to 0.97 higher)
Remission. Adults - CBT_ITT	152 (3 studies)	⊕⊖⊖⊖ VERY LOW1,3,13,14 due to risk of bias, inconsistency, imprecision, publication bias	RR 1.14 (0.32 to 4.13)	185 per 1000	26 more per 1000 (from 126 fewer to 580 more)
Remission. Adults - Focal/psychoeducation_ITT	44 (1 study)	⊕⊖⊖⊖ VERY LOW3,10,14 due to risk of bias, imprecision, publication bias	RR 0.67 (0.12 to 3.61)	136 per 1000	45 fewer per 1000 (from 120 fewer to 356 more)
Quality of life - Adults - CBT	37 (1 study)	⊕⊖⊖⊖ VERY LOW3,6,9 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean quality of life - adults - cbt in the intervention groups was 0.43 standard deviations lower (1.08 lower to 0.22 higher)
General symptoms - Guided SH	49 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,6 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general symptoms - guided sh in the intervention groups was 0.37 standard deviations higher (0.2 lower to 0.93 higher)
General symptoms - CBT	48 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,12 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general symptoms - cbt in the intervention groups was 0.18 standard deviations higher

					(0.23 lower to 0.59 higher)
General symptoms - Focal psychoeducation	44 (1 study)	⊕⊖⊖⊖ VERY LOW3,5,10 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean general symptoms - focal psychoeducation in the intervention groups was 0 standard deviations higher (0.59 lower to 0.59 higher)
No side-effects. Adults - CBT	123 (2 studies)	⊕⊖⊖⊖ VERY LOW3,15,16 due to risk of bias, imprecision, publication bias	RR 1.12 (1.01 to 1.25)	114 per 1000	14 more per 1000 (from 1 more to 29 more)
Binge frequency FU. Adults - CBT	56 (1 study)	⊕⊖⊖⊖ VERY LOW3,9,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean binge frequency fu. adults - cbt in the intervention groups was 0.05 standard deviations lower (0.58 lower to 0.48 higher)
Laxative FU abuse - CBT	87 (1 study)	⊕⊖⊖⊖ VERY LOW3,9,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean laxative fu abuse - cbt in the intervention groups was 0.06 standard deviations lower (0.5 lower to 0.38 higher)
Vomit frequency FU. Adults - CBT	56 (1 study)	⊕⊖⊖⊖ VERY LOW3,9,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean vomit frequency fu. adults - cbt in the intervention groups was 0.13 standard deviations lower (0.66 lower to 0.4 higher)
Depression FU. Adults - CBT	87 (2 studies)	⊕⊖⊖⊖ VERY LOW2,3,17 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression fu. adults - cbt in the intervention groups was 0.18 standard deviations higher (0.25 lower to 0.62 higher)
EDI Body dissatisfaction FU. Adults - CBT	56 (1 study)	⊕⊖⊖⊖ VERY LOW3,9,17 due to risk of bias,		Not calculable for SMD values	The mean edi body dissatisfaction fu. adults - cbt in the intervention groups

		imprecision, publication bias			was 0.36 standard deviations lower (0.89 lower to 0.18 higher)
Did not achieve Remission-FU. Adults - CBT_ITT	53 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,19 due to risk of bias, imprecision, publication bias	RR 0.86 (0.71 to 1.05)	34 per 1000	5 fewer per 1000 (from 10 fewer to 2 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

1 Unclear how random sequence was generated and if allocation concealment was conducted. Across studies, it was unclear if either participants, investigators or assessors were blind. High dropouts were reported >20%,
2 95% CI crossed 1 MID (0.5)
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.
4 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were not blind in one study and it was unclear if investigators or assessors were blind. In the other study it was unclear if any were blind. High dropouts were reported >20%,
5 95% CI crossed 2 MIDs (-0.5 and 0.5)
6 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were not blind but it was unclear if investigators or assessors were blind. High dropouts were reported >20%,
7 For a continuous outcome there were fewer than 400 participants.
8 Heterogeneity detected I2 >80%
9 95% CI crossed 1 MID (-0.5).
10 Unclear how random sequence was generated and if allocation concealment was conducted. Participants and investigators were blind but it was unclear if assessors were blind. High dropouts were reported >20%,
11 Unclear how random sequence was generated and if allocation concealment was conducted. It was unclear if participants, investigators or assessors were blind. High dropouts were reported >20%,
12 Unclear how random sequence was generated and if allocation concealment was conducted. Participants may have been blind to pills taken, but it was unclear if investigators or assessors were blind. High dropouts were reported >20%,
13 Heterogeneity was detected 12>50%
14 95% CI crossed 2 MIDs (0.75 and 1.25)
15 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were not blind in one study but it was unclear if investigators or assessors were blind. In the other study, the participants were blind but it was unclear if either the investigators or assessors were blind, High dropouts were reported >20%,
16 For a dichotomous outcome there were fewer than 300 events.
17 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were blind in one study and it was unclear if

investigators or assessors were blind. In the other study it was unclear if any were blind. High drop outs were reported >20%,
18 Unclear how random sequence was generated and if allocation concealment was conducted. Participants were blind in one study, and investigators were not blind. But it was unclear if assessors were blind.
19 95% CI crossed 1 MID (0.75)

Table 231: Summary of findings table for an anticonvulsant versus placebo at end of treatment in adults with bulimia nervosa

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with placebo	Risk difference with Anticonvulsant (95% CI)
Clinical Global Impressions-Severity of Illness Scale (CGI-S). Adults	64 (1 study)	⊕⊖⊖⊖ VERY LOW 2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean clinical global impressions-severity of illness scale (cgi-s). adults in the intervention groups was 0.47 standard deviations lower (0.97 lower to 0.02 higher)
Clinical Global Impressions-Improvement Scale (CGI-I). Adults	64 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean clinical global impressions-improvement scale (cgi-i). adults in the intervention groups was 0.68 standard deviations lower (1.19 to 0.18 lower)
EDI - Drive for thinness. Adults	64 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - drive for thinness. adults in the intervention groups was 0.86 standard deviations lower (1.37 to 0.34 lower)
EDI - Bulimia. Adults	64 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - bulimia. adults in the intervention groups was 0.66 standard deviations lower (1.17 to 0.16 lower)
EDI - Body dissatisfaction. Adults	64 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean edi - body dissatisfaction. adults in the intervention groups was 0.7 standard deviations lower (1.21 to 0.19 lower)
General health perceptions - SF-36. Adults	60 (1 study)	⊕⊖⊖⊖ VERY LOW1,3,5		Not calculable	The mean general health perceptions - sf-36. adults in the intervention groups was

		due to risk of bias, imprecision, publication bias		for SMD values	1.22 standard deviations higher (0.67 to 1.78 higher)
No side-effects. Adults	67 (1 study)	⊕⊖⊖⊖ VERY LOW3,4,6 due to risk of bias, imprecision, publication bias	RR 1.03 (0.93 to 1.15)	61 per 1000	2 more per 1000 (from 4 fewer to 9 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 95% CI crossed 1 MID (0.5). 2 95% CI crossed 1 MID (-0.5). 3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded. 4 Unclear how random sequence was generated and if allocation concealment was conducted. Study was an open trial and it was unclear if investigators or assessors were blind. High dropouts were reported >20%. 5 Unclear how random sequence was generated and if allocation concealment was conducted. Participants and investigators were blind but it was unclear if assessors were blind. 6 For a dichotomous outcome there were fewer than 300 events.					

1 **Table 232: Summary of findings table for an antiemetics versus placebo at end of treatment in adults with bulimia nervosa.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Other medication (not antidepressants) vs, placebo (95% CI)
Did not drop out due to adverse events. Adults - Antiemetics	26 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias	RR 1.00 (0.87 to 1.15)	Not calculable for SMD values	Not estimable because of zero events in raw data.
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 It was unclear if assessors were blind.					

2 For a dichotomous outcome there were fewer than 300 events.
3 High risk of publication bias from studies sponsored by the pharmaceutical industry. There is a risk in the 1980's, 1990's and early 2000's that only positive findings are being published, there is selective outcome reporting and outliers are being excluded.

1

2

3

7.6.3 Economic Evidence

No economic evidence on the cost effectiveness of pharmacological interventions for the treatment of bulimia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Pharmacological interventions showed no benefit in the NMA conducted for this guideline and they were not considered as options in the economic model. Details on the methods used for the systematic search of the economic literature are described in Chapter 3. Details and findings of the NMA are provided in the Appendix R.

7.6.4 Clinical Evidence Statements

7.6.4.1 Antidepressant versus placebo in adults with bulimia nervosa at the end of treatment

SSRIs

Very low quality evidence from one study (n=42 to 46) showed no difference in the effect of SSRI's on bingeing, vomiting, EDI-total, EDI-drive for thinness, EDI-body dissatisfaction and EDI-bulimia compared with placebo.

Very low quality evidence from two studies (n=88) showed SSRI's may have a benefit on depression compared with placebo.

Very low quality evidence from three studies (n=234) showed SSRI's may have a benefit on global clinical scores compared with placebo.

Very low quality evidence from five studies (n=610) showed SSRIs may lead to more adverse events compared with placebo.

Very low quality evidence from one study (n=46) showed no difference in the effect of SSRI's on drop-out rates and remission compared with placebo.

TCAs

Very low quality evidence from one study (n=78) showed no difference in the effect of TCA's on purging and global clinical score compared with placebo.

Very low quality evidence from two studies (n=101) showed no difference in the effect of TCA's on depression compared with placebo.

Very low quality evidence from two studies (n=165) showed TCA's may lead to more adverse events compared with placebo.

MAOIs

Very low quality evidence from one study (n=77) showed MAOI's may have a harmful effect on EDI-total compared with placebo.

Very low quality evidence from two studies (n=127) showed no difference in the effect of MAOI's on depression compared with placebo.

Very low quality evidence from two studies (n=139) showed MAOI's may lead to more adverse events compared with placebo.

7.6.4.2 Antidepressant versus placebo in adults with bulimia nervosa at follow up

Very low quality evidence from one study (n=38) showed no difference in the effect of TCA's on bingeing, laxative use, vomiting, depression and EDI-body dissatisfaction compared with placebo.

7.6.4.3 Antidepressant versus another antidepressant in adults with bulimia nervosa at end of treatment.

Very low quality evidence from one study (n=28) showed no difference in the effect of SSRIs (Citalopram) on depression, EDI-drive for thinness, EDI-body dissatisfaction, EDI-bulimia and clinical global impression (adverse events) compared with another SSRI (Fluroxetine).

Very low quality evidence from one study (n=37) showed no difference in the effect of SSRIs (Citalopram) on drop-outs compared with another SSRI (Fluroxetine).

Very low quality evidence from one study (n=28) showed a benefit of SSRIs (Citalopram) on excessive exercise compared with another SSRI (Fluroxetine).

7.6.4.4 Antidepressant versus combined antidepressant and psychotherapy in adults with bulimia nervosa at end of treatment.

Antidepressant and self-help

Very low quality evidence from one study (n=44) showed no difference in the effect of antidepressants on laxative use, bingeing, depression, vomiting, general psychiatric symptoms compared with combined self-help and antidepressant.

Antidepressant and CBT

Very low quality evidence from one study (n=44) showed no difference in the effect of antidepressants on vomiting compared with combined CBT and antidepressant.

Very low quality evidence from one study (n=34 to 51) showed antidepressants are less effective on bingeing, EDI –bulimia, EDI body dissatisfaction and EDE-Global compared with combined CBT and antidepressant.

Very low quality evidence from three studies (n=109) showed antidepressants are less effective on purging compared with combined CBT and antidepressant.

Very low quality evidence from four studies (n=126) showed antidepressants are less effective on depression compared with combined CBT and antidepressant.

Very low quality evidence from two studies (n=85) showed no difference in the effect of antidepressants on general psychiatric symptoms compared with combined CBT and antidepressant.

Very low quality evidence from one study (n=24 to 34) showed no difference in the effect of antidepressants on EDE-shape concerns, EDE-weight concerns, EDE-drive for thinness and quality of life compared with combined CBT and antidepressant.

Very low quality evidence from two studies (n=140) showed no difference in the number of dropouts in those treated with antidepressants compared with combined CBT and antidepressant.

Very low quality evidence from three studies (n=155) showed antidepressants are less effective on remission (binge free) compared with combined CBT and antidepressant.

Very low quality evidence from three studies (n=155) showed no difference in the effect of antidepressants on remission (purge free) compared with combined CBT and antidepressant.

Antidepressant and supportive psychotherapy

Very low quality evidence from one study (n=50) showed no difference in the effect of antidepressants on purging, depression and general psychiatric symptoms compared with combined supportive psychotherapy and antidepressant.

Very low quality evidence from one study (n=45) showed no difference in the effect of antidepressants on remission (binge and purge free) compared with combined supportive psychotherapy and antidepressant.

7.6.4.5 Antidepressant versus combined antidepressant and psychotherapy in adults with bulimia nervosa at follow up.

Antidepressant and CBT

Very low quality evidence from one study (n=52) showed no difference in the effect of antidepressants on bingeing, EDI-body dissatisfaction, vomiting, laxative use and remission (either binge or purge free) compared with combined CBT and antidepressant.

Very low quality evidence from two studies (n=92) showed no difference in the effect of antidepressants on depression compared with combined CBT and antidepressant.

Antidepressant and supportive psychotherapy

Very low quality evidence from one study (n=52) showed antidepressants have no effect on remission compared with combined supportive psychotherapy and antidepressant.

7.6.4.6 Antidepressant and nutritional counselling versus combined placebo and nutritional counselling in adults with bulimia nervosa at end of treatment

Very low quality evidence from one study (n=67) showed antidepressants and nutritional counselling have no effect on EDE-eating concern or dropouts compared with combined placebo and nutritional counselling.

Very low quality evidence from one study (n=67) showed antidepressants and nutritional counselling may have a benefit on EDE-weight concern and EDE-shape concern compared with combined placebo and nutritional counselling.

7.6.4.7 Antidepressant and nutritional counselling versus combined placebo and nutritional counselling in adults with bulimia nervosa at follow up

Very low quality evidence from one study (n=67) showed antidepressants and nutritional counselling have no effect on EDE-weight concern, EDE-shape concern or EDE-eating concern compared with combined placebo and nutritional counselling.

Very low quality evidence from one study (n=67) showed antidepressants and nutritional counselling have more drop-outs due to adverse events than combined placebo and nutritional counselling.

7.6.4.8 Psychotherapy versus antidepressant in adults with bulimia nervosa at end of treatment

Guided self-help

Very low quality evidence from one study (n=45) showed guided self-help may be less effective on laxative use and vomiting compared with an antidepressant.

Very low quality evidence from one study (n=45) showed no difference in the effect of guided self-help on bingeing, depression and general psychiatric symptoms compared with an antidepressant.

CBT

Very low quality evidence from two studies (n=88 to 123) showed CBT may be less effective on vomiting and adverse events compared with an antidepressant, but there was some uncertainty.

Very low quality evidence from two studies (n=88) showed no difference in the effect of CBT on bingeing, purging and general psychiatric symptoms compared with an antidepressant.

Very low quality evidence from two studies (n=141) showed no difference in the effect of CBT on depression compared with an antidepressant.

Very low quality evidence from one study (n=26 to 35) showed no difference in the effect of CBT on EDI-drive for thinness, EDI-weight concern, EDI-shape concern, EDE-global, EDE-bulimia, EDE-body dissatisfaction and quality of life compared with an antidepressant.

Very low quality evidence from three studies (n=146) showed CBT may be more effective on remission (purge free) compared with an antidepressant.

Very low quality evidence from four studies (n=233) showed CBT may be more effective on remission (binge free) compared with an antidepressant.

Supportive psychotherapy

Very low quality evidence from one study (n=50) showed supportive psychotherapy may be less effective on vomiting and purging compared with an antidepressant, but there was some uncertainty.

Very low quality evidence from one study (n=45 to 50) showed no difference in the effect of supportive psychotherapy on bingeing, depression, general psychiatric symptoms and remission (binge and purge free) compared with an antidepressant.

7.6.4.9 Psychotherapy versus antidepressant in adults with bulimia nervosa at follow up

Very low quality evidence from two studies (n=76) showed no difference in the effect of CBT on depression compared with an antidepressant.

Very low quality evidence from one study (n=45 to 47) showed no difference in the effect of CBT on laxative use, vomiting, EDI-body dissatisfaction and remission (purge and binge free) compared with an antidepressant.

7.6.4.10 Psychotherapy versus combined psychotherapy and antidepressant in adults with bulimia nervosa at end of treatment

Guided self-help

Very low quality evidence from one study (n=49) showed guided self-help was less effective on vomiting compared with combined guided self-help and antidepressant.

Very low quality evidence from one study (n=49) showed no difference in the effect of guided self-help on general symptoms compared with combined guided self-help and antidepressant.

CBT

Very low quality evidence from three studies (n=111) showed CBT was less effective on vomiting compared with combined CBT and antidepressant.

Very low quality evidence from one study (n=26 to 49) showed no difference in the effect of CBT on bingeing, purging, laxative use, EDE-global, EDE-shape concern, EDI-body dissatisfaction, EDE-weight concern, EDI-drive for thinness, EDI-bulimia, general symptoms and quality of life compared with combined CBT and antidepressant.

Very low quality evidence from three studies (n=108 to 152) showed no difference in the effect of CBT on depression and remission compared with combined CBT and antidepressant.

Very low quality evidence from two studies (n=123) showed CBT may have fewer side-effects compared with combined CBT and antidepressant.

Supportive psychotherapy

Very low quality evidence from 1 study (n=44) showed no difference in the effect of supportive psychotherapy self-help on vomiting, depression, general symptoms and remission compared with an antidepressant.

7.6.4.11 Psychotherapy versus combined psychotherapy and antidepressant in adults with bulimia nervosa at follow up

Very low quality evidence from one study (n=56) showed no difference in the effect of CBT on bingeing, vomiting and EDI body dissatisfaction compared with combined CBT and antidepressant.

Very low quality evidence from two study (n=87) showed no difference in the effect of CBT on laxative use and depression compared with combined CBT and antidepressant.

Very low quality evidence from one study (n=53) showed CBT may have a benefit on remission compared with combined CBT and antidepressant, but there was some uncertainty.

7.6.4.12 Anticonvulsant versus placebo in adults with bulimia nervosa at end of treatment

Very low quality evidence from one study (n=60 to 64) showed an anticonvulsant may have a benefit on global clinical score, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction and general health perceptions (SF-36) compared with placebo.

Very low quality evidence from one study (n=67) showed no difference in the number of drop-outs in the anticonvulsant group compared with placebo.

7.6.5 Economic Evidence statements

No economic evidence on the cost effectiveness of pharmacological interventions for people with bulimia nervosa was available.

7.6.6 Recommendations and link to evidence for the review on: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

Medication

	113. Do not offer medication as the sole treatment for bulimia nervosa.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of pharmacotherapies for treating bulimia nervosa in children, young people and adults. For this population, binge eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with bulimia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	When considering whether to recommend a pharmacological agent for treating bulimia nervosa, it is important to firstly establish whether the agent is more effective compared with placebo and then consider head to head trials. Two studies were found comparing TCAs with placebo in adults with bulimia nervosa and showed no difference in purge frequency or global clinical score at the end of treatment. However, depression scores were lower in the TCA-treated group (with some uncertainty). At 10 years follow up, none of the outcomes were different between the TCA- and placebo-treated groups. The outcomes included: binge frequency, laxative use, vomit frequency, depression and EDI- body dissatisfaction (showed a trend only to be improved). At both time points, no data was reported on remission, all-cause mortality, adverse events, quality of life, resource use, relapse general psychopathology, general functioning, family functioning or service user experience A few studies were identified that compared SSRIs with placebo and showed at the end of treatment depression was improved in the SSRI-treated group, but not global clinical scores. No difference in binge frequency or vomiting frequency, EDI-global, EDI- drive for thinness, EDI- body dissatisfaction or EDI-bulimia were detected. Remission data was not available, nor was all-cause mortality, adverse events, quality of life, resource use, relapse, general functioning, family functioning or service user experience Some evidence was found comparing MAOI versus placebo and showed at the end of treatment no difference in depression or EDI-global scores. There were, however, more adverse events reported in the MAOI-treated group. Remission data was not available, nor was all-cause mortality, quality of life, resource use, relapse, general functioning, family functioning or service user experience. A head-to-head trial of two SSRIs, citalopram versus fluoxetine, showed excessive exercise appeared to be higher in citalopram arm but no difference in any of the outcomes including: depression, global clinical score, EDI-drive for thinness, EDI-body dissatisfaction, EDI-bulimia and drop-outs. Remission was not reported. nor was all-cause mortality, quality of life, resource use, relapse, general functioning, family functioning or service user experience Combining an antidepressant with CBT-ED showed favourable results on binge and purge frequency, and a trend to improve depression, remission, EDI-bulimia,

	and EDE-global scores compared with an antidepressant alone. All other outcomes were similar between the two groups including: vomiting frequency, adverse events, EDE and EDI subscales, quality of life and general psychiatric features. At 6 months to 1 year follow up, remission rates were no longer different, nor was any other outcome. All-cause mortality, quality of life, resource use, relapse, general functioning, family functioning or service user experience were not reported When combining an antidepressant with self-help there was no difference in binge frequency, laxative use, vomiting frequency, depression or general psychiatric features compared with an antidepressant alone. Similar results were found when an antidepressant was combined with supportive psychotherapy compared with an antidepressant alone and this included remission as an outcome. All-cause mortality, quality of life, resource use, relapse, adverse events, general functioning, family functioning or service user experience were not reported. Combining an antidepressant with nutritional therapy appeared to have some benefit on EDE-weight, EDE-shape concern compared with nutritional therapy alone. But no difference was found in EDE eating concern or dropouts due to adverse events or for any reason. At 12 weeks follow up, no difference was found in EDE weight concern or EDE-shape concern. There was a trend to favour EDE-eating concern in the combined treatment group. No data was available on remission, nor all-cause mortality, quality of life, resource use, relapse, general functioning, family functioning or service user experience. A head-to-head of self-help with an antidepressant showed some favourable results for the antidepressant on vomiting frequency and a trend to reduce laxative use. No differences were found for depression, binge frequency, purges and general psychiatric features. No data was available on remission, nor all-cause mortality, quality of life, resource use, relapse, adverse events, general functioning, family functioning or service user experience. CBT compared with an antidepressant showed some benefit of CBT-ED on remission rates, but no other benefits on depression, vomiting frequency (a trend for a negative effect), binge frequency, purges, general psychiatric features, quality of life, EDI and EDE subscales. No difference in harms were detected. The benefit of CBT-ED on remission rates were no longer evident at follow up. No data was available on all-cause mortality, resource use, relapse, adverse events, general functioning, family functioning or service user experience. Supportive psychotherapy compared with an antidepressant showed no difference in the effect of the two treatments on vomiting frequency (a trend for a negative effect), binge frequency, purges (a trend for a negative effect), general psychiatric features or remission rates. No data was available on all-cause mortality, resource use, relapse, adverse events, general functioning, family functioning or service user experience. Combining an antidepressant with psychotherapy and comparing it with psychotherapy alone showed guided self-help combined with an antidepressant had no additional benefit on binge frequency or general psychiatric features compared with guided self-help alone. A greater reduction in vomiting frequency was found in the guided self-help group alone. No data was available on remission, all-cause mortality, quality of life, resource use, relapse, adverse events, general functioning, family functioning or service user experience. Combining CBT-ED with an antidepressant showed a greater benefit on vomiting frequency compared with CBT-ED alone, and a trend for improving laxative use, but no additional benefit on depression, binge frequency, quality of life, EDE and EDI subscales or general psychiatric features. Fewer adverse events were detected in the CBT-ED alone group. No benefits were sustained long-term in the CBT-ED and antidepressant group. No data was available on remission, all-cause mortality, resource use, relapse, adverse events, general functioning, family functioning or service user experience. Supportive psychotherapy combined with an antidepressant did not add any benefit on depression, vomit frequency, laxative use, general psychiatric features or remission rates at the end of treatment compared with supportive psychotherapy alone. No data was available on remission, all-cause mortality, resource use,

	relapse, adverse events, general functioning, family functioning or service user experience. Anticonvulsants appeared to show positive effects on EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction, EDI bulimia, clinical global impressions (trend for one sub-scale) and general health perceptions compared with placebo. No difference was found on documented side-effects. No data was available on remission, all-cause mortality, resource use, relapse, general functioning, family functioning or service user experience. Finally, one small study reported no difference in adverse events in a group treated with antiemetics compared with placebo, but no other relevant outcomes were reported. The committee highlighted there are risks associated with prescribing medication to people with bulimia nervosa and a comorbidity because of potential physical problems. Depending on the severity and duration of the eating disorder, they may also have cardiovascular and renal problems, gastrointestinal disturbance, fluid and electrolyte abnormalities and dental abnormalities. For this reason, the committee wanted to express caution when prescribing or discontinuing treatment, such as serotonin reuptake inhibitors (SSRIs) that are prescribed as an antidepressant. The network meta-analysis showed pharmacological interventions were not effective at treating people with bulimia nervosa. For this reason, the committee agreed not to recommend any of the medications for treating bulimia nervosa. There was some discussion that the development of new drugs in the future may be effective, but they are a long way off and may be more effective for treating binge eating disorder. Adverse events or all-cause mortality were not reported in any of the studies.
Trade-off between net health benefits and resource use	The network meta-analysis found no evidence for the effectiveness of pharmacological interventions for the management of people with bulimia nervosa. As a result, such treatments are also likely to be not cost effective.
Quality of evidence	The quality of the evidence was considered very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised, if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded and high dropouts were detected >20%. . All outcomes were downgraded for risk of publication bias since the studies were sponsored by pharmaceutical companies and in the 1980s to early 2000s there is a risk that mostly positive findings were being published (Lexchin 2003(Lexchin et al., 2003)). Imprecision was often detected because the 95% confidence interval crossed one to two minimal important differences or the outcome did not meet the optimal information size. Remission was not often reported for the studies comparing a medication with placebo or directly comparing one medication with another.
Other considerations	The committee agreed that the evidence was too low quality and few benefits were found that would justify recommending a medication for treating bulimia nervosa. However, they agreed that medication may be needed for treating comorbidities such as major depressive disorders and anxiety disorders since high prevalence rates have been reported in people with bulimia nervosa. None of the evidence in this review included people with a comorbidity. The committee decided to exclude two drugs in the network meta-analysis: desiprarmine (TCA) because it is not available in the UK and imipramine (TCA) because it is not licensed for treating eating disorders in the UK. Nevertheless, these medications were kept in the pharmacological review to assess whether they are effective and if yes, whether a research recommendations may be made and whether licencing may need to be reconsidered. The results of the review suggest neither. The committee also wanted to include these drugs in the review in case they would

	make a recommendation about TCAs for people with an eating disorder and a mental health comorbidity. Unfortunately, people with comorbidities were not included in the studies, so it was not possible to make a specific recommendation about TCAs or any other medication for this population.

7.7 Management of long- and short-term complications

7.7.1 Review question: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 233. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all interventions that may be delivered to manage or reduce the short- or long-term physical complications of eating disorders in children, young people and adults and includes recovered as well as current service users. The interventions were categorised according to type of physical complication and intervention, the age of the participants and the type of eating disorder, and were compared to the control arm as reported in the relevant studies.

Table 233: Clinical review protocol summary for the review of: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

Component	Description
Review question(s)	What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) • recovered or current service users • **Strata:** • Children (≤12), young people (13-≤17 years), adults ≥18 years. • Eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder).
Intervention(s)	• Interventions to address the following: • Low bone mineral density (risk of fracture) • Growth (physical development) • Pubertal development • Tooth wear • Low body weight • Interventions to address the **long-term physical complications** may include: • GH/IGF-I • Calcium with and without Vitamin D • Bisphosphonates (age dependent and exclude pregnancy) • Exercise (low impact)/Physiotherapy • Oestrogen (patches/exogenous/pills other)

Component	Description
	• Testosterone (males/females) • Weight gain vs. Weight restoration (brain size) • Interventions to address the **short-term physical complications** may include • Phosphates supplementation (refeeding) • Potassium • Thiamine (refeeding) • Laxatives (for when underweight patients are constipated) • Salbutamol (reduce food intake)
Comparison	• Control arm as defined by study
Critical outcomes	• Primary outcome as reported by study
Important outcomes	• Secondary outcome as reported by study
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs) (if no RCTs)

7.7.2 Clinical Evidence for: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

No studies in people with bulimia nervosa were found that met the eligibility criteria for this review.

7.7.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for the management of short and long-term physical complications of bulimia nervosa was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

7.7.4 Clinical evidence statements

No studies in people with bulimia nervosa were found that met the eligibility criteria for this review.

7.7.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the management of short and long-term physical complications of bulimia nervosa was available.

7.7.6 Recommendations and linking evidence for the review: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

	The committee agreed no recommendation was needed for this review question on those with bulimia nervosa.
Relative value of different	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of nutritional interventions for treating bulimia nervosa in children, young people and adults. For this population, bingeing and remission

outcomes	are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	No relevant clinical evidence was identified.
Trade-off between net health benefits and resource use	No relevant existing economic evidence was identified.
Quality of evidence	Not applicable
Other considerations	The committee agreed that no recommendation was needed on how to treat or manage people with bulimia nervosa who have short or long-term complications.

7.8 Management of comorbidities

7.8.1 Review question: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 234. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers whether any intervention used to treat eating disorders in children, young people and adults needs to be modified in the presence of a common long-term health condition (i.e. comorbidity).The interventions were categorised according to their type, the type of eating disorder and comorbidity examined and the age of the participants. The comparison arm was the same intervention delivered to participants with the relevant eating disorder but without the relevant comorbidity.

Table 234: Clinical review protocol summary for the review of: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Component	Description
Review question(s)	Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) and a common comorbidity (e.g. diabetes, hypothyroidism). • Mental comorbidities may include: • Depression

Component	Description
	• Anxiety • Social anxiety • Autism • Obsessive Compulsive Disorder • Personality Disorder • Learning disability • ADHD (Bulimia) • Self-harm • Substance misuse • Physical comorbidities may include: • Celiac disease • Diabetes (type II – relevant to obesity) • Irritable Bowel Disease • Cystic Fibrosis **Strata:** • Children (≤12), young people (13-≤17 years), adults ≥18 years. • Eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder).
Intervention(s)	Trials will be included that address the ED as primary or secondary aim to treating the comorbidity. Interventions may include: Psychotherapy (including psychoeducation) Pharmacological Nutritional Physical Combination of any listed above
Comparison	• The same intervention but delivered to people with an eating disorder without a comorbidity.
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global

Component	Description
	Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs)

7.8.2 Clinical Evidence for Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Diabetes

One observational study in people with bulimia nervosa and type I diabetes was found that compared the effectiveness of an integrated inpatient care programme versus treatment as usual inpatient care (n=18) (Takii et al., 2003)). An overview of the trial can be found in Table 235: Study information of the observational study that compared different inpatient care for adults with bulimia nervosa and type I diabetes. Table 235.

History of substance misuse

One observational study in people with bulimia nervosa (n=100) who had a history of substance misuse met the eligibility criteria for this review (Mitchell et al., 1990a).

Although this review question includes people with any eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder, EDNOS), the committee wanted to firstly consider the evidence for individual eating disorders to see if specific recommendations could be made. If none was available, or it was deemed insufficient, then they agreed to make a general recommendation for treating people with any eating disorder and a common long-term health condition.

Summary of findings for those on bulimia nervosa can be found in Table 238. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 235: Study information of the observational study that compared different inpatient care for adults with bulimia nervosa and type I diabetes.

Study	Eating Disorder	Age mean (SD)	BMI	N	Stage of illness	Intervention	Comparison	Duration
Takii 2003 Japan	Bulimia Nervosa + Type I Diabetes	23.8 (5)	21.4 (3.1)	19	Binge Eating duration=4.9 (3.9) years Diabetes T1DM duration=7.6 (5.1) years	Integrated Inpatient Therapy – CBT + addressed diabetes + family relationships	No integrated inpatient care	Variable

Abbreviations: T1DM, type I diabetes melitis; CBT, cognitive behavioural therapy

Table 236: Study information of the observational study included in the review of interventions in young people with bulimia nervosa and common long-term health condition at follow up.

Study ID	N	Mean age	Female	Setting	Group 1	Group 2	Intervention	Duration
Mitchell 1990	100	17.5	not reported	OP	BN + History of chemical dependence	BN only	Group CBT	12 weeks

Abbreviations: BN, bulimia nervosa; CBT, cognitive behavioural therapy; OP, outpatients

Table 237: Summary of findings table on the effective of integrated inpatient care versus inpatient care for adults with bulimia nervosa and type I diabetes.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with IP therapy v No IP Therapy for BN+Diabetes1 (95% CI)
Did not achieve remission (no diagnosis of BN)	18 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision	RR 0.13 (0.02 to 0.8)	111 per 1000	97 fewer per 1000 (from 22 fewer to 109 fewer)
Depression	17 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias,		Not calculable for SMD	The mean depression in the intervention groups was 1.42 standard deviations lower

		imprecision		values	(2.52 to 0.32 lower)
General Psychopathology	17 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 1.25 standard deviations lower (2.31 to 0.18 lower)
Inappropriate Compensatory Behaviours to prevent weight gain past 3 months	18 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision	RR 4 (1.15 to 13.88)	778 per 1000	1000 more per 1000 (from 117 more to 1000 more)
Insulin Omission	18 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision	RR 2 (0.93 to 4.3)	556 per 1000	556 more per 1000 (from 39 fewer to 1000 more)
Calorific Value of Binge Episodes (Kcal)	18 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean calorific value of binge episodes (kcal) in the intervention groups was 1.52 standard deviations lower (2.6 to 0.44 lower)
EDI Total	17 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total in the intervention groups was 1.16 standard deviations lower (2.21 to 0.11 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 The patients were selected from the same recruitment site and showed no difference in their characteristics, except for binge frequency that was significantly higher in the inpatient group. The follow up was different for the two groups: 36 months for IP group and 24 months for non-IP group. Investigators were not blind to treatment allocation.
2 There were fewer than 10 per arm.
3 For a dichotomous outcome, there were fewer than 300 events.
4 95% CI crossed 1 MID (-0.5)
5 For a continuous outcome, there were fewer than 400 participants.

Table 238: Summary of findings table for Group CBT in young people with bulimia nervosa and history of substance misuse compared with those who had no history of substance misuse at follow-up.

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with Control	Risk difference with Group CBT (95% CI)
Remission FU	87 (1 study) 3.5 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision	RR 1.01 (0.72 to 1.4)	677 per 1000	7 more per 1000 (from 190 fewer to 271 more)
Treatment Failures FU	87 (1 study) 3.5 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision	RR 1.11 (0.5 to 2.48)	246 per 1000	27 more per 1000 (from 123 fewer to 364 more)
Hospitalised for substance abuse FU	87 (1 study) 3.5 years	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision	RR 0.98 (0.11 to 8.99)	46 per 1000	1 fewer per 1000 (from 41 fewer to 369 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Mitchell 1990: Sample is those with and without history of substance abuse; current substance abuse comorbidity not included; selection bias (history of substance abuse group significantly older); performance bias (no info about intervention etc.); attrition bias (insufficient info about intervention); high detection bias.
2 CI crosses both 0.75 and 1.25.

1

2

7.8.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for bulimia nervosa in the presence of common long-term conditions was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

7.8.4 Clinical evidence statements

7.8.4.1 Diabetes and eating disorder treatment (integrated care) in hospital in adults with bulimia nervosa

Very low quality evidence from one observational study (n=17 to 18) showed integrated inpatient care is more effective on remission, depression, calorie intake with bingeing, EDI-total and general psychopathology compared with treatment as usual inpatient care.

Very low quality evidence from one observational study (n= 18) showed integrated inpatient care is less effective on inappropriate compensatory behaviours to prevent weight gain compared with treatment as usual inpatient care.

Very low quality evidence from one observational study (n= 18) showed no difference in the effect of integrated inpatient care on insulin omission compared with treatment as usual inpatient care.

7.8.4.2 Substance misuse in young people with bulimia nervosa

Very low quality evidence from one observational study (n=87) showed no difference in the effect of group CBT-general in people with bulimia nervosa and a history of substance misuse on remission and treatment failures compared with those with bulimia and no history of substance abuse.

Very low quality evidence from one observational study (n=87) showed no difference in the effect of group CBT-general in people with bulimia nervosa and a history of substance misuse on the number of people were hospitalised for substance abuse during the follow up period compared with those with bulimia and no history of substance abuse.

7.8.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for bulimia nervosa in the presence of common long-term conditions was available.

7.8.6 Recommendations and link to evidence for the review on: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Diabetes

	114. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder. **115. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each**

	condition and the potential impact they have on each other. **Diabetes** **116. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.** **117. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.** **118. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes.** **119. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.** **120. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.** **121. Offer people with an eating disorder who are misusing insulin the following treatment plan:** • a low carbohydrate diet, so that insulin can be started at a low level • gradually increasing insulin doses to reduce blood glucose levels • adjusted total glycaemic load and carbohydrate distribution to meet their individual needs and prevent rapid weight gain • carbohydrate counting when adjusting their insulin dose (including via pumps) • a diabetic educational intervention such as DAFNE • education about the problems caused by misuse of diabetes medication **122. For more guidance on managing diabetes, refer to the NICE guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem. In the case of diabetes, HbA1c levels and insulin omission days were considered critical outcomes. The other critical outcomes depended on the eating disorder included in the study. Remission is of greatest concern for any eating disorder. In addition, for those with anorexia nervosa body weight or BMI are of greatest concern. For bulimia nervosa and binge eating disorder, binge eating is a critical outcome. For any eating disorder, other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore

	extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with an eating disorder that are of lesser importance but are clearly still important outcomes include general psychopathology, eating disorder psychopathology, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	The ideal study design to answer the question of whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem would be to randomise people with an eating disorder and diabetes to two different treatment groups: one modified to address both the eating disorder and diabetes and one non-modified eating disorder treatment. **Bulimia nervosa (reviewed in this chapter)** One observational study compared the effectiveness of inpatient integrated care with treatment as usual in adults with bulimia nervosa and type I diabetes. The integrated care provided CBT-ED, family based therapy and addressed control of diabetes. Whilst treatment as usual included outpatient counselling sessions on diabetes but not inpatient care or treatment for the eating disorder. This study showed better outcomes for the integrated care including, remission, general psychopathology, depression, EDI-total, the size of the binges, few compensatory behaviours but no difference in insulin omission. No data was available on HbA1c levels, all-cause mortality, adverse events, quality of life, resource use, relapse, general functioning, family functioning and service user experience. **Anorexia nervosa (as reviewed in chapter 6)** No published evidence was found on people with anorexia nervosa and diabetes, however there was a sub-group analysis from a study described below on any eating disorder that showed those with anorexia nervosa and type I diabetes are equally responsive to treatment as those with anorexia nervosa alone. No data was available on HbA1c levels, insulin omission days, remission, all-cause mortality, adverse events, quality of life, resource use, relapse, general functioning, family functioning and service user experience. **Binge eating disorder (as reviewed in chapter 8)** One study randomised adults with type II diabetes and binge eating to either group CBT-ED or a non-prescriptive control therapy (NPT). The results showed no difference in remission or binge frequency at the end of treatment. BMI showed a trend to be higher in the group CBT-ED arm, however EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction and quality of life were no different. At follow up, remission rates were higher in the CBT-ED arm, but again no difference in any of the other outcomes and BMI showed a trend to be higher in the group CBT-ED arm compared with controls. No data was available on HbA1c levels, insulin omission, all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience. An observational study compared the same diabetes prevention programme but in two populations, one with bulimia nervosa and a major depressive disorder and one with just any eating disorder. The results showed no difference in the degree of weight loss between the two populations. No data was available on HbA1c levels, insulin omission, remission, bingeing, all-cause mortality, adverse events, resource use, relapse, general functioning, eating disorder psychopathology, family functioning and service user experience. **Any eating disorder (as reviewed in chapter 9)** One randomised control trial compared group psychoeducation (combined with treatment as usual) with treatment as usual (diabetes treatment only) in people with type I diabetes and disturbed eating behaviours and showed no difference at the end of treatment on bingeing, EDE-restraint, EDE-shape concern, EDE-eating concern, EDE-weight concern, EDI-drive for thinness, EDI-bulimia, insulin omission days and HbA1c (%). One outcome, EDI-body dissatisfaction, favoured group psychoeducation over treatment as usual but there was some uncertainty. At follow up some benefit was found in response to group psychoeducation on bingeing but there was some uncertainty. No data was available on remission, all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience.

	An observational study was identified that compared the same CBT-ED intervention but in two populations, one with any eating disorder and type I diabetes, and one with just any eating disorder. Thus, this design allowed us to see whether those with a comorbidity would respond equally well to treatment as those with just an eating disorder. The results showed adults with any eating disorder and a comorbidity are less likely to recover than those with just an eating disorder. No difference was found in dropouts. No data was available on all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience.
Trade-off between net health benefits and resource use	The committee considered that providing care for eating disorders in the presence of a long-term health problems, such as diabetes, may have resource implications in terms of extra time required to provide collaborative and comprehensive care in line with the principles outlined in the recommendations 108-114. However, the committee expressed the view that if such care arrangements (that is, multidisciplinary approach, involvement of family members and carers, and the use of treatment plans) lead to better and appropriate treatment and management of health problems (including other long-term health problems such as diabetes) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating such care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence was mostly low quality from the RCT studies and very low quality from the observational studies. In both types of study designs the sample size was generally small and only one study was available for most outcomes, thus imprecision was often detected due to the 95% confidence interval crossing a minimal important difference or the outcome did not meet the optimal information size. **Binge eating disorder** In one RCT where they compared group CBT-ED with a control therapy in the same population (people with type I diabetes and binge eating disorder) inadequate randomisation was performed and it was unclear if allocation concealment was carried out. Neither the participant nor investigator was blind, nor was it clear if the assessor was blind. It was unclear how many participants completed the intervention. The observational study included in the review provided indirect evidence since it was a diabetes prevention programme and the participants had a major depressive disorder in addition to binge eating disorder or binge eating disorder alone. The only outcome reported was weight loss. The committee did not consider this study helpful. **Bulimia nervosa** In the observational study where they compared inpatient integrated care with treatment as usual, the people were selected from the same recruitment site and showed no difference in their characteristics, except that binge frequency was significantly higher in the inpatient group. The duration of follow up was different for the two groups: 36 months versus 24 months in the inpatient care and treatment as usual groups, respectively. Investigators were not blind to treatment allocation and only 18 participants were included in the study. **Any eating disorder (including subgroup analysis on anorexia nervosa)** In the RCT where they compared group psychoeducation and management (and treatment as usual) with treatment as usual (diabetes only programme), it was unclear if allocation concealment was performed. Neither the participant, investigator or assessor were blind and it was unclear how many completed the intervention. The population was also indirect since it included people with disturbed eating. Also, the comparison did not show whether a modified eating disorder programme is more effective at treating people with diabetes and an eating disorder compared with an eating disorder programme alone. Rather the study compared a modified diabetes programme with a regular diabetes programme. In the observational study where they compared CBT-ED in people with eating

	disorder alone or with a comorbidity, the authors attempted to match the groups based on age, marital status, education, catchment area and onset of diagnosis. However, it was unclear whether the two groups were followed up for the same duration and the sample size was very small. **Overall discussion** No RCT or observational study met the criteria of what would have been the ideal study design for this review (as described above). One RCT compared the effectiveness of an intervention that addressed both the eating disorder and diabetes, but the other arm addressed the diabetes alone. In another RCT, one intervention was modified but it was compared with a control therapy. In the observational studies, one study compared the same intervention but in those with either an eating disorder and diabetes or just the eating disorder alone. Thus, it only provided insight into whether one group was more responsive to treatment than the other. In the other observational study, inpatient integrated care was compared with treatment as usual, but the treatment as usual only addressed the diabetes not the eating disorder. Therefore, it did not provide insight into whether a modified eating disorder treatment was needed for those with a comorbidity.
Other considerations	In conclusion, it was difficult for the committee to draw conclusions from these studies on whether treatment for an eating disorder needs to be modified in the presence of a comorbidity such as diabetes. The committee therefore agreed that it was best to instead provide guidance on how to manage the diabetes. Usually, the committee would refer to the diabetes NICE guideline, but because the diabetes guideline refers to this guideline, the committee needed to recommend what to do in the presence of both. The committee agreed on a series of recommendations based on their experience and knowledge on how to manage the diabetes in the presence of an eating disorder. A number of the recommendations are based on what would be considered good practice. For instance: i) establish who will monitor the physical health, ii) explain to the person that they need to monitor their diabetes during the treatment for the eating disorder, and iii) be aware of the problems caused by misuse of diabetes medication. The committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in the treatment of diabetes. They highlighted that the quality of the family environment has been shown to affect treatment compliance and metabolic control among young people with an eating disorder (Hauser 1990). Family members may also need to care for someone if they hyper or phyo (which is a case for medical emergency), so they know how to respond. There is also the possibility that eating disturbances in young girls with diabetes are associated with significantly more family dysfunction than girls with diabetes alone (i.e. 13 to 18 years of age). Specifically, they can receive less support, and have poorer communication and less trust in their relationship with their parents than diabetic girls without eating disturbances. For these reasons, the committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in their treatment. There was some indirect evidence to support the recommendation to address insulin misuse as part of any psychological treatment. One RCT (n=85) showed that a modified group psychoeducation and management programme reduced bingeing episodes at follow up compared with a programme that just addressed the diabetes alone. This study was considered with the reservation that it was indirect because: 1) it did not investigate the effectiveness of a modified eating disorder psychological treatment and 2) the population had a disturbed eating behaviour, not a specific eating disorder diagnosis. Nevertheless, it showed that a psychoeducation and management programme may help reduce eating disorder psychopathology in those who also have diabetes. The committee discussed the problem of a relatively high prevalence of EDNOS In young girls with diabetes. In girls who have body dissatisfaction, diabetes provides a unique but dangerous opportunity to control weight by deliberate insulin omission, which can lead to hyperglycaemia and glycosuria. It is therefore important that insulin misuse is addressed in any psychological intervention.

	It can be noted that the recommendations relating to diet control were contributed to by the expert opinion of a dietician on the committee, based on their experience of treating those with an eating disorder who misuse insulin. These recommendations are based upon the treatment approach of small, attainable and incremental goals. At the outset of treatment, intensive glucose management is not an appropriate goal. The first goal must be to establish medical safety for the person with diabetes by gradually increasing the doses of insulin and food intake (as described in the recommendation). Given the fear of weight gain in this population, the committee recommended that the diet is amended to prevent rapid weight gain. They also suggested an educational programme called Dose Adjustment for Normal Eating (DAFNE) that provides people with the skills necessary to estimate the carbohydrate in each meal and to inject the right dose of insulin. There was no evidence on how to treat the eating disorder in the presence of any other long-term physical health condition, such as cystic fibrosis, celiac disease, pregnancy or irritable bowel disease. Some eating disorder specialists on the committee highlighted that they would generally refer someone with an eating disorder and diabetes to a diabetologist rather than address the points raised in the recommendations on diabetes. However, the committee agreed that it should be collaborative approach for the health care professionals who treat eating disorders and diabetes. Especially for young people who may need to involve family members and carers in therapy sessions to help the person with blood glucose control. Given the lack of direct evidence to address this review question the committee agreed to make a research recommendation to ask: “Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?” See chapter 6.8

1 **Substance and medication misuse**

	123. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder. **124. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.** **Substance and medication misuse** **125. For people with an eating disorder who are misusing substances, or over the counter or prescribed medication, provide treatment for the eating disorder unless the substance misuse is interfering with this treatment.** **126. If substance misuse or medication is interfering with treatment, consider a multi-disciplinary approach with substance misuse services.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of treating people with an eating disorder and a comorbidity. For binge eating disorder and bulimia nervosa, it was agreed binge eating frequency and remission are of greatest concern. For anorexia nervosa, body weight/BMI and remission are critical and for ENDOS, remission and either binge eating or body weight/BMI depending on the eating disorder they most closely resemble. The other outcomes that are critical are the primary outcomes

	that are relevant to the physical or mental health comorbidity being treated. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	**Bulimia nervosa (included in this chapter)** An observational study compared the long-term outcomes (2 to 5 years) of women with bulimia nervosa who had a history of substance abuse with those who no history of substance abuse (or chemical dependency). Both groups had received outpatient group cognitive behavioural psychotherapy and showed no different in long-term remission rates or being hospitalised for substance abuse. No evidence was found on the critical outcome of binge eating, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. **Bulimia nervosa and EDNOS (data presented in chapter 9)** Another observational study was identified were they extracted data from a randomised control trial and compared the outcomes in those with bulimia nervosa and EDNOS who had a low or high alcohol intake. The participants were treated with either broad or focused CBT-ED. At the end of 20 weeks of treatment, there was no difference in the number who had EDE scores one standard deviation above the community norms (i.e., relatively abnormal eating psychopathology) in those with a low or high alcohol intake. However, the number who continued to have excessive alcohol intake (defined as >21 units or >14 units/week for males and females respectively) was higher in those whose alcohol intake was high compared with those whose intake was low. At 60 weeks of follow up, there continued to be no difference in EDE scores between those who had low versus high alcohol intake. No evidence was found on the critical outcomes of remission and binge eating, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. No randomised control trial evidence was found. **Anorexia nervosa or binge eating disorder** No relevant published evidence was found in those with anorexia nervosa or binge eating disorder.
Trade-off between net health benefits and resource use	The committee considered that providing care for people with an eating disorder who are misusing substances or medication may have resource implications in terms of the extra time required to facilitate care for such people (in particular the use of a multi-disciplinary approach). However, the committee expressed the view that if such care leads to better identification of health needs and this results in appropriate subsequent treatment and management of health problems (including eating disorder and substance and medication misuse) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating a multi-disciplinary care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence used to generate these recommendations was very low quality. The evidence was observational and in GRADE (the software used to assess the quality) the evidence starts at very low quality and can only be upgraded if large effect sizes are found or if a dose response is identified. Neither was the case for this review. In the absence of RCT evidence the committee still considered this evidence insightful. In one study a reasonable number of participants were included (n=119) and they underwent a currently recommended CBT-ED programme, however, there were few outcomes reported and no remission data. In the other study, again there was a reasonable number of participants included (n=81), but there was no

	data at the end of treatment (only at follow up) and again few outcomes were reported. However, they did measure remission. The evidence suggested that those with a low or high alcohol intake may be equally responsive to an eating disorder treatment. Also, in the long-term, a positive response to treatment may be found in both those with a history of substance misuse and those with no history. Thus, the committee recommended that for people with an eating disorder who are misusing substances, provide treatment for the eating disorder unless the substance misuse is demonstrably interfering with this treatment. If substance misuse is interfering with treatment, the committee recommended considering a multidisciplinary approach. Psychological treatment is unlikely to be effective if cognitive function is poor as a result of the substance misuse. Therefore, it may be best if the person first receives treatment for their substance misuse problem. A multidisciplinary approach will ensure once the person is in a better cognitive state, they can begin treatment for their eating disorder.
Other considerations	The committee discussed the scenario of when a clinician is faced with someone who has an eating disorder and a drinking problem, they need to consider two important questions: firstly, do those eating disorder patients with concurrent high alcohol consumption do less well in treatment? And secondly, do they require treatment that is different and modified to focus upon both drinking and eating problems? The evidence presented in this review answered both question with no; people with a high or low alcohol intake or a history or no history of substance have almost identical responses to CBT-ED, therefore an amended programme may not be needed. There is also evidence to suggest that patients benefit from treatment not only with respect to their eating disorder but also in terms of their alcohol intake. In the study by Karacic 2011, over half the high alcohol intake group were no longer drinking excessively (52.8%, n=19) at the end of treatment, however, 12.5% (n=10) of the low alcohol intake group were now drinking above the safe limit (this data was not extracted because change scores were not presented). Another important finding was that mean intake for the high alcohol intake group changed from a risky drinking pattern to one closer to recommended guidelines (again this data could not be extracted because no error estimates around the mean were provided). These changes happened even although the treatment did not address alcohol consumption. For these reasons, the committee were confident recommending that the person undergoes treatment for the eating disorder unless the substance misuse is interfering with the treatment. In such cases, a multidisciplinary approach may be needed. Although the evidence for this recommendation was found in those with bulimia nervosa and ENDOS, the committee were confident that the findings would translate to those with any eating disorder. For this reason, they did not specify the type of eating disorder. It was discussed in the committee meeting that comorbid alcoholism has been associated with an increased risk of mortality in people with an eating disorder.

1

8 Treatment and management of binge eating disorder

8.1 Introduction

Individuals with BED regularly binge on large amounts of food in a discrete period with accompanying loss of control. Bingeing is accompanied by significant distress and may involve high levels of guilt and shame, eating in secret and eating despite not being hungry or until feeling uncomfortably full. Recurrent binges occur may occur against a background of a general tendency to overeat, or people with BED may eat normally between binges, but do not fast or use other compensatory behaviours to a significant degree. As a result, many people with BED are overweight or obese. The demographic distribution of BED is distinctive in that the majority of patients are middle-aged and about a third are male. The course of BED is also quite different from other eating disorders. Rather than being persistent, it tends to remit and recur, with extended periods, often lasting many months, free from the eating disorder. It is generally recognised that treatment should be focused around reducing or eliminating bingeing rather than on weight loss.

8.2 Psychological interventions

8.2.1 Review question: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with binge eating disorder compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 239. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all psychological interventions that may be delivered to children, young people and adults with binge eating disorder with or without a pharmacological intervention. The interventions were categorised according to their mode of delivery, i.e. individual, group or self-help, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to any other intervention or to wait list controls.

Table 239: Clinical review protocol summary for the review of: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with binge eating disorder compared with any other intervention or controls?

Component	Description
Review question(s)	Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder. • Strata: ○ children (≤12), young people (13-17 years), adults≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. atypical eating disorder) ○ mode of delivery (i. individual ii. family iii. group iv. self-help)
Intervention(s)	Psychological intervention including:

Component	Description
	• Dialectical behaviour therapy (DBT) • Counselling (Nutritional/Other) • Integrative Cognitive-Affective Therapy for Binge Eating (ICAT) • Maudsley model for treatment of adults with anorexia nervosa (MANTRA) • Cognitive remediation therapy (CRT) • Specialist supportive clinical management for anorexia nervosa (SSCM) • Behavioural therapy (BT) • CBT (General or ED specific) • Dynamic (IPT, Psychodynamic General or ED specific) • Guided Self Help w therapist guidance • Pure self-help • E-therapies Psychological in combination with any pharmacological intervention.
Comparison	• wait list control • treatment as usual • another other intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for bulimia nervosa and binge eating disorder; and weight/body mass index (Appropriate adjustment for age) for anorexia nervosa
Important outcomes	• Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF). • Family functioning. • Service user experience • Quality of life. • All-cause mortality. • Relapse. • Adverse events • Resource use.
Study design	• Systematic reviews • RCTs

8.2.2 Clinical Evidence for Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?

8.2.2.1 Individual psychotherapy versus any other intervention or wait list control

Eight RCTs (n=760) met the eligibility criteria for this review on individual therapies for those with binge eating disorder. The majority of those found were on adults (Castelnuovo 2011 (Castelnuovo et al., 2011); DeBar 2013 (DeBar et al., 2013); Fischer 2014 (Fischer et al., 2014); Hill 2011 (Hill et al., 2011); Kristeller 2014 (Kristeller et al., 2014), McIntosh 2016

(McIntosh et al., 2016), Ricca 2010 (Ricca et al., 2010), Wilson 2010 (Wilson et al., 2010)). The study by DeBar 2013 on young people.

An overview of the trials included in the meta-analysis can be found in Table 240. Further information about both included and excluded studies can be found in Appendix J.

No studies were identified that compared a combined individual psychotherapy with a pharmacological agent with any other intervention or wait list controls.

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

8.2.2.2 Group therapy versus any other intervention or wait list control

15 RCTs (n=1504) met the eligibility criteria for this review all of which were on adults (Agras 1994 (Agras et al., 1994b), Alfonsson 2015 (Alfonsson et al., 2015), Grilo 2011 (Grilo et al., 2011), Hilbert 2004 (Hilbert and Tuschen-Caffier, 2004), Kristeller 2014 (Kristeller et al., 2014), Munsch 2007 (Munsch et al., 2007), Nauta 2000 (Nauta et al., 2000), Peterson 2009 (Peterson et al., 2009), Peterson 2001 (Peterson et al., 2001), Ricca 2010 (Ricca et al., 2010), Safer 2010 (Safer et al., 2010), Shapiro 2007 (Shapiro et al., 2007), Telch 1990 (Telch et al., 1990), Wilfley 1993 (Wilfley et al., 1993) and Wilfley 2002 (Wilfley et al., 2002)). Further information about both included and excluded studies can be found in Appendix J.

8.2.2.3 Self-help versus any other intervention or wait list control

16 RCTs (n=1605) met the eligibility criteria for this review all of which were on adults (Carrard 2011 (Carrard et al., 2011), Carter 1988 (Carter and Fairburn, 1998), Cassin 2008 (Cassin et al., 2008), DeBar 2011 (DeBar et al., 2011), Dunn 2006 (Dunn et al., 2006), Ghaderi 2003 (Ghaderi and Scott, 2003), Grilo 2005 (Grilo and Masheb, 2005), Grilo 2013 (Grilo et al., 2013), Jones 2008 (Jones et al., 2008), Loeb 2000 (Loeb et al., 2000), Masson 2013 (Masson et al., 2013), Peterson 2009 (Peterson et al., 2009), Peterson 2001 (Peterson et al., 2001), Shapiro 2007 (Shapiro et al., 2007), Striegel-Moore 2010 (Striegel-Moore et al., 2010), Wilson 2010 (Wilson et al., 2010)). Further information about both included and excluded studies can be found in Appendix J.

8.2.2.4 Family therapy versus any other intervention or wait list control

One RCT (n=94), which was on adults, met the eligibility criteria for this review (Gorin 2003 (Gorin et al., 2003)). The study compared group CBT-ED with spouse involvement, group CBT-ED without spouse involvement and wait list control. An overview of this trial can be found in Table 243.

Table 240: Study information for trials included in the meta-analysis of individual psychotherapy versus any other intervention or wait list controls.

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison Arm	Sessions N	Treatment length	Long-term FU
Castelnuovo 2011	46.2 (10.5)	mean weight: 107 ± 6.9 kg	100%	Admission to hospital	60	Hybrid 1 (CBT-ED and weight loss diet, exercise, counselling) group. Inpatient and outpatient treatment.	Hybrid 2 (brief strategic thinking and weight loss group) Inpatient and outpatient treatment.	16	7 months	None reported
DeBar 2013	15.1 (1.9)	26.6 (5.7)	100%	At least 3 months	26	CBT-ED	Treatment as usual	8	6 months	None reported
Fischer 2014	45.6 (11.2).	34.3 (8.2)	88%	NR	41	CBT-ED	Wait list control	8 + 5 =13	8 weeks (active treatment) and 12 month	12 months FU
Hill 2011	22.7 (5.9)	23.2 (5.2)	100%	At least one binge eating and one vomit episode per week/3 month	32	Dialectical behaviour therapy	Wait list control	12	12 weeks	None reported
Kristeller 2014	46.6 (20–74)	40.3 (26-78)	88%	20 years of BED	140	Mindfulness	Wait list control CBT-ED	12	12 weeks	4 months FU
McIntosh 2016	35.3 (12.0)	29.9 (7.8)	100%	Mean duration of illness 14.6 (13.2) years	112	CBT-ED	CBT-general Behavioural therapy	32	12 months	12 months FU
Ricca 2010	46.5 (12.4)	NR	86%	Minimum duration of	144	CBT-ED	CBT-ED Group	22	24 weeks	3 year FU

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N random-ised	Intervention	Comparison Arm	Sessions N	Treatment length	Long-term FU
				six consecutive months						
Wilson 2010	50.3 (14.6)	36.2 (4.3)	85%	NR	205	IPT	Guided SH-ED Behavioural weight loss	20	24 weeks	None

Abbreviations: BED - binge eating disorder; BN - bulimia nervosa: BMI - body mass index; BT - behavioural therapy; CBT-ED -cognitive behavioural therapy with eating disorder focus; ED -eating disorder; EDNOS - eating disorder not otherwise specified; FU - follow up; IPT - interpersonal patient therapy; N - number; NR - not reported; SD - standard deviation; SH-ED - self-help with eating disorder focus.

Table 241: Study information for trials included in the meta-analysis of group psychotherapy versus any other intervention or wait list controls.

Study_ID	Mean age (SD) years	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention	Comparison Arm	Sessions N	Treatment Length	Long-term FU
Agras 1994	45.0 (10)	38.6 (6.6)	100%	Average duration of binge eating was 26 years	108	CBT-ED Group	Self-help with Group therapy CBT-ED Group + Pharmacotherapy	30	30 weeks	None reported
Alfonsson 2015	44.3 (10.7)	41.2 (5.3)	94%	NR	100	BT Group	Wait list controls	10	10 weeks	None reported
Grilo 2011	44.9 (9.5)	38.7 (5.7)	62%	20 years of BED	125	CBT-ED Group	BT (Group) CBT- ED + BT	16	6 months	12 months FU
Hilbert 2004	42.1 (12.1)	34.0 (10.2)	100%	Mix of BED and Binge eaters. Duration mean 13 years	28	CBT-ED Group (Body exposure)	CBT-ED Group (cognitive)	24	5 months + 9 weeks	4 months FU
Kristeller 2014	46.6, range of 20–74	39.6 (8.0)	88%	20 years of BED	140	Nutritional Group Mindfulness	Wait list controls	12	12 weeks	4 months FU

Study_ID	Mean age (SD) years	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention	Comparison Arm	Sessions N	Treatment Length	Long-term FU
						Group				
Munsch 2007	44.4 (11.5)	33.7 (4.3)	91%	NR	80	CBT-ED Group	BT Group	16	16 weeks + 6 months	12 months FU
Nauta 2000	38.3 (7.1)	33.1 (4.3)	100%	NR	37	CBT-ED Group	BT Group	15	15 weeks	6 months FU
Peterson 2009	47.1 (10.4)	39.0 (7.8)	88%	NR	259	Group psychoeducation	Group guided self-help Wait list controls Group self-help	15	20 weeks	12 mo FU
Peterson 2001	42.9 (10.1)	34.1 (7.4)	100%	NR	51	Group psychoeducation	Group guided self-help Group self-help	14	8 weeks	12 mo FU
Ricca 2010	46.5 (12.4)	NR	86%	Minimum duration of six consecutive months	144	CBT-ED Group	CBT-ED	22	24 weeks	3 year FU
Safer 2010	51.9 (11.6)	35.8 (9.4)	86%	Duration of bingeing 32.8 years	101	BT - ED	Counselling (Group)	20	21 weeks	12 months FU
Shapiro 2007	39.1 (11.4)	37.7 (11.4)	93%	NR	66	CBT-ED Group	Wait list controls Guided self-help (ED)	10	10 weeks	2 months FU
Telch 1990	42.6 (8.4)	32.6 (5.1)	100%	Subjects reported binge eating for 22.9 years (11.9)	44	CBT-ED Group	Wait list controls	10	10 weeks	None reported
Wilfley 1993	44.3 (8.3)	32.8 (5.2)	100%	Subjects reported binge eating for an	56	IPT Group	CBT-ED Group Wait list controls	16	16 weeks	None reported

Study_ID	Mean age (SD) years	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention	Comparison Arm	Sessions N	Treatment Length	Long-term FU
				average of 23.7 years (13.4)						
Wilfley 2002	45.6 (9.6)	NR	83%	Duration of disorder approximately 21 years	162	CBT-ED Group	Guided Group IPT (ED)	20	20 weeks	12 months FU

Abbreviations: BED - binge eating disorder; BN - bulimia nervosa: BMI - body mass index; BT - behavioural therapy; CBT-ED -cognitive behavioural therapy with eating disorder focus; ED -eating disorder; EDNOS - eating disorder not otherwise specified; FU - follow up; IPT - interpersonal patient therapy; N - number; NR - not reported; SD - standard deviation; SH-ED - self-help with eating disorder focus

Table 242: Study information for trials included in the meta-analysis of self-help versus any other intervention or wait list controls.

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention	Comparison Arm	Sessions N	Treatment Length	Long-term FU
Carrard 2011	34.4 (11.0)	29.8 (5.9)	100%	Binge eating for at least 3 months	74	Guided Self-help-ED	Wait list controls	26	6 months	6 months FU
Carter 1988	39.7 (10.0)	31.6 (6.6)	100%	Approximately 16 years of BED	93	Guided Self-help ED	Self-help ED Wait list controls	6 to 8	12 weeks	6 months FU
Cassin 2008	42.5 (12.7)	33.2 (7.8)	100%	Duration of illness 15.1 (11.6)	108	Guided Self-help ED	Self-help ED	1	16 weeks	None reported
DeBar 2011	39.1 (6.7)	31.5 (5.9)	100%	NR	160	Guided Self-help ED	Treatment as usual	8	3 months	None reported
Dunn 2006	19.0 (2.6)	23.8 (4.1)	89%	NR BN (23.3%) and BED (27.8%), subthreshold BN (6.7%) subthreshold	90	Guided Self-help ED	Self-help ED	1	4 months	None reported

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention	Comparison Arm	Sessions N	Treatment Length	Long-term FU
				BED (8.9%) and partial criteria for BN or BED and						
Ghaderi 2003	29 (10.7)	24.7 (5.5)	NR	Approximately 9 years of BED. Sub-threshold BN (n=11, BN (n=9) and BED (n=11)	31	Guided Self-help ED	Self-help ED	6 to 8	16 weeks	None reported
Grilo 2005	46.3 (9.0)	35.5 (6.7)	79%	Duration of BED: 17.1 years	90	Guided Self-help ED	Guided Self-help ED 2. Control	6	12 weeks	None reported
Grilo 2013	45.8 (11.0)	37.6 (4.8)	79%	Duration approximately 20 years. Age of onset, 25.8	48	Self-help ED	Usual care		4 months	None reported
Jones 2008	15.1 (1.0)	30.58 (4.9)	73%	Bingeing at least 3 months	105	Self-help ED Internet	Wait list controls		16 weeks	9 months FU
Loeb 2000	41.5 (9.4)	35.8 (9.0)	100%	Sub-threshold BN n=2; BN=2; subthreshold BED n=3, full criteria BED n=33; ;	40	Guided Self-help ED	Self-help ED	6	10 weeks	None reported
Masson 2013	42.8 (10.5)	37.1 (8.8)	88%	NR	60	Guided Self-help ED	Wait list controls	6	13 weeks	12 months FU
Peterson 2009	47.1 (10.4)	39.0 (7.8)	88%	NR	259	Group Guided Self-help ED	Wait list controls Group	15	20 weeks	12 months

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention	Comparison Arm	Sessions N	Treatment Length	Long-term FU
							Psychoeducation Group Self-help ED			FU
Peterson 2001	42.9 (10.1)	34.1 (7.4)	100%	NR	51	Group Guided Self-help ED	Group Self-help ED Group Psychoeducation	14	8 weeks	12 months FU
Shapiro 2007	40.4 (10.6)	39.6 (7.9)	92%	NR BED 71%; no BED 29%	66	Guided Self-help (ED) CD-ROM	Group CBT Wait list controls	10	10 weeks	2 months FU
Striegel-Moore 2010	37.2 (7.8)	31.3 (6.2)	92%	NR BED (53%) and BN	123	Guided Self-help ED	Usual care	8	12 weeks	40 weeks FU
Wilson 2010	50.3 (14.6)	36.2 (4.3)	85%	NR	205	Guided Self-help ED	IPT Behavioural weight loss	20	24 weeks	None reported

Abbreviations: BED - binge eating disorder; BN - bulimia nervosa: BMI - body mass index; BT - behavioural therapy; CBT-ED -cognitive behavioural therapy with eating disorder focus; ED -eating disorder; EDNOS - eating disorder not otherwise specified; FU - follow up; IPT - interpersonal patient therapy; N - number; NR - not reported; SD - standard deviation; SH-ED - self-help with eating disorder focus.

Table 243: Study information for trials included in the analysis of family therapy versus another intervention or wait list control.

Study_ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Sample	N Initially Random-ised	Intervention	Comparison	Sessions N	Treatment Length	Long-term FU
Gorin 2003	45.2 (10.0)	39.4 (7.7)	NR	Satisfies DSM-IV criteria BED.	94	Group CBT-ED with Spouse	Group CBT-ED without Spouse Wait list controls	12	12 weeks	6 months

Abbreviations: BED - binge eating disorder; CBT-ED -cognitive behavioural therapy for eating disorders; NR – not reported

Table 244: Summary table of findings for hybrid therapy (CBT-ED and weight loss) compared to another hybrid therapy (brief strategic thinking and weight loss) at the end of treatment in adults with binge eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with other Hybrid 2	Risk difference with Binge Hybrid (95% CI)
Global clinical score	60 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean global clinical score in the intervention groups was 1.09 standard deviations lower (1.64 to 0.55 lower)
% weight loss	60 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean % weight loss in the intervention groups was 0.34 standard deviations higher (0.17 lower to 0.85 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Unclear if allocation concealment was performed. Unclear if the participants, assessors or investigators were blind.
2 Fewer than 400 participants
3 95% CI crossed 1 MID (0.5).

Table 245: Summary table of findings for CBT-ED compared to another intervention at the end of treatment in adults and young people with binge eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with Binge CBT-ED (95% CI)
BMI Young people	26 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi young people in the intervention groups was 0.02 standard deviations higher (0.75 lower to 0.79 higher)
Depression Young people	26 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression young people in the intervention groups was 1.08 standard deviations lower (1.91 to 0.25 lower)

Depression Adults	141 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression adults in the intervention groups was 0.00 standard deviations higher (33 lower to 0.33 higher)
EDE - Dietary restraint Young people	26 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - dietary restraint young people in the intervention groups was 0.65 standard deviations lower (1.44 lower to 0.15 higher)
EDE- Dietary restraint Adults	253 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- dietary restraint adults in the intervention groups was 0.52 standard deviations lower (0.78 to 0.26 lower)
EDE - Eating concerns Young people	26 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - eating concerns young people in the intervention groups was 1.41 standard deviations lower (2.29 to 0.54 lower)
EDE- Eating concerns Adults	256 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concerns adults in the intervention groups was 0.51 standard deviations lower (0.76 to 0.25 lower)
EDE - Shape concerns Young people	26 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - shape concerns young people in the intervention groups was 0.11 standard deviations higher (0.66 lower to 0.88 higher)
EDE- Shape concerns Adults	256 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concerns adults in the intervention groups was 0.56 standard deviations lower (0.80 to 0.28 lower)
EDE-Weight concerns Adults	256 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concerns adults in the intervention groups was 0.07 standard deviations higher (0.18 lower to 0.32 higher)
EDE - Weight concerns Young people	26 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - weight concerns young people in the intervention groups was 0.30 standard deviations lower (1.07 lower to 0.48 higher)
EDE- Global score Adults	346	⊕⊕⊖⊖		Not calculable for	The mean ede- global score adults in the

	(2 studies)	LOW1,5 due to risk of bias, imprecision		SMD values	intervention groups was 0.99 standard deviations lower (1.24 to 0.74 lower)
Social adjustment - Young people	26 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean social adjustment - young people in the intervention groups was 0.52 standard deviations lower (1.3 lower to 0.27 higher)
Binge eating Adults	253 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating adults in the intervention groups was 0.05 standard deviations higher (0.20 lower to 0.30 higher)
Remission Young people_ITT	26 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR 2.00 (0.95 to 4.23)	385 per 1000	385 more per 1000 (from 19 fewer to 1000 more)
Remission Adults	112 (1 study)	⊕⊕⊖⊖ LOW1,7 due to risk of bias, imprecision	RR 0.63 (0.39 to 1.03)	541 per 1000	200 fewer per 1000 (from 330 fewer to 16 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted. Assessors were blind but it was unclear if participants or investigators were blind.
2 95% CI crossed 2 MIDs (-0.5 and 0.5).
3 95% CI crossed 1 MID (-0.5)
4 It was unclear if allocation concealment was conducted. Assessors were blind but it was unclear if participants or investigators were blind. High drop outs were reported >20%
5 For a continuous outcome there were fewer than 400 participants.
6 95% CI crossed 1 MID (1.25).
7 95% CI crossed 1 MID (0.75)

1 **Table 246: Summary table of findings for CBT-ED compared to another intervention at follow up in adults with binge eating disorder.**

Outcomes	No of Participants (studies) Follow up	**Quality of the evidence (GRADE)**	**Relative effect (95% CI)**	**Anticipated absolute effects**	
				Risk with another intervention FU	**Risk difference with Binge CBT-ED (95% CI)**

BMI FU	346 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi fu in the intervention groups was 0.19 standard deviations lower (0.41 lower to 0.03 higher)
Depression FU	141 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.00 standard deviations higher (0.33 lower to 0.33 higher)
Binge eating FU	258 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating fu in the intervention groups was 0.10 standard deviations higher (0.15 lower to 0.34 higher)
EDE- Global scale FU	346 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- global scale fu in the intervention groups was 1.02 standard deviations lower (1.27 to 0.77 lower)
EDE- Dietary restraint FU	231 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- dietary restraint fu in the intervention groups was 0.39 standard deviations lower (0.66 to 0.13 lower)
EDE- Weight concerns FU	231 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concerns fu in the intervention groups was 1.53 standard deviations lower (1.86 to 1.20 lower)
EDE- Shape concerns FU	231 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values Not calculable for SMD values	The mean ede- shape concerns fu in the intervention groups was 1.67 standard deviations lower (2.0 to 1.33 lower)
EDE- Eating concerns FU	231 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision			The mean ede- eating concerns fu in the intervention groups was 1.28 standard deviations lower (1.59 to 0.97 lower)
Remission FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 0.84 (0.57 to 1.24)	632 per 1000	101 fewer per 1000 (from 272 fewer to 152 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%

confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Across studies it was unclear if allocation concealment was conducted. In Wilson, it was unclear if either the participants or investigators were blind, assessors were blind. In Ricca participants were not blind and assessors were only blind at baseline. Investigators were not blind. High drop outs were reported in Ricca >20%.
2 For a continuous outcome there were fewer than 400 participants.
3 95% CI crossed 1 MID (-0.5).
4 95% CI crossed 1 MID (0.75)

Table 247: Summary table of findings for interpersonal psychotherapy versus any other intervention at the end of treatment in adults of binge eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Another intervention	Risk difference with Binge IPT (95% CI)
BMI	205 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.02 standard deviations higher (0.26 lower to 0.31 higher)
Binge eating	205 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating in the intervention groups was 0.05 standard deviations lower (0.33 lower to 0.24 higher)
Remission ITT	205 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision	RR 1.05 (0.94 to 1.2)	815 per 1000	41 more per 1000 (from 49 fewer to 163 more)
BMI FU	205 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.01 standard deviations higher (0.27 lower to 0.3 higher)
Binge eating FU	205 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating fu in the intervention groups was 0.07 standard deviations lower (0.35 lower to 0.22 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how the random sequence was generated or if allocation concealment was performed. It was unclear if participants and investigators were blind to treatment, however, assessors were blind. High dropout rates were reported >20%
2 For a continuous outcome, there were fewer than 400 participants.
3 For a dichotomous outcome, there were fewer than 300 events.

Table 248: Summary table of findings for dialectical behaviour therapy versus wait list control at the end of treatment in adults with binge eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Wait list control	Risk difference with Binge DBT (95% CI)
Binge eating (objective	32 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating (objective in the intervention groups was 0.14 standard deviations lower (1.2 lower to 0.22 higher)
Vomiting episodes	32 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting episodes in the intervention groups was 0.72 standard deviations lower (1.44 lower to 0 higher)
EDE-Global Score	32 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-global score in the intervention groups was 1.02 standard deviations lower (1.77 to 0.27 lower)
Depression	32 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.90 standard deviations lower (1.63 to 0.16 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear if allocation concealment was performed. It was also unclear if participants and investigators were blind, however, assessors were.

2 95% CI crossed 1 MID (-0.5)

Table 249: Summary of findings table for BT compared to another intervention at the end of treatment and follow up in adults with BED

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with Binge BT (95% CI)
Bulimic episodes	112 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic episodes in the intervention groups was 0.03 standard deviations higher (0.37 lower to 0.42 higher)
Purging	112 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging in the intervention groups was 0.19 standard deviations higher (0.21 lower to 0.58 higher)
Symptom checklist	112 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist in the intervention groups was 0.16 standard deviations higher (0.24 lower to 0.55 higher)
EDE-Dietary restraint	112 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-dietary restraint in the intervention groups was 0.01 standard deviations higher (0.38 lower to 0.41 higher)
EDE-weight concern	112 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern in the intervention groups was 0.06 standard deviations lower (0.46 lower to 0.33 higher)
EDE-shape concern	112 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concern in the intervention groups was 0.06 standard deviations lower (0.46 to 0.33 higher)
EDE-eating concern	112 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern in the intervention groups was 0.26 standard deviations higher (0.14 lower to 0.65 higher)

EDI-bulimia	112 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-bulimia in the intervention groups was 0.18 standard deviations lower (0.57 lower to 0.22 higher)
EDI-body dissatisfaction	112 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-body dissatisfaction in the intervention groups was 0.16 standard deviations lower (0.55 lower to 0.24 higher)
EDI-drive for thinness	112 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-drive for thinness in the intervention groups was 0.18 standard deviations lower (0.58 lower to 0.22 higher)
Remission	148 (1 study)	⊕⊕⊖⊖ LOW1 due to risk of bias, imprecision	RR 0.64 (0.41 to 1.01)	434 per 1000	156 fewer per 1000 (from 256 fewer to 4 more)
Bulimic episodes FU	86 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean bulimic episodes fu in the intervention groups was 0.11 standard deviations lower (0.56 lower to 0.34 higher)
Purging FU	87 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging fu in the intervention groups was 0.34 standard deviations higher (0.12 lower to 0.79 higher)
Symptom checklist FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean symptom checklist fu in the intervention groups was 0.29 standard deviations higher (0.16 lower to 0.74 higher)
EDE-Dietary restraint FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-dietary restraint fu in the intervention groups was 0.07 standard deviations lower (0.52 lower to 0.38 higher)
EDE-weight concern FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern fu in the intervention groups was 0.08 standard deviations lower (0.53 lower to 0.37 higher)
EDE-shape concern	87	⊕⊕⊖⊖		Not calculable for	The mean ede-shape concern fu in the

FU	(1 study)	LOW1,2 due to risk of bias, imprecision		SMD values	intervention groups was 0.03 standard deviations higher (0.42 lower to 0.49 higher)
EDE-eating concern FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern fu in the intervention groups was 0.16 standard deviations lower (0.61 lower to 0.29 higher)
EDI-bulimia FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-bulimia fu in the intervention groups was 0.29 standard deviations lower (0.74 lower to 0.17 higher)
EDI-body dissatisfaction FU	87 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-body dissatisfaction fu in the intervention groups was 0.05 standard deviations lower (0.50 lower to 0.40 higher)
EDI-drive for thinness FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-drive for thinness fu in the intervention groups was 0.20 standard deviations lower (0.65 lower to 0.25 higher)
Remission FU	112 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 1.22 (0.81 to 1.82)	434 per 1000	96 more per 1000 (from 82 fewer to 356 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how randomisation was conducted or if allocation concealment was performed. Assessors were blind but it was unclear if investigators or participants were blind. High drop outs were reported >20%.
2 For a continuous outcome there were fewer than 400 participants
3 95% CI crossed 1 MID (0.5)
4 95% CI crossed 1 MID (-0.5)
5 95% CI crossed 1 MID (1.25)

Table 250: Summary of findings table for CBT-general compared to another intervention at the end of treatment and follow up in adults with BED

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with another interventi on	Risk difference with CBT-General vs another intervention (95% CI)
Purging	112 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculabl e for SMD values	The mean purging in the intervention groups was 0.16 standard deviations lower (0.56 lower to 0.23 higher)
Bingeing	112 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculabl e for SMD values	The mean bingeing in the intervention groups was 0.14 standard deviations lower (0.53 lower to 0.25 higher)
EDE-Weight concern	112 (1 study)	⨁⨁⨀⨀ LOW1 due to risk of bias, imprecision		Not calculabl e for SMD values	The mean ede-weight concern in the intervention groups was 0.22 standard deviations higher (0.17 lower to 0.61 higher)
EDE-Shape concern	112 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculabl e for SMD values	The mean ede-shape concern in the intervention groups was 0.21 standard deviations higher (0.18 lower to 0.50 higher)
EDE-Eating concern	112 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculabl e for SMD values	The mean ede-eating concern in the intervention groups was 0.11 standard deviations lower (0.51 lower to 0.28 higher)
EDE- Restraint	112 (1 study)	⨁⨁⨀⨀ LOW1,4 due to risk of bias, imprecision		Not calculabl e for SMD values	The mean ede- restraint in the intervention groups was 0.01 standard deviations higher (0.38 lower to 0.40 higher)
EDI-Body dissatisfaction	112 (1 study)	⨁⨁⨀⨀ LOW1,4 due to risk of bias,		Not calculabl e for SMD	The mean edi-body dissatisfaction in the intervention groups was 0.33 standard deviations higher

		imprecision		values	(0.06 lower to 0.72 higher)
EDI-Drive for thinness	112 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-drive for thinness in the intervention groups was 0.74 standard deviations higher (0.15 lower to 0.64 higher)
EDI- Bulimia	112 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- bulimia in the intervention groups was 0.07 standard deviations higher (0.33 lower to 0.46 higher)
SCL-90-R Global severity index	112 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean scl-90-r global severity index in the intervention groups was 0.07 standard deviations lower (0.46 lower to 0.32 higher)
Remission ITT	112 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 1.28 (0.87 to 1.89)	434 per 1000	122 more per 1000 (from 56 fewer to 386 more)
Purging FU	87 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean purging fu in the intervention groups was 0.23 standard deviations lower (0.68 lower to 0.22 higher)
Bingeing FU	87 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.05 standard deviations lower (0.50 lower to 0.40 higher)
EDE-Weight concern FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern fu in the intervention groups was 0.24 standard deviations higher (0.20 lower to 0.69 higher)
EDE-Shape concern FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4		Not calculabl	The mean ede-shape concern fu in the intervention groups was

		due to risk of bias, imprecision		e for SMD values	0.34 standard deviations higher (0.11 lower to 0.78 higher)
EDE-Eating concern FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern fu in the intervention groups was 0.16 standard deviations higher (0.29 lower to 0.60 higher)
EDE- Restraint FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint fu in the intervention groups was 0.14 standard deviations higher (0.58 lower to 0.57 higher)
EDI-Body dissatisfaction FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-body dissatisfaction fu in the intervention groups was 0.32 standard deviations higher (0.13 lower to 0.76 higher)
EDI-Drive for thinness FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-drive for thinness fu in the intervention groups was 0.32 standard deviations higher (0.13 lower to 0.77 higher)
EDI- Bulimia FU	87 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi- bulimia fu in the intervention groups was 0.29 standard deviations higher (0.16 lower to 0.74 higher)
SCL-90-R Global severity index FU	87 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean scl-90-r global severity index fu in the intervention groups was 0.00 standard deviations higher (0.64 lower to 0.64 higher)
Remission IT FU	112 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 1.89 (1.45 to 2.46)	473 per 1000	421 more per 1000 (from 213 more to 691 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear how randomisation was conducted or if allocation concealment was performed. Assessors were blind but it was unclear if investigators or participants were blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (-0.5).
3 For a continuous outcome there were fewer than 400 participants.
4 95% CI crossed 1 MID (0.5).
5 95% CI Crossed 1 MID (1.25)

8.2.2.51 Group therapy

2 **Table 251: Summary table of findings for group mindfulness versus another group intervention in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other Group	Risk difference with BED Group Mindfulness (95% CI)
BMI	103 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.07 standard deviations higher (0.32 lower to 0.45 higher)
Binge eating days	103 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating days in the intervention groups was 0.06 standard deviations lower (0.45 lower to 0.32 higher)
Depression	103 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.1 standard deviations lower (0.49 lower to 0.29 higher)
BMI FU	103 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.12 standard deviations higher (0.26 lower to 0.51 higher)
Depression FU	103 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.06 standard deviations lower

		imprecision			(0.45 lower to 0.32 higher)
Binge eating days FU	103 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating days fu in the intervention groups was 0.26 standard deviations lower (0.64 lower to 0.13 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Unclear if allocation concealment was performed. Participants were not blind, and it was unclear if investigators and assessors were blind. High dropouts were reported >20%. 2 For a continuous outcome, there were fewer than 400 participants. 3 95% CI crossed 1 MID (0.5). 4 95% CI crossed 1 MID (-0.5).					

1 **Table 252: Summary table of group mindfulness compared to wait list controls in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Wait list control	Risk difference with BED Group Mindfulness (95% CI)
Binge eating days	100 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating days in the intervention groups was 1.08 standard deviations lower (1.5 to 0.66 lower)
Depression	100 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.85 standard deviations lower (1.26 to 0.44 lower)
BMI	100 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.19 standard deviations higher (0.2 lower to 0.59 higher)
Binge eating scale	100 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 1.24 standard deviations lower (1.67 to 0.81 lower)

Binge eating days FU	100 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating days fu in the intervention groups was 1.02 standard deviations lower (1.44 to 0.6 lower)
Depression FU	100 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.44 standard deviations lower (0.83 to 0.04 lower)
BMI FU	100 (1 study)	⨁⨁⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.2 standard deviations higher (0.19 lower to 0.59 higher)
Binge eating scale FU	100 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale fu in the intervention groups was 1.39 standard deviations lower (1.83 to 0.95 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Unclear if allocation concealment was performed. Participants were not blind, and it was unclear if investigators and assessors were blind. High dropouts were reported >20%.
2 For a continuous outcome, there were fewer than 400 participants.
3 95% CI crossed 1 MID (-0.5).
4 95% CI crossed 1 MID (0.5).

1 **Table 253: Summary table of group CBT (ED) compared to another intervention in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BED Group CBT (ED) (95% CI)
Weight	530 (6 studies)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean weight in the intervention groups was 0.23 standard deviations higher (0.03 lower to 0.49 higher)
Bingeing	795 (9 studies)	⨁⨁⨁⊖ MODERATE1		Not calculable	The mean bingeing in the intervention groups was

		due to risk of bias		for SMD values	0.13 standard deviations lower (0.27 lower to 0.01 higher)
Depression	588 (7 studies)	⊕⊕⊕⊖ MODERATE1 due to risk of bias		Not calculable for SMD values	The mean depression in the intervention groups was 0.03 standard deviations higher (0.13 lower to 0.19 higher)
Anxiety	53 (1 study)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean anxiety in the intervention groups was 0.13 standard deviations lower (0.69 lower to 0.42 higher)
EDE Global clinical score	266 (2 studies)	⊕⊖⊖⊖ VERY LOW1,6,7 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede global clinical score in the intervention groups was 1.08 standard deviations higher (0.79 to 1.37 higher)
EDE- Shape concerns	241 (3 studies)	⊕⊖⊖⊖ VERY LOW2,8,9 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- shape concerns in the intervention groups was 0.14 standard deviations lower (0.4 lower to 0.11 higher)
EDE-Dietary restraint	384 (4 studies)	⊕⊖⊖⊖ VERY LOW6,8,9 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-dietary restraint in the intervention groups was 0.02 standard deviations higher (0.19 lower to 0.22 higher)
EDE-Weight concern	384 (4 studies)	⊕⊖⊖⊖ VERY LOW2,8,9 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-weight concern in the intervention groups was 0.19 standard deviations lower (0.39 lower to 0.02 higher)
EDE-Eating concern	384 (4 studies)	⊕⊖⊖⊖ VERY LOW6,8,9 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-eating concern in the intervention groups was 0.18 standard deviations higher (0.03 lower to 0.38 higher)
Global symptom score	158 (1 study)	⊕⊕⊖⊖ LOW9,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean global symptom score in the intervention groups was 0.06 standard deviations higher (0.25 lower to 0.37 higher)
Remission_ITT	404 (4 studies)	⊕⊕⊖⊖ LOW1,11 due to risk of bias, imprecision	RR 1.22 (1.03 to 1.45)	502 per 1000	111 more per 1000 (from 15 more to 226 more)

Weight FU	514 (6 studies)	⊕⊕⊕⊖ MODERATE1 due to risk of bias		Not calculable for SMD values	The mean weight fu in the intervention groups was 0.09 standard deviations higher (0.08 lower to 0.27 higher)
Bingeing FU	651 (7 studies)	⊕⊕⊕⊖ MODERATE1 due to risk of bias		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.03 standard deviations lower (0.19 lower to 0.12 higher)
Depression FU	587 (6 studies)	⊕⊕⊕⊖ MODERATE1 due to risk of bias		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.04 standard deviations higher (0.13 lower to 0.2 higher)
Anxiety FU	185 (2 studies)	⊕⊖⊖⊖ VERY LOW5,6,8 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean anxiety fu in the intervention groups was 0.86 standard deviations higher (0.55 to 1.17 higher)
EDE Global clinical score FU	266 (3 studies)	⊕⊖⊖⊖ VERY LOW1,3,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede global clinical score fu in the intervention groups was 1.01 standard deviations higher (0.73 to 1.3 higher)
EDE-Dietary restraint FU	350 (4 studies)	⊕⊕⊖⊖ LOW2,8 due to risk of bias, inconsistency		Not calculable for SMD values	The mean ede-dietary restraint fu in the intervention groups was 0.16 standard deviations higher (0.05 lower to 0.37 higher)
EDE- Shape concerns FU	350 (4 studies)	⊕⊖⊖⊖ VERY LOW3,6,8 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- shape concerns fu in the intervention groups was 0.74 standard deviations higher (0.5 to 0.98 higher)
EDE-Weight concern FU	540 (5 studies)	⊕⊖⊖⊖ VERY LOW6,8 due to risk of bias, inconsistency		Not calculable for SMD values	The mean ede-weight concern fu in the intervention groups was 0.24 standard deviations higher (0.05 to 0.43 higher)
EDE-Eating concern FU	540 (5 studies)	⊕⊖⊖⊖ VERY LOW3,6,8 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-eating concern fu in the intervention groups was 0.26 standard deviations higher (0.08 to 0.45 higher)
Global symptom index	138	⊕⊕⊖⊖		Not	The mean global symptom index fu in the

FU	(1 study)	LOW9,10 due to risk of bias, imprecision		calculable for SMD values	intervention groups was 0.13 standard deviations higher (0.2 lower to 0.47 higher)
Remission FU_ITT	279 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,12 due to risk of bias, inconsistency, imprecision	RR 1.25 (0.85 to 1.85)	549 per 1000	137 more per 1000 (from 82 fewer to 467 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Across studies, in some or all studies, it was unclear what methods were used for randomisation or if allocation concealment was performed. Across studies, in some or all, it was unclear if participants, investigators, and assessors were blind. High dropout rates were detected >20%.
2 Heterogeneity was detected I2 >50%
3 95% CI crossed 1 MID (0.5)
4 Unclear what methods were used for randomisation or if allocation concealment was performed. Neither the participants nor investigators were blind. The assessors were not blinded. High dropouts were reported >20%.
5 95% CI crossed 1 MID (-0.5)
6 Heterogeneity was detected I2 >80%.
7 95% CI crossed 2 MIDs (-0.5 and 0.5)
8 Across studies, in some or all studies, it was unclear what methods were used for randomisation or if allocation concealment was performed. Across studies, in some or all, it was unclear if participants, investigators, and assessors were blind. One study by Musch the assessors were blind. High dropout rates were detected >20%.
9 For a continuous outcome, there were fewer than 400 participants.
10 Unclear what methods were used for randomisation or if allocation concealment was performed. Neither the participants nor investigators were blind. The assessors were not blinded.
11 For a dichotomous outcomes, there were fewer than 300 events.
12 95% CI crossed 1 MID (1.25)

1 **Table 254: Summary table of group CBT (ED) compared with wait list controls in adults with BED.**

Outcomes	**No of Participants (studies) Follow up**	**Quality of the evidence (GRADE)**	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Wait list control	Risk difference with BED Group CBT (ED) (95% CI)
Weight (BMI)	181 (3 studies)	⊕⊕⊖⊖ LOW1 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (bmi) in the intervention groups was 0.14 standard deviations higher (0.15 lower to 0.43 higher)

Binge eating days	141 (2 studies)	⨁⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean binge eating days in the intervention groups was 0.36 standard deviations lower (1.45 lower to 0.72 higher)
Depression	160 (3 studies)	⨁⊖⊖⊖ VERY LOW2,3,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.19 standard deviations higher (0.5 lower to 0.11 higher)
BMI-FU	130 (2 studies)	⨁⨁⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi-fu in the intervention groups was 0.12 standard deviations higher (0.22 lower to 0.47 higher)
Depression FU	137 (2 studies)	⨁⊖⊖⊖ VERY LOW3,4,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.04 standard deviations higher (1.06 lower to 1.15 higher)
Binge eating days FU	130 (2 studies)	⨁⨁⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating days fu in the intervention groups was 0.62 standard deviations lower (0.97 to 0.26 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; FU: Follow up

1 For a continuous outcome, there were fewer than 400 participants.
2 It was unclear if allocation concealment was performed. Across the studies, either the participants, investigators and assessors were not blinded or it was unclear. High drop outs were reported >20% and greater than 10% difference in drop outs were detected between the two groups.
3 Heterogeneity was detected, I2 >80%
4 95% CI crossed 2 MIDs (-0.5 and 0.5)
5 95% CI crossed 1 MID (-0.5)
6 It was unclear if allocation concealment was performed. The participants were not blind, however, it was unclear if the investigators and assessors were blinded. High drop outs were reported >20%.

1 **Table 255: Summary table of group behavioural therapy (ED) compared with wait list controls in adults with BED.**

Outcomes	No of Participants	Quality of the evidence (GRADE)	Relative effect	Anticipated absolute effects

	(studies) Follow up		(95% CI)	Risk with WLC	Risk difference with BED Group BT(ED) (95% CI)
Bingeing frequency	72 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing frequency in the intervention groups was 0.24 standard deviations lower (0.7 lower to 0.23 higher)
EDE- Total	72 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- total in the intervention groups was 0.1 standard deviations higher (0.37 lower to 0.56 higher)
Anxiety	72 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean anxiety in the intervention groups was 0.03 standard deviations lower (0.49 lower to 0.44 higher)
Depression	72 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.5 standard deviations lower (0.97 to 0.03 lower)
Remission_ITT	100 (1 study)	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision	RR 1.00 (0.46 to 2.19)	200 per 1000	0 fewer per 1000 (from 108 fewer to 238 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was performed, Neither the participants, investigators or assessors were blind.
2 95% CI crossed 1 MID (-0.5).
3 95% CI crossed 1 MID (0.5).
4 For a continuous outcome, there were fewer than 400 participants.
5 95% CI crossed 2 MIDs (0.75 and 1.25).

1 **Table 256: Summary table of group behavioural therapy (ED) compared with another group intervention in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with BED Group BT (ED)

				Other Group	(95% CI)
Depression	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.21 standard deviations lower (0.61 lower to 0.19 higher)
BMI	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.18 standard deviations lower (0.58 lower to 0.22 higher)
Weight loss (pounds)	98 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss (pounds) in the intervention groups was 0.18 standard deviations higher (0.22 lower to 0.57 higher)
Remission_ITT	101 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 1.81 (1.18 to 2.78)	353 per 1000	286 more per 1000 (from 64 more to 628 more)
EDE-Eating concern	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern in the intervention groups was 0.54 standard deviations lower (0.95 to 0.14 lower)
EDE-Dietary restraint	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-dietary restraint in the intervention groups was 0.54 standard deviations lower (0.94 to 0.14 lower)
EDE- Shape concerns	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concerns in the intervention groups was 0.32 standard deviations lower (0.72 lower to 0.07 higher)
EDE-Weight concern	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern in the intervention groups was 0.38 standard deviations lower (0.78 lower to 0.02 higher)
BMI FU	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.14 standard deviations lower (0.56 lower to 0.28 higher)

Weight loss (pounds) FU	88 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss (pounds) fu in the intervention groups was 0.05 standard deviations higher (0.37 lower to 0.47 higher)
Depression FU	88 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.03 standard deviations higher (0.39 lower to 0.46 higher)
EDE-Dietary restraint FU	88 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-dietary restraint fu in the intervention groups was 0.6 standard deviations lower (1.03 to 0.17 lower)
EDE-Weight concern FU	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern fu in the intervention groups was 0.4 standard deviations lower (0.82 lower to 0.03 higher)
EDE- Shape concerns FU	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concerns fu in the intervention groups was 0.12 standard deviations lower (0.54 lower to 0.3 higher)
EDE-Eating concern FU	88 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern fu in the intervention groups was 0.18 standard deviations higher (0.24 lower to 0.6 higher)
Remission_ITT FU	101 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR 1.44 (0.98 to 2.11)	431 per 1000	190 more per 1000 (from 9 fewer to 479 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear methods for randomisation or if allocation concealment was performed. It was unclear if participants and investigators were blind, however, assessors were blind. High dropouts were reported >20% and a greater than 10% difference in dropout rates were detected between the two groups.
2 95% CI crossed 1 MID (-0.5).
3 95% CI crossed 1 MID (0.5).
4 For a dichotomous outcome, there were fewer than 300 events.

5 For a continuous outcome, there were fewer than 400 participants.
6 95% CI crossed 1 MID (1.25).

Table 257: Summary table of group CBT-ED (body exposure) compared with CBT-ED (cognitive) in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT (cognitive).	Risk difference with BED CBT (body exposure). (95% CI)
EDE- Restraint	24 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0 standard deviations higher (0.8 lower to 0.8 higher)
EDE- Eating concern	24 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.41 standard deviations lower (1.22 lower to 0.4 higher)
EDE- Weight concern	24 (1 study)	⨁⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0 standard deviations higher (0.8 lower to 0.8 higher)
EDE- Shape concern	24 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.19 standard deviations higher (0.62 lower to 0.99 higher)
BMI	24 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi in the intervention groups was 0.38 standard deviations lower (1.19 lower to 0.43 higher)
Depression	24 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.01 standard deviations higher (0.79 lower to 0.81 higher)
Bingeing episodes	24 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing episodes in the intervention groups was 0.27 standard deviations lower (1.07 lower to 0.53 higher)

Remission_ITT	28 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 0.44 (0.18 to 1.11)	643 per 1000	360 fewer per 1000 (from 527 fewer to 71 more)
EDE- Restraint FU	24 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint fu in the intervention groups was 0.08 standard deviations lower (0.88 lower to 0.72 higher)
EDE- Eating concern FU	24 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern fu in the intervention groups was 0 standard deviations higher (0.8 lower to 0.8 higher)
EDE- Weight concern FU	24 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern fu in the intervention groups was 0.18 standard deviations higher (0.62 lower to 0.98 higher)
EDE- Shape concern FU	24 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern fu in the intervention groups was 0.45 standard deviations higher (0.37 lower to 1.26 higher)
BMI FU	24 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi fu in the intervention groups was 0.25 standard deviations lower (1.05 lower to 0.56 higher)
Depression FU	24 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.2 standard deviations higher (0.61 lower to 1 higher)
Bingeing episodes FU	24 (1 study)	⊕⊖⊖⊖ VERY LOW1,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing episodes fu in the intervention groups was 0.43 standard deviations higher (0.38 lower to 1.24 higher)
Remission_ITT FU	28 (1 study)	⊕⊖⊖⊖ VERY LOW1,7 due to risk of bias, imprecision	RR 0.75 (0.35 to 1.6)	571 per 1000	143 fewer per 1000 (from 371 fewer to 343 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%					

confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 There was unclear methods of randomisation and allocation concealment. The participants, investigators and assessors were not blinded.
2 95% CI crossed 2 MIDs (-0.5 and 0.5)
3 95% CI crossed 1 MID (-0.5)
4 For a continuous outcome, there were fewer than 400 participants.
5 95% CI crossed 1 MID (0.75)
6 95% CI crossed 1 MID (0.5)
7 95% CI crossed 2 MIDs (0.75 and 1.25)

Table 258: Summary table of group interpersonal psychotherapy (IPT) compared with another intervention in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BED Group IPT (ED) (95% CI)
Bingeing	158 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.16 standard deviations higher (0.15 lower to 0.48 higher)
Remission_ITT	162 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 0.92 (0.77 to 1.1)	790 per 1000	63 fewer per 1000 (from 182 fewer to 79 more)
Depression	194 (2 studies)	⊕⊕⊖⊖ LOW4,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.22 standard deviations lower (0.5 lower to 0.06 higher)
EDE-Restraint	158 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-restraint in the intervention groups was 0.59 standard deviations higher (0.27 to 0.91 higher)
EDE-Shape concern	158 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concern in the intervention groups was 0.08 standard deviations higher (0.23 lower to 0.39 higher)

EDE-Eating concern	158 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern in the intervention groups was 0.12 standard deviations higher (0.19 lower to 0.44 higher)
EDE-Weight concern	158 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern in the intervention groups was 0.08 standard deviations higher (0.23 lower to 0.39 higher)
Global symptom index	158 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean global symptom index in the intervention groups was 0.06 standard deviations lower (0.37 lower to 0.25 higher)
BMI	158 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.06 standard deviations lower (0.37 lower to 0.26 higher)
Bingeing FU	138 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.14 standard deviations lower (0.48 lower to 0.19 higher)
EDE-Restraint FU	138 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-restraint fu in the intervention groups was 0.25 standard deviations higher (0.09 lower to 0.58 higher)
EDE-Shape concern FU	138 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concern fu in the intervention groups was 0 standard deviations higher (0.33 lower to 0.33 higher)
EDE-Eating concern FU	138 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern fu in the intervention groups was 0 standard deviations higher (0.33 lower to 0.33 higher)
EDE-Weight concern FU	138 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern fu in the intervention groups was 0 standard deviations higher (0.33 lower to 0.33 higher)
Global symptom index	138	⊕⊕⊖⊖		Not	The mean global symptom index fu in the

FU	(1 study)	LOW1,2 due to risk of bias, imprecision		calculable for SMD values	intervention groups was 0.13 standard deviations lower (0.47 lower to 0.2 higher)
Remission FU_ITT	162 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 1.04 (0.81 to 1.34)	593 per 1000	24 more per 1000 (from 113 fewer to 201 more)
Depression FU	138 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.1 standard deviations lower (0.43 lower to 0.24 higher)
BMI FU	138 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.17 standard deviations lower (0.5 lower to 0.16 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 There were unclear methods for randomisation and if allocation concealment was performed. It was unclear if participants, investigators and assessors were blind.
2 For a continuous outcome, there were fewer than 400 participants.
3 For a dichotomous outcome, there were fewer than 300 events.
4 There were unclear methods for randomisation and if allocation concealment was performed. The participants, investigators and assessors were either not blinded or it was unclear if they were. High dropouts were detected in Wilfley 1993 >20% and high difference in dropouts between the two groups >10%.
5 95% CI crossed 1 MID (-0.5).
6 95% CI crossed 1 MID (0.5).

1 **Table 259: Summary table of group counselling compared with another intervention in adults with BED at the end of treatment.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention	Risk difference with BED Group Counselling (95% CI)
BMI	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean BMI in the intervention groups was 0.14 standard deviations higher

		imprecision			(0.28 lower to 0.56 higher)
EDE - Dietary restraint	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - dietary restraint in the intervention groups was 0.54 standard deviations higher (0.14 to 0.94 higher)
EDE- Shape concerns	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concerns in the intervention groups was 0.32 standard deviations higher (0.07 lower to 0.72 higher)
EDE- Weight concerns	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concerns in the intervention groups was 0.38 standard deviations higher (0.02 lower to 0.78 higher)
EDE - Eating concerns	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - eating concerns in the intervention groups was 0.54 standard deviations higher (0.14 to 0.95 higher)
Remission_ITT	101 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 8.33 (2.03 to 34.21)	40 per 1000	293 more per 1000 (from 41 more to 1000 more)
Depression	98 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.21 standard deviations higher (0.19 lower to 0.61 higher)
Weight loss	98 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss in the intervention groups was 0.18 standard deviations lower (0.57 lower to 0.22 higher)
Patient's preference for treatment	98 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean patient's preference for treatment in the intervention groups was 0.37 standard deviations lower (0.77 lower to 0.03 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 There were unclear methods for randomisation and if allocation concealment was performed. It was unclear if participants and investigators were blind, but the assessors were blind. High dropouts were reported in one arm >20% and a greater than 10% difference was detected for dropouts between the two groups.
2 95% CI crossed 1 MID (0.5).
3 95% CI crossed 1 MID (-0.5)
4. Fewer than 300 events

1 **Table 260: Summary table of group counselling compared with another intervention in adults with BED at follow up.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention FU	Risk difference with BED Group Counselling (95% CI)
BMI FU	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.14 standard deviations higher (0.28 lower to 0.56 higher)
Depression FU	88 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.03 standard deviations lower (0.46 lower to 0.39 higher)
EDE - Dietary restraint FU	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - dietary restraint fu in the intervention groups was 0.6 standard deviations higher (0.17 to 1.03 higher)
EDE- Shape concerns FU	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concerns fu in the intervention groups was 0.12 standard deviations higher (0.3 lower to 0.54 higher)
EDE- Weight concerns FU	88 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concerns fu in the intervention groups was 0.4 standard deviations higher (0.03 lower to 0.82 higher)
EDE - Eating concerns FU	88 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede - eating concerns fu in the intervention groups was 0.18 standard deviations lower (0.6 lower to 0.24 higher)
Remission_ITT FU	101	⊕⊕⊖⊖	RR 0.70	620 per 1000	186 fewer per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with another intervention FU	Risk difference with BED Group Counselling (95% CI)
	(1 study)	LOW1,5 due to risk of bias, imprecision	(0.47 to 1.02)		(from 329 fewer to 12 more)
Weight loss FU	98 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss fu in the intervention groups was 0.18 standard deviations lower (0.57 lower to 0.22 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 There were unclear methods for randomisation and if allocation concealment was performed. It was unclear if participants and investigators were blind, but the assessors were blind. High dropouts were reported in one arm >20% and a greater than 10% difference was detected for dropouts between the two groups.
2 95% CI crossed 1 MID (0.5).
3 For a continuous outcome, there were fewer than 400 participants.
4 95% CI crossed 1 MID (-0.5).
5 95% CI crossed 1 MID (0.75).

1 **Table 261: Summary table of group diet compared with another group intervention in adults with BED**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other Group	Risk difference with BED Group Diet (95% CI)
Weight	242 (3 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight in the intervention groups was 0.54 standard deviations lower (0.81 to 0.28 lower)
Bingeing	241 (3 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.24 standard deviations higher (0.02 lower to 0.5 higher)

EDE- Shape concern	85 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.26 standard deviations higher (0.17 lower to 0.7 higher)
EDE- Weight concern	85 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.19 standard deviations higher (0.24 lower to 0.63 higher)
EDE-Eating concern	85 (2 studies)	⊕⊕⊖⊖ LOW1,3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern in the intervention groups was 0.26 standard deviations higher (0.17 lower to 0.7 higher)
EDE- Restraint	85 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.14 standard deviations higher (0.29 lower to 0.57 higher)
Depression	327 (4 studies)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.19 standard deviations higher (0.03 lower to 0.42 higher)
Global EDE	125 (1 study)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean global ede in the intervention groups was 0.17 standard deviations higher (0.2 lower to 0.54 higher)
Remission_ITT	242 (3 studies)	⊕⊕⊖⊖ LOW1,8 due to risk of bias, imprecision	RR 0.64 (0.46 to 0.88)	503 per 1000	181 fewer per 1000 (from 60 fewer to 272 fewer)
Weight FU	229 (3 studies)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight fu in the intervention groups was 0.17 standard deviations lower (0.44 lower to 0.1 higher)
Bingeing FU	241 (3 studies)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.21 standard deviations higher (0.05 lower to 0.47 higher)
EDE- Shape concern FU	71	⊕⊕⊖⊖		Not	The mean ede- shape concern fu in the

	(2 studies)	LOW1,2 due to risk of bias, imprecision		calculable for SMD values	intervention groups was 0.03 standard deviations lower (0.5 lower to 0.44 higher)
EDE- Weight concern FU	71 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern fu in the intervention groups was 0.11 standard deviations higher (0.36 lower to 0.59 higher)
EDE-Eating concern FU	71 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern fu in the intervention groups was 0.06 standard deviations lower (0.53 lower to 0.41 higher)
EDE- Restraint FU	71 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint fu in the intervention groups was 0.16 standard deviations lower (0.63 lower to 0.3 higher)
Global EDE FU	125 (1 study)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean global ede fu in the intervention groups was 0.17 standard deviations higher (0.19 lower to 0.54 higher)
Depression FU	205 (3 studies)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.03 standard deviations lower (0.32 lower to 0.25 higher)
Remission-ITT FU	117 (2 studies)	⊕⊕⊖⊖ LOW1,8 due to risk of bias, imprecision	RR 0.67 (0.47 to 0.95)	662 per 1000	218 fewer per 1000 (from 33 fewer to 351 fewer)
EDE- Shape concern < 18 binges per month	48 (1 study)	⊕⊕⊖⊖ LOW2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern < 18 binges per month in the intervention groups was 0.13 standard deviations lower (0.69 lower to 0.44 higher)
EDE- Shape concern > 18 binges per month	37 (1 study)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern > 18 binges per month in the intervention groups was 0.83 standard deviations higher (0.15 to 1.51 higher)
EDE- Restraint <18 binges per month	48 (1 study)	⊕⊕⊖⊖ LOW1,2		Not calculable	The mean ede- restraint <18 binges per month in the intervention groups was

		due to risk of bias, imprecision		for SMD values	0.29 standard deviations lower (0.86 lower to 0.28 higher)
EDE- Restraint > 18 binges per month	37 (1 study)	⊕⊕⊖⊖ LOW3,4,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint > 18 binges per month in the intervention groups was 0.90 standard deviations higher (0.21 to 1.58 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Across studies it was unclear in somehow randomisation was performed and in all studies if allocation concealment was performed. Across the studies, either it was unclear of the participants, investigators or assessors were not blinded. Only in Munsch 2007 were the assessors blind. High dropout rates were detected >20%.
2 95% CI crossed 1 MID (-0.5)
3 95% CI crossed 1 MID (0.5)
4 Heterogeneity was detected I2 >50%
5 Heterogeneity was detected I2 >80%
6 For a continuous outcome, there were fewer than 400 participants.
7 It was unclear how randomisation was performed and if allocation concealment was performed. The participants were not blinded, and it was unclear if investigators and assessors were blinded. High dropout rates were detected >20%.
8 95% CI crossed 1 MID (0.75)

1 **Table 262: Summary table of group self-help (ED) compared with another group in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other Group	Risk difference with BED Group SH(ED) (95% CI)
BMI	234 (2 studies)	⊕⊕⊖⊖ LOW1,2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.19 standard deviations lower (0.46 lower to 0.08 higher)
Bingeing	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.30 standard deviations higher (0.01 to 0.6 higher)
Depression	44	⊕⊕⊖⊖ LOW1,4		Not calculable	The mean depression in the intervention groups was

	(1 study)	due to risk of bias, imprecision		for SMD values	0.23 standard deviations higher (0.43 lower to 0.89 higher)
EDE Q Global Score	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q global score in the intervention groups was 0.33 standard deviations higher (0.03 to 0.62 higher)
EDE Q Restraint	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q restraint in the intervention groups was 0.46 standard deviations higher (0.16 to 0.76 higher)
EDE Q Eating Concern	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q eating concern in the intervention groups was 0.31 standard deviations higher (0.01 to 0.6 higher)
EDE Q Shape Concern	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q shape concern in the intervention groups was 0.22 standard deviations higher (0.08 lower to 0.52 higher)
EDE Q Weight Concern	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q weight concern in the intervention groups was 0.27 standard deviations higher (0.03 lower to 0.57 higher)
Quality of life	167 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.00 standard deviations lower (0.32 lower to 0.32 higher)
Remission_ITT	51 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 2.83 (1.29 to 6.23)	200 per 1000	366 more per 1000 (from 58 more to 1000 more)
BMI FU	231 (2 studies)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.08 standard deviations lower (0.35 lower to 0.2 higher)
Bingeing FU	190 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias,		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.10 standard deviations lower

		imprecision			(0.4 lower to 0.19 higher)
Depression FU	44 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.23 standard deviations higher (0.43 lower to 0.89 higher)
EDE Q Restraint FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q restraint fu in the intervention groups was 0.46 standard deviations higher (0.16 to 0.76 higher)
EDE Q Eating Concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q eating concern fu in the intervention groups was 0.08 standard deviations lower (0.38 lower to 0.22 higher)
EDE Q Shape Concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q shape concern fu in the intervention groups was 0.07 standard deviations higher (0.23 lower to 0.37 higher)
EDE Q Weight Concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q weight concern fu in the intervention groups was 0.07 standard deviations higher (0.23 lower to 0.37 higher)
EDE Q Global Score FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q global score fu in the intervention groups was 0.06 standard deviations higher (0.24 lower to 0.35 higher)
Quality of life FU	167 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life fu in the intervention groups was 0.02 standard deviations higher (0.3 lower to 0.34 higher)
Remission_ITT FU	51 (1 study)	⊕⊖⊖⊖ VERY LOW1,7 due to risk of bias, imprecision	RR 0.67 (0.22 to 2.09)	286 per 1000	94 fewer per 1000 (from 223 fewer to 311 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear how they generated random sequence for randomisation and if allocation concealment was performed. It is unclear if either the participants, investigators or assessors were blind.
2 95% CI crossed 1 MID (-0.5).
3 For a continuous outcome, there were fewer than 400 participants.
4 95% CI crossed 1 MID (0.5).
5 For a dichotomous outcome, there were fewer than 300 events.
6 95% CI crossed 2 MIDs (-0.5 and 0.5).
7 95% CI crossed 2 MIDs (0.75 and 1.25).

1 **Table 263: Summary table of group guided self-help (ED) compared with other group in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other Group	Risk difference with BED Group Guided SH(ED) (95% CI)
BMI	234 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.16 standard deviations higher (0.11 lower to 0.44 higher)
Bingeing	183 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.35 standard deviations lower (0.66 to 0.04 lower)
Depression	44 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.53 standard deviations lower (1.15 lower to 0.09 higher)
EDE Q Global Score	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q global score in the intervention groups was 0.07 standard deviations higher (0.24 lower to 0.38 higher)
EDE Q Restraint	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q restraint in the intervention groups was 0.22 standard deviations lower (0.52 lower to 0.09 higher)
EDE Q Eating concern	190 (1 study)	⊕⊕⊖⊖ LOW1,2		Not calculable for SMD values	The mean ede q eating concern in the intervention groups was

		due to risk of bias, imprecision			0.08 standard deviations lower (0.39 lower to 0.22 higher)
EDE Q Weight concern	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q weight concern in the intervention groups was 0.26 standard deviations higher (0.05 lower to 0.57 higher)
EDE Q Shape concern	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q shape concern in the intervention groups was 0.09 standard deviations higher (0.21 lower to 0.4 higher)
Quality of life	176 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.01 standard deviations higher (0.31 lower to 0.32 higher)
Remission_ITT	51 (1 study)	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision	RR 0.57 (0.21 to 1.52)	375 per 1000	161 fewer per 1000 (from 296 fewer to 195 more)
BMI FU	231 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.02 standard deviations lower (0.29 lower to 0.26 higher)
Bingeing FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.23 standard deviations higher (0.02 lower to 0.48 higher)
Depression FU	41 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.49 standard deviations lower (1.13 lower to 0.14 higher)
EDE Q Global Score FU	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q global score fu in the intervention groups was 0.40 standard deviations lower (0.71 to 0.09 lower)
EDE Q Restraint FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias,		Not calculable for SMD values	The mean ede q restraint fu in the intervention groups was 0.21 standard deviations higher

		imprecision			(0.1 lower to 0.52 higher)
EDE Q Eating concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q eating concern fu in the intervention groups was 0.29 standard deviations higher (0.02 lower to 0.6 higher)
EDE Q Weight concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q weight concern fu in the intervention groups was 0.07 standard deviations higher (0.23 lower to 0.37 higher)
EDE Q Shape concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede q shape concern fu in the intervention groups was 0.42 standard deviations higher (0.11 to 0.73 higher)
Quality of life FU	167 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life fu in the intervention groups was 0.01 standard deviations higher (0.31 lower to 0.33 higher)
Remission_ITT FU	51 (1 study)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR 1.97 (0.78 to 4.99)	188 per 1000	182 more per 1000 (from 41 fewer to 748 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear how they generated random sequence for randomisation and if allocation concealment was performed. It is unclear if either the participants, investigators or assessors were blind.
2 For a continuous outcome, there were fewer than 400 participants.
3 95% CI crossed 1 MIDs (-0.5).
4 95% CI crossed 1 MIDs (0.5).
5 95% CI crossed 2 MIDs (0.75 and 1.25).
6 95% CI crossed 1 MIDs (1.25).

Table 264: Summary table of group self-help (ED) compared with wait list controls in adults with BED.

Outcomes	No of Participants	Quality of the evidence (GRADE)	Relative effect	Anticipated absolute effects

	(studies) Follow up		(95% CI)	Risk with WLC	Risk difference with BED Group SH (ED) (95% CI)
BMI	136 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.09 standard deviations higher (0.25 lower to 0.42 higher)
Bingeing	136 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.41 standard deviations lower (0.75 lower to 0.07 higher)
EDE-Q Global Score	136 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global score in the intervention groups was 0.00 standard deviations higher (0.34 lower to 0.34 higher)
EDE-Q Restraint	136 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q restraint in the intervention groups was 0.08 standard deviations higher (0.26 lower to 0.42 higher)
EDE-Q Eating concern	136 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concern in the intervention groups was 0.09 standard deviations higher (0.25 lower to 0.42 higher)
EDE-Q Shape concern	136 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concern in the intervention groups was 0.00 standard deviations higher (0.34 lower to 0.34 higher)
EDE-Q Weight concern	136 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concern in the intervention groups was 0.00 standard deviations higher (0.34 lower to 0.34 higher)
Quality of life	136 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.08 standard deviations higher (0.27 lower to 0.45 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Unclear how they generated random sequence for randomisation and if allocation concealment was performed. It is unclear if either the participants, investigators or assessors were blind.
2 For a continuous outcome, there were fewer than 400 participants.
3 95% CI crossed 1 MID (-0.5).

1 **Table 265: Summary table of group guided self-help (ED) compared with wait list controls in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BED Group Guided SH (ED) (95% CI)
BMI	129 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.26 standard deviations higher (0.09 lower to 0.61 higher)
Bingeing	129 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.83 standard deviations lower (1.19 to 0.47 lower)
EDE-Q Global Score	129 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global score in the intervention groups was 0.22 standard deviations lower (0.57 lower to 0.13 higher)
EDE-Q Restraint	129 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q restraint in the intervention groups was 0.34 standard deviations lower (0.69 to 0.01 lower)
EDE-Q Eating concern	129 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concern in the intervention groups was 0.18 standard deviations lower (0.53 lower to 0.17 higher)
EDE-Q Shape concern	129 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concern in the intervention groups was 0.09 standard deviations lower (0.43 lower to 0.26 higher)
EDE-Q Weight	129	⊕⊕⊖⊖		Not	The mean ede-q weight concern in the

concern	(1 study)	LOW1,4 due to risk of bias, imprecision		calculable for SMD values	intervention groups was 0.00 standard deviations higher (0.35 lower to 0.35 higher)
Quality of life	129 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.09 standard deviations higher (0.28 lower to 0.47 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Unclear how they generated random sequence for randomisation and if allocation concealment was performed. It is unclear if either the participants, investigators or assessors were blind.
2 95% CI crossed 1 MID (0.5).
3 95% CI crossed 1 MID (-0.5).
4 For a continuous outcome, there were fewer than 400 participants.

1 **Table 266: Summary table of group psychoeducation compared with another group in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other Group	Risk difference with BED Group Psychoeducation (95% CI)
BMI	234 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.02 standard deviations higher (0.25 lower to 0.29 higher)
Bingeing	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.05 standard deviations higher (0.25 lower to 0.35 higher)
Depression	44 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.48 standard deviations higher (0.17 lower to 1.13 higher)
EDE-Q Global Score	253 (1 study)	⊕⊕⊖⊖ LOW1,4		Not calculable for SMD	The mean ede-q global score in the intervention groups was

		due to risk of bias, imprecision		values	0.45 standard deviations lower (0.7 to 0.2 lower)
EDE-Q Restraint	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q restraint in the intervention groups was 0.22 standard deviations lower (0.52 lower to 0.09 higher)
EDE-Q Eating Concern	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concern in the intervention groups was 0.22 standard deviations lower (0.52 lower to 0.09 higher)
EDE-Q Shape Concern	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concern in the intervention groups was 0.30 standard deviations lower (0.6 lower to 0.01 higher)
EDE-Q Weight Concern	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concern in the intervention groups was 0.55 standard deviations lower (0.86 to 0.24 lower)
Did not Achieve Remission_ITT	51 (1 study)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision	RR 1.32 (0.94 to 1.85)	371 per 1000	119 more per 1000 (from 22 fewer to 316 more)
Quality of life	176 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.01 standard deviations lower (0.32 lower to 0.3 higher)
BMI FU	243 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.06 standard deviations higher (0.21 lower to 0.33 higher)
Bingeing FU	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing fu in the intervention groups was 0.03 standard deviations higher (0.27 lower to 0.34 higher)
Depression FU	41 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias,		Not calculable for SMD values	The mean depression fu in the intervention groups was 1.01 standard deviations lower

		imprecision			(1.83 to 0.18 lower)
EDE-Q Global Score FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global score fu in the intervention groups was 0.37 standard deviations lower (0.67 to 0.06 lower)
EDE-Q Restraint FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q restraint fu in the intervention groups was 0.29 standard deviations lower (0.59 lower to 0.02 higher)
EDE-Q Eating Concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concern fu in the intervention groups was 0.20 standard deviations lower (0.51 lower to 0.1 higher)
EDE-Q Shape Concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concern fu in the intervention groups was 0.37 standard deviations lower (0.68 to 0.07 lower)
EDE-Q Weight Concern FU	190 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concern fu in the intervention groups was 0.51 standard deviations lower (0.82 to 0.2 lower)
Quality of life FU	167 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life fu in the intervention groups was 0.03 standard deviations lower (0.35 lower to 0.3 higher)
Did not Achieve Remission_ITT FU	51 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.13 (0.83 to 1.55)	286 per 1000	37 more per 1000 (from 49 fewer to 157 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear how they generated random sequence for randomisation and if allocation concealment was performed. It is unclear if either the participants, investigators or assessors were blind.
2 For a continuous outcome, there were fewer than 400 participants.
3 95% CI crossed 1 MID (0.5).

4 95% CI crossed 1 MID (-0.5).
5 95% CI crossed 1 MID (1.25).

8.2.2.61 Self-help

2 **Table 267: Summary table of guided self-help (ED) (self-help with support) compared with any other intervention in adults with BED.**

Outcomes	No of Participants (studies) Follow up	**Quality of the evidence (GRADE)**	Relative effect (95% CI)	**Anticipated absolute effects**	
				Risk with Other	**Risk difference with BED Guided SH (ED) (95% CI)**
Bingeing	490 (7 studies)	⨁⨁⨁⨀ MODERATE1,2 due to risk of bias			The mean bingeing in the intervention groups was 0.28 standard deviations lower (0.47 to 0.09 lower)
Vomiting	90 (1 study)	⨁⨀⨀⨀ VERY LOW2,3,4 due to risk of bias, indirectness, imprecision			The mean vomiting in the intervention groups was 0.81 standard deviations lower (1.24 to 0.38 lower)
Use of laxatives	90 (1 study)	⨁⨀⨀⨀ VERY LOW2,3,4 due to risk of bias, indirectness, imprecision			The mean use of laxatives in the intervention groups was 0.21 standard deviations higher (0.21 lower to 0.62 higher)
BMI	690 (7 studies)	⨁⨁⨁⨀ MODERATE1 due to risk of bias			The mean bmi in the intervention groups was 0.04 standard deviations higher (0.11 lower to 0.2 higher)
Depression	394 (5 studies)	⨁⨁⨀⨀ LOW5,6 due to risk of bias, indirectness			The mean depression in the intervention groups was 0.29 standard deviations lower (0.5 to 0.08 lower)
Remission_ITT	661 (9 studies)	⨁⨀⨀⨀ VERY LOW6,7,8,9 due to risk of bias, inconsistency, imprecision	RR 1.76 (1.42 to 2.19)	242 per 1000	184 more per 1000 (from 102 more to 288 more)
EDE-Global severity	389 (4 studies)	⨁⨀⨀⨀ VERY LOW2,7,10			The mean ede-global severity in the intervention groups was 0.14 standard deviations lower

		due to risk of bias, indirectness, imprecision			(0.35 lower to 0.07 higher)
EDE- Shape concern	740 (7 studies)	⊕⊖⊖⊖ VERY LOW2,4,6,7,8 due to risk of bias, inconsistency, indirectness, imprecision			The mean ede- shape concern in the intervention groups was 0.27 standard deviations lower (0.53 to 0.02 lower)
EDE- Weight concern	740 (7 studies)	⊕⊖⊖⊖ VERY LOW2,4,6,7,8 due to risk of bias, inconsistency, indirectness, imprecision			The mean ede- weight concern in the intervention groups was 0.22 standard deviations lower (0.52 lower to 0.08 higher)
EDE- Restraint	740 (7 studies)	⊕⊖⊖⊖ VERY LOW2,4,6,7,8 due to risk of bias, inconsistency, indirectness, imprecision			The mean ede- restraint in the intervention groups was 0.37 standard deviations lower (0.6 to 0.13 lower)
EDE- Eating concern	650 (6 studies)	⊕⊖⊖⊖ VERY LOW6,7,10,11 due to risk of bias, inconsistency, indirectness, imprecision			The mean ede- eating concern in the intervention groups was 0.27 standard deviations lower (0.43 to 0.11 lower)
Excessive exercise	90 (1 study)	⊕⊖⊖⊖ VERY LOW2,3,4 due to risk of bias, indirectness, imprecision			The mean excessive exercise in the intervention groups was 0.28 standard deviations lower (0.7 lower to 0.13 higher)
Satisfaction with life	284 (2 studies)	⊕⊕⊕⊖ MODERATE12 due to risk of bias			The mean satisfaction with life in the intervention groups was 0.12 standard deviations higher (0.13 lower to 0.36 higher)
Bingeing FU	300 (4 studies)	⊕⊕⊖⊖ LOW10,13 due to risk of bias, imprecision			The mean bingeing fu in the intervention groups was 0.09 standard deviations higher (0.15 lower to 0.33 higher)
BMI FU	409 (4 studies)	⊕⊖⊖⊖ VERY LOW6,7,10 due to risk of bias, indirectness, imprecision			The mean bmi fu in the intervention groups was 0.02 standard deviations higher (0.18 lower to 0.22 higher)
EDE- Weight concern FU	368 (3 studies)	⊕⊖⊖⊖ VERY LOW6,8,10,14 due to risk of bias, inconsistency,			The mean ede- weight concern fu in the intervention groups was 0.12 standard deviations higher

		indirectness, imprecision			(0.31 lower to 0.56 higher)
EDE- Restraint FU	368 (3 studies)	⊕⊖⊖⊖ VERY LOW6,8,10,14 due to risk of bias, inconsistency, indirectness, imprecision			The mean ede- restraint fu in the intervention groups was 0.12 standard deviations lower (0.52 lower to 0.27 higher)
EDE- Shape concern FU	368 (3 studies)	⊕⊖⊖⊖ VERY LOW6,8,10,14 due to risk of bias, inconsistency, indirectness, imprecision			The mean ede- shape concern fu in the intervention groups was 0.00 standard deviations higher (0.42 lower to 0.42 higher)
EDE- Eating concern FU	368 (3 studies)	⊕⊖⊖⊖ VERY LOW6,8,10,14 due to risk of bias, inconsistency, indirectness, imprecision			The mean ede- eating concern fu in the intervention groups was 0.06 standard deviations lower (0.47 lower to 0.36 higher)
EDE-Q-Global score-FU	260 (2 studies)	⊕⊕⊖⊖ LOW10 due to risk of bias, imprecision			The mean ede-q-global score-fu in the intervention groups was 0.32 standard deviations lower (0.58 to 0.06 lower)
Remission FU_ITT	229 (3 studies)	⊕⊖⊖⊖ VERY LOW6,7,15 due to risk of bias, indirectness, imprecision	RR 1.40 (1.06 to 1.85)	374 per 1000	150 more per 1000 (from 22 more to 318 more)
Quality of life FU	167 (1 study)	⊕⊕⊖⊖ LOW10,16 due to risk of bias, imprecision			The mean quality of life fu in the intervention groups was 0.01 standard deviations higher (0.31 lower to 0.33 higher)
Depression FU	150 (2 studies)	⊕⊖⊖⊖ VERY LOW4,6,13 due to risk of bias, indirectness, imprecision			The mean depression fu in the intervention groups was 0.39 standard deviations lower (0.71 to 0.06 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

1 Across studies it was unclear if allocation concealment was conducted (or adequately). In Peterson 2001 neither the investigator nor assessor were blind and in Dunn 2005 the participants were not blind. In Grilo 2013 the assessors were blind, but it was unclear if the others were blind. In Carter, randomisation and allocation concealment was adequate, however, participants, investigators and assessors were not blind. In other studies, it was unclear if either the participants, assessors or investigators were blind. High drop outs were reported >20%.
2 Dunn 2006 included a mixed population of BN and BED
3 in Dunn 2005, no details were provided on how the random sequence was generated and it was unclear if allocation concealment was performed. The pariticipants were not blind and it was unclear if investigators or assessors were blind. High drop outs were reported >20%.
4 95% CI crossed 1 MID (-0.5).
5 In Carrard, allocation concealment was not condcuted. It was unclear in all other studies. Across studies, it was unclear if all or either the participants, assessors or investigators were blind. In Carrard, assessors were not blind, whilst in Striegel-Moore assessors were blind. High drop outs were reported >20%.
6 Striegel-Moore 2010 included a mixed population of BED (53%) and BN (47%)
7 Across studies it was unclear if allocation concealment was conducted (or adequately). It was also unclear if either or all of the participants, assessors or investigators were blind. High drop outs were reported >20%.
8 Heterogeneity was detected, I2 >50%
9 For a dichotomous outcome, there were fewer than 300 events.
10 For a continuous outcome, there are fewer than 400 participants.
11 Heterogeneity was detected, I2 >80%,
12 No details were provided on how random sequence was generated and it was unclear if allocation concealment was conducted. In Cassin, only assessors were blind, and in Peterson neither the assessors nor investigators were blind. High drop outs were detected >20%.
13 It was unclear if allocation concealment was conducted. In Peterson 2009, neither the assessors or investigators were blind, Whilst in the other study, it was unclear if either the participants, investigators or assessors were blind. High dropout rates were detected >20%.
14 It was unclear how random sequence was generated and it was unclear if allocation concealment was conducted. In Peterson, neither the assessors nor investigators were blind. Whilst in Striegel-Moore 2001, assessors were blind but it was unclear if either investigators or participants were blind. In Carter, randomisation and allocation concealment was adequate, however, participants, investigators and assessors were not blind. High dropout rates were detected in Peterson 2009.
15 95% CI crossed 1 MID (1.25).
16 No details were provided on how random sequence was generated and it was unclear if allocation concealment was conducted. Neither the assessors nor investigators were blind. High drop outs were detected >20%.

1 **Table 268: Summary table of guided self-help (ED) (self-help with support) compared with wait list controls in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	**Anticipated absolute effects**	
				Risk with WLC	**Risk difference with BED Guided SH (ED) (95% CI)**
Bingeing	218 (3 studies)	⨁⨁⊖⊖ LOW1,2		Not calculable for SMD	The mean bingeing in the intervention groups was 0.85 standard deviations lower

		due to risk of bias, imprecision		values	(1.14 to 0.56 lower)
BMI	188 (2 studies)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi in the intervention groups was 0.17 standard deviations higher (0.12 lower to 0.46 higher)
EDE- Weight concern	248 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.48 standard deviations lower (1.04 lower to 0.08 higher)
EDE- Shape concern	248 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.58 standard deviations lower (1.16 lower to 0 higher)
EDE- Restraint	252 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.43 standard deviations lower (0.96 lower to 0.11 higher)
EDE- Eating concern	248 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.90 standard deviations lower (1.83 lower to 0.03 higher)
EDE-Global	248 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-global in the intervention groups was 0.71 standard deviations lower (1.34 to 0.08 lower)
Quality of life	109 (1 study)	⊕⊕⊖⊖ LOW7,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.09 standard deviations higher (0.28 lower to 0.47 higher)
Did not achieve Remission	59 (1 study)	⊕⊕⊖⊖ LOW9,10 due to risk of bias, imprecision	RR 0.54 (0.38 to 0.78)	80 per 1000	37 fewer per 1000 (from 18 fewer to 50 fewer)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 It was unclear how the random sequence was generated and if allocation concealment was conducted (except in Carter). In Masson, the assessors were blind but it was unclear if participants or investigators were blind. In Peterson 2009, neither the investigators or assessors were blind nor was it clear if participants were. In Carter, participants, assessors and investigators were not blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (-0.5).
3 It was unclear how the random sequence was generated and if allocation concealment was conducted (except in Carter 1988). Peterson 2009, neither the investigators nor assessors were blind and it was unclear if participants were. In Carter, participants, assessors and investigators were not blind. High dropouts were reported >20%.
4 For a continuous outcome, there were fewer than 400 participants.
5 Heterogeneity was detected, I2 >50%
6 Heterogeneity was detected, I2 >80%
7 It was unclear in either study if allocation concealment was conducted. Neither the assessors or investigators were blind nor was it unclear if participants were. High dropouts were detected >20%.
8 95% CI crossed 1 MID (0.5).
9 Allocation concealment was conducted but neither the participants, investigators or assessors were blind. It was unclear how many participants were randomised.
10 For a dichotomous outcome, there were fewer than 300 participants.

1 **Table 269: Summary table of self-help (ED) (self-help without support) compared with another intervention in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Other	Risk difference with BED Self-help (ED) (95% CI)
Bingeing	475 (6 studies)	⊕⊕⊕⊖ MODERATE1 due to risk of bias		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.25 standard deviations higher (0.06 to 0.43 higher)
Vomiting	90 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.81 standard deviations higher (0.38 to 1.24 higher)
Use of laxatives	90 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD	The mean use of laxatives in the intervention groups was 0.21 standard deviations lower

				values	(0.62 lower to 0.21 higher)
BMI	417 (4 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.13 standard deviations lower (0.33 lower to 0.06 higher)
Depression	236 (4 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.07 standard deviations higher (0.19 lower to 0.33 higher)
Remission_ITT	345 (6 studies)	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision	RR 0.84 (0.68 to 1.04)	494 per 1000	79 fewer per 1000 (from 158 fewer to 20 more)
EDE- Restraint	389 (4 studies)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.39 standard deviations higher (0.19 to 60 higher)
EDE- Shape concern	389 (4 studies)	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.24 standard deviations higher (0.04 to 0.44 higher)
EDE- Weight concern	389 (4 studies)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.30 standard deviations higher (0.1 to 0.51 higher)
EDE- Eating concern	389 (3 studies)	⊕⊕⊖⊖ LOW3,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.34 standard deviations higher (0.14 to 0.55 higher)
EDE- Global severity	437 (5 studies)	⊕⊕⊖⊖ LOW3,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- global severity in the intervention groups was 0.30 standard deviations higher (0.11 to 0.5 higher)
Excessive exercise	90 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean excessive exercise in the intervention groups was 0.28 standard deviations higher (0.13 lower to 0.7 higher)
Satisfaction with life	284	⊕⊕⊖⊖		Not	The mean satisfaction with life in the

	(2 studies)	LOW5,9 due to risk of bias, imprecision		calculable for SMD values	intervention groups was 0.13 standard deviations lower (0.35 to 0.13 higher)
Bingeng FU	227 (3 studies)	⊕⊕⊖⊖ LOW5,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean binging fu in the intervention groups was 0.06 standard deviations lower (0.34 lower to 0.21 higher)
BMI FU	296 (3 studies)	⊕⊕⊖⊖ LOW5,10 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.10 standard deviations lower (0.34 lower to 0.14 higher)
Depression FU	37 (1 study)	⊕⊖⊖⊖ VERY LOW10,11 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.18 standard deviations higher (0.51 lower to 0.88 higher)
Remission FU_ITT	118 (2 studies)	⊕⊕⊖⊖ LOW6,12 due to risk of bias, imprecision	RR 0.78 (0.5 to 1.2)	458 per 1000	101 fewer per 1000 (from 229 fewer to 92 more)
EDE- Restraint FU	259 (2 studies)	⊕⊕⊖⊖ LOW13,14 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint fu in the intervention groups was 0.20 standard deviations higher (0.05 lower to 0.45 higher)
EDE- Shape concern FU	259 (2 studies)	⊕⊕⊖⊖ LOW5,12 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern fu in the intervention groups was 0.07 standard deviations higher (0.18 lower to 0.32 higher)
EDE- Weight concern FU	259 (2 studies)	⊕⊖⊖⊖ VERY LOW5,12,15 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- weight concern fu in the intervention groups was 0.04 standard deviations higher (0.22 lower to 0.29 higher)
EDE- Eating concern FU	259 (2 studies)	⊕⊕⊖⊖ LOW5,12 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern fu in the intervention groups was 0.01 standard deviations higher (0.24 lower to 0.27 higher)
EDE-Q Global Score FU	260 (2 studies)	⊕⊕⊖⊖ LOW5,12 due to risk of bias, imprecision		Not calculable for SMD	The mean ede-q global score fu in the intervention groups was 0.08 standard deviations higher

				values	(0.17 lower to 0.33 higher)
Quality of life FU	167 (1 study)	⊕⊕⊖⊖ LOW5,12 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life fu in the intervention groups was 0.02 standard deviations higher (0.3 lower to 0.34 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Across studies it was unclear if allocation concealment was conducted (except for Carter). In addition, it was unclear if all or either the participants, investigators or assessors were blind. In Dunn, the participants were not blind, in Peterson 2009 the investigators and assessors were not blind, whilst in Grilo assessors were blind. High dropouts were reported >20%.
2 It was unclear if allocation concealment was conducted. In addition, the participants were not blind but it was unclear if investigators and assessors were blind. High dropouts were reported >20%.
3 95% CI crossed 1 MID (0.5).
4 95% CI crossed 1 MID (-0.5).
5 For a continuous outcome, there were fewer than 400 participants.
6 95% CI crossed 1 MID (0.75).
7 Across studies it was unclear if allocation concealment was conducted (except in Carter). In Loeb 2000 it was unclear if all or either the participants, investigators or assessors were blind. In Dunn, the participants were not blind, in Peterson 2009 the investigators and assessors were not blind, In Carter, participants, investigators, assessors were not blind. High dropouts were reported >20%.
8 Across studies it was unclear if allocation concealment was conducted (except in Carter). In Loeb 2000 it was unclear if all or either the participants, investigators or assessors were blind. In Dunn, the participants were not blind, in Peterson 2009 the investigators and assessors were not blind. In Grilo the assessors were blind. In Carter, the investigators, participants, assessors were not blind. High dropouts were reported >20%.
9 It was unclear if allocation concealment was conducted. In Cassin 2008 the assessors were blind, but it was unclear if investigators and participants were blind. In Peterson, the investigators and assessors were not blind but it was unclear if participants were blind. High dropouts were reported >20%.
10 It was unclear if allocation concealment was conducted (except in Carter). In Peterson 2009, the investigators and assessors were not blind but it was unclear if participants were blind. In Peterson 2001, it was unclear if any were blind. It was unclear if investigators, assessors and participants were not blind. High dropouts were reported >20%.
11 95% CI crossed 2 MIDs (-0.5 and 0.5).
12 It was unclear if allocation concealment was conducted (except in Carter). In Peterson 2009, the investigators and assessors were not blind but it was unclear if participants were blind. In Carter, participants, assessors and participants were not blind. High dropouts were reported >20%.
13 It was unclear if allocation concealment was conducted (except in Carter). In Peterson 2001, it was unclear if either the participants, investigator or assessors were blind. In Carter, participants, assessors and investigators were not blind.
14 For a dichotomous outcome, there were fewer than 300 events.
15 Heterogeneity was detected I2 >50%.

1 **Table 270: Summary table of self-help (ED) (self-help without support) compared with wait list controls in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BED Self-help (ED) (95% CI)
Bingeing	196 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.40 standard deviations lower (0.68 to 0.11 lower)
BMI	205 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.01 standard deviations higher (0.27 lower to 0.28 higher)
Remission_ITT	60 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 5.36 (1.34 to 21.36)	80 per 1000	349 more per 1000 (from 27 more to 1000 more)
EDE- Restraint	196 (2 studies)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.05 standard deviations lower (0.33 lower to 0.23 higher)
EDE- Shape concern	196 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.19 standard deviations lower (0.47 lower to 0.09 higher)
EDE- Weight concern	196 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0.14 standard deviations lower (0.42 lower to 0.15 higher)
EDE- Eating concern	196 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.25 standard deviations lower (0.54 lower to 0.04 higher)
EDE-Q- Global severity	196 (2 studies)	⊕⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-q- global severity in the intervention groups was 0.20 standard deviations lower (0.49 lower to 0.08 higher)

Quality of life	110 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.08 standard deviations higher (0.29 lower to 0.45 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was conducted, except in Carter. In Peterson 2009, the investigators and assessors were not blind but it was unclear if participants were blind. In Carter, participants, assessors, investigators were not blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (-0.5).
3 For a continuous outcome, there were fewer than 400 participants.
4 For a dichotomous outcome, there were fewer than 300 events.
5 Heterogeneity detected I2 >50%.
6 Heterogeneity detected, I2 >80%.

1 **Table 271: Summary table of internet self-help (ED) compared with wait list controls in young people and adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with BED Internet SH (ED) (95% CI)
Bingeing - Adults	118 (2 studies)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing - adults in the intervention groups was 0.03 standard deviations lower (0.4 lower to 0.34 higher)
BMI - Young people	93 (1 study)	⨁⨁⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI - young people in the intervention groups was 0.21 standard deviations lower (0.62 lower to 0.2 higher)
BMI - Adults	118 (2 studies)	⨁⨁⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI - adults in the intervention groups was 0.38 standard deviations higher (0.02 to 0.75 higher)
Depression - Young people	93 (1 study)	⨁⨁⊖⊖ LOW3,4 due to risk of bias,		Not calculable for SMD	The mean depression - young people in the intervention groups was 0.32 standard deviations lower

		imprecision		values	(0.72 lower to 0.09 higher)
Depression - Adults	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - adults in the intervention groups was 0.38 standard deviations lower (0.84 lower to 0.08 higher)
EDI Drive for thinness	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness in the intervention groups was 0.38 standard deviations lower (0.84 lower to 0.08 higher)
EDI Bulimia	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia in the intervention groups was 0.85 standard deviations lower (1.33 to 0.37 lower)
EDI Body dissatisfaction	74 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction in the intervention groups was 0.01 standard deviations higher (0.44 lower to 0.47 higher)
Remission_ITT	74 (1 study)	⊕⊕⊖⊖ LOW6,7 due to risk of bias, imprecision	RR 4.33 (1.35 to 13.96)	81 per 1000	270 more per 1000 (from 28 more to 1000 more)
EDE-Total	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-total in the intervention groups was 0.38 standard deviations lower (0.84 lower to 0.08 higher)
EDE- Restraint	74 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.01 standard deviations lower (0.47 lower to 0.45 higher)
EDE- Shape concern	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.3 standard deviations lower (0.76 lower to 0.15 higher)
Global severity index	74 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean global severity index in the intervention groups was 0.44 standard deviations lower (0.9 lower to 0.02 higher)

Quality of life	74 (1 study)	⊕⊕⊖⊖ LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.01 standard deviations lower (0.46 lower to 0.45 higher)
Bingeing FU - Adults	109 (2 studies)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing fu - adults in the intervention groups was 0.05 standard deviations higher (0.33 lower to 0.42 higher)
BMI FU - Young people	93 (1 study)	⊕⊕⊖⊖ LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu - young people in the intervention groups was 0.27 standard deviations lower (0.67 lower to 0.14 higher)
BMI FU - Adults	109 (2 studies)	⊕⊕⊖⊖ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu - adults in the intervention groups was 0.33 standard deviations higher (0.05 lower to 0.71 higher)
Depression FU - Young people	93 (1 study)	⊕⊕⊖⊖ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu - young people in the intervention groups was 0.17 standard deviations higher (0.24 lower to 0.58 higher)
Depression FU - Adults	74 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu - adults in the intervention groups was 0.4 standard deviations lower (0.86 lower to 0.06 higher)
EDE- Restraint FU	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint fu in the intervention groups was 0.08 standard deviations higher (0.37 lower to 0.54 higher)
EDE- Shape concern FU	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern fu in the intervention groups was 0.23 standard deviations lower (0.69 lower to 0.23 higher)
EDE-Total FU	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-total fu in the intervention groups was 0.3 standard deviations lower (0.76 lower to 0.16 higher)
EDI Drive for thinness FU	74	⊕⊕⊖⊖		Not	The mean edi drive for thinness fu in the

	(1 study)	LOW4,6 due to risk of bias, imprecision		calculable for SMD values	intervention groups was 0.44 standard deviations lower (0.9 lower to 0.02 higher)
EDI Bulimia FU	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia fu in the intervention groups was 0.32 standard deviations lower (0.78 lower to 0.14 higher)
EDI Body dissatisfaction FU	74 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction fu in the intervention groups was 0.13 standard deviations higher (0.33 lower to 0.58 higher)
Global severity index- FU	74 (1 study)	⊕⊕⊖⊖ LOW4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean global severity index- fu in the intervention groups was 0.33 standard deviations lower (0.79 lower to 0.13 higher)
Quality of life-FU	74 (1 study)	⊕⊕⊖⊖ LOW5,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life-fu in the intervention groups was 0.12 standard deviations higher (0.33 lower to 0.58 higher)
Remission FU_ITT	74 (1 study)	⊕⊕⊖⊖ LOW6,8 due to risk of bias, imprecision	RR 2 (0.98 to 4.09)	216 per 1000	216 more per 1000 (from 4 fewer to 668 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 In Carrard, allocation concealment was not conducted and it was unclear in Shapiro if it was performed. In Carrard assessors were not blind and it was unclear if either participants or investigators were blind. In Shapiro assessors were only bind at baseline measurement it was unclear if participants or investigators were blind. High dropouts were reported >20%.
2 For a continuous outcome, there were fewer than 400 participants.
3 In Jones 2008 it was unclear if allocation concealment was performed. Assessors were not blind and it was unclear if either participants or investigators were blind.
4 95% CI Crossed 1 MID (-0.5).
5 95% CI Crossed 1 MID (0.5).
6 In Carrard, allocation concealment was not conducted, Assessors were not blind and it was unclear if either participants or investigators were blind. High dropouts were reported >20%.
7 For a dichotomous outcome, there were fewer than 300 events.

8 95% CI crossed 1 MID (1.25).

Table 272: Summary table of guided self-help (ED) (self-help with support) compared with another guided self-help (ED) in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with BED Guided SH (ED) vs. Guided SH (95% CI)
Bingeing	75 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bingeing in the intervention groups was 0.48 standard deviations lower (0.94 to 0.02 lower)
BMI	75 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.19 standard deviations lower (0.64 lower to 0.27 higher)
Depression	75 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.25 standard deviations lower (0.71 lower to 0.2 higher)
Remission_ITT	75 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 2.51 (1.34 to 4.71)	237 per 1000	358 more per 1000 (from 81 more to 879 more)
EDE- Restraint	75 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- restraint in the intervention groups was 0.38 standard deviations lower (0.84 lower to 0.08 higher)
EDE- Shape concern	75 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- shape concern in the intervention groups was 0.12 standard deviations lower (0.57 lower to 0.33 higher)
EDE- Weight concern	75 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- weight concern in the intervention groups was 0 standard deviations higher (0.45 lower to 0.45 higher)

EDE- Eating concern	75 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede- eating concern in the intervention groups was 0.44 standard deviations lower (0.9 lower to 0.02 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 It was unclear if allocation concealment was performed. It was unclear if either the participants, assessors or investigators were blind. High dropouts were detected >20%.
2 95% CI crossed 1 MID (-0.5).
3 For a dichotomous outcome, there were fewer than 300 events.
4 For a continuous outcome there were fewer than 400 participants.

8.2.2.71 Family therapy

2 **Table 273: Summary table of findings for family therapy-ED versus wait list control in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Wait list control	Risk difference with Family Therapy-ED (95% CI)
Weight (kg)	62 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) in the intervention groups was 0.08 standard deviations higher (0.42 lower to 0.58 higher)
Binge Frequency EDE-Q-OBE	62 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q objective binge episode in the intervention groups was 0.56 standard deviations lower (1.07 to 0.05 lower)
Depression Beck Depression Inventory (BDI)	62 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.52 standard deviations lower (1.02 to 0.01 lower)
Family Functioning Dyadic Adjustment Scale	62 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean family functioning in the intervention groups was 0.04 standard deviations lower (0.54 lower to 0.46 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Wait list control	Risk difference with Family Therapy-ED (95% CI)
		imprecision			

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Gorin 2003: Dropout rate>20% (34% for whole sample), inadequate randomization method (used blocks by binge eating frequency), unclear allocation concealment, participant and assessor blinding.
2 CI crosses either 0.5 or -0.5 (SMD).

1 **Table 274: Summary table of findings for family therapy-ED versus group CBT-ED in adults with BED at end of treatment.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT	Risk difference with Family Therapy-ED (95% CI)
Weight (kg)	63 (1 study) 6 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) in the intervention groups was 0.2 standard deviations higher (0.29 lower to 0.7 higher)
Binge Frequency EDE-Q-OBE	63 (1 study) 6 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.24 standard deviations higher (0.26 lower to 0.73 higher)
Depression Beck Depression Inventory (BDI)	63 (1 study) 6 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.31 standard deviations lower (0.81 lower to 0.19 higher)
Family Functioning Level of Expressed Emotion (LEE)	63 (1 study) 6 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean family functioning in the intervention groups was 0.09 standard deviations lower (0.59 lower to 0.4 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT	Risk difference with Family Therapy-ED (95% CI)
CI: Confidence interval;					
1 Gorin 2003: Dropout rate>20% (34% for whole sample), inadequate randomization method (used blocks by binge eating frequency), unclear allocation concealment, participant and assessor blinding. 2 CI crosses either 0.5 or -0.5 (SMD).					

1 **Table 275: Summary table of findings for family therapy-ED versus group CBT-ED in adults with BED at follow up.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT	Risk difference with Family Therapy-ED (95% CI)
Weight (kg) FU	63 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight (kg) fu in the intervention groups was 0.22 standard deviations higher (0.28 lower to 0.71 higher)
Binge Frequency FU EDE-Q-OBE	63 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 0.52 standard deviations higher (0.01 to 1.02 higher)
Depression FU Beck Depression Inventory (BDI)	63 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.07 standard deviations lower (0.57 lower to 0.42 higher)
Family Functioning FU Level of Expressed Emotion (LEE)	63 (1 study) 6 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean family functioning fu in the intervention groups was 0.01 standard deviations lower (0.5 lower to 0.49 higher)
1 Gorin 2003: Dropout rate>20% (34% for whole sample), inadequate randomization method (used blocks by binge eating frequency), unclear allocation concealment, participant and assessor blinding. 2 CI crosses either 0.5 or -0.5 (SMD). 3 <400 participants.					

1

2

8.2.3 Economic Evidence

8.2.3.1 Systematic literature review

The systematic search of the economic literature undertaken for the guideline identified one study on the cost utility of CBT guided self-help in adults with recurrent binge eating disorder (Lynch et al., 2010). The study was conducted in the US.

References to all included studies and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix P. Completed methodology checklists of the studies are provided in Appendix O. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix Q.

Lynch and colleagues (2010) evaluated the cost effectiveness and cost-utility of CBT guided self-help (CBT-GSH) when compared with treatment as usual (TAU) in adults with recurrent binge eating disorder. The economic analysis was undertaken alongside an RCT (Striegel-Moore 2010) (N=123) conducted in the US. The intervention involved 8 brief coaching sessions provided by a master's-level therapist. The first session lasted 60 min and each subsequent session lasted 20 to 25 min. TAU was defined as people seeking help from primary care providers or nutrition care providers and self-referral to the mental health department. The analysis was conducted from a health and social care perspective (plus out-of-pocket expenses). The results could be calculated excluding out-of-pocket expenses. The study considered a range of costs including weight and eating disorder services, other medical services, psychiatric medications and peoples' expenses (time and expenditure for health care services, non-health services, over the counter medications and other weight loss products). The resource use estimates were based on the RCT (N=123). The unit costs were obtained from published studies, local market unit costs and wages in order to value participants' time spent receiving interventions. The measures of outcome for the economic analysis included the number of binge free days and quality adjusted life years (QALYs). However, the health related quality of life (HRQoL) weights used to estimate QALYs were derived by three experts. The weights were based on 52 disease categories from the ICD-9 using the person trade-off method of valuation. The time horizon of the analysis was 12 months.

The intervention resulted in a greater number of mean binge-free days at 12 month follow up compared with TAU (330.7 versus 305.5, respectively; a difference of 25.2, p-value = 0.002). The intervention also resulted in a greater number of QALYs at 12 month follow up (0.932 versus 0.863, respectively; a difference of 0.069, p-value not reported).

From a health and social care sector perspective the mean total costs per participant over 12 months were $3,527 for the intervention and $3,806 for TAU, a difference of -$279 (p-value not reported) in 2006 US dollars. Similarly, when considering health and social care plus out-of-pocket expenses the mean total costs per participant over 12 months were $3,671 for the intervention and $4,098 for TAU, a difference of -$427 (p = 0.3).

From both perspectives CBT-GSH was dominant (that is, it was less costly and more effective). Bootstrapping indicated that CBT-GSH had better outcomes and lower costs (health and social care plus out-of-pocket expenses) in 69% of replications when compared with TAU.

From both perspectives, at a willingness-to-pay (WTP) of $40 per additional binge free day, the probability that the intervention was cost effective was 90% and at a WTP of $100 per additional binge free day the probability was 98%. Deterministic sensitivity analysis from a health and social care perspective plus out-of-pocket expenses indicated that when removing one high-cost outlier and when using only cases with complete data the results did not change.

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it was conducted in the US. The authors estimated QALYs but the weights were based on expert opinion. However, this was not a major limitation since the intervention was found to be dominant when a natural outcome (i.e. the number of binge eating days) was used, and therefore the interpretation of the results was straightforward and did not require further judgments. Overall, this was a well conducted study and was judged by the committee to have only minor methodological limitations.

8.2.3.2 Economic modelling

Two decision-analytical models were developed to assess the relative cost effectiveness of interventions for adults with BED. The choice of treatments assessed in the economic analyses was determined by the availability of respective clinical data (that is, full remission at the end of treatment) included in the guideline systematic literature review. The economic analyses considered effective treatments, as demonstrated by the systematic review of clinical evidence, that were deemed appropriate by the committee as treatment options for people with BED in the UK. The study population in both models comprised of adults with BED.

Clinical data were derived from studies included in the guideline systematic review of clinical evidence and other published literature. Clinical data (that is, full remission at the end of treatment) were analysed using mixed treatment comparison technique. Full remission was defined as cessation of BED-related symptoms over and above two weeks. The comparisons between psychological interventions in the area of BED created two separate networks that could not be linked. Consequently, two separate NMAs were undertaken, which informed two different economic models. Interventions across the two economic models could not be compared due to the lack of a common comparator between the interventions that would allow the relative effects across interventions to be assessed. Details on the methods and clinical data utilised in the NMAs that were undertaken to estimate full remission for each treatment option considered in the economic analyses are presented in Appendix R.

The economic models assessed the following interventions:

- model one included individual therapies: Interpersonal Psychotherapy (IPT)-general (that is, not specific to eating disorders), behavioural weight loss, self-help ED with support, self-help ED no support, and no treatment (wait list);
- model two included group therapies: behavioural weight loss, CBT-ED, and IPT-ED.

Pharmacological interventions created a separate (3rd), limited network in the NMA, had small numbers randomised and generally showed no effectiveness. As a result, these were not considered in a separate NMA and economic analysis.

The rationale for economic modelling, the methodology adopted, the results and the conclusions from economic analyses are described in detail in Appendix R. The section 7.2.7.2.2 provides a summary of the methods employed. This section provides only the results of the 2 NMAs and associated economic analyses.

8.2.3.2.1 NMA and economic modelling results - individual therapies

The results of the NMA indicated that wait list had the lowest probability of full remission (mean 0.20 over 16 weeks), followed by self-help ED with no support (0.56), behavioural weight loss (0.72), self-help ED with support (0.73), and IPT-general (0.78). All treatments showed a significant effect compared with wait list. Also, self-help ED with no support was significantly worse than self-help ED with support with an OR of 0.46 (95% CrI: 0.25 to 0.76)

According to the deterministic analysis, wait list was dominated by self-help ED with no support (that is, self-help ED with no support resulted in lower costs and also was more effective). Similarly, behavioural weight loss was dominated by self-help ED with support

(that is, self-help ED with support resulted in lower costs and also was more effective). Both wait list and behaviour weight loss options were thus excluded from further analysis. When calculating incremental cost-effectiveness ratios (ICERs) for all consecutive pairs of options self-help ED with support versus self-help ED with no support resulted in the ICER of £7,381 per QALY. IPT-general was not cost effective (that is, it resulted in a cost per QALY versus self-help ED with support that was above upper NICE cost-effectiveness threshold of £30,000 per QALY.

The ICER of self-help ED with support (vs. self-help ED with no support) was sensitive to the utility value of remission and the cost of remission associated with self-help ED with support. The ICER of IPT individual (vs. self-help ED with support) was above the upper NICE cost-effectiveness threshold of £30,000 per QALY in all considered scenarios.

Conclusions of probabilistic analysis were the same as those of deterministic analysis. Self-help ED with support had the highest probability of being the most cost-effective treatment option, at any level of willingness-to-pay per additional QALY gained above £7,000 per QALY. At the lower NICE cost-effectiveness threshold of £20,000 per QALY (NICE, 2008b) the probability of self-help ED with support being cost effective was 0.83.

8.2.3.2.2 NMA and economic modelling results - group therapies

The results of the NMA indicated that group behavioural weight loss had the lowest probability of full remission (mean 0.27 over 16 weeks), followed by IPT-ED group (0.37) and CBT-ED group (0.45). Only CBT-ED group showed a significant effect compared with behavioural weight loss OR 2.31 (95% CrI: 1.16 to 4.19).

According to the deterministic analysis, IPT-ED group was extendedly dominated by behavioural weight loss group and CBT-ED group (that is, IPT-ED group was less effective and more costly than a linear combination of group behavioural weight loss and CBT-ED group). CBT-ED group (vs. behavioural weight loss group) resulted in an ICER of £3,834 per QALY and was the preferred treatment option.

According the deterministic sensitivity analyses the ICER of CBT-ED group (vs. behavioural weight loss group) was robust to changes in all model inputs. Under none of the scenarios examined IPT-ED group or the behavioural weight loss were the preferred treatment options.

Conclusions of the probabilistic analysis were the same as those of deterministic analysis. CBT-ED group had the highest probability of being the most cost-effective treatment option, at any level of willingness-to-pay per additional QALY gained above £3,500 per QALY. At the lower NICE cost-effectiveness threshold of £20,000 per QALY the probability of CBT-ED group being cost effective was 0.74.

8.2.3.2.3 Strengths and limitations

Clinical data on remission were synthesised using network meta-analytic techniques. Such methods enabled evidence synthesis from both direct and indirect comparisons between treatments. The base-case economic analysis considered only data on remission at the end of treatment. Due to the lack of suitable data the cost estimates during the follow up were based on the committee expert opinion. Also, due to the lack of suitable data utility values for BED were derived from people with an eating disorder not otherwise specified.

8.2.4 Clinical evidence statements

8.2.4.1 Individual Therapy

8.2.4.1.1 Hybrid therapy versus any other hybrid therapy in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=60) showed hybrid therapy is more effective on EDE-global compared with another hybrid therapy.

Low quality evidence from one RCT (n=60) showed no difference in the effect of hybrid therapy on weight loss compared with another hybrid therapy.

8.2.4.1.2 Dialectical behaviour therapy (DBT) versus wait list control in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=32) showed DBT is more effective on EDE-global and depression compared with wait list controls.

Low quality evidence from one RCT (n=32) showed no difference in the effect of DBT on binge eating and vomiting compared with wait list controls.

8.2.4.1.3 CBT-ED versus any other intervention in young people with binge eating disorder at end of treatment

Low to very low quality evidence from one RCT (n=26) showed no difference in the effect of CBT-ED on BMI, EDE-dietary restraint, EDE-shape concern, social adjustment, remission compared with any other intervention.

Low quality evidence from two RCTs (n=144) showed no difference in the effect of CBT-ED on EDE-weight concern compared with any other intervention.

Low quality evidence from one RCT (n=26) showed CBT-ED is more effective on depression and EDE-eating concern compared with any other intervention.

8.2.4.1.4 CBT-ED versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from two RCTs (n=253 to 256) showed no difference in the effect of CBT-ED on bingeing and EDE-weight concern compared with any other intervention.

Low quality evidence from one RCT (n=141) showed no difference in the effect of CBT-ED on depression compared with any other intervention.

Low quality evidence from two RCTs (n=253 to 256) showed CBT-ED is more effective on EDE-dietary concern, EDE-shape concern and EDE-eating concern compared with any other intervention.

Low quality evidence from one RCT (n=26) showed CBT-ED is more effective EDE-global score compared with any other intervention.

Low quality evidence from one RCT (n=26) showed CBT-ED is less effective on remission compared with any other intervention but there was some uncertainty

8.2.4.1.5 CBT-ED versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from two RCTs (n=346) showed CBT-ED is more effective on BMI compared with any other intervention.

Low quality evidence from two RCTs (n=231) showed CBT-ED is more effective on EDE-weight concern, EDE-dietary restraint, EDE-shape concern, EDE-eating concern compared with any other intervention.

Low quality evidence from two RCTs (n=346) showed CBT-ED is more effective on EDE-global score compared with any other intervention.

Low quality evidence from two RCTs (n=258) showed no difference in the effect of CBT-ED on bingeing compared with any other intervention.

Low quality evidence from one RCT (n=141) showed no difference in the effect of CBT-ED on depression compared with any other intervention.

Low quality evidence from one RCT (n=87) showed no difference in the effect of CBT-ED on remission compared with any other intervention.

8.2.4.1.6 Interpersonal psychotherapy (IPT) versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=205) showed no difference in the effect of IPT on BMI, bingeing and remission compared with any other intervention.

Low quality evidence from one RCT (n=205) showed no difference in the effect of IPT on BMI and bingeing compared with any other intervention.

8.2.4.2 Behavioural therapy versus any other intervention in adults with binge eating disorder at end of treatment.

Low quality evidence from one RCT (n=112) showed no difference in the effect of behavioural therapy on purging, bingeing, symptom checklist, EDE-weight concern, EDE-dietary restraint, EDI- body dissatisfaction, EDI-bulimia and EDI-drive for thinness compared with any other intervention.

Low quality evidence from one RCT (n=112) showed behavioural therapy may be less effective on EDE-weight concern, EDE-shape concern and EDI-bulimia compared with any other intervention.

Low quality evidence from one RCT (n=148) showed behavioural therapy may be less effective on remission compared with any other intervention but there was some uncertainty.

8.2.4.3 Behavioural therapy versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=87) showed no difference in the effect of behavioural therapy on purging, bingeing, symptom checklist, EDE-weight concern, EDE-shape concern, EDE-eating concern, EDE-dietary restraint, EDI- body dissatisfaction, EDI-bulimia and EDI-drive for thinness compared with any other intervention.

Low quality evidence from one RCT (n=112) showed no difference in the effect of behavioural therapy on remission compared with any other intervention.

8.2.4.4 CBT general therapy versus any other intervention in adults with binge eating disorder at end of treatment.

Low quality evidence from one RCT (n=112) showed no difference in the effect of general CBT on purging, bingeing, EDE-weight concern, EDE-shape concern, EDE-eating concern, EDE-dietary restraint, EDI- body dissatisfaction, EDI-bulimia, EDI-drive for thinness, global severity index and remission compared with any other intervention.

8.2.4.5 CBT general therapy versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=87) showed no difference in the effect of general CBT on purging, bingeing, EDE-weight concern, EDE-shape concern, EDE-eating concern, EDE-dietary restraint, EDI- body dissatisfaction, EDI-bulimia, EDI-drive for thinness and global severity index compared with any other intervention.

Low quality evidence from one RCT (n=112) showed no difference in the effect of general CBT on remission compared with any other intervention.

8.2.4.6 Group therapy

8.2.4.6.1 Group mindfulness versus any another group intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=103) showed no difference in group mindfulness on BMI, binge eating, depression compared with another group intervention.

8.2.4.6.2 Group mindfulness versus any another group intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=103) showed no difference in group mindfulness on BMI, binge eating, depression compared with another group intervention.

8.2.4.6.3 Group mindfulness versus wait list controls in adults with binge eating disorder at the end of treatment

Low quality evidence from one RCT (n=100) showed no difference in group mindfulness on BMI compared with wait list controls.

Low quality evidence from one RCT (n=100) showed group mindfulness is more effective on binge eating and depression compared with wait list controls.

8.2.4.6.4 Group mindfulness versus wait list controls in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=100) showed no difference in group mindfulness on BMI compared with wait list controls.

Low quality evidence from one RCT (n=100) showed group mindfulness is more effective on binge eating and depression compared with wait list controls.

8.2.4.6.5 Group CBT-ED versus any other intervention in adults with binge eating disorder at end of treatment

Moderate quality evidence from seven RCTs (n=588) showed no difference in group CBT-ED on depression compared with any other intervention.

Low quality evidence from one RCT (n=53) showed no difference in group CBT-ED on anxiety compared with any other intervention.

Very low quality evidence from three RCTs (n=241) showed no difference in group CBT-ED on EDE-shape concerns compared with any other intervention.

Very low quality evidence from four RCTs (n=384) showed no difference in group CBT-ED on EDE-dietary restraint compared with any other intervention.

Very low quality evidence from one RCT (n=158) showed no difference in group CBT-ED on global symptom score compared with any other intervention.

Moderate quality evidence from six RCTs (n=530) showed group CBT-ED may be less effective on decreasing weight compared with any other intervention but there was some uncertainty

Very low quality evidence from two RCTs (n=266) showed group CBT-ED may be less effective on EDE-global clinical score compared with any other intervention.

Very low quality evidence from four RCTs (n=384) showed group CBT-ED may be less effective on EDE-eating concern compared with any other intervention but there was some uncertainty.

Very low quality evidence from four RCTs (n=384) showed group CBT-ED may be more effective on EDE-weight concern compared with any other intervention but there was some uncertainty.

Low quality evidence from four RCTs (n=404) showed group CBT-ED is more effective on remission compared with any other intervention.

8.2.4.6.6 Group CBT-ED versus any other intervention in adults with binge eating disorder at follow up

Moderate quality evidence from six RCTs (n=514) showed group CBT-ED may be less effective on reducing weight compared with any other intervention but there was some uncertainty.

Very low quality evidence from two RCTs (n=185) showed group CBT-ED may be less effective on anxiety compared with any other intervention.

Very low quality evidence from three RCTs (n=266) showed group CBT-ED may be less effective on EDE-global clinical score compared with any other intervention.

Very low quality evidence from three RCTs (n=266) showed group CBT-ED may be less effective on EDE-dietary restraint compared with any other intervention but there was some uncertainty.

Very low quality evidence from five RCTs (n=540) showed group CBT-ED may be less effective on EDE-weight concern and eating concern compared with any other intervention

Moderate quality evidence from seven RCTs (n=651) showed no difference in the effect of group CBT-ED on bingeing compared with any other intervention.

Moderate quality evidence from six RCTs (n=587) showed no difference in the effect of group CBT-ED on depression compared with any other intervention.

Very low quality evidence from four RCTs (n=350) showed no difference in the effect of group CBT-ED on EDE-shape concern compared with any other intervention.

Low quality evidence from one RCT (n=138) showed no difference in the effect of group CBT-ED on global symptom index compared with any other intervention.

Low quality evidence from three RCTs (n=279) showed no difference in the effect of group CBT-ED on remission compared with any other intervention.

8.2.4.6.7 Group CBT-ED versus wait list controls in adults with binge eating disorder at end of treatment

Low quality evidence from three RCTs (n=181) showed no difference in the effect of group CBT-ED on weight compared with wait list controls.

Very low quality evidence from three RCTs (n=160) showed no difference in the effect of group CBT-ED on depression compared with wait list controls.

Low quality evidence from two RCTs (n=141) showed no difference in the effect of group CBT-ED on binge eating compared with wait list controls.

8.2.4.6.8 Group CBT-ED versus wait list controls in adults with binge eating disorder at follow up

Low quality evidence from two RCTs (n=130) showed no difference in the effect of group BT-ED on BMI compared with wait list controls.

Low quality evidence from two RCTs (n=130 to 137) showed group BT-ED is more effective on depression and binge eating compared with wait list controls.

8.2.4.6.9 Group BT-ED versus wait list controls in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=72) showed no difference in the effect of group BT-ED on bingeing, EDE-total and anxiety compared with wait list controls.

Low quality evidence from one RCT (n=100) showed no difference in the effect of group BT-ED on remission compared with wait list controls.

Low quality evidence from one RCT (n=72) showed group BT-ED is more effective on depression compared with wait list controls.

8.2.4.6.10 Group BT-ED versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=98) showed no difference in the effect of group BT-ED on depression, BMI, weight loss, EDE-shape concerns and EDE-weight concern compared with any other intervention.

Low quality evidence from one RCT (n=98) showed group BT-ED is more effective on EDE-eating concern, EDE-dietary restraint and remission compared with any other intervention.

8.2.4.6.11 Group BT-ED versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=88) showed no difference in the effect of group BT-ED on depression, BMI, weight loss, EDE-dietary restraint compared with any other intervention.

Low quality evidence from one RCT (n=88) showed no difference in the effect of group BT-ED on EDE-shape concern and EDE-eating concern compared with any other intervention but there was some uncertainty.

Low quality evidence from one RCT (n=88 to 101) showed group BT-ED may be more effective on EDE-weight concern and remission compared with any other intervention but there was some uncertainty.

8.2.4.6.12 Group CBT-ED (body exposure) versus CBT-ED (cognitive) in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=24 to 28) showed no difference in the effect of group CBT-ED (body exposure) on EDE-restraint, EDE-eating concern, EDE-weight concern, EDE-shape concern, BMI, depression, bingeing and remission compared with CBT-ED (cognitive).

8.2.4.6.13 Group CBT-ED (body exposure) versus CBT-ED (cognitive) in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=24 to 28) showed no difference in the effect of group CBT-ED (body exposure) on EDE-restraint, EDE-eating concern, EDE-weight concern, EDE-shape concern, BMI, depression, bingeing and remission compared with CBT-ED (cognitive).

8.2.4.6.14 Group interpersonal psychotherapy (IPT) versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=158 to 162) showed no difference in the effect of group IPT on bingeing, EDE-shape concern, EDE-eating concern, EDE-weight concern, global symptom index, depression and BMI compared with any other intervention.

Low quality evidence from one RCT (n=158) showed group IPT is less effective on EDE-restraint compared with any other intervention.

8.2.4.6.15 Group interpersonal psychotherapy (IPT) versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=138) showed no difference in the effect of group IPT on bingeing, EDE-shape concern, EDE-eating concern, EDE-weight concern, EDE-restraint, global symptom index, depression and BMI compared with any other intervention.

8.2.4.6.16 Group counselling versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=88 to 98) showed no difference in the effect of group counselling on BMI, depression and weight loss compared with any other intervention.

Low quality evidence from one RCT (n=98 to 101) showed group counselling may be less effective on EDE-dietary restraint, EDE-eating concern and remission compared with any other intervention.

Low quality evidence from one RCT (n=98) showed group counselling may be less effective on EDE-shape concern and EDE-weight concern compared with any other intervention.

Low quality evidence from one RCT (n=98) showed group counselling may be less effective on patient preference compared with any other intervention.

8.2.4.6.17 Group counselling versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=88) showed no difference in the effect of group counselling on BMI, depression, EDE-shape concerns, EDE-eating concern and weight loss compared with any other intervention.

Low quality evidence from one RCT (n=88 to 101) showed group counselling may be less effective on remission and EDE-weight concern compared with any other intervention but there was some uncertainty.

Low quality evidence from one RCT (n=88) showed group counselling is less effective on EDE-dietary restraint compared with any other intervention.

8.2.4.6.18 Group diet versus any other group intervention in adults with binge eating disorder at end of treatment

Low quality evidence from three RCTs (n=242) showed no difference in the effect of group diet on weight compared with any other intervention.

Very low to low quality evidence from two RCTs (n=85) showed no difference in the effect of group diet on EDE-shape concerns, EDE-eating concern, EDE-restraint and EDE-weight concern compared with any other intervention.

Low quality evidence from one RCT (n=48) showed no difference in the effect of group diet on EDE-shape concern and EDE-restraint compared with any other intervention in people who binged less than 18 times per month.

Low quality evidence from one RCT (n=37) showed group diet is less effective on EDE-shape concern EDE-restraint compared with any other intervention in people who binged more than 18 times per month.

Low quality evidence from one RCT (n=125) showed no difference in the effect of group diet on EDE-global compared with any other intervention.

Low quality evidence from four RCTs (n=327) showed group diet may be less effective on depression compared with any other intervention but there was some uncertainty.

Low quality evidence from three RCTs (n=242) showed group diet is less effective on remission compared with any other intervention.

8.2.4.6.19 Group diet versus any other group intervention in adults with binge eating disorder at follow up

Low quality evidence from three RCTs (n=229) showed no difference in the effect of group diet on weight compared with any other intervention.

Low quality evidence from two RCTs (n=71) showed no difference in the effect of group diet on EDE-shape concerns, EDE-eating concern, EDE-restraint and EDE-weight concern compared with any other intervention.

Low quality evidence from one RCT (n=125) showed no difference in the effect of group diet on EDE-global compared with any other intervention.

Low quality evidence from three RCTs (n=205) showed no difference in the effect of group diet on depression compared with any other intervention.

Low quality evidence from three RCTs (n=241) showed group diet may be less effective on bingeing compared with any other intervention but there was some uncertainty.

Low quality evidence from two RCTs (n=117) showed group diet is less effective on bingeing compared with any other intervention.

8.2.4.6.20 Group self-help (ED) versus any other group intervention in adults with binge eating disorder at end of treatment

Low quality evidence from two RCTs (n=234) showed group self-help is more effective on BMI compared with another group intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=190) showed group diet is less effective on bingeing, EDE-restraint and EDE-eating concern compared with any other intervention.

Low quality evidence from one RCT (n=190) showed group diet is less effective on EDE-global score, EDE-shape concern and EDE-weight concern compared with any other intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=167) showed no difference in the effect of group self-help on depression and quality of life compared with another group intervention.

Low quality evidence from one RCT (n=51) showed group self-help is more effective on remission compared with another group intervention.

8.2.4.6.21 Group self-help (ED) versus any other group intervention in adults with binge eating disorder at follow up

Low quality evidence from two RCTs (n=231) showed no difference in the effect of group self-help on BMI compared with another group intervention.

Low quality evidence from one RCTs (n=167 to 190) showed no difference in the effect of group self-help on bingeing, EDE-eating concern, EDE-shape concern, EDE-weight concern, EDE-global score and quality of life compared with another group intervention.

Low quality evidence from one RCT (n=44 to 51) showed no difference in the effect of group self-help on depression and remission compared with another group intervention.

Low quality evidence from one RCT (n=190) showed group self-help is less effective on EDE-restraint compared with another group intervention.

8.2.4.6.22 Group guided self-help (ED) versus any other group intervention in adults with binge eating disorder at end of treatment

Low quality evidence from two RCTs (n=234) showed no difference in the effect of group guided self-help (ED) on BMI compared with another group intervention.

Very low to low quality evidence from one RCT (n=176 to 190) showed no difference in the effect of group guided self-help (ED) on EDE-global, EDE-shape concern, EDE-weight concern, EDE-eating concern and quality of life compared with another group intervention.

Very low evidence from one RCT (n=51) showed no difference in the effect of group guided self-help (ED) on remission compared with another group intervention.

Low quality evidence from one RCT (n=183) showed group guided self-help (ED) is more effective on bingeing compared with another group intervention.

Low quality evidence from one RCT (n=190) showed group guided self-help (ED) is more effective on EDE-restraint compared with another group intervention but there was some uncertainty.

Low quality evidence from one RCT (n=44) showed group guided self-help (ED) is more effective on weight compared with another group intervention but there was some uncertainty.

8.2.4.6.23 Group guided self-help (ED) versus any other group intervention in adults with binge eating disorder at follow up

Low quality evidence from two RCTs (n=231) showed no difference in the effect of group guided self-help (ED) on BMI compared with another group intervention.

Low quality evidence from one RCT (n=167 to 190) showed no difference in the effect of group guided self-help (ED) on EDE-weight concern, EDE-restraint and quality of life compared with another group intervention.

Low quality evidence from one RCT (n=44) showed no difference in the effect of group guided self-help (ED) on depression compared with another group intervention.

Low quality evidence from one RCT (n=44) showed group guided self-help (ED) is less effective on EDE-eating concern compared with another group intervention but there was some uncertainty.

Low quality evidence from one RCT (n=190) showed group guided self-help (ED) is less effective on EDE-shape concern compared with another group intervention.

Low quality evidence from one RCT (n=190) showed group guided self-help (ED) is less effective on bingeing compared with another group intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=190) showed group guided self-help (ED) is more effective on EDE-global core compared with another group intervention.

Low quality evidence from one RCT (n=51) showed no difference in the effect of group guided self-help (ED) on remission compared with another group intervention.

8.2.4.6.24 Group self-help (ED) versus wait list controls in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=136) showed no difference in the effect of group guided self-help (ED) on BMI, EDE-global, EDE-weight concern, EDE-eating concern, EDE-shape concern, EDE-restraint and quality of life compared with wait list controls.

Low quality evidence from one RCT (n=136) showed group guided self-help (ED) is more effective on bingeing compared with wait list controls, but there was some uncertainty.

8.2.4.6.25 Group guided self-help (ED) versus wait list controls in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=129) showed group guided self-help (ED) is more effective on bingeing and EDE-restraint compared with wait list controls.

Low quality evidence from one RCT (n=129) showed no difference in the effect of group guided self-help (ED) on EDE-weight concern, EDE-eating concern, EDE-shape concern, EDE-global and quality of life compared with wait list controls.

Low quality evidence from one RCT (n=129) showed group guided self-help (ED) is less effective on BMI compared with wait list controls but there was some uncertainty.

8.2.4.6.26 Group psychoeducation versus another group intervention in adults with binge eating disorder at end of treatment

Low quality evidence from two RCTs (n=234) showed no difference in group psychoeducation on BMI compared with another group intervention.

Low quality evidence from one RCT (n=174 to 190) showed no difference in group psychoeducation on bingeing and quality of life compared with another group intervention.

Low quality evidence from one RCT (n=44) showed no difference in group psychoeducation on depression compared with another group intervention.

Low quality evidence from one RCT (n=253) showed group psychoeducation is more effective on EDE-global score compared with another group intervention.

Low quality evidence from one RCT (n=190) showed group psychoeducation is more effective on EDE-weight concern, EDE-eating concern, EDE-shape concern and EDE-restraint compared with another group intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=51) showed group psychoeducation is more effective on remission compared with another group intervention, but there was some uncertainty.

8.2.4.6.27 Group psychoeducation versus another group intervention in adults with binge eating disorder at follow up

Low quality evidence from two RCTs (n=243) showed no difference in group psychoeducation on BMI compared with another group intervention.

Low quality evidence from one RCT (n=167 to 190) showed no difference in group psychoeducation on bingeing, EDE-eating concern, EDE-weight concern and quality of life compared with another group intervention.

Low quality evidence from one RCT (n=51) showed no difference in group psychoeducation on remission compared with another group intervention.

Low quality evidence from one RCT (n=41) showed group psychoeducation is more effective on depression compared with another group intervention.

Low quality evidence from one RCT (n=190) showed group psychoeducation is more effective on EDE-global and EDE-shape concern compared with another group intervention.

Low quality evidence from one RCT (n=190) showed group psychoeducation is more effective on EDE-restraint compared with another group intervention but there was some uncertainty.

8.2.4.7 Self-help therapy

8.2.4.7.1 Guided self-help (ED) versus another other intervention in adults with binge eating disorder at end of treatment

Very low quality evidence from seven RCTs (n=740) showed guided self-help (ED) is more effective on EDE-weight concern, EDE-shape concern and EDE-restraint compared with any other intervention.

Very low quality evidence from seven RCTs (n=740) showed guided self-help (ED) is more effective on EDE-weight concern compared with any other intervention but there was some uncertainty

Very low quality evidence from six RCTs (n=650) showed guided self-help (ED) is more effective on EDE-eating concern compared with any other intervention.

Very low quality evidence from one RCT (n=90) showed guided self-help (ED) is more effective on vomiting compared with any other intervention.

Moderate quality evidence from seven RCTs (n=490) showed guided self-help (ED) is more effective bingeing compared with any other intervention.

Low quality evidence from five RCTs (n=394) showed guided self-help (ED) is more effective depression compared with any other intervention, but there was some uncertainty.

Very low quality evidence from nine RCTs (n=661) showed guided self-help (ED) is more effective remission compared with any other intervention.

Very low quality evidence from four RCTs (n=389) showed guided self-help (ED) is more effective EDE-global severity compared with any other intervention, but there was some uncertainty

Very low quality evidence from one RCT (n=90) showed no difference in the effect of guided self-help (ED) on use of laxatives and excessive exercise compared with any other intervention.

Moderate quality evidence from seven RCTs (n=690) showed no difference in the effect of guided self-help (ED) on BMI compared with any other intervention.

Moderate quality evidence from two RCTs (n=284) showed no difference in the effect of guided self-help (ED) on satisfaction with life compared with any other intervention.

8.2.4.7.2 Guided self-help (ED) versus another other intervention in adults with binge eating disorder at follow up

Low quality evidence from two RCTs (n=260) showed guided self-help (ED) is more effective on EDE-global score compared with any other intervention.

Very low quality evidence from three RCTs (n=368) showed guided self-help (ED) is more effective on EDE-eating concern compared with any other intervention.

Very low quality evidence from three RCTs (n=368) showed no difference in the effect of guided self-help (ED) on EDE-shape concern compared with any other intervention.

Very low quality evidence from three RCTs (n=229) showed guided self-help (ED) is more effective on remission compared with any other intervention.

Very low quality evidence from two RCTs (n=150) showed guided self-help (ED) is more effective on depression compared with any other intervention.

Very low quality evidence from three RCTs (n=368) showed no difference in the effect of guided self-help (ED) on EDE-restraint compared with any other intervention.

Low quality evidence from four RCTs (n=300) showed no difference in the effect of guided self-help (ED) on bingeing compared with any other intervention.

Very low quality evidence from four RCTs (n=409) showed no difference in the effect of guided self-help (ED) on BMI compared with any other intervention.

Very low quality evidence from three RCTs (n=368) showed no difference in the effect of guided self-help (ED) on EDE-weight concern compared with any other intervention.

Low quality evidence from one RCT (n=167) showed no difference in the effect of guided self-help (ED) on quality of life compared with any other intervention.

8.2.4.7.3 Guided self-help (ED) versus wait list controls in adults with binge eating disorder at end of treatment

Low quality evidence from three RCTs (n=218) showed guided self-help (ED) is more effective on bingeing compared with wait list controls.

Very low quality evidence from two RCTs (n=248) showed guided self-help (ED) is more effective on EDE-shape concern compared with wait list controls.

Very low quality evidence from two RCTs (n=248) showed guided self-help (ED) is more effective on EDE-global compared with wait list controls but there was some uncertainty.

Very low quality evidence from two RCTs (n= 252) showed no difference in guided self-help (ED) EDE-restraint compared with wait list controls

Very low quality evidence from two RCTs (n=248 to 252) showed guided self-help (ED) is more effective on EDE-global compared with wait list controls

Very low quality evidence from two RCTs (n=248 to 252) showed guided self-help (ED) is more effective on EDE-weight concern and EDE-eating concern compared with wait list controls but there was some uncertainty

Low quality evidence from one RCTs (n=59) showed guided self-help (ED) is more effective on remission compared with wait list controls.

Low quality evidence from two RCTs (n=188) showed no difference in the effect of guided self-help (ED) on BMI compared with wait list controls.

Low quality evidence from one RCT (n=109) showed no difference in the effect of guided self-help (ED) on quality of life compared with wait list controls.

8.2.4.7.4 Self-help (ED) versus any other intervention in adults with binge eating disorder at end of treatment

Moderate quality evidence from six RCTs (n=475) showed self-help (ED) is less effective on bingeing compared with any other intervention.

Low quality evidence from one RCT (n=90) showed self-help (ED) is less effective on vomiting compared with any other intervention.

Low quality evidence from one RCT (n=90) showed no difference in the effect of self-help (ED) on the use of laxatives and excessive exercise compared with any other intervention.

Low quality evidence from four RCTs (n=417) showed self-help (ED) is more effective on BMI compared with any other intervention, but there was some uncertainty.

Low quality evidence from four RCTs (n=236) showed no difference in the effect of self-help (ED) on depression compared with any other intervention.

Low quality evidence from six RCTs (n=345) showed self-help (ED) is less effective on remission compared with any other intervention but there was some uncertainty.

Low quality evidence from four RCTs (n=389) showed self-help (ED) is less effective on EDE-weight concern, EDE-shape concern, EDE-restraint and EDE-eating concern compared with any other intervention.

Low quality evidence from five RCTs (n=437) showed self-help (ED) is less effective on EDE-global compared with any other intervention.

Low quality evidence from two RCTs (n=284) showed self-help (ED) is more effective on the satisfaction with life compared with any other intervention.

8.2.4.7.5 Self-help (ED) versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from three RCTs (n=227 to 296) showed no difference in the effect of self-help (ED) on bingeing and BMI compared with any other intervention.

Very low to low quality evidence from two RCTs (n=259) showed no difference in the effect of self-help (ED) on EDE-weight concern, EDE-shape concern and EDE-eating concern compared with any other intervention.

Low quality evidence from one RCT (n=37) showed no difference in the effect of self-help (ED) on depression compared with any other intervention.

Low quality evidence from two RCTs (n=118) showed no difference in the effect of self-help (ED) on remission compared with any other intervention.

Low quality evidence from one RCT (n=167) showed no difference in the effect of self-help (ED) on the quality of life compared with any other intervention.

8.2.4.7.6 Self-help (ED) versus wait list controls in adults with binge eating disorder at end of treatment

Low quality evidence from two RCTs (n=196) showed self-help (ED) is more effective on bingeing compared with any other intervention but there was some uncertainty.

Low quality evidence from one RCT (n=60) showed self-help (ED) is more effective on remission compared with any other intervention.

Very low to low quality evidence from two RCTs (n=196 to 205) showed no difference in the effect of self-help (ED) on BMI, EDE-weight concern, EDE-shape concern and EDE-restraint compared with wait list controls.

Very low quality evidence from two RCTs (n=196) showed self-help (ED) is more effective on EDE-eating concern and EDE-global severity compared with any other intervention, but there was some uncertainty.

Low quality evidence from one RCT (n=110) showed no difference in the effect of self-help (ED) on quality of life compared with wait list controls.

8.2.4.7.7 Internet self-help (ED) versus wait list controls in adults with binge eating disorder at end of treatment

Low quality evidence from two RCTs (n=118) showed no difference in the effect of internet self-help (ED) on bingeing compared with wait list controls

Low quality evidence from two RCTs (n=118) showed internet self-help (ED) may be less effective on BMI compared with wait list controls, but there was some uncertainty.

Low quality evidence from one RCT (n=74) showed internet self-help (ED) is more effective on depression, EDE-total, EDI-drive for thinness, global severity index compared with wait list controls, but there was some uncertainty.

Low quality evidence from one RCT (n=74) showed internet self-help (ED) is more effective on EDI-bulimia and remission compared with wait list controls.

Low quality evidence from one RCT (n=74) showed no difference in the effect of internet self-help (ED) on EDI-body dissatisfaction, EDE-restraint, EDE-shape concern and quality of life compared with wait list controls.

8.2.4.7.8 Internet self-help (ED) versus wait list controls in young people with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=93) showed no difference in the effect of internet self-help (ED) on bingeing compared with wait list controls.

Low quality evidence from one RCT (n=93) showed internet self-help (ED) is more effective on depression compared with wait list controls, but there was some uncertainty

8.2.4.7.9 Internet self-help (ED) versus wait list controls in adults with binge eating disorder at follow up

Low quality evidence from two RCTs (n=118) showed internet self-help (ED) is less effective on BMI compared with wait list controls, but there was some uncertainty.

Low quality evidence from one RCT (n=74) showed internet self-help (ED) is more effective on depression, EDI-drive for thinness and remission compared with wait list controls but there was some uncertainty.

Low quality evidence from one RCT (n=74) showed no difference in the effect of internet self-help (ED) on EDE-restraint, EDE-shape concern, EDE-total, EDI-body dissatisfaction, EDI-bulimia, global severity index and quality of life compared with wait list controls.

8.2.4.7.10 Internet self-help (ED) versus wait list controls in young people with binge eating disorder at follow up

Low quality evidence from one RCT (n=93) showed no difference in the effect of internet self-help (ED) on BMI and depression compared with wait list controls.

8.2.4.7.11 Guided self-help (ED) versus another guided self-help (ED) in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=74) showed no difference in the effect of guided self-help (ED) on BMI, depression and EDE-shape concern compared with another guided self-help.

Low quality evidence from one RCT (n=74) showed guided self-help (ED) is more effective on bingeing, EDE-restraint, EDE-weight concern, EDE-eating concern compared with another guided self-help, but there was some uncertainty.

Low quality evidence from one RCT (n=74) showed guided self-help (ED) is more effective on remission compared with another guided self-help.

8.2.4.7.12 Internet self-help (ED) versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=44) showed no difference in the effect of internet self-help (ED) on BMI and binge eating compared with any other intervention.

8.2.4.7.13 Internet self-help (ED) versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=44) showed no difference in the effect of internet self-help (ED) on BMI and binge eating compared with any other intervention.

8.2.4.8 Family therapy

8.2.4.8.1 Family therapy-ED versus wait list control in adults with binge eating disorder

Low quality evidence from one RCT (n=62) showed no difference in the effect of family therapy-ED on Weight and family functioning compared with wait list controls.

Low quality evidence from one RCT (n=63) showed family therapy-ED is more effective on binge frequency and depression compared with wait list controls.

8.2.4.8.2 Family therapy-ED versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=63) showed no difference in the effect of family therapy-ED on Weight, binge frequency, depression and family functioning compared with any other intervention.

8.2.4.8.3 Family therapy-ED versus any other intervention in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=63) showed no difference in the effect of family therapy-ED on Weight, depression and family functioning compared with any other intervention.

Low quality evidence from one RCT (n=63) showed family therapy-ED was less effective on binge frequency compared with any other intervention.

8.2.5 Economic Evidence statements

There was evidence from one US study (N=123) which found guided self-help to be dominant when compared with treatment as usual. The reviewed study was only partially applicable and was characterised by minor methodological limitations.

In the economic analysis conducted for this guideline, self-help ED with support appeared to be more cost-effective option for adults with BED when compared with other individual therapies. Self-help ED with support had the highest probability of being the most cost-effective treatment option, at any level of willingness-to-pay per additional QALY gained above £7,000 per QALY. At the lower NICE cost-effectiveness threshold of £20,000 per QALY the probability of self-help ED with support being cost effective was 0.83. Similarly, CBT-ED group when compared with other group therapies was the most cost-effective option for adults with BED. CBT-ED group had the highest probability of being the most cost-effective treatment option, at any level of willingness-to-pay per additional QALY gained above £3,500 per QALY. At the lower NICE cost-effectiveness threshold of £20,000 per QALY the probability of CBT-ED group being cost effective was 0.74. The evidence from the guideline economic analysis was directly applicable to the UK context and it was characterised by potentially serious methodological limitations.

No economic evidence on the cost effectiveness of interventions for children and young people with BED was available.

8.2.6 Recommendations and link to evidence for the review on: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with binge eating disorder compared with any other intervention or controls?

Psychological treatment (guided self-help) for adults with BED

	127. Offer a binge-eating-focused guided self-help programme to adults with binge eating disorder. **128. Binge-eating-focused guided self-help programmes for adults should:** • use a cognitive behavioural self-help book • focus on adherence to the self-help programme • supplement the self-help programme with brief supportive sessions (for example four to nine sessions lasting 20 minutes each over 16 weeks that are first run weekly): o delivered by a practitioner who is competent in delivering the treatment o that focus exclusively on helping the person follow the programme.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating binge eating disorder. For this population, it was agreed binge eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between	Adults with binge eating disorder To firstly see the benefits of guided self-help it is best to compare its effects to

clinical benefits and harms	waitlist controls. The results showed clear benefits of the treatment in adults on remission rates and binge eating compared with waitlist controls. Other outcomes that showed favourable results included EDE-global score, EDE-eating concern, EDE-shape concern and EDE-restraint. No difference was found in BMI and quality of life. Long-term follow up data was not available, nor was any data reported at any time point on all-cause mortality, adverse events, resource use, relapse, general psychopathology, body weight, general functioning, family functioning and service user experience. If you compare guided self-help with another intervention, the benefits in adults are still evident. Positive effects were found on remission rates and binge eating. In addition to vomiting, depression, EDE-shape concern, EDE-restraint, EDE-eating concern and EDE-weight concern. A trend for a benefit was found on EDE-global severity but no difference in excessive exercise, satisfaction with life, laxative use and BMI. Six to 12 months follow up data showed remission rates still favoured guided self-help compared with any other intervention but binge frequency was no longer different. Nor was BMI, EDE-eating concern, EDE-shape concern, EDE-weight concern, EDE-global score, quality of life or depression. In addition to remission, only EDE-restraint favoured guided self-help but there was some uncertainty. No data at any time point was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience. When comparing self-help (ED) to any other intervention, the results are not so favourable. Both remission and binge eating favoured any other intervention but there was some uncertainty. However, it clearly favoured any other intervention for vomiting, EDE-global and EDE-subscales. BMI showed a trend to be lower in the self-help-treated group, but no difference was found in and excessive exercise laxative use, depression, or satisfaction of life. At 6 to 12 months follow up, no differences were found in remission, binge eating or any other outcome, except EDE-restraint showed a trend to be lower in the self-help group compared with any other intervention. No data was reported at any time point on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience. Self-help (ED) compared with wait list controls showed favourable results in adults on remission rates and binge eating, in addition to show a trend to favour EDE-Global and EDE-eating concern. No difference was found on the effect on BMI, EDE-restraint, EDE-shape concern, EDE-weight concern and quality of life. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning, general psychopathology and service user experience. Other interventions investigated in adults included internet self-help. Compared with wait list controls, a benefit was found on remission but no difference was found on bingeing. Other favourable results included EDI-bulimia and a trend to favour EDI-drive for thinness, EDE-total, global severity index. No other differences were found between depression, EDI-body dissatisfaction, EDE-restraint, EDE-shape concern and quality of life. There appeared to be negative effect on BMI. At 2 to 6 months follow up, there was a trend to still favour internet self-help over wait list controls for remission but no difference was found in bingeing, depression, EDE-restraint, EDE-shape concern, EDE-total, EDI-bulimia, EDI-body dissatisfaction, global severity index, quality of life. There was a trend to favour EDI-drive for thinness and depression. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience. One study compared two different types of guided self-help. One type based on the manual by Fairburn 1995 showed the most favourable results on remission and bingeing compared with a behavioural weight loss guided self-help manual called LEARN. The latter focuses on making a gradual and moderate lifestyle change with goal of calorie restriction and increased physical activity. The Fairburn manual focuses on developing a regular pattern of moderate eating, self-control strategies and problem-solving. No data was reported on all-cause mortality, adverse events,

resource use, relapse, general functioning, body weight, general psychopathology, family functioning and service user experience.

Individual therapy

Other interventions that were considered but did not show convincing evidence included individual hybrid therapy. Comparing one type of hybrid therapy to another in adults showed it favoured hybrid 1 (CBT-ED and weight loss [diet, exercise, counselling]) over hybrid 2 (combined brief strategic thinking and weight loss group) for a global clinical score but no difference was found in weight loss. Remission data could not be included because it was unclear over what duration the symptoms were measured. No data was reported on bingeing, all-cause mortality, adverse events, resource use, relapse, general functioning, body weight, family functioning and service user experience.

Comparing individual CBT-ED with any other intervention in adults showed no difference in binge eating behaviour and that it has less favourable effects on remission (with some uncertainty). Positive results were found on EDE-dietary restraint, EDE-shape concern, EDE-eating concern and EDE-Global score. No difference was found in depression or EDE-weight concern.

At 12 months follow up, CBT-ED showed no effect on bingeing or remission compared with any other intervention. However, benefits were found on BMI (with some uncertainty), EDE-global and EDE-subscales. No difference in depression was found. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, body weight, family functioning and service user experience.

Interpersonal psychotherapy (IPT), which focuses on our relationship with others and how this affects our moods and vice versa, had no additional benefit on remission, bingeing and BMI compared with any other intervention at the end of treatment and at follow up. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, eating disorder psychopathology, family functioning, general psychopathology and service user experience.

Dialectical behaviour therapy had no additional benefit on bingeing compared with wait list controls. However, vomiting episodes, EDE-global and depression scores favoured dialectical behaviour therapy. No data was reported on remission, all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience.

Behavioural therapy compared with any other treatment showed remission was improved at the end of treatment (with some uncertainty), but not at follow up. All other outcomes showed no difference including bingeing, purging, EDE and EDI subscales and symptom checklist. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience.

CBT general showed no significant effect on bingeing and most of the other outcomes (including EDE and EDI subscales) measured at the end of the treatment and follow up compared with any other treatment. However, remission at follow up did favour CBT general compared with any other treatment. No data was reported on all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience.

Family therapy

One study on family therapy in adults with binge eating disorders was identified. Compared with wait list controls, benefits on binge frequency and depression were found at the end of treatment but not on weight or family functioning.

Family therapy showed similar effects as another intervention on weight, bingeing, depression and family functioning at the end of treatment. At 6 months follow up, similar results were found except family functioning was less effective in bingeing compared with another intervention.

No evidence was found on the critical outcome of remission, nor the important outcomes of general functioning, service user experience, resource use, adverse events, quality of life, all-cause mortality and relapse.

Young people with binge eating disorder

	In young people, individual CBT-ED also failed to show much benefit compared with any other intervention. Both binge eating and remission were similar between the two arms (with a trend for remission to favour any other intervention), as was BMI, EDE-weight concern, EDE-dietary restraint, EDE-shape concern and social adjustment. Depression and EDE-eating concern favoured CBT-ED at the end of treatment. One study was identified on the effects of internet self-help compared with wait list controls on young people with binge eating disorder. It showed a benefit on depression compared with wait list control but there was some uncertainty. No other differences were detected at the end of treatment or at follow up. Adverse events or all-cause mortality were not reported in any of the RCTs on adults or young people. Refer to following LETR for results on group therapy
Trade-off between net health benefits and resource use	Existing economic evidence pertaining to the psychological therapies for adults with binge eating disorder was very limited and the committee could not draw any conclusions from it. The guideline economic analysis demonstrated that self-help for an eating disorder with support is the most cost-effective individual treatment option for people with binge eating disorder. At the lower NICE cost-effectiveness threshold of £20,000 per QALY (NICE, 2008b) the probability of self-help for an eating disorder with support being cost effective was 0.83. Results were robust under all other scenarios examined in one-way sensitivity analyses. The guideline economic analysis demonstrated that CBT-ED group is the most cost-effective group treatment option for people with binge eating disorder. Also, at the lower NICE cost-effectiveness threshold of £20,000 per QALY (NICE, 2008b) the probability of CBT-ED group being cost effective was 0.74. It was impossible to compare self-help for an eating disorder with support and CBT-ED group in 1 analysis (due to the lack of common comparator between the treatments). However, the intervention costs are £238 and £317 per participant, for self-help for an eating disorder with support and CBT-ED group, respectively (in 2014/15 prices). Consequently, the committee expressed the view that self-help for an eating disorder with support should be offered as a first line treatment and CBT-ED group only if self-help for an eating disorder with support is ineffective or is unacceptable.
Quality of evidence	The evidence for psychotherapies for binge eating disorder was mostly low quality. The evidence was downgraded for imprecision and risk of bias for reasons such it was unclear how they randomised, if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. Moderate quality evidence was available for guided self-help compared with any other intervention. This comparison had some of the highest number of studies and participants available for the meta-analysis. However, some outcomes were downgraded because of indirectness (depending on the % contribution the study made to the effect size) owing to the inclusion of a mixture of people with binge eating disorder and EDNOS. Some outcomes were downgraded for heterogeneity. Nevertheless, the committee were most confident with the data on guided self-help, hence recommending it a first line treatment for adults with binge eating disorder. For the comparison of hybrid versus hybrid, remission data was excluded because the patient's symptoms were measured over one week, not over a minimum of two weeks (at the committee's request). The study was described as being combined because it included inpatient and outpatient care, in combination with diet, exercise and psychotherapy. Thus, it was difficult to isolate what the important component of the programme was. The evidence for the other treatments DBT, IPT, hybrid and family therapies was low quality and because they failed to report data on remission rates (except IPT), the committee did not consider them worth recommending. Heterogeneity was detected for remission at end of treatment for guided self-help compared with any other treatment. One study had zero events but changing the outcome to a non-event, did not significantly reduce the heterogeneity. This same

	study had a mixed population of binge eating disorder and people with EDNOS, and if removed from the analysis heterogeneity is reduced to an acceptable level. Another possible explanation for the heterogeneity in remission, is the inclusion of participants with mental health comorbidities. Both DeBar 2011 and Striegel-Moore 2010 both included participants with depression and anxiety (33 to 39%), and if they are separated from the other studies where it is unclear if the participants had significant comorbidities (not just a history of) then heterogeneity is reduced from 62% to 6%. Studies that included participants with comorbidities appeared to respond better to the treatment compared with those without. If either or both of these studies are removed from other outcomes where heterogeneity is detected, including EDE-eating concern, EDE-restraint, EDE-shape concern, EDE-weight concern at end of study and or follow up, then heterogeneity is reduced to acceptable levels (except EDE-restraint remains unexplained). Heterogeneity was not detected in any other comparison or outcome
Other considerations	Overall the evidence on treatments for binge eating disorder was limited. The majority of the evidence found was on guided self-help (7 studies, n=490) and self-help (6 studies, n=345). For other treatments, there was generally only one study for each comparison, with a small number of participants and remission was often not reported. The evidence clearly showed that guided self-help had a benefit on remission compared with wait list controls (at end of treatment) and compared with any other treatment (at both end of treatment and follow up). The committee used detail of the guided self-help programmes used in the studies to generate the recommendation on what the treatment should include. The studies generally used the same manuals. In fact, the type of manuals used was how committee decided to group the studies to begin with. The committee were hoping for stronger evidence to support a recommendation for internet self-help given the move towards internet and smart-phone based treatments. Compared with wait list controls, internet self-help showed greater improvements in remission rates in adults at the end of treatment and at follow up. However, this was based on one study with only 74 participants. Compared with any other intervention, internet self-help had a similar effect on binge frequency but remission was not reported. The committee noted the lack of evidence on individual CBT-ED in this population. In adults, only one study provided data on remission at the end of treatment and follow up. There was also some discussion that individual CBT-ED is a useful treatment for those with complex binge eating disorder. However, no RCT evidence on this population was found. Owing to the lack of evidence, the committee generated a research recommendation to “compare the clinical and cost effectiveness of individual eating-disorder focused cognitive behavioural therapy (CBT-ED) with guided self-help and group CBT-ED for adults with binge eating disorder, including complex binge eating disorder

9. Research recommendation: Compare the clinical and cost-effectiveness of individual eating-disorder focused cognitive behavioural therapy (CBT-ED) with guided self-help and group CBT-ED for adults with binge eating disorder, including complex binge eating disorder.

Second-line psychological treatments for adults with BED

	129. If guided self-help is ineffective after four weeks or is not acceptable, offer group eating-disorder-focused cognitive behavioural therapy (CBT-ED) **130. Group CBT-ED programmes for adults with binge eating disorder**

	should: • use a CBT-ED manual • consist of 16 weekly 90-minute group sessions over four months • focus on psychoeducation, self-monitoring of the eating behaviour and helping the person analyse their problems and goals • include making a daily food intake plan and identifying binge eating cues • include body exposure training and helping the person to identify and change negative beliefs about their body • help with avoiding relapses and coping with current and future risks and triggers. **131. Explain to people with binge eating disorder that psychological treatments aimed at treating binge eating have a limited effect on body weight and that weight loss is a post-therapy target. Refer to the NICE guideline on obesity identification, assessment and management for guidance on weight loss and bariatric surgery.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating binge eating disorder. For this population, it was agreed binge eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	When considering the effectiveness of group therapy it is best to first observe how it compares to no treatment or wait list controls. Group CBT-ED reduces binge eating compared with waitlist controls, but it has no effect on weight or depression at the end of treatment. At follow up, benefits were still found on binge eating but also on depression. No difference in weight was found. Remission data was not available. Group mindfulness compared with waitlist controls showed favourable results on remission and bingeing, in addition to depression. But no difference was found in BMI between the two arms. The same trends were found at long-term follow up. When comparing group mindfulness to any other intervention there was no difference in the outcomes between the two arms for remission, bingeing and BMI at the end of treatment and at follow up. Comparing group CBT-ED with any other intervention, showed positive results on remission rates and a trend for improving binge eating at the end of treatment. EDE-weight concern also showed favourable results. Some outcomes showed less favourable results, EDE-global, EDE-eating concern (trend) and weight. However, the remainder were no different between the two arms including: depression, anxiety, EDE-dietary restraint, EDE-shape concern, and EDE-global score. At follow up, the benefits on remission and binge eating were no longer evident. Weight, anxiety, EDE-global and the EDE-subscales all favoured the other interventions. No difference in depression or global symptom score was found. Group behavioural therapy is more effective on depression compared with wait list controls, but there is no difference in the effect on bingeing, remission, EDE-total

	or anxiety. Compared with another group treatment, group behavioural therapy (ED) is more effective on EDE- eating concern, EDE-dietary restraint and remission at the end of treatment. No difference between the two treatment arms was found on depression, BMI and weight loss and only a trend to favour EDE-shape concerns and EDE-weight concern. At follow up, group behavioural therapy (ED) may be more effective on EDE-dietary restraint and remission compared with another group therapy but there was some uncertainty. Other outcomes showed no difference between the two treatment arms. They included: depression, BMI, weight loss, EDE-shape concern and EDE-eating concern compared with any other intervention. EDE-dietary restraint favoured group behavioural therapy. A head-to-head of group CBT-ED with another group CBT-ED that focused more body exposure (looking at your body under different conditions) showed no difference in their effectiveness at the end of treatment on remission or bingeing behaviour, nor on depression, BMI or any of the EDE-subscales. The same results were found at long-term follow up. Group interpersonal psychotherapy (IPT) showed no additional benefit on remission or binge eating compared with any other intervention. Similarly, no difference was found in BMI, EDE-subscales or global symptom index but a trend to improve depression but not EDE-restraint. At follow up no differences were found (except EDE-restraint that showed a negative effect). Other comparisons included group counselling compared with any other intervention that showed favourable effects on remission at end of treatment but no other benefit. EDE-subscales all favoured any other intervention, as did the preference of the person receiving treatment. At long-term follow up, interestingly remission now favoured any other treatment (but there was some uncertainty) and no difference was found in BMI, weight loss or depression. Again, EDE-subscales (except EDE-dietary restraint and EDE-weight concerns) and preference of the person receiving treatment favoured any other intervention. Group diet compared with any other group intervention showed negative results for group diet. Remission rates and bingeing (with some uncertainty) favoured the other intervention, as did depression. EDE-subscales showed no difference between the two treatment arms. However, BMI was lower. At follow up, remission still favoured the other group and bingeing also showed a trend to favour the other arm, but no differences were found in all other outcomes. Group self-help (ED) compared with any other intervention showed a benefit on remission and a trend to reduce BMI, but no benefit on bingeing and most of the other outcomes favoured the other arm. At follow up, most outcomes showed no difference. When compared with any other intervention, group guided self-help (ED) still showed a benefit on bingeing but no difference on remission was found. Depression and EDE-restraint also favoured group guided self-help. No difference in EDE-global, EDE-shape concern, EDE- eating concern or quality of life was found. EDE-weight concern favoured any other intervention. At follow up all the benefits were no longer evident and all outcomes were similar between the two treatment arms, except EDE-shape concern, EDE-eating concern and EDE-weight concern that favoured the other treatment arm. Group self-help (ED) compared with wait list controls showed a trend for a reduction in bingeing but no difference in BMI, quality of life, EDE-global or EDE-subscales. When compared with any other intervention, it shows mixed results. Remission rates were improved in group self-help but bingeing behaviour. BMI, depression, quality of life, EDE-shape concern and EDE-eating concern all showed no difference between the two arms. EDE-global, EDE-restraint, EDE-weight concern all favoured any other intervention. Group guided self-help (ED) compared with waitlist control showed a benefit on bingeing and EDE-restraint but no other outcome including BMI (trend to be higher), quality of life, EDE-global and the remaining EDE subscales. Group psychoeducation compared with any other intervention showed no difference in the effect on remission or binge eating at the end of treatment. A

	positive effect was found on EDE-global but no other outcomes including the EDE-subscales, BMI, depression and quality of life. At follow up, still no difference in remission or bingeing was detected. However, some positive effects were found for depression, EDE-global and EDE-subscales (except eating concern that showed no difference). No difference in BMI or quality of life was found. Adverse events or all-cause mortality were not reported in any of the studies. Refer to previous LETR for results on individual, family and self-help therapies.
Trade-off between net health benefits and resource use	Existing economic evidence pertaining to the psychological therapies for adults with binge eating disorder was very limited and the committee could not draw any conclusions from it. The guideline economic analysis demonstrated that self-help for an eating disorder with support is the most cost-effective individual treatment option for people with binge eating disorder. At the lower NICE cost-effectiveness threshold of £20,000 per QALY (NICE, 2008b) the probability of self-help for an eating disorder with support being cost effective was 0.83. Results were robust under all other scenarios examined in one-way sensitivity analyses. The guideline economic analysis demonstrated that CBT-ED group is the most cost-effective group treatment option for people with binge eating disorder. Also, at the lower NICE cost-effectiveness threshold of £20,000 per QALY (NICE, 2008b) the probability of CBT-ED group being cost effective was 0.74. It was impossible to compare self-help for an eating disorder with support and CBT-ED group in one analysis (due to the lack of common comparator between the treatments). However, the intervention costs are £238 and £317 per participant, for self-help for an eating disorder with support and CBT-ED group, respectively (in 2014/15 prices). Consequently, the committee expressed the view that self-help for an eating disorder with support should be offered as a first line treatment and CBT-ED group in cases where self-help with support is ineffective or unacceptable.
Quality of evidence	The evidence for psychotherapies for binge eating disorder was mostly low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as it was unclear how they randomised, if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. When comparing the different therapies for binge eating disorder, the results show after guided self-help, group CBT-ED is the most effective. It was one of the few interventions that showed convincing results on remission at the end of treatment (22% increase in remission rates) and a trend for benefits to be sustained at long-term follow up (14%). (Note: the most convincing effects on remission at follow up was found in the guided self-help-treated groups). Remission was not always reported or the definition did not meet the criteria set by the committee (symptoms measured over a minimum of two weeks). Studies by Agras 1994, Wilfrey 1993, Klein 2014 and Telch 1990 on group therapy were excluded from the analysis for this reason. Heterogeneity was detected for remission at follow up in the group CBT-ED versus any other intervention. Only three studies contributed to the outcome and none appeared to show a high risk of bias. One study (Wilfley 2002) clearly had a population with comorbid mental health conditions (30% Axis I conditions) while it was unclear if any were included in Nuata 2000, and in Munsch 2007 they were excluded. When Wilfley 2002 was separated into another sub-group, the heterogeneity was reduced to acceptable levels. Heterogeneity was also detected for a number of EDE-subscale outcomes. In all of these studies, Ricca 2010 was clearly an outlier and was the only study where the other active arm was individual CBT-ED. All other treatment arms were different types of group therapy, for example group interpersonal psychotherapy or group behavioural weight loss. For the network meta-analysis it was thought that there would be a bias in the literature towards individual CBT-ED as the preferred outcome. Therefore, when Ricca 2010 was removed on the grounds of a bias towards individual over group CBT-ED, the heterogeneity was reduced to acceptable levels for the following outcomes: EDE-Global clinical score (0%), EDE-

	dietary restraint (60%), EDE-eating concern (0%), but not EDE-weight concern. At follow up: EDE-Global clinical score (0%), EDE-dietary restraint (4%), EDE-shape concern (0%) EDE-weight concern (3%), EDE-weight concern (0%). For group CBT-ED versus waitlist controls, heterogeneity was found in two outcomes: depression and bingeing. For depression, there was a risk of bias in the study by Telch 1990 because a number of symptom-related outcomes were measured over only seven days. When it was removed from the analysis for depression, heterogeneity was reduced to an acceptable level (30%). For bingeing, the study by Shapiro 2007 had a high dropout rate (32-40%) and because of the risk of bias it may explain the heterogeneity. Both studies had a similar severity of illness based on the number of binges per month, so this could not explain the heterogeneity. Comorbidies were similar and duration of illness was not reported. Heterogeneity was not detected in any other group-related comparisons.
Other considerations	Taking into account all the evidence, the committee decided to first recommend guided self-help (ED) and if it is ineffective after four weeks of treatment or is not acceptable to adults with binge eating disorder, offer group cognitive behavioural therapy (CBT-ED). A number of committee members were not in agreement with recommending group CBT-ED, because the studies had almost exclusively been conducted in the USA with people who had comorbid obesity and who were recruited by advertising. This raised the question of whether the findings could be generalised to people with binge eating disorder in NHS settings. It was discussed that people who respond to group CBT-ED were also likely to respond to guided self-help, so the recommendations do not address the needs of people who present with more complex conditions and require one-to-one treatment rather than a generic group treatment. The other concern raised was the difficulty that would be faced in forming groups for the therapy because most eating disorder services find that people with binge eating disorder sporadically register for treatment. In most of the studies, weight or BMI did not differ at the end of treatment between two active treatment arms or even with waitlist controls. Thus, highlighting that psychotherapy is not very effective in producing weight loss, thus supporting the view that only focusing on weight loss is an inappropriate goal for obese people. For these reasons, the committee wanted to highlight the importance of explaining to people with binge eating disorder that psychological treatment alone has a limited effect on weight and that weight loss is a post-therapy target. **Assessment after 4 weeks** The GC discussed how there is RCT evidence that shows an assessment during the early stages of a psychological treatment can predict the likelihood of a full-response at the end of treatment. For this reason they felt it was important to have an early assessment after 4 weeks and if the person is not showing signs of responding then they should be offered CBT-ED. There is no justification for ending treatment early, after only 4 weeks, because the magnitude of the response at this stage is not the same as that seen at the end of treatment, i.e. 16 weeks. For this reason, it is important that if the person is showing early signs of responding to treatment that they continue-on to the end of the programme. Furthermore, there is no evidence on people with an eating disorder that 4 weeks of treatment is sufficient to show long-term benefits. Also, the programme is delivered in a way that different elements are explored over the full 16 weeks and this can not be shortened. The GC agreed that it is important to have this recommendation because there is a concern that people with a severe case of binge eating disorder may be better suited to CBT-ED rather than guided self-help. An early assessment should detect these patients and ensure that they receive a treatment that may be more cost-effective.

1 **Psychological treatment for young people with BED**

	132. For young people with binge eating disorder, offer the same treatments recommended for adults with binge eating disorder.

Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating binge eating disorder. For this population, it was agreed binge eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	In young people with binge eating disorder, individual CBT-ED showed no difference in its effectiveness compared with any other treatment (treatment as usual). Both binge eating and remission were similar between the two arms (with a trend for remission to favour any other treatment), as was BMI, EDE-weight concern, EDE-dietary restraint, EDE-shape concern and social adjustment. CBT-ED did show favourable results on depression and EDE-eating concern at the end of treatment. No data was available at follow up. One study was identified on the effects of internet self-help compared with wait list controls on young people with binge eating disorder. It showed no additional benefit on bingeing compared with wait list controls but a trend to improve depression. No difference were found at follow up. Remission was not reported. Adverse events or all-cause mortality were not reported in any of the studies.
Trade-off between net health benefits and resource use	The committee noted that there is a lack of evidence on the effectiveness and cost effectiveness of interventions for young people with BED. Consequently, the committee extrapolated the cost effectiveness of interventions from the economic analyses conducted for this guideline for adults with BED. According to the economic analyses, for adults with BED self-help with support was the most cost effective individual treatment and CBT-ED group was the most cost effective group treatment. It was impossible to compare self-help with support and CBT-ED group in one analysis (due to the lack of common comparator between the treatments). However, the intervention costs are £238 and £317 per participant, for self-help for an eating disorder with support and CBT-ED group, respectively (in 2014/15 prices). Consequently, the committee expressed the view that self-help for an eating disorder with support should be offered as a first line treatment and CBT-ED group in cases where self-help with support is ineffective or is unacceptable.
Quality of evidence	The evidence for psychotherapies for binge eating disorder was mostly low quality. The evidence was downgraded for imprecision and risk of bias for reasons such it was unclear how they randomised, if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. In both comparisons, only one study was identified and a small number of participants were included (n=26 to 93).
Other considerations	The evidence on treatments for young people with binge eating disorder was very limited. Only one study on CBT-ED with 26 participants was identified, only just reaching the minimum number required to be included in the review. In addition, only one study on internet self-help was found with 93 participants and no data on remission was reported. Because the evidence on young people was not strong enough to generate a recommendation, the committee agreed to extrapolate the evidence from adults with binge eating disorder to young people. They felt it was important to include a recommendation on how to treat this population, rather than a research recommendation alone, given the number of young people with this disorder and the rising problem of obesity in young people. The committee discussed there may be some cohort studies on this population, but the protocol only allowed the inclusion of observational studies in the absence of RCT evidence. They also discussed that historically this population may have been labelled as having “loss of control eating” behaviour. And that it wasn’t until DSM-5 in 2013 that the definition of binge eating disorder and EDNOS became clear, so

	the lack of studies on this population may in part be due to the legacy of how they were classified in older studies. Due to the lack of evidence on children and young people with binge eating disorder the committee decided to generate a research recommendation to: "investigate the clinical and cost effectiveness of psychological treatments for children and young people with binge eating disorder".

10. Research recommendation: Investigate the clinical and cost effectiveness of psychological treatments for children and young people with binge eating disorder.

8.3 Carer interventions

8.3.1 Review question: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 276. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all psychological interventions for the parents or carers of children or young people with an eating disorder. The interventions were categorised according to their mode of delivery (e.g. group, self-help), the age of the people with the eating disorder, and the type of eating disorder and were compared to wait list controls, treatment as usual or any other intervention.

Table 276: Clinical review protocol summary for the review of: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Component	Description
Review question(s)	Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?
Population	• Family or carers of people with an eating disorder (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder)
Intervention(s)	Psychological interventions may include: • Family-based • Parent only (not necessarily focussed on eating disorder) • Parent-focused therapy (PFT) • Group Parent-Training (GPT) • Separated family therapy • Parents with person with ED (greater focus on eating disorder) • Behavioural Family Therapy (BFT) • Behavioural family systems therapy (BFST). • Family Based Treatment (FBT) • Family Day Workshops (FDW) • Family Therapy (FT) • Family therapy for anorexia nervosa (FT-AN) • Multi-Family Group Day Treatment (MFGDT) • Multi-Family Group Therapy (MFGT)

Component	Description
	• Systemic Family Therapy (SFT) • Systemic Family Therapy for AN (SFT-AN) • Multifamily therapy (MFT) is synonymous with (MFGT; MFGDT). • Uniting couples in the treatment of AN (UCAN • Conjoint family therapy
Comparison	Wait list control Treatment as usual Another intervention
Critical outcomes	• Parent's or carer's general psychopathology (including mood/depression/anxiety) • Family functioning • Quality of life • Other primary outcomes commonly reported in studies that just target the family/carer • The following outcomes will be included if the family or carer intervention includes the child or person with an eating disorder: • Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• General functioning • Resource use. • Service user experience • All-cause mortality. • Adverse events • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion)
Study design	• Systematic Reviews • RCTs

8.3.2 Clinical Evidence for Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

No studies were found that met the eligibility criteria for this review.

8.3.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with binge eating disorder was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.3.4 Clinical evidence statements

No studies were found that met the eligibility criteria for this review.

8.3.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with binge eating disorder was available.

8.3.6 Recommendations and link to evidence for the review on: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Working with family members and carers

	133. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes, when assessing whether any interventions help the parents and carers of children and young people with an eating disorder. The critical outcomes for the parents and carers were: general psychopathology, family functioning, quality of life and other primary outcomes reported by the study. Other outcomes that are critical for the child or young person with the eating disorder include remission and bingeing or body weight, depending on the eating disorder. Other outcomes that are of lesser importance but clearly important outcomes include, general functioning, service user experience, all-cause mortality, adverse events and eating disorder psychopathology.
Trade-off between clinical benefits and harms	**Bulimia nervosa or binge eating disorder (chapter 7 and 8)** No relevant published evidence was found on parents or carers of children and young people with bulimia nervosa or binge eating disorder. **Any eating disorder (chapter 9)** Randomised control trials investigating interventions for the carers of young people with any eating disorder failed to show many favourable outcomes. Psychoeducation compared with waitlist control showed a positive effect on carer self-efficacy and a trend to improve carer knowledge of eating disorders at the end of treatment. Long-term follow up (unclear duration) showed favourable results in both but carer burden (only measured at follow up) was not different compared with wait list controls. No evidence was found on the critical outcomes of carer general psychopathology, family functioning, and quality of life, nor on the other important outcomes. Comparing guided self-help with self-help showed no difference in any of the carer-related outcomes at the end of treatment. No evidence was found on the other important outcomes. **Anorexia nervosa (evidence in chapter 6)** One randomised controlled trial (RCT), aimed at carers of young people with anorexia nervosa, and compared the effectiveness of guided self-help or self-help (and treatment as usual) with treatment as usual alone. After 12 months there was no difference in carer general psychopathology. No evidence was found on the critical outcomes of carer general psychopathology, carer family functioning, carer quality of life, nor the important outcomes of eating psychopathology, carer general functioning, service user experience, resource use, adverse events, and all-cause mortality. Another study compared self-help (with treatment as usual) with treatment as usual and showed no difference in the carer's general psychopathology or carer skills after 6 to 12 months but a trend for poorer results for family functioning but there was some uncertainty. In the young people with anorexia nervosa whom they care for, there was no difference in BMI, weight, severity index (SEED), general psychopathology, clinical improvement, peer related problems between the 2 treatment arms. However, there was a trend for poorer outcomes in prosocial behaviour in the self-help group but there was some uncertainty. No evidence was found on the critical outcomes of remission, carer general psychopathology, nor the important outcomes of service user experience, resource use, adverse events, and all-cause mortality. Comparing guided self-help (and treatment as usual) with treatment as usual

	showed at 12 months a trend for positive outcomes in the combined treatment group on carer burden and quality of life, but no difference in family functioning, carer skills or carer psychopathology. There was a trend for poorer outcomes in carer accommodation and enabling. At 24 months, there was a trend for a positive result on carer burden, quality of life, carer accommodation and enabling, and carer psychopathology. And a trend for poorer outcomes in family functioning and time spent caring. No evidence was found on the critical outcome of carer general psychopathology, nor the important outcomes of service user experience and resource use. In the same intervention, the guided self-help for the carers did not translate to many benefits in the young people with anorexia nervosa whom they care for. At 12 months, no differences were found in any of the outcomes including mortality, admission to hospital, patient relapse, BMI, EDE-global, severity index (SEED), general psychopathology, clinical improvement. However, there was improvement in peer problems but a trend for a negative result in prosocial behaviour. At 24 months, there was a trend for positive increase in BMI and EDE-global, no difference in general psychopathology and a trend for a negative result in quality of life. No evidence was found on the critical outcome of remission, nor the important outcomes of adverse events, and all-cause mortality. Comparing two active treatments generally showed no difference in effectiveness in the carer-related outcomes. Guided self-help compared with self-help were equally effective on all outcomes 6 to 12 months after the young people with anorexia nervosa had been discharged from inpatient care, except there was a trend for carer accommodation to favour guided self-help. No evidence was found on the critical outcome of remission, nor the important outcomes of adverse events, and all-cause mortality. In the young people with anorexia nervosa, there was a trend for poorer results on BMI and peer problems in the guided self-help group compared with self-help. No difference was found in clinical severity (SEED), general psychopathology, clinical improvement, prosocial behaviour but there was a trend for better results in peer problems. Web-based guided self-help also failed to show convincing benefits for the carers of young people with anorexia nervosa compared with treatment as usual. At the end of treatment, a poorer outcome in distress was found but there was some uncertainty. The other outcomes, such as carer accommodation and enabling, family functioning, carer burden and caregiving experience showed no difference. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. Web-based guided self-help compared with web-based self-help showed no difference in the outcomes for carers at the end of treatment. At follow up, favourable results were found on family functioning in the guided web-based self-help group, but no difference in carer experience, quality of life, and general psychopathology. There was a trend for poorer results in carer burden, but there was some uncertainty. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. Adverse events or all-cause mortality were not reported in any of the studies.
Trade-off between net health benefits and resource use	The committee expressed the view that offering family members and carers an assessment of their own needs may incur additional healthcare resources (that is, time required to perform such assessment). However, the committee considered the cost of providing such assessment to be small, taking into account the potential reduction in family and carers' burden, potential depression and other health vulnerabilities which may be costly to other parts of the healthcare system, especially considering that the burden on family and carers can last for many years and increase their morbidity and stress. Consequently, the committee judged that assessment that aims to improve family and carers' experience are likely to represent good value for money.
Quality of evidence	The quality of the evidence was mostly very low. The outcomes were downgraded because it was unclear how they randomised, if allocation concealment was performed or if participants were blinded. In some, not all, assessors were blinded. High dropout rates were also detected in some groups >20%.

	Imprecision was detected in most outcomes due to the 95% confidence interval crossing one or two minimal important differences or because it did not meet the optimal information size. Outcomes were not always measured at the end of treatment or at follow up. It is not known if any improvements in the carer's general psychopathology also translated to benefits in the children with the eating disorder.
Other considerations	Given the very low quality of the data with very few positive findings favouring one arm over the other, the committee came to the consensus that there was not enough evidence to support a recommendation on any specific treatment for parents or carers of people with an eating disorder. Nevertheless, the committee acknowledged the stress and burden that a person with an eating disorder, in particular anorexia nervosa, can have on family members over a long period of time. Therefore, they agreed that offering family members and carers an assessment of their own needs, including: personal, social and emotional support available to them, the need for support in the caring role for example if the child should need urgent care and there are other children to take care of and to offer advice on where they can get some practical support. The extent to which the family need to be involved in treatment depends on the age and developmental needs of the person with the eating disorder, the severity of the illness, the risk from harm and the person receiving treatment's wishes. In general, parents and other family members will want to be involved in the treatment. If a parent or carer cannot attend a meeting the healthcare professional should provide written information on the outcome of an assessment or treatment where appropriate. The committee acknowledged the importance of consent and confidentiality and their discussion can be found in the LETRs relating to this. They also discussed that although the evidence found was for carers and parents of people with anorexia nervosa or any eating disorder, the recommendation is relevant for parents and carers of people with bulimia nervosa and binge eating disorder. This is mostly because no specific intervention was recommended, rather to offer an assessment of their needs and to help them find the necessary support. In absence of good evidence, the committee agreed to generate a research recommendation to address the question "What is the effectiveness of a carer-focused psychological intervention in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?" See chapter 6.2.

8.4 Pharmacological interventions

8.4.1 Review question: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 277. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all pharmacological interventions that may be delivered to children, young people and adults with an eating disorder with or without a psychological intervention. The interventions were categorised according to the type of pharmacological intervention, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to placebo, wait list controls TAU or any other intervention.

Table 277: Clinical review protocol summary for the review of: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

Component	Description

Component	Description
Review question(s)	Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). • Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder) ○
Intervention(s)	• Pharmacological interventions may include: • Antidepressants, e.g. SSRIs, fluoxetine (Prozac) • Anxiolytics (antianxiety) • Antipsychotics • Antiemetic medication, e.g. ondansetron • Antiepileptic/anticonvulsant, e.g. topiramate (Topomax) • Appetite suppressant, e.g. lisdexamfetamine dimesylate • Pharmacological in combination with any psychological intervention
Comparison	Placebo Wait list control Treatment as usual Another intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN • Adverse events
Important outcomes	• All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Relapse • Resource use • Quality of life • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs

8.4.2 Clinical Evidence Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

23 RCTs (n=2544) met the eligibility criteria for this review (Arnold 2002 (Arnold et al., 2002), Grilo 2005 (Grilo et al., 2005)/2012(Grilo et al., 2012), Guerdjikova 2008 (Guerdjikova et al., 2008), Guerdjikova 2009 (Guerdjikova et al., 2009), Guerdjikova 2012 (Guerdjikova et al., 2012), Hudson 1998 (Hudson et al., 1998), McElroy & Casuto 2000 (McElroy et al., 2000), McElroy, Arnold & Shapira 2003 (McElroy et al., 2003a), McElroy, Guerdjikova & Kotwal

2007 (McElroy et al., 2007a), McElroy, Guerdjikova & Blom 2013 (McElroy et al., 2013), McElroy, Guerdjikova & Winstanley 2011 (McElroy et al., 2011), McElroy & Guerdjikova 2015 (McElroy et al., 2015a), McElroy, Hudson & Capece 2007 (McElroy et al., 2007b), McElroy, Hudson & Malhotra 2003(McElroy et al., 2003b), McElroy, Hudson & Mitchell 2015 (McElroy et al., 2015b), McElroy & Hudson 2016 (McElroy et al., 2016), McElroy, Kotwal & Guerdjikova 2006(McElroy et al., 2006), White 2013 (White and Grilo, 2013)). All of the studies were in adults and the majority of participants were obese females. The majority of studies examined the efficacy of antidepressants versus placebo in treating binge eating disorder and used flexible rather than fixed doses.

Eight studies (n=470) compared an antidepressant with placebo (Arnold 2002, Grilo 2005/2012, Guerdjikova 2008, Guerdjikova 2012, Hudson 1998, McElroy, Hudson & Malhotra 2003, McElroy 2000 (McElroy et al., 2000), White 2013). Two studies (n=150; Leombruni 2008 (Leombruni et al., 2008), Ricca 2001) compared one antidepressant with another, fluoxetine with sertraline and fluvoxamine, respectively.

Two studies (n=173) compared an antidepressant with individual therapy (Molinari 2005 (Molinari et al., 2005), Ricca 2001). Four studies (n=579), all of which used obese participants, compared an antiepileptic with placebo (Guerdjikova 2009, McElroy & Arnold 2003 (McElroy et al., 2003a), McElroy 2006 (McElroy et al., 2006), McElroy & Hudson 2007 (McElroy et al., 2007b)).

Three studies (n=1033) compared an appetite suppressant, lisdexamfetamine dimesylate, a dextroamphetamine prodrug approved for the treatment of ADHD in children, with placebo (McElroy & Hudson 2015 (Hudson et al., 2015), McElroy & Hudson 2016 (McElroy et al., 2016)). McElroy & Hudson 2016 examined the efficacy of fixed doses of 30, 50 and 70 mg with placebo. The two studies reported in McElroy & Hudson 2016 used flexible doses, starting from 30 mg/day and increasing up to 70 mg/day if tolerated and clinically needed/ Two studies (n=109) compared a substance abuse treatment agent with placebo (McElroy 2011 (McElroy et al., 2011), McElroy 2013 (McElroy et al., 2013)).

Two studies were found that did not fit the above categories (McElroy & Guerdjikova 2007 (McElroy et al., 2007a), McElroy & Guerdjikova 2015 (McElroy et al., 2015a)). McElroy & Guerdjikova 2007 (n=40) compared atomoxetine, a norepinephrine reuptake inhibitor, with placebo. McElroy & Guerdjikova 2015 (n=60) compared armodafinil, an active isomer of modfinil, with placebo.

8.4.2.1 Combined pharmacological and psychological interventions

Six studies (n=434) were identified that compared a combined pharmacological and psychological intervention with any other intervention (Claudino 2007 (Claudino et al., 2007), Cristina 2014, Grilo 2005 (Grilo et al., 2005)/2012(Grilo et al., 2012), Grilo & Masheb 2005, Molinari 2005, Ricca 2001 (Ricca et al., 2001b)). The majority of the participants were adult females. Three of the 6 studies (n=203) examined adjunctive antidepressant treatment to CBT with CBT alone (Cristina 2014 (Cristina et al., 2014), Molinari 2005 (Molinari et al., 2005), Ricca 2001 (Ricca et al., 2001a)). One study (n=108) compared fluoxetine and CBT-ED with placebo and CBT-ED (Grilo 2005; Grilo 2012). Two studies (n=138) compared a combined antidepressant and CBT intervention with another combined antidepressant and CBT intervention (Cristina 2014, Ricca 2001). One study (n=73) compared an antiepileptic combined with group CBT-ED with placebo and group CBT-ED (Claudino 2007 (Claudino et al., 2007)). One study (n=30) compared adjunctive topiramate treatment to sertraline, 1700 kcal diet and Group CBT with sertraline, 1700 kcal diet and Group CBT only (Brambilla 2009 (Brambilla et al., 2009)). One study (n=50) compared a combined antiobesity agent orlistat and guided self-help CBT-ED with placebo and guided self-help CBT-ED (Grilo & Masheb 2005).

An overview of the trials included in the meta-analysis can be found in Table 278. Summary of findings tables can be found in Table 289, Table 290, Table 291, Table 292, Table 293,

Table 294, Table 295, Table 296, Table 297 and Table 298. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 278: Study information for trials included in the meta-analysis of a pharmacological intervention versus placebo, wait list control, or another intervention for adults with binge eating disorder.

Study ID	N Randomized	Female (%)	Mean BMI (kg/m2) or Weight (kg)	Sample	Type(s) of pharmacological intervention	Intervention Age at onset and/or duration of illness (years)	Max dose/ day of intervention	Duration (weeks)	Comparison(s) Age at onset and/or duration of illness (years)	Duration or max dose/ day of comparison(s)
Arnold 2002	60	93	BMI=38.15 (7)	DSM-IV	AD - SSRI	Fluoxetine Duration of illness: 19.9 (12.5)	80 mg	6	Placebo Duration of illness: 16.7 (9.5)	na
Brambilla 2009	30	100	BMI=35.7 (6.0)	DSM-IV and obesity	AD – SSRI, AE,	Sertraline + Topiramate + Group CBT-ED + 1700 kcal Diet Duration of illness: 15 (10)	150 mg for both	26	Sertraline + Group CBT-ED + 1700 kcal diet Duration of illness: 9 (5) Group CBT-ED Duration of illness: 13 (6)	150 mg
Claudino 2007	73	96	BMI=37.4 (4.2)	DSM-IV-TR with obesity	AE	Topiramate + Group CBT-ED	200 mg	21	Placebo + Group CBT-ED	na
Cristina 2014	30	Not reported	Not reported	Not reported	AD – SSRI, SNRI	Paroxetine + CBT-ED	Not reported	52	1. CBT-ED 2. Venlafaxine + CBT-ED	2. not reported
Grilo & Masheb 2005	50	88	Not reported	DSM-IV BED with obesity	AO	Orlistat + Guided Self-Help CBT-ED	360 mg	12	Placebo + Guided Self-Help CBT-ED	na
Grilo 2005/2012	108	78	BMI=36.3 (7.9)	DSM-IV	AD - SSRI	Fluoxetine Age at onset: 24.5 (11.9)	60 mg	16	Placebo Age at onset: 23.8 (19) Fluoxetine (SSRI) + CBT-ED Age at onset: 22.4 (13)	60 mg

Study ID	N Randomized	Female (%)	Mean BMI (kg/m2) or Weight (kg)	Sample	Type(s) of pharmacological intervention	Intervention Age at onset and/or duration of illness (years)	Max dose/ day of intervention	Duration (weeks)	Comparison(s) Age at onset and/or duration of illness (years)	Duration or max dose/ day of comparison(s)
									Placebo + CBT-ED Age at onset: 25.9 (18.1)	
Guerdjikova 2008	44	98	BMI=40.20 (5.77)	DSM-IV with obesity	AD - SSRI	Escitalopram	30 mg	12	Placebo	na
Guerdjikova 2009	51	76	BMI=40.09 (6.45)	DSM-IV with obesity	AE	Lamotrigine Age at onset: 29.77 (16.06)	200 mg	16	Placebo Age at onset: 21.44 (15.32)	na
Guerdjikova 2012	40	88	BMI=40.6 (7.4)	DSM-IV with comorbid depressive disorder	AD - SNRI	Duloxetine	120 mg	12	Placebo	na
Hudson 1998	85	91	BMI=35.(7.3)	Draft criteria for DSM-IV BED	AD - SSRI	Fluvoxamine	300 mg	9	Placebo	na
Leombruni 2008	42	100	BMI=39.3 (3.5)	DSM-IV with obesity	AD - SSRI	Fluoxetine	80 mg	24	Sertraline	200 mg
McElroy 2000	34	94	BMI=36.12 (7.3)	DSM-IV	AD - SSRI	Sertraline	200 mg	6	Placebo	na
McElroy & Arnold 2003	61	87	BMI=43.1 (6.9)	DSM-IV-TR with obesity	AE	Topiramate	600 mg	14	Placebo	na
McElroy 2006	60	88	BMI=41.7 (8.6)	DSM-IV-TR with obesity	AE	Zonisamide Duration of illness: 19 (13.8)	600 mg	16	Placebo Duration of illness: 17.9 (12.9)	na
McElroy & Guerdjikova 2007	40	83	BMI=39.4 (7.8)	DSM-IV-TR	NRI	Atomoxetine	120 mg	10	Placebo	na

Study ID	N Randomized	Female (%)	Mean BMI (kg/m2) or Weight (kg)	Sample	Type(s) of pharmacological intervention	Intervention Age at onset and/or duration of illness (years)	Max dose/ day of intervention	Duration (weeks)	Comparison(s) Age at onset and/or duration of illness (years)	Duration or max dose/ day of comparison(s)
McElroy & Hudson 2007	407	86	BMI=38.5 (5.3)	DSM-IV with obesity	AE	Topiramate Age at onset: 25.4 (13.5) Duration of illness: 18.6 (14.3)	400 mg	16	Placebo Age at onset: 23.9 (14.3) Duration of illness: 20.6 (14.5)	na
McElroy 2011	40	85	BMI=39.5 (0.30)	DSM-IV-TR	SATA	Acamprosate	2997 mg	10	Placebo	na
McElroy 2013	69	90	BMI=38.95 (5.8)	DSM-IV-TR with obesity	SATA	ALK-33 (Samidorphan)	10 mg	6	Placebo	na
McElroy & Guerdjikova 2015	60	85	BMI=40.1 (8.0)	DSM-IV TR + EDE-Q	WPA	Armodafinil	250 mg	10	Placebo	na
McElroy & Hudson 2015	260	82	BMI=34.9 (5.3)	DSM-IV-TR	AS	Lisdexamfetamine dimesylate	30, 50 or 70 mg	11	Placebo	na
McElroy 2016 Study 1	383	86	BMI=33.5 (6.3)	DSM-IV-TR + CGI-S≥4	AS	Lisdexamfetamine dimesylate	70 mg	12	Placebo	na
McElroy 2016 Study 2	390	80	BMI=33.5 (6.3)	DSM-IV-TR + CGI-S≥4	AS	Lisdexamfetamine dimesylate	70 mg	12	Placebo	na
Molinari 2005	65	Not reported	BMI=38.4 (3.8) (n=60)	DSM-IV	AD - SSRI	Fluoxetine	60 mg	54	CBT Fluoxetine + CBT	60 mg
Ricca 2001	108	59	BMI=32.3 (5.8)	DSM-IV	AD - SSRI	Fluoxetine Duration of illness: 5.1 (4.1)	60 mg	24	CBT-ED Duration of illness: 6.4 (6) Fluvoxamine Duration of illness: 5.3 (4.8) Fluoxetine (SSRI) +	300 mg 60 mg 300 mg

Study ID	N Randomized	Female (%)	Mean BMI (kg/m2) or Weight (kg)	Sample	Type(s) of pharmacological intervention	Intervention Age at onset and/or duration of illness (years)	Max dose/ day of intervention	Duration (weeks)	Comparison(s) Age at onset and/or duration of illness (years)	Duration or max dose/ day of comparison(s)
									CBT-ED Duration of illness: 4.9 (5.1) Fluvoxamine + CBT-ED Duration of illness: 4.8 (4.4)	
White 2013	61	100	BMI=35.8 (6.8)	DSM-IV-TR, overweight and obese	AD - NDRI	Buproprion	300 mg	8	Placebo	na

Abbreviations: AD, antidepressant; AE, antiepileptic/anticonvulsant; AO, antiobesity; AS, appetite suppressant; BED, Binge Eating Disorder; BMI, Body Mass Index; CBT, Cognitive Behavioural Therapy; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; DSM-IV-TR, Diagnostic and Statistical Manual of Mental Disorders, 4th Edition-Text Revision; na, not available; NRI, Norepinephrine Reuptake Inhibitor; NDRI, Norepinephrine-Dopamine Reuptake Inhibitor; SATA, substance abuse treatment agent; SSRI, Selective Serotonin Reuptake Inhibitor; SNRI, Serotonin-Norepinephrine Reuptake Inhibitor.

Table 279: Summary of table for antidepressants compared to placebo at the end of treatment in adults with BED

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antidepressants (95% CI)
Remission >=2 weeks assessment period (e.g. EDE OBE)	199 (4 studies) 12 months	⊕⊖⊖⊖ VERY LOW1,2,3,4,5,6 due to risk of bias, indirectness, imprecision	RR 1.39 (0.92 to 2.09)	270 per 1000	105 more per 1000 (from 22 fewer to 294 more)
Binge Frequency binge episodes/week or month, binge days/week	196 (4 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4,5,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.18 standard deviations lower (0.42 lower to 0.06 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antidepressants (95% CI)
BMI or Weight	379 (8 studies)	⨁⨀⨀⨀ VERY LOW1,2,3,4,5,7,8,9,10,11,12,13 due to risk of bias, inconsistency, indirectness, imprecision, publication bias		Not calculable for SMD values	The mean bmi or weight in the intervention groups was 0.06 standard deviations lower (0.33 lower to 0.21 higher)
Withdrawn due to Adverse Events	255 (5 studies)	⨁⨀⨀⨀ VERY LOW2,3,5,7,8,9,10,13 due to risk of bias, indirectness, imprecision, publication bias	RR 2.35 (0.91 to 6.08)	32 per 1000	43 more per 1000 (from 3 fewer to 161 more)
EDE-Q Global Scale from: 0 to 6.	115 (2 studies)	⨁⨁⨀⨀ LOW1,4,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global in the intervention groups was 0.03 standard deviations higher (0.34 lower to 0.39 higher)
EDE-Q Dietary Restraint Scale from: 0 to 6.	115 (2 studies)	⨁⨀⨀⨀ VERY LOW1,4,12,14 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-q dietary restraint in the intervention groups was 0.07 standard deviations higher (0.51 lower to 0.66 higher)
EDE-Q Eating Concerns Scale from: 0 to 6.	115 (2 studies)	⨁⨁⨀⨀ LOW1,4,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concerns in the intervention groups was 0.15 standard deviations higher (0.22 lower to 0.52 higher)
EDE-Q Weight Concerns Scale from: 0 to 6.	115 (2 studies)	⨁⨁⨀⨀ LOW1,4,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concerns in the intervention groups was 0.1 standard deviations higher (0.27 lower to 0.46 higher)
EDE-Q Shape Concerns Scale from: 0 to 6.	115 (2 studies)	⨁⨁⨀⨀ LOW1,4,7 due to risk of bias, imprecision		Not calculable for SMD	The mean ede-q shape concerns in the intervention groups was 0.11 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antidepressants (95% CI)
				values	(0.47 lower to 0.26 higher)
Depression HRSD, BDI, IDS-C	382 (8 studies)	⨁⊖⊖⊖ VERY LOW1,2,3,4,5,7,8,9,10,11,13 due to risk of bias, indirectness, imprecision, publication bias		Not calculable for SMD values	The mean depression in the intervention groups was 0.2 standard deviations lower (0.4 lower to 0.01 higher)
Clinical Global Impressions - Severity of Illness Scale from: 1 to 7.	267 (6 studies)	⨁⊖⊖⊖ VERY LOW2,3,5,6,8,9,10,11,13 due to risk of bias, indirectness, imprecision, publication bias		Not calculable for SMD values	The mean clinical global impressions - severity of illness in the intervention groups was 0.71 standard deviations lower (0.96 to 0.46 lower)
Clinical Global Impressions - Severity of Illness for depressive disorders Scale from: 1 to 7.	38 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean clinical global impressions - severity of illness for depressive disorders in the intervention groups was 0.51 standard deviations lower (1.16 lower to 0.14 higher)
Clinical Global Impressions - Improvement of Illness for depressive disorders Scale from: 1 to 7.	38 (1 study)	⨁⊖⊖⊖ VERY LOW3,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean clinical global impressions - improvement of illness for depressive disorders in the intervention groups was 0.54 standard deviations lower (1.19 lower to 0.11 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

1 Grilo 2005/2012: Randomization method and allocation concealment unclear. Assessor blinding unclear. Intervention group dropout rate>20%.
2 Guerdjikova 2008: Randomisation method unclear. Intervention group dropout rate>20%.
3 Guerdjikova 2012: Duloxetine group significantly older than placebo group. Randomization method unclear. Dropout rate for both groups>20%.
4 White 2013: Randomization method and allocation concealment unclear. Assessor blinding unclear.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antidepressants (95% CI)

5 Population for Guerdjikova 2012 were BED patients with comorbid depressive disorder.
6 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
7 <300 events (dichotomous outcome) or <400 participants (continuous outcome).
8 Hudson 1998: fluvoxamine group had significantly higher number of patients with lifetime history of major depression. Randomization method and allocation concealment unclear. Intervention group dropout rate>20%.
9 McElroy and Hudson 2003: Randomisation method and allocation concealment unclear. Assessor blinding unclear. Dropout rate for both groups>20%.
10 Arnold 2002: Randomization method and allocation concealment unclear. Assessor blinding unclear. Dropout rate for both groups>20%.
11 McElroy 2000: Randomization method and allocation concealment unclear. Assessor blinding unclear. Intervention group dropout rate>20%.
12 I2>50%.
13 One study (Hudson 1998) published before 2000.
14 CI crosses both 0.5 and -0.5 (SMD).

1 **Table 280: Summary of table for antidepressant-1 versus antidepressant-2 at end of treatment in adults with BED**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with Antidepressant-1 v Antidepressant-2 (95% CI)
Binge Frequency Mean binge episodes/month	43 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.33 standard deviations higher (0.27 lower to 0.94 higher)
BMI	31 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.40 standard deviations lower (1.11 lower to 0.31 higher)
#>5% Weight Loss	37 (1 study)	⊕⊖⊖⊖ VERY LOW3,4 due to risk of bias, imprecision	RR 1.05 (0.52 to 2.1)	450 per 1000	22 more per 1000 (from 216 fewer to 495 more)
Withdrawn due to Adverse	43	⊕⊖⊖⊖	RR 0.52	182 per	87 fewer per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with Antidepressant-1 v Antidepressant-2 (95% CI)
Events	(1 study) 12 months	VERY LOW1,4 due to risk of bias, imprecision	(0.11 to 2.56)	1000	(from 162 fewer to 284 more)
# Binge Eating Scale score < 17	39 (1 study)	⊕⊖⊖⊖ VERY LOW3,4 due to risk of bias, imprecision	RR 0.91 (0.44 to 1.88)	455 per 1000	41 fewer per 1000 (from 255 fewer to 400 more)
Binge Eating Scale	31 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.32 standard deviations higher (0.39 lower to 1.03 higher)
EDI-2 Drive for Thinness	31 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 drive for thinness in the intervention groups was 0.26 standard deviations lower (0.97 lower to 0.45 higher)
EDI-2 Bulimia	31 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 bulimia in the intervention groups was 0.24 standard deviations higher (0.46 lower to 0.95 higher)
EDI-2 Body Dissatisfaction	31 (1 study)	⊕⊖⊖⊖ VERY LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi-2 body dissatisfaction in the intervention groups was 0.1 standard deviations lower (0.81 lower to 0.6 higher)
Depression	31 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.24 standard deviations lower (0.95 lower to 0.47 higher)
Clinical Global Impression -	31	⊕⊕⊖⊖		Not	The mean clinical global impression - severity of

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with Antidepressant-1 v Antidepressant-2 (95% CI)
Severity of Illness	(1 study)	LOW2,3 due to risk of bias, imprecision		calculable for SMD values	illness in the intervention groups was 0.32 standard deviations higher (0.39 lower to 1.03 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Ricca 2001: inadequate randomization method, treatment allocation unclear. No participant, investigator nor assessor blinding. Dropout rate of both treatment groups >20%.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 Leombruni 2008: Randomization method and allocation concealment unclear. Investigator and assessor blinding unclear. Dropout rate both groups>20%, reasons not stated.
4 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).

1 **Table 281: Summary of table for antidepressant-1 versus antidepressant-2 at follow up in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with	Risk difference with Antidepressant-1 v Antidepressant-2 (95% CI)
Binge Frequency 12-mo FU	32 (1 study) 12 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency 12-mo fu in the intervention groups was 1.17 standard deviations higher (0.41 to 1.93 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Ricca 2001: inadequate randomization method, treatment allocation unclear. No participant, investigator nor assessor blinding. Dropout rate of both treatment groups>20%.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 282: Summary of table for antidepressant versus another intervention at end of treatment in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any individual therapy	Risk difference with Antidepressants (95% CI)
Binge Frequency Mean binge episodes/month	103 (2 studies) 12 months	⊕⊖⊖⊖ VERY LOW1,2,3,4,5 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 2.57 standard deviations higher (2.02 to 3.13 higher)
% Weight Loss	40 (1 study)	⊕⊖⊖⊖ VERY LOW2,4,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean % weight loss in the intervention groups was 2.26 standard deviations lower (3.07 to 1.45 lower)
EDI-2 Bulimia	40 (1 study)	⊕⊖⊖⊖ VERY LOW2,4,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi-2 bulimia in the intervention groups was 2.52 standard deviations higher (1.67 to 3.38 higher)
Depression MMPI-2 Depression	40 (1 study)	⊕⊖⊖⊖ VERY LOW2,4,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 1.17 standard deviations higher (0.5 to 1.85 higher)
Family Functioning MMPI-2 Family Problems	40 (1 study)	⊕⊖⊖⊖ VERY LOW2,4,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean family functioning in the intervention groups was 0.14 standard deviations higher (0.48 lower to 0.76 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Ricca 2001: Randomization method inadequate (allocated to treatment groups enrolment day, allocation concealment unclear. No participant, investigator, assessor blinding. Dropout rate for both arms>20%.
2 Molinari 2005: Randomization method and allocation concealment unclear. Participant, investigator and assessor blinding unclear.
3 I2>=50%.
4 Molinari 2005: both Fluoxetine+CBT and CBT only groups also had Group Nutritional Counselling + Diet.
5 <400 participants.
6 CI crosses either 0.5 or -0.5 (SMD).

1 **Table 283: Summary of table for antidepressant versus another intervention at follow up in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any CBT	Risk difference with Antidepressants (95% CI)
Binge Frequency FU Mean binge episodes/month	49 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 3.08 standard deviations higher (2.19 to 3.97 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; FU: Follow up					
1 Ricca 2001: Randomization method inadequate (allocated to treatment groups enrolment day, allocation concealment unclear. No participant, investigator, assessor blinding. Dropout rate for both arms >20%. 2 <400 participants.					

2 **Table 284: Summary of table for antiepileptics/anticonvulsants versus placebo in adults with BED**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antiepileptics (95% CI)
Remission	111 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, imprecision	RR 0.88 (0.62 to 1.25)	564 per 1000	68 fewer per 1000 (from 214 fewer to 141 more)
Binge Frequency binge episodes/week or binge days/week	111 (2 studies)	⊕⊕⊖⊖ LOW1,2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.23 standard deviations lower (0.49 lower to 0.03 higher)
Withdrawn due to Adverse Events	573 (4 studies)	⊕⊕⊖⊖ LOW1,2,4,5,6 due to risk of bias, imprecision	RR 1.94 (1.22 to 3.08)	83 per 1000	78 more per 1000 (from 18 more to 173 more)
BMI	565	⊕⊕⊖⊖		Not	The mean BMI in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antiepileptics (95% CI)
	(4 studies)	LOW1,2,4,5,6 due to risk of bias, imprecision		calculable for SMD values	0.45 standard deviations lower (0.62 to 0.29 lower)
EDE-Q Global	51 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global in the intervention groups was 0.44 standard deviations lower (0.99 lower to 0.12 higher)
EDE-Q Restraint	51 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q restraint in the intervention groups was 0.12 standard deviations lower (0.67 lower to 0.43 higher)
EDE-Q Weight Concerns	51 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concerns in the intervention groups was 0.48 standard deviations lower (1.04 lower to 0.08 higher)
EDE-Q Eating Concerns	51 (1 study)	⊕⊖⊖⊖ VERY LOW2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concerns in the intervention groups was 0.03 standard deviations lower (0.58 lower to 0.51 higher)
EDE-Q Shape Concerns	51 (1 study)	⊕⊕⊖⊖ LOW2,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concerns in the intervention groups was 0.48 standard deviations lower (1.04 lower to 0.08 higher)
Depression	565 (4 studies)	⊕⊕⊖⊖ LOW1,2,3,5,6 due to risk of bias, inconsistency		Not calculable for SMD values	The mean depression in the intervention groups was 0.07 standard deviations higher (0.1 lower to 0.23 higher)
Clinical Global Impressions - Severity of Illness	172 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean clinical global impressions - severity of illness in the intervention groups was 0.65 standard deviations lower (0.83 to 0.47 lower)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Antiepileptics (95% CI)
General functioning Sheehan Disability Scale Total	445 (2 studies)	⨁⨁⨁⨀ MODERATE2,6 due to risk of bias		Not calculable for SMD values	The mean general functioning in the intervention groups was 0.24 standard deviations lower (0.43 to 0.05 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 McElroy 2006: Randomization method and allocation concealment unclear. Dropout rate for both groups >20%.
2 Guerdjikova 2009: Randomization method unclear. Dropout rate for both groups >20%.
3 I2>50%.
4 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
5 McElroy and Arnold 2003: Randomization method and allocation concealment unclear. Dropout rate for both groups >20%.
6 McElroy and Hudson 2007: Randomization method and allocation concealment unclear. Dropout rate for both groups >20%.
7 CI crosses both 0.5 and -0.5 (SMD).

1 **Table 285: Summary of table for appetite suppressants (lisdexamfetamine dimesylate) versus placebo in adults with BED**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Appetite Suppressants (95% CI)
Remission (ITT) 100% reduction binge episodes in past 4 weeks	1032 (3 studies)	⨁⨁⨁⨀ MODERATE1,2 due to risk of bias	RR 2.6 (2.02 to 3.36)	138 per 1000	220 more per 1000 (from 141 more to 325 more)
BMI (change scores)	983 (3 studies)	⨁⨁⨀⨀ LOW1,2,3 due to risk of bias, inconsistency		Not calculable for SMD values	The mean bmi (change scores) in the intervention groups was 1.24 standard deviations lower (1.51 to 0.98 lower)
Withdrawn due to Adverse Events	1004 (3 studies)	⨁⨁⨀⨀ LOW1,2,4 due to risk of bias, imprecision	RR 2.05 (1.01 to 4.18)	21 per 1000	22 more per 1000 (from 0 more to 66 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Appetite Suppressants (95% CI)
Binge Eating Scale from: 0 to 46.	255 (1 study)	⨁⨁⨀⨀ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 4.11 standard deviations lower (4.59 to 3.63 lower)
Depression MADRS. Scale from: 0 to 60.	255 (1 study)	⨁⨁⨀⨀ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.28 standard deviations higher (0.01 lower to 0.57 higher)
General Physical Functioning SF-12 Physical. Scale from: 0 to 100.	255 (1 study)	⨁⨁⨀⨀ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean general physical functioning in the intervention groups was 0.27 standard deviations higher (0.01 lower to 0.56 higher)
General Mental Functioning SF-12 Mental. Scale from: 0 to 100.	255 (1 study)	⨁⨁⨀⨀ LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean general mental functioning in the intervention groups was 0.03 standard deviations higher (0.26 lower to 0.32 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

1 McElroy 2015: Dropout rate for all arms>=20%.
2 McElroy and Hudson 2016 Study 1 and 2: unclear whether assessor blinded. McElroy and Hudson 2016 Study 2: dropout rate for both groups>=20%.
3 I2>50%.
4 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
5 <300 events (dichotomous outcome) or <400 participants (continuous outcome).

1 **Table 286: Summary of table for substance abuse treatment agents versus placebo in adults with BED**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Substance Abuse Treatment Agents (95% CI)
Remission	109	⨁⨁⨀⨀	RR 0.72	404 per	113 fewer per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Substance Abuse Treatment Agents (95% CI)
	(2 studies)	LOW1,2,3 due to risk of bias, imprecision	(0.42 to 1.24)	1000	(from 234 fewer to 97 more)
BMI (raw and change scores)	86 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.22 standard deviations lower (0.66 lower to 0.22 higher)
Weight (raw and change scores)	86 (2 studies)	⊕⊕⊖⊖ LOW1,2,5 due to risk of bias, imprecision		Not calculable for SMD values	The weight in the intervention groups was 0.05 standard deviations lower (0.48 lower to 0.38 higher)
Binge episode Frequency Mean binge episodes/week (raw and change scores)	86 (2 studies)	⊕⊕⊖⊖ LOW1,2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge episode frequency in the intervention groups was 0.15 standard deviations lower (0.58 lower to 0.28 higher)
Binge Day Frequency binge days/week (raw and change scores)	86 (2 studies)	⊕⊕⊖⊖ LOW1,2,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge day frequency in the intervention groups was 0.07 standard deviations higher (0.36 lower to 0.5 higher)
Withdrawn due to Adverse Event	108 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,4,5 due to risk of bias, inconsistency, imprecision	RR 10.69 (2.13 to 53.57)	18 per 1000	170 more per 1000 (from 20 more to 922 more)
Depression MADRS	24 (1 study)	⊕⊕⊖⊖ VERY LOW2,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.08 standard deviations lower (0.9 lower to 0.75 higher)
Depression – change scores BDI	62 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression – change scores in the intervention groups was 0.43 standard deviations higher (0.08 lower to 0.95 higher)
General physical functioning SF-12 Physical	24 (1 study)	⊕⊕⊖⊖ VERY LOW2,6		Not calculable	The mean general physical functioning in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Substance Abuse Treatment Agents (95% CI)
		due to risk of bias, imprecision		for SMD values	0.25 standard deviations higher (0.58 lower to 1.08 higher)
General mental functioning SF-12 Mental	24 (1 study)	⊕⊕⊖⊖ LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean general mental functioning in the intervention groups was 0.39 standard deviations higher (0.45 lower to 1.22 higher)
Clinical Global Impressions - Severity of Illness	86 (2 studies)	⊕⊕⊖⊖ LOW1,2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical global impressions - severity of illness in the intervention groups was 0.17 standard deviations higher (0.26 lower to 0.61 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 McElroy 2013: Unclear randomization method and treatment allocation. Intervention group dropout rate>=50%.
2 McElroy 2011: Unclear randomization method. Dropout rate for both groups >20%.
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
4 I2>80%.
5 <300 events (dichotomous outcome) or <400 participants (continuous outcome).

1 **Table 287: Summary of table for atomoxetine versus placebo in adults with BED**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Atomoxetine (95% CI)
Remission 100% decrease frequency binge episodes from baseline	40 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 2.33 (1.13 to 4.83)	316 per 1000	420 more per 1000 (from 41 more to 1000 more)
BMI	40 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD	The mean BMI or weight in the intervention groups was 0.74 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo	Risk difference with Atomoxetine (95% CI)
		imprecision		values	(1.38 to 0.1 lower)
Weight loss (kg)	40 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss (kg) in the intervention groups was 0.77 standard deviations higher (0.12 to 1.41 higher)
Binge Frequency Binge episodes/week or binge days/week	40 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.72 standard deviations lower (1.17 to 0.27 lower)
Withdrawn due to Adverse Events	40 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 3 (0.34 to 26.45)	50 per 1000	100 more per 1000 (from 33 fewer to 1000 more)
Depression HDRS	40 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.05 standard deviations higher (0.57 lower to 0.67 higher)
Clinical Global Impressions - Severity of Illness	40 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical global impressions - severity of illness in the intervention groups was 1.1 standard deviations lower (1.77 to 0.44 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 McElroy 2007: Randomization method and allocation concealment unclear. Dropout rate for both arms >20%.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 CI crosses both 0.5 and -0.5.

1 **Table 288: Summary of table for armodafinil versus placebo in adults with BED.**

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Control	Risk difference with Armodafinil v Placebo (95% CI)
Remission	55 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.21 (0.47 to 3.14)	214 per 1000	45 more per 1000 (from 114 fewer to 459 more)
BMI - Change scores	55 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI - change scores in the intervention groups was 0.67 standard deviations lower (1.22 to 0.13 lower)
Withdrawn due to adverse events	60 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision	RR 1 (0.15 to 6.64)	67 per 1000	0 fewer per 1000 (from 57 fewer to 376 more)
Binge Frequency - Change scores	55 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency - change scores in the intervention groups was 0.46 standard deviations lower (0.84 to 0.09 lower)
Clinical Global Impressions Severity - Change scores	55 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean clinical global impressions severity - change scores in the intervention groups was 0.49 standard deviations lower (1.03 lower to 0.04 higher)
Depression - Change scores	55 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - change scores in the intervention groups was 0.01 standard deviations higher (0.52 lower to 0.54 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 McElroy & Guerdjikova 2015: Dropout rate of both groups >=47%.
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or 0.5 and -0.5 (SMD).
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or 0.5 or -0.5 (SMD).
4 <300 events.

Table 289: Summary table of findings for antidepressant and CBT-ED versus CBT-ED at end of treatment in adults with BED

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT	Risk difference with Antidepressant+CBT (95% CI)
Binge Frequency	105 (2 studies) 12 months	⨁⊖⊖⊖ VERY LOW1,2,3,4,5 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.14 standard deviations higher (0.6 lower to 0.89 higher)
% Weight Loss	40 (1 study)	⨁⊖⊖⊖ VERY LOW2,4,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean % weight loss in the intervention groups was 0.2 standard deviations lower (0.82 lower to 0.43 higher)
EDI-2 Bulimia	40 (1 study)	⨁⊖⊖⊖ VERY LOW2,4,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi-2 bulimia in the intervention groups was 1.25 standard deviations higher (0.57 to 1.94 higher)
Not withdrawn due to Adverse Events	105 (2 studies) 12 months	⨁⊖⊖⊖ VERY LOW1,2,4,7 due to risk of bias, indirectness, imprecision	RR 0.92 (0.84 to 1.02)	1000 per 1000	80 fewer per 1000 (from 160 fewer to 20 more)
Binge Eating Scale Scale from: 0 to 46.	30 (1 study) 12 months	⨁⨁⊖⊖ LOW6,8 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.42 standard deviations lower (1.19 lower to 0.35 higher)
Depression MMPI-2 Depression	70 (2 studies) 12 months	⨁⊖⊖⊖ VERY LOW2,3,4,6,8 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.18 standard deviations higher (0.31 lower to 0.68 higher)
Family Functioning MMPI-2 family problems	40 (1 study) 12 months	⨁⊖⊖⊖ VERY LOW2,4,6 due to risk of bias, indirectness,		Not calculable for	The mean family functioning in the intervention groups was 0.28 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT	Risk difference with Antidepressant+CBT (95% CI)
		imprecision		SMD values	(0.34 lower to 0.91 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

1 Ricca 2001: Inadequate randomization method. Allocation concealment unclear. No participant, investigator and assessor blinding. Dropout rate of four of five groups>20%.
2 Molinari 2005: Randomization method and allocation concealment unclear. Participant, investigator and assessor blinding unclear.
3 I2>50%.
4 Molinari 2005: Treatment was carried out in both in-patient (4 weeks) and out-patient setting (50 weeks); both Fluoxetine+CBT and CBT only groups also had Group Nutritional Counselling + Diet.
5 CI crosses both 0.5 and -0.5 (SMD).
6 CI crosses either 0.5 or -0.5 (SMD).
7 <300 events (dichotomous outcome) or <400 participants (continuous outcome).
8 Cristina 2014: Randomization method and allocation concealment unclear. Participant, investigator and assessor blinding unclear. No details provided regarding dropouts.

1 **Table 290: Summary table of findings for antidepressant and CBT-ED versus CBT-ED at follow up in adults with BED**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT	Risk difference with Antidepressant+CBT (95% CI)
Binge Frequency FU	50 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 4.42 standard deviations lower (5.53 to 3.3 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Ricca 2001: Inadequate randomization method. Allocation concealment unclear. No participant, investigator and assessor blinding. Dropout rate of four of five groups>20%.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CBT	Risk difference with Antidepressant+CBT (95% CI)
2 CI crosses either 0.5 or -0.5 (SMD).					

Table 291: Summary table of findings for antidepressant and CBT-ED versus placebo and CBT-ED at end of treatment in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+CBT	Risk difference with Antidepressant+CBT (95% CI)
Remission EDE-Q No OBE/28 days	54 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.08 (0.67 to 1.73)	536 per 1000	43 more per 1000 (from 177 fewer to 391 more)
BMI	54 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.1 standard deviations higher (0.43 lower to 0.63 higher)
Binge Frequency Mean binge episodes/month	54 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.33 standard deviations higher (0.21 lower to 0.87 higher)
EDE-Q Global	54 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global in the intervention groups was 0.08 standard deviations higher (0.46 lower to 0.61 higher)
EDE-Q Dietary Restraint	54 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q dietary restraint in the intervention groups was 0 standard deviations higher (0.53 lower to 0.53 higher)
EDE-Q Eating Concerns	54 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concerns in the intervention groups was 0.19 standard deviations lower (0.73 lower to 0.34 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+CBT	Risk difference with Antidepressant+CBT (95% CI)
EDE-Q Weight Concerns	54 (1 study) 12 months	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concerns in the intervention groups was 0.16 standard deviations lower (0.69 lower to 0.38 higher)
EDE-Q Shape Concerns	54 (1 study) 12 months	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concerns in the intervention groups was 0.06 standard deviations lower (0.6 lower to 0.47 higher)
Depression BDI	54 (1 study) 12 months	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.38 standard deviations higher (0.16 lower to 0.92 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Grilo 2005/2012: randomization method and allocation concealment unclear. Assessor blinding unclear. Dropout rate of three of four groups >20%.
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

1 **Table 292: Summary table of findings for antidepressant and CBT-ED versus placebo and CBT-ED at follow up in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+CBT	Risk difference with Antidepressant+CBT (95% CI)
Remission FU EDE-Q No OBE/28 days	54 (1 study)	⨁⨀⨀⨀ VERY LOW1,2 due to risk of bias, imprecision	RR 0.75 (0.34 to 1.69)	357 per 1000	89 fewer per 1000 (from 236 fewer to 246 more)
BMI FU	41 (1 study)	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.43 standard deviations higher (0.19 lower to 1.05 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+CBT	Risk difference with Antidepressant+CBT (95% CI)
Binge Frequency FU Mean binge episodes/month	41 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 0 standard deviations higher (0.61 lower to 0.62 higher)
EDE-Q Global FU	41 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q global fu in the intervention groups was 0.29 standard deviations lower (0.91 lower to 0.33 higher)
EDE-Q Dietary Restraint FU	41 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q dietary restraint fu in the intervention groups was 0.36 standard deviations lower (0.98 lower to 0.26 higher)
EDE-Q Eating Concerns FU	41 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q eating concerns fu in the intervention groups was 0.04 standard deviations lower (0.65 lower to 0.58 higher)
EDE-Q Weight Concerns FU	41 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concerns fu in the intervention groups was 0.32 standard deviations lower (0.94 lower to 0.3 higher)
EDE-Q Shape Concerns FU	41 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concerns fu in the intervention groups was 0.45 standard deviations lower (1.07 lower to 0.17 higher)
Depression FU BDI	41 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.04 standard deviations lower (0.65 lower to 0.58 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Grilo 2005/2012: randomization method and allocation concealment unclear. Assessor blinding unclear. Dropout rate of three of four groups >20%.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+CBT	Risk difference with Antidepressant+CBT (95% CI)
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD). 3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).					

Table 293: Summary table of findings for antidepressant-1 and CBT-ED versus antidepressant-2 and CBT-ED at end of treatment in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant-2+CBT	Risk difference with Antidepressant-1+CBT (95% CI)
Binge Frequency Binge episodes/month	45 (1 study) 12 months	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.5 standard deviations lower (1.09 lower to 0.1 higher)
Withdrawn due to Adverse Events	45 (1 study) 12 months	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 1.05 (0.24 to 4.64)	130 per 1000	7 more per 1000 (from 99 fewer to 475 more)
Binge Eating Scale	20 (1 study) 12 months	⨁⊖⊖⊖ VERY LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.25 standard deviations higher (0.63 lower to 1.13 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Ricca 2001: Randomization method inadequate. Allocation concealment unclear. No participant, investigator and assessor blinding. Dropout rate for groups all >20%. 2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD). 3 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD). 4 Cristina 2014: Randomization method and allocation concealment unclear. Participant, investigator and assessor blinding unclear. No details provided regarding dropouts.					

Table 294: Summary table of findings for antidepressant-1 and CBT-ED versus antidepressant-2 and CBT-ED at follow up in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant-2+any CBT	Risk difference with Antidepressant-1+any CBT (95% CI)
Binge Frequency FU Binge episodes/month	34 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 0.34 standard deviations lower (1.01 lower to 0.34 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Ricca 2001: Randomization method inadequate. Allocation concealment unclear. No participant, investigator and assessor blinding. Dropout rate for groups all >20%.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 295: Summary table of findings for antiepileptic and group CBT-ED versus placebo and group CBT-ED in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+gCBT-ED	Risk difference with Antiepileptic+gCBT-ED (95% CI)
BMI	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.41 standard deviations lower (0.94 lower to 0.12 higher)
# patients achieving Weight Loss>10%	56 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 3.18 (0.99 to 10.17)	115 per 1000	252 more per 1000 (from 1 fewer to 1000 more)
Not withdrawn due to Adverse Events	73 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 0.97 (0.9 to 1.05)	1000 per 1000	30 fewer per 1000 (from 100 fewer to 50 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+gCBT-ED	Risk difference with Antiepileptic+gCBT-ED (95% CI)
Binge Eating Scale	56 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.17 standard deviations lower (0.69 lower to 0.36 higher)
Depression BDI	56 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.24 standard deviations higher (0.29 lower to 0.77 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Claudino 2007: topiramate group significantly older and report more depression than placebo group. Dropout rate for placebo group>20%.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 <300 events (dichotomous outcome).

Table 296: Summary table of findings for antidepressant, antiepileptic, group behavioural weight loss therapy and group CBT versus antidepressant, group behavioural weight loss therapy and group CBT in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant+gBWLT+gCBT	Risk difference with Antidepressant+Antiepileptic+gBWLT+gCBT+ (95% CI)
BMI	20 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi in the intervention groups was 0.41 standard deviations higher (0.48 lower to 1.29 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Brambilla 2009: Randomization method and allocation concealment unclear. Weight and BMI significantly higher at baseline in 1700kcal Group

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Antidepressant+gBWLT+gCBT	Risk difference with Antidepressant+Antiepileptic+gBWLT+gCBT+ (95% CI)
BWLT+Topiramate+Sertraline+CBT group compared to 1700kcal Group BWLT+Sertraline+CBT group. 2 CI crosses either 0.5 or -0.5.					

Table 297: Summary table of findings for antiobesity agent and guided self-help CBT-ED versus placebo and guided self-help CBT-ED in adults with binge eating disorder at end of treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antiobesity+gSH CBT-ED v Placebo+gSH CBT-ED (95% CI)
Remission (ITT)	50 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.78 (0.98 to 3.24)	360 per 1000	281 more per 1000 (from 7 fewer to 806 more)
Binge frequency EDE OBE in past 28 days	50 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.07 standard deviations lower (0.63 lower to 0.48 higher)
Weight loss>=5% (ITT)	50 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 4.5 (1.08 to 18.77)	80 per 1000	280 more per 1000 (from 6 more to 1000 more)
Weight loss (kg)	50 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss (kg) in the intervention groups was 0.62 standard deviations higher (0.05 to 1.19 higher)
Mean percentage weight loss	50 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean percentage weight loss in the intervention groups was 0.58 standard deviations higher (0.01 to 1.15 higher)
EDE Global	50 (1 study)	⨁⨁⊖⊖ LOW1,2		Not calculable for	The mean ede global in the intervention groups was 0.34 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antiobesity+gSH CBT-ED v Placebo+gSH CBT-ED (95% CI)
		due to risk of bias, imprecision		SMD values	(0.9 lower to 0.22 higher)
EDE Dietary restraint	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede dietary restraint in the intervention groups was 0.05 standard deviations higher (0.5 lower to 0.61 higher)
EDE Eating concern	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concern in the intervention groups was 0.1 standard deviations lower (0.65 lower to 0.46 higher)
EDE Weight concern	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.21 standard deviations lower (0.77 lower to 0.34 higher)
EDE Shape concern	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.39 standard deviations lower (0.95 lower to 0.17 higher)
Depression BDI	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.54 standard deviations lower (1.11 lower to 0.02 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Grilo, Masheb & Salent 2005: high risk of bias (unclear allocation concealment, dropout rate of both groups >=20%).
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 298: Summary table of findings for antiobesity agent and guided self-help CBT-ED versus placebo and guided self-help CBT-ED in adults with binge eating disorder at follow up

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Control	Risk difference with Antiobesity+gSH CBT v Placebo+gSH CBT at 3-mo FU (95% CI)
Remission (ITT)	50 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1 (0.59 to 1.7)	520 per 1000	0 fewer per 1000 (from 213 fewer to 364 more)
Binge frequency EDE OBE in past 28 days	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.1 standard deviations higher (0.46 lower to 0.65 higher)
Weight loss>=5% (ITT)	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 4 (0.94 to 17)	80 per 1000	240 more per 1000 (from 5 fewer to 1000 more)
Weight loss (kg)	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss (kg) in the intervention groups was 0.5 standard deviations higher (0.07 lower to 1.06 higher)
Mean percentage weight loss	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean percentage weight loss in the intervention groups was 0.48 standard deviations higher (0.09 lower to 1.04 higher)
EDE Global	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.09 standard deviations lower (0.65 lower to 0.46 higher)
EDE Dietary restraint	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede dietary restraint in the intervention groups was 0.15 standard deviations lower (0.71 lower to 0.4 higher)
EDE Eating concern	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concern in the intervention groups was 0.07 standard deviations lower (0.63 lower to 0.48 higher)
EDE Weight concern	50 (1 study)	⊕⊕⊖⊖ LOW1,3		Not calculable	The mean ede weight concern in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Antiobesity+gSH CBT v Placebo+gSH CBT at 3-mo FU (95% CI)
		due to risk of bias, imprecision		for SMD values	0.08 standard deviations higher (0.47 lower to 0.64 higher)
EDE Shape concern	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.07 standard deviations lower (0.62 lower to 0.49 higher)
Depression BDI	50 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.47 standard deviations lower (1.03 lower to 0.09 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Grilo, Masheb & Salent 2005: high risk of bias (unclear allocation concealment, dropout rate of both groups >=20%).
2 CI crosses both 0.75 and 1.25 (Risk Ratio).
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

8.4.3 Economic Evidence

The systematic search of the economic literature undertaken for the guideline identified:

- One US study on the cost utility of lisdexamfetamine dimesylate (LDX) in adults with binge eating disorder in the US (Agh et al., 2016).

References to included studies and evidence tables for all economic evaluations included in the systematic literature review are provided in Appendix P. Completed methodology checklists of the studies are provided in Appendix O. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix Q.

Agh and colleagues (2016) evaluated the cost utility of LDX compared with no drug treatment in adults with BED in the US. BED population comprised women with mild, moderate and severe and extreme binge eating behaviour. This was a modelling study (a decision analytical Markov cohort model). Treatment with LDX (50 or 70 mg/day) was for 12 weeks. The analysis was conducted from a health care perspective and considered a range of costs including drug treatment, general internist, family doctor, psychiatrist, psychologist, psychotherapist, nurse practitioner, gynaecologist, emergency room and hospital admissions. The effectiveness data was from 2 RCTs (RCT 1, N=383; RCT 2, N=390; McElroy 2015) and the resource use was obtained from a national survey (N=22,397). The unit costs were obtained from a national source (Medical Expenditure Panel Survey). The measure of outcome for the economic analysis was QALYs. Utility values were based on the EQ-5D-5L individual patient level data of the clinical trials (McElroy 2015) and were valued using US population norms. The time horizon of the analysis was 52 weeks.

Treatment with LDX resulted in a greater number of QALYs at 52 weeks compared with no drug treatment (0.917 versus 0.911, respectively; a difference of 0.006). The mean total costs per participant over 52 weeks were $7,042 for the LDX and $6,867 for no drug treatment, a difference of $175 (in favour of no drug treatment group) in 2013 US dollars. The ICER of LDX when compared with no treatment was $27,618 per QALY gained. Bootstrapping indicated that at willingness-to-pay of $50,000 per QALY LDX had an 82% chance of being cost-effective.

Deterministic sensitivity analyses indicated that the model was most sensitive to the utility of remission (that is, non-symptomatic BED). The results were robust to changes in other model inputs.

The analysis was judged by the committee to be partially applicable to the NICE decision-making context, as it has been conducted in the US. This was a well conducted study and was judged by the committee to have only minor methodological limitations.

8.4.4 Clinical evidence statements

8.4.4.1 Antidepressants versus placebo in adults with BED

Very low quality evidence from four RCTs (n=199) showed antidepressants may be more effective on increasing remission compared with placebo, although there was some uncertainty.

Very low quality from four RCTs (n=196) showed antidepressants may be more effective on reducing binge frequency compared with placebo, although there was some uncertainty.

Very low quality evidence from eight RCTs (n=379) showed no difference on BMI/Weight compared to those who took placebo.

Very low quality evidence from five RCTs (n=255) showed antidepressants may be less effective on the number of people who withdrew from the RCTs due to adverse effects compared with those who took placebo, although there was some uncertainty.

Very low to low quality evidence from two RCTs (n=115) showed no difference in the effect of antidepressants on reducing scores on EDE-Global, EDE-dietary restraint, EDE-eating concerns, EDE-weight concerns and EDE-shape concerns compared with placebo.

Very low quality evidence from eight RCTs (n=382) showed antidepressants may be more effective on reducing depression compared with placebo, although there was some uncertainty.

Very low quality evidence from six RCTs (n=267) showed antidepressants are more effective on reducing the severity of BED compared with placebo.

Very low quality evidence from one RCT (n=38) showed no difference in the effect of antidepressants on reducing the severity of, and ameliorating, comorbid depressive disorders compared with placebo.

8.4.4.2 Antidepressant-1 versus antidepressant-2 at end of treatment in adults with BED

Low quality evidence from one RCT (n=43) showed no difference in the effect of one antidepressant (fluoxetine) on binge frequency and the number of people who withdrew from the RCTs due to adverse events compared with another antidepressant (fluvoxamine).

Low quality evidence from one RCT (n=31) showed no difference in the effect of one antidepressant (fluoxetine) on BMI, Binge Eating Scale, EDI-2-drive for thinness, EDI-2-bulimia, EDI-2-body dissatisfaction, depression and severity of illness compared with another antidepressant (sertraline).

Very low quality evidence from one RCT (n=37) showed no difference in the effect of one antidepressant (fluoxetine) on the number of people who achieved ≥5% weight loss compared with another antidepressant (sertraline).

Very low quality evidence from one RCT (n=39) showed no difference in the effect of one antidepressant (fluoxetine) on the number of people who achieved a Binge Eating Scale score <17 compared with another antidepressant (sertraline).

8.4.4.3 Antidepressant-1 versus antidepressant-2 at follow up in adults with BED

Low quality evidence from one RCT (n=32) showed one antidepressant (fluoxetine) is less effective on binge frequency compared with another antidepressant (fluvoxamine).

8.4.4.4 Antidepressant versus any other intervention at end of treatment in adults with BED

Very low quality evidence from two RCTs (n=103) showed antidepressants are less effective at reducing binge frequency compared with any other intervention.

Very low quality evidence from one RCT (n=40) showed antidepressants are less effective on weight loss, improving scores on EDI-2-bulimia and reducing depression compared with any other intervention.

Very low quality evidence from one RCT (n=40) showed no difference in the effect of antidepressants on family functioning compared with any other intervention.

8.4.4.5 Antidepressant versus any other intervention at follow up in adults with BED

Low quality evidence from one RCT (n=49) showed antidepressants are less effective in reducing binge frequency compared with any other intervention.

8.4.4.6 Antiepileptics (anticonvulsants) versus placebo in adults with binge eating disorder

Very low quality evidence from two RCTs (n=111) showed no difference in the effect of antiepileptics/anticonvulsants on increasing remission compared with placebo.

Low quality evidence from two RCTs (n=111) showed antiepileptics/anticonvulsants may be more effective in reducing binge frequency compared with placebo, although there was some uncertainty.

Low quality evidence from four RCTs (n=573) showed that more people who took antiepileptics/anticonvulsants withdrew from the studies due to adverse effects compared with those who took placebo.

Low quality evidence from four RCTs (n=565) showed antiepileptics/anticonvulsants are more effective on BMI compared with placebo.

Low quality evidence from four RCTs (n=565) showed no difference in the effect of antiepileptics/anticonvulsants on depression compared with placebo.

Very low to low quality evidence from one RCT (n=51) showed no difference in the effect of antiepileptics/anticonvulsants in reducing scores on EDE-Q-global, EDE-Q-dietary restraint and EDE-Q-eating concerns compared with placebo.

Low quality evidence form one RCT (n=51) showed antiepileptics/anticonvulsants may be more effective in reducing scores on EDE-Q-weight concerns and EDE-Q-shape concerns compared with placebo, although there was some uncertainty.

Very low quality evidence from three RCTs (n=172) showed no difference in the effect of antiepileptics/anticonvulsants in reducing the severity of BED compared with placebo.

Moderate quality evidence from two RCTs (n=445) showed no difference in the effect of antiepileptics/anticonvulsants in improving general functioning compared with placebo.

8.4.4.7 Appetite suppressants (lisdexamfetamine dimesylate) versus placebo in adults with binge eating disorder

Moderate quality of evidence from three RCTs (n=1032) showed appetite suppressants are more effective in increasing remission compared with placebo.

Low quality evidence from three RCTs (n=983) showed appetite suppressants are more effective in increasing change in BMI compared with placebo.

Low quality evidence from three RCTs (n=1004) showed that more people who took appetite suppressants withdrew from the studies due to adverse effects compared to those taking placebo.

Low quality evidence from one RCT (n=255) showed appetite suppressants are more effective in reducing BES scores compared with placebo.

Low quality evidence from one RCT (n=255) showed no difference in the effect of appetite suppressants in improving general mental functioning compared with placebo.

Low quality evidence from one RCT (n=255) showed appetite suppressants may be more effective in improving general physical functioning compared with placebo, although there was some uncertainty.

Low quality evidence from one RCT (n=255) showed appetite suppressants may be less effective on reducing depression compared with placebo, although there was some uncertainty.

8.4.4.8 Substance abuse treatment agents versus placebo in adults with binge eating disorder

Very low quality evidence from two RCTs (n=109) showed no difference in the effect of substance abuse treatment agents on remission compared with placebo.

Very low quality evidence from two RCTs (n=86) showed no difference in the effect of substance abuse treatment agents in reducing BMI and Weight compared with placebo.

Low quality evidence from two RCTs (n=86) showed no difference in the effect of substance abuse treatment agents in reducing binge episode, binge day frequency and the severity of BED compared to placebo.

Very low quality evidence from one RCT (n=62) showed substance abuse treatment agents may be less effective on change in depression compared to placebo, although there was some uncertainty.

Very low quality evidence from two RCTs (n=108) showed that more people who took substance abuse treatment agents withdrew from the studies due to adverse effects compared with those who took placebo.

Very low to low quality evidence from one RCT (n=24) showed no difference in the effect of substance abuse treatment agents on depression, nor in improving general mental or physical functioning compared to placebo.

8.4.4.9 Other pharmacological interventions

Atomoxetine (norepinephrine reuptake inhibitor) versus placebo in adults with BED

Low quality evidence from one RCT (n=40) showed that atomoxetine is more effective on weight loss, at increasing remission, reducing BMI and binge frequency and ameliorating the severity of BED compared with placebo.

Very low quality evidence from one RCT (n=40) showed no difference in the number of people who withdrew from the trials due to adverse effects and who took atomoxetine compared to those who took placebo, although there was some uncertainty.

Very low quality evidence from one RCT (n=40) showed no difference in the effect of atomoxetine in reducing depression compared with placebo.

Armodafinil versus placebo in adults with BED

Very low quality evidence from one RCT (n=55) showed no difference in the effect of armodafinil in increasing remission and reducing depression compared with placebo.

Low quality evidence from one RCT (n=55) showed armodafinil is more effective in reducing BMI and binge frequency compared with placebo.

Low quality evidence from one RCT (n=60) showed armodafinil may be less effective on the number of people who withdrew from the trials due to adverse effects compared to those who took placebo, although there was some uncertainty.

Low quality of evidence from one RCT (n=55) showed armodafinil may be more effective in ameliorating the severity of BED compared with placebo, although there was some uncertainty.

8.4.4.10 Combined pharmacological and psychological interventions

Antidepressants and CBT-ED versus CBT-ED in adults with BED at end of treatment

Very low quality evidence from two RCTs (n=105) showed no difference in the effect of antidepressant and CBT-ED on binge frequency compared with CBT-ED alone.

Very low quality evidence from one RCT (n=40) showed no difference in the effect of antidepressant and CBT-ED on weight loss and family functioning compared with CBT-ED alone.

Very low quality evidence from one RCT (n=40) showed antidepressant and CBT-ED is less effective on EDI-2-bulimia compared with CBT_ED alone.

Very low quality evidence from two RCTs (n=105) showed antidepressant and CBT-ED may be less effective on the number of people who did not withdraw due to adverse events compared with CBT-ED alone, although there was some uncertainty.

Very low quality evidence from two RCTs (n=70) showed no difference in the effect of antidepressant and CBT-ED on depression compared with CBT-ED alone.

Antidepressants and CBT-ED versus CBT-ED in adults with BED at follow up

Low quality evidence from one RCT (n=50) showed antidepressant and CBT is more effective on binge frequency compared with CBT-ED alone.

Antidepressant and CBT-ED versus placebo and CBT-ED in adults with BED at end of treatment

Very low to low quality evidence from one RCT (n=54) showed no difference in the effect of Antidepressants and CBT-ED on remission, BMI, binge frequency, EDE-Q-global, EDE-Q-dietary restraint, EDE-Q-eating concerns, EDE-Q-weight concerns, EDE-Q-shape concerns and depression compared with placebo and CBT-ED.

Antidepressant and CBT-ED versus placebo and CBT-ED in adults with BED at follow up

Very low quality evidence from one RCT (n=54) showed no difference in the effect of Antidepressants and CBT-ED on remission compared with placebo and CBT-ED.

Very low to low quality evidence from one RCT (n=41) showed no difference in the effect of antidepressants and CBT-ED, BMI, binge frequency, EDE-Q-global, EDE-Q-dietary restraint, EDE-Q-eating concerns, EDE-Q-weight concerns, EDE-Q-shape concerns and depression compared with placebo and CBT-ED.

Antidepresssant-1 and CBT-ED versus antidepressant-2 and CBT-ED in adults with BED at end of treatment

Low quality evidence from one RCT (n=45) showed no difference in the effect of one antidepressant and CBT-ED on binge frequency compared with another antidepressant and CBT-ED.

Very low quality evidence from one RCT (n=45) showed one antidepressant and CBT-ED may be less effective on the number of people who withdrew from the RCT due to adverse events score compared with another antidepressant and CBT-ED, although there was some uncertainty.

Very low quality evidence from one RCT (n=20) showed no difference in the effect of one antidepressant and CBT-ED on Binge Eating Scale score compared with another antidepressant and CBT-ED.

Antidepresssant-1 and CBT-ED versus antidepressant-2 and CBT-ED in adults with BED at follow up

Low quality evidence from one RCT (n=34) showed no difference in the effect of one antidepressant and CBT-ED on binge frequency compared with another antidepressant and CBT-ED.

Antiepileptics and group CBT-ED versus placebo and group CBT-ED in adults with BED

Low quality evidence from one RCT (n=56) showed no difference in the effect of antiepileptics and group CBT-ED in reducing BMI, BES score and depression compared with placebo and group CBT-ED.

Low quality evidence from one RCT (n=56) showed antiepileptics and group CBT-ED may be more effective in the number of people achieving more than 10% weight loss compared with placebo and group CBT-ED, although there was some uncertainty.

Very low quality from one RCT (n=73) showed antiepileptics and group CBT-ED may be less effective on the number of people who withdrew due to adverse effects compared to those who received placebo and group CBT-ED, although there was some uncertainty.

Antidepressant, antiepileptic, group behavioural weight loss therapy and group CBT versus antidepressant, group behavioural weight loss therapy and group CBT

Low quality evidence from one RCT (n=20) showed no difference in the effect of antidepressant, antiepileptic, group behavioural weight loss therapy and group CBT on BMI compared with a combined antidepressant, group behavioural weight loss therapy and group CBT.

Antiobesity agent and guided self-help CBT-ED versus placebo and guided self-help (CBT-ED) in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=50) showed an antiobesity agent and guided self-help CBT-ED may be more effective on remission (ITT) and depression compared with placebo and guided self-help CBT-ED, although there was some uncertainty.

Low quality evidence from one RCT (n=50) showed antiobesity agent and guided self-help CBT-ED is more effective on number of people losing 5% or more weight, weight loss and mean percentage weight loss compared with placebo and guided self-help CBT-ED.

Low quality evidence from one RCT (n=50) showed no difference in the effect of antiobesity agent and guided self-help CBT-ED on binge frequency, EDE-global, EDE-dietary restraint, EDE-eating concern, EDE-weight concern and EDE-shape concern compared with placebo and guided self-help CBT-ED.

Antiobesity agent and guided self-help CBT-ED versus placebo and guided self-help (CBT-ED) in adults with binge eating disorder at follow up

Low quality evidence from one RCT (n=50) showed no difference in the effect of antiobesity agent and guided self-help CBT-ED on remission (ITT), binge frequency, EDE-global, EDE-dietary restraint, EDE-eating concern, EDE-weight concern and EDE-shape concern compared with placebo and guided self-help CBT-ED.

Low quality evidence from one RCT (n=50) showed an antiobesity agent and guided self-help CBT-ED may be more effective on the number of people who lost 5% or more weight, weight loss, mean percentage weight loss and depression, compared with placebo and guided self-help CBT-ED, although there was some uncertainty.

8.4.5 Economic Evidence statements

The existing economic evidence on pharmacological interventions for people with binge eating disorder was very limited and not directly applicable to the NICE decision-making context. The reviewed modelling study found that lisdexamfetamine dimesylate was potentially cost effective when compared with no drug treatment in the US. However, this drug is not licensed for the use in the UK. The reviewed study was characterised by minor methodological limitations.

Generally, the data on pharmacological interventions for people with BED was very scarce with very small numbers randomised. Pharmacological interventions showed no benefit in the NMA conducted for this guideline and they were not considered in the economic model. Details on the methods used for the systematic search of the economic literature are described in Chapter 3. Details and findings of the NMA are provided in the Appendix R.

8.4.6 Recommendations and link to evidence for the review on: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

Medication for binge eating disorder

	134. Do not offer medication as the sole treatment for binge eating disorder.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of pharmacotherapies for treating binge eating disorder. For this population, it was agreed binge eating frequency and remission were of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	In adults with binge eating disorder, antidepressants versus placebo showed little to no effect on the critical outcomes remission or binge eating frequency. There was also no difference EDE-global, EDE-subscales and a trend to reduce depression but there was some uncertainty. Global severity of illness favoured the antidepressant arms, but more people withdrew due to adverse events. There was also no difference in the severity of comorbid depressive disorders nor in their improvement. No evidence was found on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. A head-to-head of two different antidepressants showed no difference in binge frequency at the end of treatment. All other outcomes also showed no difference between the two arms including: body weight or BMI, binge frequency, binge eating scale, EDE-subscales, depression or global severity of illness. At 12 months' follow up binge frequency favoured fluvoxamine over fluoxetine. No other outcomes were reported. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. When antidepressants are compared with another intervention (CBT-ED), the results mostly favoured CBT-ED for binge frequency, weight loss, depression and EDI-2 Bulimia. Family functioning showed no difference. At 12 months' follow up, binge frequency favoured CBT-ED over antidepressants. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of

	life, all-cause mortality, relapse, general functioning, cost effectiveness, resource use, and service user experience. When an antiepileptic was compared with placebo no difference was found in remission rates and only a trend to reduce binge frequency (there was some uncertainty). Antiepileptics increase weight loss and improve the severity of illness but more people withdrew due to side-effects compared with placebo. No difference was found in EDE-total, EDE-restraint, EDE-eating concern and depression and a trend to improve EDE-weight concern and EDE-shape concern (but there was some uncertainty). No evidence was found on the important outcomes of quality of life, all-cause mortality, relapse, family functioning, cost effectiveness, resource use, and service user experience. Appetite suppressants (lisdexamphetamine) showed favourable results compared with placebo on remission, change in BMI and binge eating. There was also a trend of improvement in general physical functioning (though there was some uncertainty) in the appetite suppressant group, but no difference on general mental functioning. However, more people withdrew due to adverse events, and there was a trend towards higher depression scores in the appetite suppression arm compared with placebo. No evidence was found on the important outcomes of quality of life, all-cause mortality, relapse, cost effectiveness, resource use, and service user experience. Substance abuse treatment agents appeared to be less effective on remission (though there was some uncertainty) but have no effect on binge eating, BMI, weight, depression, global severity of illness and general function compared with placebo. More participants withdrew due to adverse events in the active treatment arm. No evidence was found on the important outcomes of quality of life, all-cause mortality, relapse, family functioning, cost effectiveness, resource use, and service user experience. Comparing a norepinephrine reuptake inhibitor (atomoxetine) with placebo in adults with binge eating disorder showed it may benefit remission, binge eating, loss in body weight and global severity of illness. No difference in depression or the number who withdrew due to adverse events. No evidence was found on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Armodafinil showed some benefit on reducing BMI, but had no effect on remission compared with placebo. There was a trend for it to improve binge frequency and global impressions severity, but it had no effect on depression nor withdrawals due to adverse events. No evidence was found on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Combining antidepressants with CBT-ED and comparing it with CBT-ED alone showed very little difference in any of the outcomes at the end of treatment. Binge frequency and binge eating scale were similar, as was depression, weight loss and adverse events. EDI-2 bulimia favoured CBT-ED alone. At 12 months' follow up, binge frequency was reduced in the combined treatment arm compared with CBT-ED alone. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, cost effectiveness, resource use, and service user experience. When antidepressants are combined with CBT-ED and compared with CBT-ED but this time with a placebo, the results again showed very little difference between the two groups at the end of treatment. The outcomes included remission, binge frequency, BMI, EDE-global, EDE-subscales, general psychopathology and depression. The same results were found at 12 months' follow up. No evidence was found on the critical outcome of adverse events, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Comparing two different antidepressants combined with CBT-ED showed no difference in any of the outcomes between the two treatment groups at the end of treatment or at follow up. The antidepressants compared included: fluoxetine

	versus fluvoxamine and paroxetine and venlafaxine. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. No differences were found between antiepileptics and group CBT-ED compared with placebo and group CBT-ED on binge eating, weight and depression. There was a trend to favour the number of people who achieved weight loss exceeding 10%. More people also withdrew due to adverse events although there was some uncertainty. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. No difference was found on weight between antidepressants combined with an antiepileptic, CBT-ED and diet compared with an antidepressant, CBT and diet. The only difference between the two arms was the antiepileptic treatment and thus it appears to have no effect. No evidence was found on the critical outcomes of remission and adverse events, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. An antiobesity agent (orlistat, a lipase inhibitor) combined with guided self-help CBT-ED compared with placebo and guided self-help CBT-ED showed no difference in remission rates and depression at the end of treatment but there was some uncertainty. However, more people lost weight if they were taking the antiobesity agent. No difference was found on EDE-subscales or binge frequency. At follow up, no difference between the groups was found in any of the outcomes. There was a trend for greater weight loss in those who were also taking the antiobesity agent but there was some uncertainty. No evidence was found on the critical outcome of adverse events, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Results presented in nutrition review chapter 8 Antidepressant and group behavioural weight control therapy compared with placebo and group behavioural weight control therapy showed no difference in any of the outcomes at the end of treatment. These outcomes included weight, binge frequency, general psychopathology and depression. No evidence was found on the critical outcomes of remission and adverse events, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. If an antidepressant, CBT-ED and group behavioural weight loss therapy is compared with placebo, CBT-ED and group behavioural weight loss therapy no difference was found in weight, binge frequency, general psychopathology and depression at the end of treatment. No evidence was found on the critical outcomes of remission and adverse events, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. If CBT-ED was first delivered prior to starting an antidepressant and group behavioural therapy and compared with CBT-ED followed by group behavioural therapy no difference was found on weight, binge frequency, general psychopathology and depression at the end of treatment. No evidence was found on the critical outcomes of remission and adverse events, nor on the important outcomes of quality of life, all-cause mortality, relapse, eating psychopathology, general functioning, family functioning, cost effectiveness, resource use, and service user experience. An antiobesity agent orlistat combined with a mild hypocaloric diet compared with placebo and a mild hypocaloric diet showed benefits on weight loss and general psychopathology, but no difference on the number of people who still satisfied DSM-IV criteria for BED, and quality of life. The antiobesity agent and mild hypocaloric diet group also favoured recovery from generalised anxiety and major

	depressive disorders, although there was some uncertainty. No evidence was found on the critical outcomes of remission and adverse events, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. When combined with behavioural weight loss therapy and compared with placebo, orlistat showed no difference on remission, BMI, and depression, nor on EDE-global and its subscales. At six months follow up, orlistat and behavioural weight loss therapy favoured depression. However there was no difference on the remaining outcomes. No evidence was found on the critical outcome of adverse events, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience.
Trade-off between net health benefits and resource use	There was no evidence for the effectiveness of pharmacological treatments for people with BED. As a result, the committee expressed the view that such treatments are also likely to be not cost-effective. There was some positive economic evidence on lisdexamphetamine. However, this drug is not licenced for the use in the UK. Also, there is a lack of data suggesting that the drug effect is enduring and as such it could not be recommended in people with BED.
Quality of evidence	The evidence on pharmacological agents for binge eating disorder was mostly very low quality. The evidence was downgraded for imprecision and risk of bias because it was unclear if there was adequate randomisation, if allocation concealment was performed, and if participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. For many comparisons only one or two studies were available with few participants. Imprecision was often detected because the 95% confidence interval crossed one or two minimal important differences or the outcome did not meet the optimal information size of 300 events or 400 participants. All of the outcomes were downgraded for risk of publication bias. In the 1980s, 1990s and early 2000s there was a risk that pharmaceutical companies only published positive findings, only reported outcomes that showed positive results and that outliers were excluded from the analysis (Lexchin 2003). Heterogeneity was detected in only a few outcomes. Some of the heterogeneity was eliminated when a sensitivity analysis was conducted on the studies that carried a high risk of bias, or it could be explained by variations in the severity of illness. For further explanation see the forest plots in the appendices.
Other considerations	There were few treatments that showed positive results on both remission and binge eating at the end of treatment and follow up. And follow up data was not always reported thus making it difficult for the committee to ascertain the long-term effects. Appetite suppressants appeared to show positive results on remission and weight loss at the end of treatment, however depression and general functioning appeared to favour the placebo arm at the end of treatment. No data was available at follow up. A study on an appetite suppressant (lisdexamphetamine) included in this review included a large sample size (n=1032) and it showed positive results on remission compared with placebo. However the committee wold not recommend it given that it is not licensed for treating eating disorders in the UK. There was one study that compared antidepressants with CBT-ED and showed better outcomes in binge frequency in the CBT-ED group at end of treatment and follow up. Granted it was only one small study of n=40, but it does provide further support for psychotherapy over pharmacotherapy. No evidence of drug treatment for children and young people with binge eating disorder was identified. In conclusion, because of the few studies identified for each comparison and the small sample size the committee agreed that the evidence was not strong enough to recommend any of the drug treatments as the sole treatment for binge eating disorder.

8.5 Nutritional interventions

8.5.1 Review question: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 299. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all nutritional interventions that may be delivered to children, young people and adults with an eating disorder with or without a pharmacological intervention. The interventions were categorised according to type of nutritional intervention, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to wait list controls, placebo, TAU or any other intervention.

Table 299: Clinical review protocol summary for the review of: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

Component	Description
Review question(s)	Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). **Strata:** • children (≤12), young people (13-≤17 years), adults ≥18 years • eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Nutritional intervention • Method of feeding • Nutritional in combination with any pharmacological intervention • Examples of nutritional interventions are nutritional counselling (with or without educational and supportive groups) and supplements (e.g. zinc)
Comparison	• Placebo • Wait list control • Treatment as usual • Another intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality

Component	Description
	• Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs

8.5.2 Clinical Evidence for Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

13 studies (n=1217) met the eligibility criteria for this review, all of which were in adults (Agras 1994 (Agras et al., 1994b), Barnes 2014 (Barnes et al., 2014); Devlin 2005 (Devlin et al., 2005), Golay 2005, Goodrick 1998 (Goodrick et al., 1998), Grilo 2005 (Grilo and Masheb, 2005), Masheb 2007 (Masheb and Grilo, 2007), Grilo 2011 (Grilo et al., 2011), Grilo 2013, Masheb 2011 (Masheb et al., 2011), Munsch 2007 (Munsch et al., 2007), Nauta 2000 (Nauta et al., 2000), Nauta 2001 (Nauta et al., 2001), Reeves 2001 (Reeves et al., 2001), Wilson 2010 (Wilson et al., 2010), Hilbert 2015 (Hilbert et al., 2015)). The majority of studies were in females and used individual or group behavioural weight loss therapy, which involves a substantial exercise component. In this sense, these studies can be considered combined interventions as they are not restricted to purely nutritional components. An overview of the trials included in the meta-analysis can be found in Table 300. Further information about both included and excluded studies can be found in Appendix J.

Eight studies (n=943) met the eligibility criteria for nutritional interventions compared with any other intervention, all of which were in adults (Barnes 2014; Goodrick 1998, Grilo 2005/Masheb 2007, Grilo 2011, Munsch 2007, Nauta 2000/2001, Reeves 2001, Wilson 2010/Hilbert 2015).

Six of the 13 studies (n=531) investigated the effectiveness of a nutritional treatment with another treatment compared with another intervention or wait list controls, some of which provided a number of different pairwise comparisons. Combined nutritional interventions included behavioural weight loss therapy and online motivational interviewing (Barnes 2014), CBT-ED with either general nutritional counselling or low energy density diet (Masheb 2011), an antidepressant fluoxetine with group behavioural weight control therapy (Devlin 2005), an antidepressant fluoxetine with CBT-ED and group behavioural weight control therapy (Agras 1994) and an antiobesity agent orlistat with either behavioural weight loss therapy (Grilo 2013) or a mild hypocaloric diet (Golay 2005).

Summary of findings for those on binge eating disorder can be found in Table 301, Table 302, Table 303, Table 304, Table 305, Table 306, Table 307, Table 308, Table 309, Table 310, Table 311, Table 312, Table 313, Table 314, Table 315, Table 316, Table 317, Table 318, Table 319, Table 320, Table 321, Table 322 and Table 323, Table 325 and Table 326. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 300: Study information for trials included in the meta-analysis of nutritional interventions versus any other intervention, wait list control or TAU for adults with binge eating disorder.

Study ID	N random-ised	Female (%)	Mean BMI (SD), kg/m2	Sample	Type of nutritional intervention	Intervention	Comparison	Duration
Agras 1994	108	100	38.6 (6.6)	Adult BED	Diet	CBT then Weight Loss Therapy + Desipramine	CBT then Weight Loss Therapy Weight Loss Therapy	9 months + 3-mo FU
Barnes 2014	89	76	35.3 (7.0)	Overweight and obese adults (BMI 25-55 kg/m2) with and without BED	Diet/Nutritional Counselling	Behavioural Weight Loss Therapy + Online Motivational Interviewing	Online Nutritional Counselling TAU	12 weeks + 3-mo FU
Devlin 2005	116	78	40.9 (6.9)	Adult BED ≥6 months	Group Diet	Group Behavioural Weight Control Therapy + Fluoxetine + CBT-ED	Group Behavioural Weight Control Therapy + Fluoxetine Group Behavioural Weight Control Therapy + Placebo Group Behavioural Weight Control Therapy + Placebo + CBT-ED	20 weeks
Golay 2005	89	91	36.5 (4.5)	Adult BED with obesity	Diet	Individualized hypocaloric diet + Orlistat (360 mg/day)	Individualized hypocaloric diet + placebo	24 weeks
Goodrick 1998	219	100	33.0 (3.3)	Adult 14-41 kgs overweight based on 1983 MLIC Height/Weight tables; BES score>21	Diet/Nutritional Counselling	Group Dieting Treatment	Group Nutritional Counselling WLC	6 months + 12-mo FU
Grilo 2005/	90	79	35.5	Adult BED,	Diet	Guided Self-Help	Guided Self-Help CBT-ED	12 weeks

Study ID	N random-ised	Female (%)	Mean BMI (SD), kg/m2	Sample	Type of nutritional intervention	Intervention	Comparison	Duration
Masheb 2007			(6.7)	BMI>27 kg/m2		Behavioural Weight Loss Therapy	Guided Self-Help	
Grilo 2011	125	67	38.8 (5.8)	Adult BED with obesity	Group Diet	Group Behavioural Weight Loss Therapy Age at onset: 44.6 (10.5)	Group CBT-ED Group CBT-ED then Group Behavioural Weight Loss Therapy	6 months
Grilo 2013	79	78	38.1 (6.2)	Adult BED with obesity	Diet	Behavioural Weight Loss treatment (adapted Diabetes Prevention Program) + Orlistat (320mg/day)	Behavioural Weight Loss treatment (adapted Diabetes Prevention Program) + Placebo	4 months + 6 months FU
Masheb 2011	50	76	39.1 (6.6)	Adult BED	Diet	Low-Energy Density Diet + CBT-ED Age at onset: 25.4 (12.2)	General Nutritional Counselling + CBT-ED Age at onset 23.1 (11.6)	6 months + 6 months FU
Munsch 2007	80	89	34.0 (4.0)*	Adult BED	Group Diet	Group Behavioural Weight Loss Therapy	Group CBT-ED	16 weeks + 6 assessment a month for 12 months
Nauta 2000/2001	37	100	33.1 (4.3)**	Obese adults with and without BED	Group Diet	Group Healthy Eating Program	Group Cognitive Therapy	15 weeks + 12 months FU
Reeves 2001	98	100	194.4 (22) lbs	Obese adults + BES score>20	Diet	Behavioural Self-Management	WLC	6 months
Wilson 2010/Hilbert 2015	205	85	36.4 (5.0)	Adult BED, BMI 27-45 kg/m2	Diet	Behavioural Weight Loss Therapy	Guided Self-Help CBT-ED Interpersonal Psychotherapy	6 months + 12 months

Study ID	N random-ised	Female (%)	Mean BMI (SD), kg/m2	Sample	Type of nutritional intervention	Intervention	Comparison	Duration
								and 24-months FU

*Notes: *n=75; **, figure is for whole sample including obese adults without BED. Abbreviations: BES, Binge Eating Scale; FU, follow up; MLIC, Metropolitan Life Insurance Company; TAU, treatment as usual; WLC, wait list controls.*

Table 301: Summary table of findings for online nutritional counselling versus TAU in adult with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Online Nutritional Counselling (95% CI)
Weight Change	59 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight change in the intervention groups was 0.72 standard deviations lower (1.25 to 0.19 lower)
EDE Global	59 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.4 standard deviations lower (0.92 lower to 0.11 higher)
Depression BDI	59 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.34 standard deviations lower (0.86 lower to 0.17 higher)
General functioning Treatment Self-Regulation Questionnaire - Autonomous Motivation	59 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general functioning in the intervention groups was 0.23 standard deviations higher (0.28 lower to 0.74 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%					

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Online Nutritional Counselling (95% CI)
confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Barnes 2014: Randomization method unclear (stratified by BED diagnosis), allocation concealment unclear. No participant nor investigator blinding. EDE Global scores significantly different at baseline. 2 Sample is adults BMI>25 and <55, overweight and obese eaters with (n=23) and without BED (n=66). 3 CI crosses either 0.5 or -0.5 (SMD).					

1 **Table 302: Summary table of findings for online nutritional counselling versus TAU in adult with BED at follow up.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Online Nutritional Counselling (95% CI)
Weight Change FU	59 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight change fu in the intervention groups was 0.74 standard deviations lower (1.27 to 0.21 lower)
EDE Global FU	59 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.24 standard deviations lower (0.76 lower to 0.27 higher)
Depression FU BDI	59 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to indirectness, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.35 standard deviations lower (0.86 lower to 0.17 higher)
General functioning Treatment Self-Regulation Questionnaire -	59 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3		Not calcula	The mean general functioning in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with Online Nutritional Counselling (95% CI)
Autonomous Motivation FU		due to risk of bias, indirectness, imprecision		ble for SMD values	0.11 standard deviations lower (0.62 lower to 0.4 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Barnes 2014: Randomization method unclear (stratified by BED diagnosis), allocation concealment unclear. No participant nor investigator blinding. EDE Global scores significantly different at baseline.
2 Sample is adults BMI>25 and <55, overweight and obese eaters with (n=23) and without BED (n=66).
3 CI crosses either 0.5 or -0.5 (SMD).

1 **Table 303: Summary table of findings for group nutritional counselling versus wait list control in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Group Nutritional Counselling (95% CI)
BMI	120 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.22 standard deviations higher (0.14 lower to 0.57 higher)
Binge Eating Scale	120 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.83 standard deviations lower (1.2 to 0.46 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Goodrick 1998: randomization method and allocation concealment unclear. No participant nor assessor blinding. Investigator blinding unclear. Reasons for dropout not clear. Participants paid fee to participate in study to be returned only if they attended>19 first 26 meetings and completion of 6- and 12-mo

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Group Nutritional Counselling (95% CI)
FU assessments. 2 Goodrick 1998: Women only. Participants were selected on basis of 14-41 kg overweight based on 1983 Metropolitan Life Insurance Company Height/Weight tables and having Binge Eating Scale score >21. 3 CI crosses either 0.5 or -0.5 (SMD).					

1 **Table 304: Summary table of findings for group behavioural weight loss therapy versus wait list control in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with GBWLT (95% CI)
BMI/Weight	205 (2 studies)	⨁⊖⊖⊖ VERY LOW1,2,3,4,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI/weight in the intervention groups was 0.20 standard deviations higher (0.07 lower to 0.48 higher)
Binge Eating Scale	123 (1 study)	⨁⊖⊖⊖ VERY LOW1,3,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 1.07 standard deviations lower (1.45 to 0.69 lower)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Goodrick 1998: randomization method and allocation concealment unclear. No participant nor assessor blinding. Investigator blinding unclear. Reasons for dropout not clear. Participants paid fee to participate in study to be returned only if they attended>19 first 26 meetings and completion of 6- and 12-mo FU assessments. 2 Reeves 2001: randomization method and allocation concealment unclear. No participant blinding. Assessor and investigator blinding unclear. Dropout rate of intervention group >20%. 3 Goodrick 1998: Women only. Participants were selected on basis of 14-41 kg overweight based on 1983 Metropolitan Life Insurance Company Height/Weight tables and having Binge Eating Scale score >21. 4 Reeves 2001: Women only. Participants were selected on basis of weight>=31 lbs or <90 lbs overweight based on 1983 Metropolitan Height/Weight tables, and Binge Eating Scale score >20. 5 <400 participants.					

Table 305: Summary table of findings for behavioural weight loss therapy versus any other intervention in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with BWLT (95% CI)
Remission	205 (1 study) 2 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.96 (0.84 to 1.11)	844 per 1000	34 fewer per 1000 (from 135 fewer to 93 more)
Rapid Response >=70% reduction binge eating by 4th week treatment	205 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision	RR 1.05 (0.88 to 1.27)	695 per 1000	35 more per 1000 (from 83 fewer to 188 more)
Binge Frequency EDE, past 28 days	205 (1 study) 2 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.07 standard deviations higher (0.22 lower to 0.37 higher)
BMI	205 (1 study) 2 years	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.12 standard deviations lower (0.41 lower to 0.18 higher)
EDE Global	205 (1 study) 2 years	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.36 standard deviations higher (0.06 to 0.66 higher)
# 5% Reduction in Weight	205 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 3 (2.08 to 4.33)	213 per 1000	426 more per 1000 (from 230 more to 709 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Wilson 2010/Hilbert 2015: adequate randomisation, unclear allocation concealment. No participant blinding, unclear investigator and assessor blinding. Dropout rates of Diet and CBT group >20%.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with BWLT (95% CI)
2 <300 events (dichotomous outcome) or <400 participants (continuous outcome). 3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).					

Table 306: Summary table of findings for behavioural weight loss therapy versus any other intervention in adults with BED at 1 year follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with BWLT (95% CI)
Binge Frequency 12-mo FU EDE Binges/past 28 days	205 (1 study) 1 years	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency 12-mo fu in the intervention groups was 0.24 standard deviations higher (0.06 lower to 0.54 higher)
BMI 12-mo FU	205 (1 study) 1 years	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI 12-mo fu in the intervention groups was 0.04 standard deviations higher (0.26 lower to 0.33 higher)
EDE Global 12-mo FU	205 (1 study) 1 years	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global 12-mo fu in the intervention groups was 0.41 standard deviations higher (0.11 to 0.71 higher)
# 5% Reduction in Weight 12-mo FU	205 (1 study) 1 years	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.26 (0.87 to 1.82)	333 per 1000	87 more per 1000 (from 43 fewer to 273 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Wilson 2010/Hilbert 2015: adequate randomisation, unclear allocation concealment. No participant blinding, unclear investigator and assessor blinding. Dropout rates of Diet and CBT group >20%.					

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with BWLT (95% CI)
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD). 3 <300 events (dichotomous outcome) or <400 participants (continuous outcome).					

Table 307: Summary table of findings for behavioural weight loss therapy versus any other intervention in adults with BED at 2 year follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with BWLT (95% CI)
Binge Frequency 24-mo FU EDE Binges/past 28 days	205 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency 24-mo fu in the intervention groups was 0.23 standard deviations higher (0.07 lower to 0.52 higher)
BMI 24-mo FU	205 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI 24-mo fu in the intervention groups was 0.07 standard deviations higher (0.22 lower to 0.37 higher)
EDE Global 24-mo FU	205 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global 24-mo fu in the intervention groups was 0.27 standard deviations higher (0.03 lower to 0.57 higher)
# 5% Reduction in Weight 24-mo FU	205 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.35 (0.92 to 1.96)	312 per 1000	109 more per 1000 (from 25 fewer to 300 more)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up					
1 Wilson 2010/Hilbert 2015: adequate randomisation, unclear allocation concealment. No participant blinding, unclear investigator and assessor blinding. Dropout rates of Diet and CBT group >20%.					

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with BWLT (95% CI)
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD). 3 <300 events (dichotomous outcome) or <400 participants (continuous outcome).					

1 **Table 308: Summary table of findings for guided self-help behavioural weight loss versus any other intervention in adults with BED.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with GSH BWL (95% CI)
Remission	90 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.52 (0.27 to 1.01)	500 per 1000	240 fewer per 1000 (from 365 fewer to 5 more)
Rapid Response >=65% reduction in binge eating by week 4 of treatment	75 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 0.76 (0.5 to 1.16)	622 per 1000	149 fewer per 1000 (from 311 fewer to 99 more)
BMI	90 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.06 standard deviations higher (0.37 lower to 0.49 higher)
Binge Frequency	90 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.29 standard deviations higher (0.14 lower to 0.72 higher)
EDE-Q Dietary Restraint	90 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q dietary restraint in the intervention groups was 0.28 standard deviations higher (0.15 lower to 0.71 higher)
EDE-Q Eating Concerns	90 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean ede-q eating concerns in the intervention groups was 0.26 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with GSH BWL (95% CI)
		imprecision			(0.17 lower to 0.69 higher)
EDE-Q Weight Concerns	90 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q weight concerns in the intervention groups was 0.03 standard deviations higher (0.4 lower to 0.46 higher)
EDE-Q Shape Concerns	90 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q shape concerns in the intervention groups was 0.05 standard deviations higher (0.38 lower to 0.48 higher)
Depression BDI	90 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.18 standard deviations higher (0.25 lower to 0.61 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Grilo 2005/Masheb 2007: No participant nor investigator blinding. Dropout rate for Guided Self-Help Behavioural Weight Loss Therapy >40%. Difference between other groups >20%.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 <400 participants.

1 **Table 309: Summary table of findings for group behavioural weight loss therapy versus any other intervention in adults with BED at**
2 **end of treatment**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with Group BWLT (95% CI)
Remission No OBEs/28 days (EDE)	207 (3 studies) 1 years	⊕⊖⊖⊖ VERY LOW1,2,3,4,5 due to risk of bias,	RR 0.99 (0.74 to 1.33)	429 per 1000	4 fewer per 1000 (from 111 fewer to 141 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with Group BWLT (95% CI)
		inconsistency, indirectness, imprecision			
Remission - subgroup analysis of severity of illness <18 binges/month No OBEs/28 days (EDE)	170 (2 studies) 1 years	⊕⊖⊖⊖ VERY LOW1,2,4,6 due to risk of bias, inconsistency, imprecision	RR 1.11 (0.79 to 1.54)	427 per 1000	47 more per 1000 (from 90 fewer to 231 more)
Remission - subgroup analysis of severity of illness >18 binges/month No OBEs/28 days (EDE)	37 (1 study) 1 years	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision	RR 0.66 (0.35 to 1.24)	667 per 1000	227 fewer per 1000 (from 433 fewer to 160 more)
No longer meets all DSM-IV BED criteria	37 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW3,6 due to risk of bias, indirectness, imprecision	RR 1.21 (0.88 to 1.65)	750 per 1000	158 more per 1000 (from 90 fewer to 487 more)
Binge Frequency Binge days or binge episodes in past 28 days	175 (3 studies) 1 years	⊕⊖⊖⊖ VERY LOW1,2,3,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.42 standard deviations higher (0.12 to 0.72 higher)
BMI or Weight	207 (3 studies) 1 years	⊕⊖⊖⊖ VERY LOW1,2,3,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean bmi or weight in the intervention groups was 0.54 standard deviations lower (0.82 to 0.26 lower)
Weight Loss (lbs)	90 (1 study) 1 years	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss (lbs) in the intervention groups was 0.53 standard deviations higher (0.11 to 0.96 higher)
EDE Global Scale from: 0 to 6.	90 (1 study) 1 years	⊕⊕⊖⊖ LOW1,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.12 standard deviations higher (0.3 lower to 0.53 higher)
EDE Restraint Scale from: 0 to 6.	175 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,7		Not calculable for SMD values	The mean ede restraint in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with Group BWLT (95% CI)
	1 years	due to risk of bias, indirectness, imprecision			0.17 standard deviations higher (0.12 lower to 0.47 higher)
EDE Shape Concern Scale from: 0 to 6.	175 (3 studies) 1 years	⊕⊖⊖⊖ VERY LOW1,2,3,4,6 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.22 standard deviations higher (0.27 lower to 0.71 higher)
EDE Shape Concern - subgroup analysis of severity of illness <18 binges/month Scale from: 0 to 6.	138 (2 studies) 1 years	⊕⊕⊖⊖ LOW1,2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern - subgroup analysis of severity of illness <18 binges/month in the intervention groups was 0.01 standard deviations higher (0.33 lower to 0.34 higher)
EDE Shape Concern - subgroup analysis of severity of illness >18 binges/month Scale from: 0 to 6.	37 (1 study) 1 years	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern - subgroup analysis of severity of illness >18 binges/month in the intervention groups was 0.83 standard deviations higher (0.15 to 1.51 higher)
EDE Weight Concern Scale from: 0 to 6.	175 (3 studies) 1 years	⊕⊖⊖⊖ VERY LOW1,2,3,4,6 due to risk of bias, inconsistency, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.16 standard deviations higher (0.44 lower to 0.77 higher)
EDE Weight Concern - subgroup analysis of severity of illness <18 binges/month Scale from: 0 to 6.	138 (2 studies) 1 years	⊕⊕⊖⊖ LOW1,2,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern - subgroup analysis of severity of illness <18 binges/month in the intervention groups was 0.1 standard deviations lower (0.43 lower to 0.23 higher)
EDE Weight Concern - subgroup analysis of severity of illness >18 binges/month Scale from: 0 to 6.	37 (1 study) 1 years	⊕⊕⊖⊖ LOW3,6 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern - subgroup analysis of severity of illness >18 binges/month in the intervention groups was 0.9 standard deviations higher (0.21 to 1.58 higher)
EDE Eating Concern	175	⊕⊖⊖⊖		Not calculable	The mean ede eating concern in the

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with Group BWLT (95% CI)
Scale from: 0 to 6.	(3 studies) 1 years	VERY LOW1,2,3,6 due to risk of bias, indirectness, imprecision		for SMD values	intervention groups was 0.22 standard deviations higher (0.07 lower to 0.52 higher)
Depression BDI. Scale from: 0 to 63.	184 (3 studies) 1 years	⨁⊖⊖⊖ VERY LOW1,2,3,7 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.12 standard deviations higher (0.17 lower to 0.41 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio;					
1 Grilo 2011: unclear allocation concealment. Participant blinding until start of treatment. Unclear investigator and assessor blinding. Group BWLT and Group CBT dropout rates both >20%. Dropout reasons not stated. 2 Munsch 2007: randomization method used permuted block design. Allocation concealment unclear. No participant, investigator nor assessor blinding. Dropout rates of both Group BWLT and Group CBT groups >20%. Dropout reasons not stated. 3 Nauta 2000/2001: randomization method and allocation concealment unclear. No investigator blinding, assessor blinding unclear. 4 I2>50%. 5 CI crosses both 0.75 and 1.25 (Risk Ratio). 6 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD). 7 <400 participants.					

Table 310: Summary table of findings for group behavioural weight loss therapy versus any other intervention in adults with BED at follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with GBWLT (95% CI)
Remission FU	108 (2 studies) 1 years	⨁⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, indirectness, imprecision	RR 0.92 (0.66 to 1.27)	613 per 1000	49 fewer per 1000 (from 208 fewer to 165 more)
Binge Frequency FU	197	⨁⊖⊖⊖		Not calculable for	The mean binge frequency fu in the

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with GBWLT (95% CI)
Binge days or episodes in past 28 days	(3 studies) 1 years	VERY LOW1,2,3,5,6 due to risk of bias, indirectness, imprecision		SMD values	intervention groups was 0.21 standard deviations higher (0.07 lower to 0.49 higher)
BMI or Weight FU	198 (3 studies) 1 years	⊕⊖⊖⊖ VERY LOW1,2,3,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI or weight fu in the intervention groups was 0.1 standard deviations lower (0.38 lower to 0.19 higher)
Weight Loss (lbs) FU	90 (1 study) 1 years	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss (lbs) fu in the intervention groups was 0.11 standard deviations higher (0.3 lower to 0.53 higher)
EDE Global FU	90 (1 study) 1 years	⊕⊕⊖⊖ LOW5,7 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.12 standard deviations higher (0.29 lower to 0.54 higher)
EDE Restraint FU	152 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede restraint fu in the intervention groups was 0.09 standard deviations higher (0.23 lower to 0.41 higher)
EDE Shape Concern FU	152 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede shape concern fu in the intervention groups was 0.03 standard deviations lower (0.35 lower to 0.3 higher)
EDE Weight Concern FU	152 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concern fu in the intervention groups was 0.1 standard deviations higher (0.23 lower to 0.42 higher)
EDE Eating Concern FU	152 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,5,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede eating concern fu in the intervention groups was 0.08 standard deviations lower (0.4 lower to 0.24 higher)
Depression FU BDI	161 (3 studies)	⊕⊖⊖⊖ VERY LOW1,2,3,5,6		Not calculable for SMD values	The mean depression fu in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any other intervention	Risk difference with GBWLT (95% CI)
		due to risk of bias, indirectness, imprecision			0.1 standard deviations higher (0.21 lower to 0.42 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Munsch 2007: randomization method used permuted block design. Allocation concealment unclear. No participant, investigator nor assessor blinding. Dropout rates of both Diet and Group CBT groups >20%. Dropout reasons not stated.
2 Nauta 2000/2001: randomization method and allocation concealment unclear. No investigator blinding, assessor blinding unclear.
3 Nauta 2000: Women only.
4 CI crosses both 0.75 and 1.25 (Risk Ratio).
5 Grilo 2011: unclear allocation concealment. Participant blinding until start of treatment. Unclear investigator and assessor blinding. Diet and Group CBT dropout rates both >20%. Dropout reasons not stated.
6 <400 participants.
7 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 311: Summary table of findings for group behavioural weight loss therapy versus group nutritional counselling in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group Nutritional Counselling	Risk difference with GBWLT (95% CI)
BMI	127 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.1 standard deviations lower (0.45 lower to 0.25 higher)
Binge Eating Scale	127 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.24 standard deviations lower (0.59 lower to 0.11 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group Nutritional Counselling	Risk difference with GBWLT (95% CI)
CI: Confidence interval;					
1 Goodrick 1998: randomization method and allocation concealment unclear. No participant nor assessor blinding. Investigator blinding unclear. Reasons for dropout not clear. Participants paid fee to participate in study to be returned only if they attended>19 first 26 meetings and completion of 6- and 12-mo FU assessments. 2 Goodrick 1998: Women only. Participants were selected on basis of 14-41 kg overweight based on 1983 Metropolitan Life Insurance Company Height/Weight tables and having Binge Eating Scale score >21. 3 <400 participants. 4 CI crosses 0.5 or -0.5 (SMD).					

1 **Table 312: Summary table of findings for group behavioural weight loss therapy versus group nutritional counselling in adults with**
2 **BED at follow up.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group Nutritional Counselling	Risk difference with GBWLT (95% CI)
BMI FU	127 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.1 standard deviations higher (0.25 lower to 0.44 higher)
Binge Eating Scale FU	127 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge eating scale fu in the intervention groups was 0.07 standard deviations lower (0.41 lower to 0.28 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; FU: Follow up					
1 Goodrick 1998: randomization method and allocation concealment unclear. No participant nor assessor blinding. Investigator blinding unclear. Reasons for dropout not clear. Participants paid fee to participate in study to be returned only if they attended>19 first 26 meetings and completion of 6- and 12-mo FU assessments. 2 Goodrick 1998: Women only. Participants were selected on basis of 14-41 kg overweight based on 1983 Metropolitan Life Insurance Company Height/Weight tables and having Binge Eating Scale score >21. 3 <400 participants.					

Table 313: Summary table of findings for behavioural weight loss therapy and online motivational interviewing versus TAU in adult with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with BWLT + Motivational Interviewing (95% CI)
Weight Change	60 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight change in the intervention groups was 0.45 standard deviations lower (0.96 lower to 0.06 higher)
EDE Global	60 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.23 standard deviations higher (0.28 lower to 0.74 higher)
Depression BDI	60 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.1 standard deviations lower (0.61 lower to 0.41 higher)
General functioning Treatment Self-Regulation Questionnaire - Autonomous Motivation	60 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general functioning in the intervention groups was 0.34 standard deviations higher (0.17 lower to 0.85 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Barnes 2014: Randomization method unclear (stratified by BED diagnosis), allocation concealment unclear. No participant nor investigator blinding. EDE Global scores significantly different at baseline.
2 Sample is adults BMI>25 and <55, overweight and obese eaters with (n=23) and without BED (n=66).
3 CI crosses either 0.5 or -0.5 (SMD).

Table 314: Summary table of findings for behavioural weight loss therapy and online motivational interviewing versus TAU in adult with BED at follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with TAU	Risk difference with BWLT + Online Motivational interviewing (95% CI)
Weight Change FU	60 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight change fu in the intervention groups was 0.37 standard deviations lower (0.88 lower to 0.14 higher)
EDE Global FU	60 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.21 standard deviations higher (0.3 lower to 0.72 higher)
Depression FU BDI	60 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.06 standard deviations lower (0.57 lower to 0.44 higher)
General functioning Treatment Self-Regulation Questionnaire - Autonomous Motivation FU	60 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general functioning in the intervention groups was 0.1 standard deviations lower (0.61 lower to 0.4 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Barnes 2014: Randomization method unclear (stratified by BED diagnosis), allocation concealment unclear. No participant nor investigator blinding. EDE Global scores significantly different at baseline.
2 Sample is adults BMI>25 and <55, overweight and obese eaters with (n=23) and without BED (n=66).
3 CI crosses either 0.5 or -0.5 (SMD).

Table 315: Summary table of findings for behavioural weight loss therapy and online motivational interviewing versus online nutritional counselling in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Online Nutritional Counselling	Risk difference with BWLT + Online Motivational Interviewing (95% CI)
Weight Change	59 (1 study) 3 weeks	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight change in the intervention groups was 0.25 standard deviations higher (0.26 lower to 0.76 higher)
EDE Global	59 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.74 standard deviations higher (0.21 lower to 1.27 higher)
Depression BDI	59 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.24 standard deviations higher (0.27 lower to 0.75 higher)
General functioning Treatment Self-Regulation Questionnaire - Autonomous Motivation	59 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general functioning in the intervention groups was 0.14 standard deviations higher (0.37 lower to 0.65 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Barnes 2014: Randomization method unclear (stratified by BED diagnosis), allocation concealment unclear. No participant nor investigator blinding. EDE Global scores significantly different at baseline.
2 Sample is adults BMI>25 and <55, overweight and obese eaters with (n=23) and without BED (n=66).
3 CI crosses either 0.5 or -0.5 (SMD).

Table 316: Summary table of findings for behavioural weight loss therapy and online motivational interviewing versus online nutritional counselling in adults with BED at follow up.

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with Online Nutritional Counselling	Risk difference with BWLT + Online Motivational Interviewing (95% CI)
Weight Change FU	59 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean weight change fu in the intervention groups was 0.35 standard deviations higher (0.17 lower to 0.86 higher)
EDE Global FU	59 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.46 standard deviations higher (0.06 lower to 0.97 higher)
Depression FU BDI	59 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.31 standard deviations higher (0.2 lower to 0.82 higher)
General functioning Treatment Self-Regulation Questionnaire - Autonomous Motivation FU	59 (1 study) 3 months	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean general functioning in the intervention groups was 0 standard deviations higher (0.51 lower to 0.51 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Barnes 2014: Randomization method unclear (stratified by BED diagnosis), allocation concealment unclear. No participant nor investigator blinding. EDE Global scores significantly different at baseline. 2 Sample is adults BMI>25 and <55, overweight and obese eaters with (n=23) and without BED (n=66). 3 CI crosses either 0.5 or -0.5 (SMD). 4 CI crosses both 0.5 and -0.5 (SMD).					

Table 317: Summary table of findings for low-energy density diet and CBT-ED versus general nutritional counselling and CBT-ED in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Nutritional Counselling + CBT-ED	Risk difference with LE Density Diet+CBT-ED (95% CI)
Remission	50	⨁⊖⊖⊖	RR 1.18	440 per 1000	79 more per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Nutritional Counselling + CBT-ED	Risk difference with LE Density Diet+CBT-ED (95% CI)
	(1 study) 6 months	VERY LOW1,2 due to risk of bias, imprecision	(0.66 to 2.11)		(from 150 fewer to 488 more)
BMI (Change scores)	50 (1 study) 6 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (change scores) in the intervention groups was 0.36 standard deviations higher (0.19 lower to 0.92 higher)
# >=5% weight loss	50 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.6 (0.61 to 4.22)	200 per 1000	120 more per 1000 (from 78 fewer to 644 more)
Mean % Weight Loss	50 (1 study) 6 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean % weight loss in the intervention groups was 0.3 standard deviations higher (0.26 lower to 0.86 higher)
EDE Global	50 (1 study) 6 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.2 standard deviations lower (0.75 lower to 0.36 higher)
EDE Weight Concern	50 (1 study) 6 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.39 standard deviations lower (0.95 lower to 0.17 higher)
EDE Shape Concern	50 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0 standard deviations higher (0.55 lower to 0.55 higher)
EDE Eating Concern	50 (1 study) 6 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concern in the intervention groups was 0.2 standard deviations higher (0.36 lower to 0.75 higher)
Depression	50 (1 study)	⊕⊕⊖⊖ LOW1,3		Not calculable for SMD values	The mean depression in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Nutritional Counselling + CBT-ED	Risk difference with LE Density Diet+CBT-ED (95% CI)
	6 months	due to risk of bias, imprecision			0.1 standard deviations higher (0.46 lower to 0.65 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Masheb 2011: Allocation concealment unclear. No participant blinding, investigator blinding unclear. Intervention group dropout rate=20%. No details of dropouts provided.
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 318: Summary table of findings for low-energy density diet and CBT-ED versus general nutritional counselling and CBT-ED in adults with BED at follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Nutritional Counselling + CBT-ED	Risk difference with LE Density diet + CBT-ED (95% CI)
BMI (change scores) FU	50 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (change scores) fu in the intervention groups was 0.26 standard deviations higher (0.3 lower to 0.81 higher)
Mean % Weight Loss FU	50 (1 study) 6 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean % weight loss fu in the intervention groups was 0.2 standard deviations higher (0.36 lower to 0.76 higher)
Binge Frequency FU EDE	50 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 0.54 standard deviations higher (0.02 lower to 1.11 higher)
# patients achieving >=5% weight loss FU	50 (1 study) 6 months	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias,	RR 1.17 (0.46 to 2.98)	240 per 1000	41 more per 1000 (from 130 fewer to 475 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with General Nutritional Counselling + CBT-ED	Risk difference with LE Density diet + CBT-ED (95% CI)
		imprecision			

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Masheb 2011: Allocation concealment unclear. No participant blinding, investigator blinding unclear. Intervention group dropout rate=20%. No details of dropouts provided.
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).

Table 319: Summary table of findings for group CBT-ED followed by group behavioural weight loss therapy versus group CBT-ED in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT-ED	Risk difference with Group Diet + Group CBT-ED (95% CI)
Remission	80 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.09 (0.68 to 1.75)	444 per 1000	40 more per 1000 (from 142 fewer to 333 more)
Binge Frequency binge episodes/month	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.18 standard deviations higher (0.26 lower to 0.62 higher)
BMI	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.07 standard deviations higher (0.37 lower to 0.51 higher)
Weight Loss	80 (1 study) 12 weeks	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss in the intervention groups was 0.44 standard deviations higher (0.01 lower to 0.88 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT-ED	Risk difference with Group Diet + Group CBT-ED (95% CI)
EDE Global	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.11 standard deviations lower (0.55 lower to 0.33 higher)
EDE Restraint	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint in the intervention groups was 0.11 standard deviations higher (0.34 lower to 0.55 higher)
EDE Eating Concern	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concern in the intervention groups was 0.32 standard deviations lower (0.77 lower to 0.12 higher)
EDE Shape Concern	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.15 standard deviations lower (0.59 lower to 0.3 higher)
EDE Weight Concern	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.17 standard deviations lower (0.61 lower to 0.27 higher)
Depression	80 (1 study) 12 months	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.04 standard deviations lower (0.49 lower to 0.4 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Grilo 2011: unclear allocation concealment. Participant blinding until start of treatment. Unclear investigator and assessor blinding. Diet+Group CBT and Group CBT groups dropout rates both >20%. Dropout reasons not stated.
2 CI crosses both 0.75 and 1.25 (Risk Ratio).
3 CI crosses either 0.5 or -0.5 (SMD).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT-ED	Risk difference with Group Diet + Group CBT-ED (95% CI)
4 <400 participants.					

Table 320: Summary table of findings for group CBT-ED followed by group behavioural weight loss therapy versus group CBT-ED in adults with BED at follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT-ED	Risk difference with Group CBT-ED then Group BWLT (95% CI)
Binge Frequency FU binge episodes/month	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency fu in the intervention groups was 0.19 standard deviations higher (0.25 lower to 0.64 higher)
BMI FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.07 standard deviations higher (0.37 lower to 0.51 higher)
Weight Loss FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss fu in the intervention groups was 0.14 standard deviations higher (0.3 lower to 0.59 higher)
EDE Global FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global fu in the intervention groups was 0.12 standard deviations lower (0.56 lower to 0.32 higher)
EDE Restraint FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint fu in the intervention groups was 0.09 standard deviations lower (0.53 lower to 0.36 higher)
EDE Eating Concern FU	80 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concern fu in the intervention groups was 0 standard deviations higher (0.44 lower to 0.44 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT-ED	Risk difference with Group CBT-ED then Group BWLT (95% CI)
EDE Shape Concern FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern fu in the intervention groups was 0.23 standard deviations lower (0.67 lower to 0.22 higher)
EDE Weight Concern FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern fu in the intervention groups was 0.09 standard deviations lower (0.54 lower to 0.35 higher)
Depression FU	80 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.07 standard deviations higher (0.37 lower to 0.51 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Grilo 2011: unclear allocation concealment. Participant blinding until start of treatment. Unclear investigator and assessor blinding. Diet+Group CBT and Group CBT groups dropout rates both >20%. Dropout reasons not stated. 2 CI crosses either 0.5 or -0.5 (SMD). 3 <400 participants.					

Table 321: Summary table of findings for antidepressant and group behavioural weight loss therapy versus placebo and group behavioural weight loss therapy in adults with BED.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo + GBWLT	Risk difference with Antidepressant + GBWLT (95% CI)
Weight	63 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable for SMD values	The mean weight in the intervention groups was 0.03 standard deviations higher

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo + GBWLT	Risk difference with Antidepressant + GBWLT (95% CI)
		imprecision			(0.46 lower to 0.53 higher)
Binge Frequency EDE OBE	63 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.16 standard deviations lower (0.66 lower to 0.33 higher)
Binge Eating Scale	63 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.13 standard deviations lower (0.62 lower to 0.37 higher)
General Psychopathology Brief symptom inventory	63 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0.07 standard deviations lower (0.56 lower to 0.43 higher)
Depression	63 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.38 standard deviations lower (0.88 lower to 0.12 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Devlin 2005: Randomization method and allocation concealment unclear. Dropout rates of all groups>20%. Dropout by groups not provided. Not clear if baseline measures for groups are similar.
2 CI crosses either 0.5 or -0.5 (SMD).

1 **Table 322: Summary table of findings for antidepressant, CBT-ED and group behavioural control therapy versus placebo, CBT-ED and**
2 **group behavioural weight control therapy in adults with BED.**

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Placebo+CBT-ED+GBWCT	Risk difference with Antidepressant+CBT-ED+GWBCT (95% CI)
Weight	53 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight in the intervention groups was 0.08 standard deviations lower (0.62 lower to 0.46 higher)
Binge Frequency EDE OBE	53 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge frequency in the intervention groups was 0.24 standard deviations lower (0.78 lower to 0.3 higher)
Binge Eating Scale	53 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 0.06 standard deviations lower (0.6 lower to 0.48 higher)
General Psychopathology	53 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0.19 standard deviations lower (0.73 lower to 0.35 higher)
Depression	53 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - fluoxetine+group behavioural weight control+cbt in the intervention groups was 0.24 standard deviations lower (0.78 lower to 0.3 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Devlin 2005: Randomization method and allocation concealment unclear. Dropout rates of all groups>20%. Dropout by groups not provided. Not clear if baseline measures for groups are similar. 2 CI crosses either 0.5 or -0.5 (SMD).					

Table 323: Summary table of findings for CBT-ED then antidepressant and group behavioural weight loss therapy versus CBT-ED then group behavioural weight loss therapy in adults with BED.

Outcomes	No of Participants	Quality of the evidence (GRADE)	Relative effect	Anticipated absolute effects

	(studies) Follow up		(95% CI)	Risk with CBT-ED then GBWLT	Risk difference with CBT-ED then Antidepressant + GBWLT (95% CI)
Weight	72 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean weight in the intervention groups was 0.28 standard deviations higher (0.18 lower to 0.74 higher)
Depression BDI	72 (1 study) 3 months	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision, publication bias		Not calculable for SMD values	The mean depression in the intervention groups was 0.14 standard deviations lower (0.6 lower to 0.32 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Agras 1994: Randomization method and allocation concealment unclear. No participant blinding, investigator and assessor blinding unclear. Dropout rate of CBT+Weight Loss+Desipramine and Weight Loss groups both >20%. Reasons for dropout not provided.
2 CI crosses either 0.5 or -0.5 (SMD).
3 Published before 2000.

1 **Table 324: Antiobesity agent and diet versus placebo and diet in adults with binge eating disorder at end of treatment**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+Diet	Risk difference with Antiobesity+Diet (95% CI)
Weight loss	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight loss in the intervention groups was 0.9 standard deviations higher (0.47 to 1.33 higher)
No longer meets BED DSM-IV criteria	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.09 (0.83 to 1.44)	706 per 1000	64 more per 1000 (from 120 fewer to 311 more)
EDI Total	89 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total in the intervention groups was 0.3 standard deviations lower (0.72 lower to 0.12 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+Diet	Risk difference with Antiobesity+Diet (95% CI)
General psychopathology HADS	89 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0.42 standard deviations lower (0.84 lower to 0 higher)
Depression BDI	89 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.40 standard deviations lower (0.82 lower to 0.02 higher)
No longer meets Generalized Anxiety disorder DSM-IV criteria	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.2 (0.87 to 1.66)	618 per 1000	124 more per 1000 (from 80 fewer to 408 more)
No longer meets Major depressive disorder DSM-IV criteria	73 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision	RR 1.1 (0.97 to 1.26)	882 per 1000	88 more per 1000 (from 26 fewer to 229 more)
Quality of Life Nottingham Health Profile	89 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean quality of life in the intervention groups was 0.2 standard deviations lower (0.62 lower to 0.21 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio;

1 Golay 2005: high risk of bias (unclear whether baseline similar, unclear randomisation method and allocation concealment; placebo+diet arm dropout rate>20%).
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 325: Antiobesity agent and behavioural weight loss therapy versus placebo and behavioural weight loss therapy in adults with binge eating disorder at end of treatment

Outcomes	No of	Quality of the evidence	Relative	Anticipated absolute effects

	Participants (studies) Follow up	(GRADE)	effect (95% CI)	Risk with Placebo+BWLT	Risk difference with Antiobesity+BWLT (95% CI)
Remission (ITT) No OBEs in past 28 days	40 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision	RR 0.86 (0.54 to 1.36)	700 per 1000	98 fewer per 1000 (from 322 fewer to 252 more)
BMI	38 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean bmi in the intervention groups was 0.31 standard deviations higher (0.33 lower to 0.95 higher)
EDE Global Scale from: 0 to 6.	38 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.49 standard deviations lower (1.13 lower to 0.16 higher)
EDE Dietary restraint Scale from: 0 to 6.	38 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede dietary restraint in the intervention groups was 0.28 standard deviations lower (0.92 lower to 0.36 higher)
EDE Eating concern Scale from: 0 to 6.	38 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede eating concern in the intervention groups was 0 standard deviations higher (0.64 lower to 0.64 higher)
EDE Shape concern Scale from: 0 to 6.	38 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.27 standard deviations lower (0.91 lower to 0.37 higher)
EDE Weight concern Scale from: 0 to 6.	38 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.51 standard deviations lower (1.15 lower to 0.14 higher)
Depression BDI	38 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.51 standard deviations lower (1.16 lower to 0.13 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+BWLT	Risk difference with Antiobesity+BWLT (95% CI)
CI: Confidence interval; RR: Risk ratio;					
1 Grilo 2013: high risk of bias (unclear randomisation method and allocation concealment, dropout rate of both groups >=20%). Participants limited to Latino/Latina patients. 2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD). 3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).					

Table 326: Antiobesity agent and behavioural weight loss therapy versus placebo and behavioural weight loss therapy in adults with binge eating disorder at follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+Diet	Risk difference with Antiobesity+Diet at 6-mo FU (95% CI)
Remission (ITT)	40 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision	RR 1 (0.54 to 1.86)	500 per 1000	0 fewer per 1000 (from 230 fewer to 430 more)
BMI	37 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean bmi in the intervention groups was 0.16 standard deviations higher (0.49 lower to 0.81 higher)
EDE Global	37 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.43 standard deviations lower (1.08 lower to 0.22 higher)
EDE Dietary restraint	37 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede dietary restraint in the intervention groups was 0.08 standard deviations lower (0.73 lower to 0.56 higher)
EDE Eating concern	37 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness,		Not calculable for SMD	The mean ede eating concern in the intervention groups was 0.54 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Placebo+Diet	Risk difference with Antiobesity+Diet at 6-mo FU (95% CI)
		imprecision		values	(1.2 lower to 0.12 higher)
EDE Shape concern	37 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.32 standard deviations lower (0.97 lower to 0.32 higher)
EDE Weight concern	37 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.29 standard deviations lower (0.94 lower to 0.36 higher)
Depression	37 (1 study)	⨁⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.94 standard deviations lower (1.62 to 0.25 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Grilo 2013: high risk of bias (unclear randomisation method and allocation concealment, dropout rate of both groups >=20%). Participants limited to Latino/Latina patients.
2 CI crosses both 0.75 and 1.25 (Risk Ratio), or both 0.5 and -0.5 (SMD).
3 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

8.5.3 Economic Evidence

No economic evidence on the cost effectiveness of nutritional interventions for people with binge eating disorder was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.5.4 Clinical evidence statements

8.5.4.1 Online nutritional counselling versus treatment as usual in adults with binge eating disorder at end of treatment

Very low quality evidence from one RCT (n=59) showed online nutritional counselling is less effective on weight change compared with treatment as usual.

Very low quality evidence from one RCT (n=59) showed no difference in the effect of online nutritional counselling on EDE-global, depression and general functioning compared with treatment as usual.

8.5.4.2 Online nutritional counselling versus treatment as usual in adults with binge eating disorder at follow up

Very low quality evidence from one RCT (n=59) showed online nutritional counselling is less effective on weight change compared with treatment as usual.

Very low quality evidence from one RCT (n=59) showed no difference in the effect of online nutritional counselling on EDE-global, depression and general functioning compared with treatment as usual.

8.5.4.3 Group nutritional counselling versus wait list control in adults with BED

Very low quality evidence from one RCT (n=120) showed no difference in the effect of group nutritional counselling on BMI and Binge Eating Scale compared with wait list control.

Very low quality evidence from one RCT (n=120) showed group nutritional counselling is more effective on BMI and Binge Eating Scale compared with wait list control.

8.5.4.4 Group behavioural weight loss therapy versus wait list control in adults with BED

Very low quality evidence from two RCTs (n=205) showed group behavioural weight loss therapy may be more effective on BMI/Weight compared with wait list control, although there was some uncertainty.

Very low quality evidence from one RCT (n=123) showed group behavioural weight loss therapy is more effective on Binge Eating Scale score compared with wait list control.

8.5.4.5 Behavioural weight loss therapy versus any other intervention in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy may be less effective on remission compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy may be more effective on remission and rapid response compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=205) showed no difference in the effect of behavioural weight loss therapy on binge frequency and BMI compared with any other intervention.

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy is more effective on the number of people achieving a 5% or more reduction in weight compared with any other intervention.

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy is less effective on EDE-global compared with any other intervention.

8.5.4.6 Behavioural weight loss therapy versus any other intervention in adults with binge eating disorder at 1 year follow up

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy may be less effective on reducing binge frequency compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy may be more effective on increasing the number of people achieving a 5% or more reduction in weight compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=205) showed no difference in the effect of behavioural weight loss therapy on BMI compared with any other intervention.

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy is less effective on EDE-global compared with any other intervention.

8.5.4.7 Behavioural weight loss therapy versus any other intervention in adults with binge eating disorder at 2 year follow up

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy may be less effective on reducing binge frequency and scores on EDE-global compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=205) showed behavioural weight loss therapy may be more effective on increasing the number of people achieving a 5% or more reduction in weight compared with any other intervention, although there was some uncertainty.

Low quality evidence from 1 one RCT (n=205) showed no difference in the effect of behavioural weight loss therapy on BMI compared with any other intervention.

8.5.4.8 Guided self-help behavioural weight loss versus any other intervention in adults with binge eating disorder

Low quality evidence from one RCT (n=90) showed guided self-help behavioural weight loss may be less effective on remission compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=75) showed guided self-help behavioural weight loss may be less effective on rapid response compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=90) showed no difference in the effect of guided self-help behavioural weight loss on BMI, binge frequency, EDE-Q-dietary restraint, EDE-Q-eating concern, EDE-Q-weight concern, EDE-Q-shape concern and depression compared with any other intervention.

8.5.4.9 Group behavioural weight loss therapy versus any other intervention in adults with binge eating disorder at end of treatment

Very low quality evidence from three RCTs (n=207) showed no difference in the effect of group behavioural weight loss therapy on remission compared with any other intervention.

Very low quality evidence from two RCTs (n=170) showed group behavioural weight loss therapy may be more effective on remission for people who engaged in less than 18 binges per month compared with any other intervention, although there was some uncertainty.

Low quality evidence from one RCT (n=37) showed no difference in the effect of group behavioural weight loss therapy on remission for people who engaged in more than 18 binges per month compared with any other intervention.

Very low quality evidence from three RCTs (n=206) showed that group behavioural weight loss therapy is more effective on BMI/Weight compared with any other intervention.

Very low quality evidence from one RCT (n=37) showed group behavioural weight loss therapy may be effective on increasing the number of people who no longer met all DSM-IV BED criteria compared with any other intervention, although there was some uncertainty.

Very low quality evidence from three RCTs (n=175) showed that group behavioural weight loss therapy is less effective on binge frequency compared with any other intervention.

Very low quality evidence from one RCT (n=90) showed that group behavioural weight loss therapy is more effective on weight loss compared with any other intervention.

Very low quality evidence from 1 one RCT (n=90) showed no difference in the effect of group behavioural weight loss therapy on EDE-global compared with any other intervention.

Very low quality evidence from three RCTs (n=175) showed no difference in the effect of group behavioural weight loss therapy on EDE-dietary restraint, EDE-shape concern, EDE-weight concern and depression compared with any other intervention.

Low quality evidence from two RCTs (n=170) showed no difference in the effect of group behavioural weight loss therapy on EDE-shape concern for people who engaged in less than 18 binges per month compared with any other intervention.

Low quality evidence from one RCT (n=37) showed group behavioural weight loss therapy was less effective on EDE-shape concern for people who engaged in more than 18 binges per month compared with any other intervention.

Low quality evidence from two RCTs (n=170) showed no difference in the effect of group behavioural weight loss therapy on EDE-weight concern for people who engaged in less than 18 binges per month compared with any other intervention.

Low quality evidence from one RCT (n=37) showed group behavioural weight loss therapy was less effective on EDE-weight concern for people who engaged in more than 18 binges per month compared with any other intervention.

Very low quality evidence from three RCTs (n=175) showed group behavioural weight loss therapy may be less effective on reducing EDE-eating concern compared with any other intervention, although there was some uncertainty.

Very low quality evidence from three RCTs (n=184) showed no difference in the effect of group behavioural weight loss therapy on depression compared with any other intervention.

8.5.4.10 Group behavioural weight loss therapy versus any other intervention in adults with binge eating disorder at follow up

Very low quality evidence from two RCTs (n=108) showed no difference in the effect of group behavioural weight loss therapy on remission compared with any other intervention.

Very low quality evidence from three RCTs (n=166) showed group behavioural weight loss therapy is less effective on binge frequency compared with any other intervention.

Very low quality evidence from three RCTs (n=198) showed no difference in the effect of group behavioural weight loss therapy on BMI/Weight compared with any other intervention.

Low quality evidence from one RCT (n=90) showed no difference in the effect of group behavioural weight loss therapy on weight loss and EDE-global compared with any other intervention.

Very low quality evidence from three RCTs (n=152) showed no difference in the effect of group behavioural weight loss therapy on EDE-dietary restraint, EDE-shape concern, EDE-weight concern and EDE-eating concern compared with any other intervention.

Very low quality evidence from three RCTs (n=161) showed no difference in the effect of group behavioural weight loss therapy on depression compared with any other intervention.

8.5.4.11 Group behavioural weight loss therapy versus group nutritional counselling in adults with BED at end of treatment

Very low quality evidence from one RCT (n=127) showed no difference in the effect of Group behavioural weight loss therapy on BMI and Binge Eating Scale score compared with group nutritional counselling.

8.5.4.12 Group behavioural weight loss therapy versus group nutritional counselling in adults with BED at follow up

Very low quality evidence from one RCT (n=127) showed no difference in the effect of group behavioural weight loss therapy on BMI and Binge Eating Scale score compared with group nutritional counselling.

8.5.4.13 Behavioural weight loss therapy and online motivational interviewing versus treatment as usual in adults with binge eating disorder at end of treatment

Very low quality evidence from one RCT (n=60) showed behavioural weight loss therapy and online motivational interviewing may be less effective on weight change compared with any treatment as usual, although there was some uncertainty.

Very low quality evidence from one RCT (n=60) showed no difference in the effect of behavioural weight loss therapy and online motivational interviewing on EDE-global, depression and general functioning compared with treatment as usual.

8.5.4.14 Behavioural weight loss therapy and online motivational interviewing versus treatment as usual in adults with binge eating disorder at follow up

Very low quality evidence from one RCT (n=60) showed no difference in the effect of behavioural weight loss therapy and online motivational interviewing on weight change, EDE-global, depression and general functioning compared with treatment as usual.

8.5.4.15 Behavioural weight loss therapy and online motivational interviewing versus online nutritional counselling at end of treatment

Very low quality evidence from one RCT (n=59) showed no difference in the effect of behavioural weight loss therapy and online motivational interviewing on weight change, EDE-global, depression and general functioning compared with online nutritional counselling.

8.5.4.16 Behavioural weight loss therapy and online motivational interviewing versus online nutritional counselling at follow up

Very low quality evidence from one RCT (n=59) showed no difference in the effect of behavioural weight loss therapy and online motivational interviewing on weight change, depression and general functioning compared with online nutritional counselling.

Very low quality evidence from one RCT (n=59) showed behavioural weight loss therapy and online motivational interviewing may be less effective on EDE-global compared with online nutritional counselling, although there was some uncertainty.

8.5.4.17 Low-energy density diet and CBT-ED versus general nutritional counselling and CBT-ED in adults with BED at end of treatment

Very low to low quality evidence from one RCT (n=50) showed no difference in the effect of low-energy density diet and CBT-ED on remission, change in BMI, mean % weight loss, EDE-global, EDE-weight concern, EDE-shape concern, EDE-eating concern and depression compared with general nutritional counselling and CBT-ED.

Very low quality evidence from one RCT (n=50) showed low-energy density diet and CBT-ED may be more effective on the number of people losing greater than 5% of their weight compared with general nutritional counselling and CBT-ED, although there was some uncertainty.

8.5.4.18 Low-energy density diet and CBT-ED versus general nutritional counselling and CBT-ED in adults with BED at follow up

Very low to low quality evidence from one RCT (n=50) showed no difference in the effect of low-energy density diet and CBT-ED on change in BMI, mean % weight loss and the number of people losing greater than 5% of their weight compared with general nutritional counselling and CBT-ED.

Low quality evidence from one RCT (n=50) showed low-energy density diet and CBT-ED may be less effective on binge frequency compared with general nutritional counselling and CBT-ED, although there was some uncertainty.

8.5.4.19 Group CBT-ED then group behavioural weight loss therapy versus group CBT-ED in adults with BED at end of treatment

Very low quality to low quality evidence from one RCT (n=80) showed no difference in the effect of group CBT-ED followed by group behavioural weight loss therapy on remission, binge frequency, BMI, EDE-global, EDE-dietary restraint, EDE-eating concern, EDE-shape concern, EDE-weight concern and depression compared with group CBT-ED only.

Low quality evidence from one RCT (n=80) showed group CBT-ED followed by group behavioural weight loss therapy may be more effective on weight loss compared with group CBT-ED only, although there was some uncertainty.

8.5.4.20 Group CBT-ED then group behavioural weight loss therapy versus group CBT-ED in adults with BED at follow up

Low quality evidence from one RCT (n=80) showed no difference in the effect of group CBT-ED followed by group behavioural weight loss therapy on binge frequency, BMI, weight loss, EDE-global, EDE-dietary restraint, EDE-eating concern, EDE-shape concern, EDE-weight concern and depression compared with group CBT-ED only.

8.5.4.21 Antidepressant and group behavioural weight control therapy versus placebo and group behavioural weight control therapy in adults with BED

Low quality evidence from one RCT (n=63) showed no difference in the effect of antidepressant and group behavioural weight control therapy on Weight, binge frequency, Binge Eating Scale score, general psychopathology and depression compared with placebo and group behavioural weight control therapy.

8.5.4.22 Antidepressant, CBT-ED and group behavioural weight control therapy versus placebo, CBT-ED and group behavioural weight control therapy in adults with BED

Low quality evidence from one RCT (n=53) showed no difference in the effect of antidepressant, CBT-ED and group behavioural weight control therapy on Weight, binge frequency, Binge Eating Scale score, general psychopathology and depression compared with placebo, CBT-ED and group behavioural weight control therapy.

8.5.4.23 CBT-ED then antidepressant and group behavioural weight loss therapy versus CBT-ED then group behavioural weight loss therapy in adults with BED

Very low quality evidence from one RCT (n=72) showed no difference in the effect of CBT-ED followed by antidepressant and group behavioural weight loss therapy on weight and depression compared with CBT-ED followed by group behavioural weight loss therapy.

8.5.4.24 Antiobesity agent and diet versus placebo and diet in adults with binge eating disorder at end of treatment

Low quality evidence from one RCT (n=73) showed no difference in the effect of antiobesity agent and diet on the number of people still meeting DSM-IV criteria for BED compared with placebo and diet.

Low quality evidence from one RCT (n=89) showed no difference in the effect of antiobesity agent and diet on EDI-total and quality of life compared with placebo and diet.

Low quality evidence from one RCT (n=89) showed antiobesity agent and diet was more effective on general psychopathology compared with placebo and diet.

Very low quality evidence from one RCT (n=89) showed antiobesity agent and diet may be more effective on depression compared with placebo and diet, although there was some uncertainty.

Low quality evidence from one RCT (n=73) showed an antiobesity agent and diet may be more effective on reducing the number of people still meeting DSM-IV criteria for generalized anxiety disorder or major depressive disorder compared with placebo and diet, although there was uncertainty.

8.5.4.25 Antiobesity agent and behavioural weight loss therapy versus placebo and behavioural weight loss therapy in adults with binge eating disorder at end of treatment

Very low quality evidence from one RCT (n=40) showed no difference in the effect of antiobesity agent and behavioural weight loss therapy on remission compared with placebo and behavioural weight loss therapy.

Very low quality evidence from one RCT (n=38) showed no difference in the effect of antiobesity agent and behavioural weight loss therapy on BMI, EDE-global, EDE-dietary restraint, EDE-eating concern, EDE-shape concern, EDE-weight concern and depression on depression compared with placebo and behavioural weight loss therapy.

8.5.4.26 Antiobesity agent and behavioural weight loss therapy versus placebo and behavioural weight loss therapy in adults with binge eating disorder at follow up

Very low quality evidence from one RCT (n=40) showed no difference in the effect of antiobesity agent and behavioural weight loss therapy on remission (ITT) compared with placebo and behavioural weight loss therapy.

Very low quality evidence from one RCT (n=37) showed no difference in the effect of antiobesity agent and behavioural weight loss therapy on BMI, EDE-global, EDE-dietary restraint, EDE-eating concern, EDE-shape concern, and EDE-weight concern compared with placebo and behavioural weight loss therapy.

Very low quality evidence from one RCT (n=37) showed antiobesity agent and behavioural weight loss therapy is more effective on depression compared with placebo and behavioural weight loss therapy.

8.5.5 Economic Evidence statements

No economic evidence on the cost effectiveness of nutritional interventions for people with binge eating disorder was available.

8.5.6 Recommendations and link to evidence for the review on: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

Nutritional counselling

	The committee expressed the view that nutritional counselling is an integral part of most eating disorder specific psychological interventions so they did not make a recommendation about this for people with binge eating disorder.
Critical and important outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of nutritional interventions for treating binge eating disorder in children, young people and adults. For this population, bingeing and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, family functioning and service user experience.
Trade off benefits and	Adults with binge eating disorder Online nutritional counselling is less effective on change in weight compared with

harms	treatment as usual showed, but equally effective on depression, EDE global and general functioning in people with binge eating disorder. No critical outcomes were measured. Similar results were found at follow up. No evidence was found on the critical outcomes of remission and binge eating, nor on the important outcomes of adverse events, quality of life, all-cause mortality, relapse, family functioning, cost effectiveness, resource use, and service user experience. Group nutritional counselling showed a benefit on weight and binge eating compared with wait list controls. No evidence was found on the critical outcome of remission, nor on the important outcomes of adverse events, quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Group behavioural weight loss therapy appeared to reduce body weight (with some uncertainty) and had a positive effect on bingeing compared with wait list controls. No evidence was found on the critical outcome of remission, nor on the important outcomes of adverse events, quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. No difference was found between behavioural weight loss therapy and any intervention on remission, binge frequency, rapid response and weight at the end of treatment. EDE-global appeared to favour any other treatment but there was some uncertainty. There was some conflict in the results since one outcome showed the number of people who achieved a 5% reduction in body weight favoured the behavioural weight loss therapy. At one year follow up, no difference was found in weight or the number who achieved a 5% reduction in body weight. The results for EDE-global favoured any other treatment and binge frequency but there was some uncertainty. No evidence was found on the important outcomes of adverse events, quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Guided self-help behavioural weight loss showed less favourable results on remission compared with any other intervention but there was some uncertainty. Other outcomes showed no difference, including bingeing, weight loss, depression and EDE-subscale results. No evidence was found on the important outcomes of adverse events, quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Group behavioural weight loss therapy had a similar effect the number meeting the DSM criteria for BED, any of the EDE-subscales or depression compared with any other intervention. It appeared to increase weight loss but was less effective on binge frequency. The benefit on weight was not maintained at follow up. With regards to remission, EDE-shape concern and EDE-weight concern, although there was no difference between group behavioural weight loss therapy and any other intervention, there was high heterogeneity. There was no obvious difference in risk of bias in the studies but a subgroup analysis according to severity of illness indicated that it favoured group behavioural weight loss therapy on remission for people who engaged in less than 18 binge episodes per month, although there was some uncertainty. However, there was no difference between group behavioural weight loss therapy and any other intervention for people who engaged in over 18 binge episodes per month. By contrast, the subgroup analysis indicated that whilst there was no difference on EDE-shape concern and EDE-weight concern for people who engaged in less than 18 binge episodes per month, group behavioural weight loss therapy was more effective for those who engaged in over 18 binge episodes per month. No evidence was found on the important outcomes of adverse events, quality of life, all-cause mortality, relapse, general functioning, family functioning, cost effectiveness, resource use, and service user experience. Other comparisons that showed no important differences (that is no critical outcomes or outcomes that would influence decision making) in their effect included behavioural weight loss versus nutritional counselling, behavioural weight loss combined with online motivational interviewing compared with

	treatment as usual, behavioural weight loss combined with online motivational interviewing compared with online nutritional counselling, low energy diet and CBT-ED versus nutritional counselling and CBT-ED, group CBT-ED stepped care versus group CBT-ED, or adding an antidepressant to group behavioural weight loss with or without CBT-ED compared with placebo and therapy. Adverse events or all-cause mortality were not reported in any of the studies.
Trade-off between net health benefits and resource use	The committee expressed the view that dietary advice is an integral part of most eating disorder specific psychological interventions and providing such supplementary advice would not incur significant extra resource implications to the healthcare system.
Quality of the evidence	The quality of the evidence very low quality. It was unclear how randomisation was conducted and if allocation concealment was performed. It was mostly unclear if either the participants, assessors or investigators were blind. High dropouts were also reported >20%. Imprecision was often detected because the 95% confidence interval crossed 1 or 2 minimal important differences or the outcome did meet the optimal information size (300 events or 400 participants). Most of the comparisons had only 1 study and a small number of participants in each arm. Very few studies measured remission. No relevant studies in children or young people were identified. Heterogeneity was not detected.
Other considerations	The committee agreed that the evidence was not strong enough to recommend nutritional counselling or a healthy weight programme as the sole treatment for adults with binge eating disorder. Only one study was available for most comparisons and very few participants were included in the studies. Readmission was rarely reported making it difficult for them to see the clinical and cost-effectiveness of the treatments in addressing the eating disorder per se. Moreover, many of the interventions included in the review, e.g. behavioural weight loss therapy, appeared to target weight loss and not necessarily the eating disorder The committee highlighted that dietary advice and counselling are an integral part of psychological treatments so it is not generally needed if the person is receiving therapy. This is normally delivered by the therapist and at times in collaboration with a dietician. For this reason they did not feel the need to make a research recommendation. No other considerations were made by the committee.

8.6 Physical interventions

8.6.1 Review Question: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 353. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix J.

This review considers all physical interventions that may be delivered to children, young people and adults with an eating disorder. The interventions were categorised according to type of physical intervention, the age of the participants and the type of eating disorder and were compared to wait list controls, placebo, TAU or any other intervention.

Table 327: Protocol summary for the clinical review of: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Component	Description
Review question(s)	Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). • Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder) ○
Intervention(s)	Physical interventions may include: • transcranial magnetic stimulation • deep brain stimulation • physiotherapy • yoga • physical exercise • acupuncture • mandometer • massage
Comparison	Placebo Wait list control Treatment as usual Another intervention
Critical outcomes	• Remission and long-term recovery (include if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs

8.6.2 Clinical Evidence for: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Two RCTs (n=248) met the eligibility criteria for this review, all of which were for adults (McIver 2010, (McIver, 2010), Pendleton 2002 (Pendleton et al., 2002)). The majority of participants in these studies were female. One study (Pendleton 2002) examined a combined physical and psychological intervention. An overview of the trials included in the analysis can be found in Table 328. Further information about both included and excluded studies can be found in Appendix J.

Table 328: Study information for trials included in the analysis of physical interventions versus any other intervention or wait list control for people with binge eating disorder.

Study ID	N Random-ized	Female (%)	Mean BMI, kg/m2 (SD)	Sample	Intervention Age at onset and/or duration of illness (years)	Comparison(s) Age at onset and/or duration of illness (years)	Duration
McIver 2009	90	100	34.1 (6.5)	Adult BMI>25, BES score>20.	Yoga	WLC	12 weeks
Pendleton 2002	114	100	36.2 (6.5)	Adult BED with obesity	Physical Exercise + gCBT-BED + Maintenance	1. Exercise + gCBT-BED 2. Exercise + Maintenance 3. CBT-BED	Non-maintenance groups: 4 months; 6-mo + 12-mo FU Maintenance groups: 10 months; 6-mo FU

Abbreviations: BED, binge eating disorder; BES, Binge Eating Scale; BMI, Body Mass Index; gCBT-BED, Group Cognitive Behavioural Therapy for binge eating disorder; FU, follow up; WLC, wait list control.

Table 329: Summary of findings table for yoga versus wait list control at the end of treatment in adults with binge eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Yoga (95% CI)
BMI	50 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups is 0.3 standard deviations higher (0.26 lower to 0.86 higher)
Binge Eating Scale	50 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge eating scale in the intervention groups was 1.77 standard deviations lower (2.43 to 1.11 lower)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Yoga (95% CI)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 McIver 2009: Allocation concealment unclear. No participant, investigator nor assessor blinding. Dropout rate for both groups>20%. 2 Sample was participants with BMI>25 and Binge Eating Scale score >20. 3 CI crosses either 0.5 or -0.5 (SMD). 4 <400 participants.					

Table 330: Summary table of findings for aerobic exercise and group CBT-ED versus group CBT-ED in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT	Risk difference with Exercise+Group CBT (95% CI)
BMI (changes scores)	37 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (changes scores) in the intervention groups was 0.93 standard deviations lower (1.61 to 0.24 lower)
Depression BDI	37 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.51 standard deviations lower (1.17 lower to 0.15 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Pendleton 2002: randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding. Droprate of Exercise+CBT group and CBT only group both >20%. 2 CI crosses either 0.5 or -0.5 (SMD).					

Table 331: Summary table of findings for aerobic exercise and group CBT-ED versus group CBT-ED in adults with BED at follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT	Risk difference with Exercise+Group CBT (95% CI)
BMI (changes scores) FU	37 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (changes scores) fu in the intervention groups was 0.91 standard deviations lower (1.6 to 0.23 lower)
Depression FU BDI	37 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.26 standard deviations lower (0.91 lower to 0.39 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Pendleton 2002: randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding. Droprate of Exercise+CBT group and CBT only group both >20%.
2 CI crosses either 0.5 or -0.5 (SMD).

Table 332: Summary table of findings for aerobic exercise and group-CBT-ED versus group CBT-ED and maintenance in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT+Maintenance	Risk difference with Exercise+Group CBT (95% CI)
BMI (Change scores)	43 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (change scores) in the intervention groups was 0.28 standard deviations lower (0.88 lower to 0.33 higher)
Depression BDI	43 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.34 standard deviations lower (0.94 lower to 0.27 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT+Maintenance	Risk difference with Exercise+Group CBT (95% CI)
CI: Confidence interval;					
1 Pendleton 2002: randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding. Droprate of Exercise+CBT group >20%. 2 CI crosses either 0.5 or -0.5 (SMD).					

1 **Table 333: Summary table of findings for aerobic exercise and group-CBT-ED versus group CBT-ED and maintenance in adults with**
2 **BED at follow up.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT+Maintenance	Risk difference with Exercise+Group CBT (95% CI)
BMI (Change scores) FU	43 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (change scores) fu in the intervention groups was 0.18 standard deviations lower (0.78 lower to 0.42 higher)
Depression FU BDI	37 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.02 standard deviations lower (0.58 lower to 0.62 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; FU: Follow up					
1 Pendleton 2002: randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding. Droprate of Exercise+CBT group >20%. 2 CI crosses either 0.5 or -0.5 (SMD). 3 CI crosses both 0.5 and -0.5 (SMD).					

Table 334: Summary table of findings for aerobic exercise, group CBT-ED and maintenance versus group CBT-ED and maintenance in adults with BED at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT+Maintenance	Risk difference with Exercise+Group CBT+Maintenance (95% CI)
BMI (change scores)	47 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (change scores) in the intervention groups was 0.53 standard deviations lower (1.11 lower to 0.05 higher)
Depression BDI	47 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.55 standard deviations lower (1.14 lower to 0.03 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Pendleton 2002: randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding.
2 CI crosses either 0.5 or -0.5 (SMD).

Table 335: Summary table of findings for aerobic exercise, group CBT-ED and maintenance versus group CBT-ED and maintenance in adults with BED at follow up.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT+Maintenance	Risk difference with Exercise+Group CBT+Maintenance (95% CI)
BMI (change scores) FU	47 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI (change scores) fu in the intervention groups was 0.57 standard deviations lower (1.15 lower to 0.02 higher)
Depression FU BDI	47 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.42 standard deviations lower (1 lower to 0.16 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group CBT+Maintenance	Risk difference with Exercise+Group CBT+Maintenance (95% CI)
confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; FU: follow up					
1 Pendleton 2002: randomization method and allocation concealment unclear. No participant blinding, unclear investigator and assessor blinding. 2 CI crosses either 0.5 or -0.5 (SMD).					

1

2

3

8.6.3 Economic Evidence

No economic evidence on the cost effectiveness of physical interventions for people with binge eating disorder was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.6.4 Clinical evidence statements

8.6.4.1 Yoga versus wait list control in adults with binge eating disorder at end of treatment

Very low quality evidence from one RCT (n=50) showed yoga is more effective in reducing scores on the Binge Eating Scale compared with wait list controls.

Very low quality evidence form one RCT (n=50) showed no difference in the effect of yoga on BMI compared with wait list controls.

8.6.4.2 Aerobic exercise and group CBT-ED versus group CBT-ED at end of treatment

Very low quality evidence from one RCT (n=37) showed exercise and group CBT-ED is more effective on change in BMI compared with group CBT-ED.

Very low quality evidence from one RCT (n=37) showed no difference in the effect of exercise and group CBT-ED on depression compared with group CBT-ED.

8.6.4.3 Aerobic exercise and group CBT-ED versus group CBT-ED at follow up

Very low quality evidence from one RCT (n=37) showed exercise and group CBT-ED is more effective on change in BMI compared with group CBT-ED.

Very low quality evidence from one RCT (n=37) showed no difference in the effect of exercise and group CBT-ED on depression compared with group CBT-ED.

8.6.4.4 Aerobic exercise and group CBT-ED versus group CBT-ED and maintenance at end of treatment

Very low quality evidence from one RCT (n=43) showed no difference in the effect of exercise, group CBT-ED on change in BMI and depression compared with group CBT-ED and maintenance.

8.6.4.5 Aerobic exercise and group CBT-ED versus group CBT-ED and maintenance at follow up

Very low quality evidence from one RCT (n=43) showed no difference in the effect of exercise, group CBT-ED and maintenance on change in BMI and depression compared with group CBT-ED and maintenance.

8.6.4.6 Aerobic exercise, group CBT-ED and maintenance versus group CBT-ED and maintenance at end of treatment

Very low quality evidence from one RCT (n=47) showed exercise, group CBT-ED and maintenance may be more effective on change in BMI and depression compared with group CBT-ED and maintenance, but there was some uncertainty.

8.6.4.7 Aerobic exercise, group CBT-ED and maintenance versus group CBT-ED and maintenance at follow up

Very low quality evidence from one RCT (n=47) showed exercise, group CBT-ED and maintenance may be more effective on change in BMI compared with group CBT-ED and maintenance, but there was some uncertainty.

Very low quality evidence from one RCT (n=47) showed no difference in the effect of exercise, group CBT-ED and maintenance on depression compared with group CBT-ED and maintenance.

8.6.5 Economic Evidence statements

No economic evidence on the cost effectiveness of physical interventions for people with binge eating disorder was available.

8.6.6 Recommendations and link to evidence for the review on: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders

Physical therapy

	135. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitization, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes for the review on the effectiveness of physical interventions, such as transcranial magnetic stimulation or physiotherapy in people with eating disorders and it was agreed that for any eating disorder remission is of greatest concern. The other critical outcomes for anorexia nervosa are body weight and BMI and for binge eating disorder and bulimia nervosa it is bingeing. Other outcomes that are important but are considered rare events or rarely measured in randomised controlled trials for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse, thus they were extracted where possible, but did not factor strongly in the decision making. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	**Adults with binge eating disorder (chapter 8)** Yoga appears to be effective at reducing scores on the binge eating scale compared with wait list controls. However, this did not translate to a benefit in BMI. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. Aerobic exercise and group CBT-ED appeared to be more effective at reducing BMI compared with group CBT-ED alone in adults with binge eating disorder. No difference was found in depression scores. Similar results were found at follow up. When a maintenance component (12 biweekly meetings over 6 months) was added to both arms to make this part of the intervention more comparable with the aerobic exercise group (because they continued to meet up), there was a trend for a reduced BMI and depression in the aerobic exercise, group CBT-ED and maintenance group compared with the group CBT-ED and maintenance group at the end of treatment and for the trend in the benefit on BMI to be maintained at follow up but not depression. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause

mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Young people with anorexia nervosa (chapter 6)

For young people with anorexia nervosa, bright light treatment and CBT showed benefits on depression compared with any other intervention. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Video feedback and nutritional counselling compared with nutritional counselling alone showed no additional benefit of the video feedback on BMI. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Resistance training and treatment as usual showed no difference on BMI and quality of life in young people with anorexia nervosa compared with treatment as usual. At 4 weeks follow up, resistance training and treatment as usual appeared to be less effective on BMI compared with treatment as usual. No evidence was found on the critical outcomes of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, family functioning, resource use, and service user experience.

Adults with anorexia nervosa (chapter 6)

Repetitive transcranial magnetic stimulation versus sham showed no difference in anorexia nervosa symptoms (urge to restrict, feeling full and feeling fat), urge to binge or side-effects from treatment. However, at one day follow up some benefits were detected on anorexia nervosa symptoms including feeling full and feeling flat compared with sham, but no difference in the symptom of urge to restrict. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Warming therapy on top of refeeding had no effect on change in BMI weight compared with refeeding alone in adults with anorexia nervosa. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Acupuncture and treatment as usual compared with acupressure, massage and treatment as usual showed acupuncture is more effective on EDE-shape concerns but no other outcome was different between the two groups including EDI-subscales, EDE-subscales, depression, general psychopathology and BMI. No evidence was found on the critical outcome of remission, nor on the important outcomes of all-cause mortality, relapse, general functioning, family functioning, resource use, and service user experience.

Adults with bulimia nervosa (chapter 7)

Repetitive transcranial magnetic stimulation versus sham showed no difference in the effect on bingeing and food cravings within 24 hours of treatment, nor on the urge to eat or the number who withdrew due to adverse events. There was a trend for hunger and the number of those who binged to be reduced but there was some uncertainty. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Aerobic exercise appeared to be less effective on EDI-drive for thinness. No difference was found on the number of people who recovered from bulimia nervosa nor who satisfied the EDNOS criteria.

Compared with wait list control, aerobic exercise was less effective on the number who had recovered (unclear definition) from bulimia nervosa but showed no

	difference on the number who satisfied the criteria for EDNOS. No evidence was found on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. **Any eating disorder (chapter 9)** One study compared eye movement desensitization and reprocessing therapy with treatment as usual in adults with any eating disorder. The results showed some improvement in the outcomes reported by the body image memory questionnaire, including the earliest memory and worst memory on body image and only a trend for the most recent memory. At 12 months follow up the worst memory on body image was still better but not the earliest or most recent. No evidence was found on the critical outcomes of remission, bingeing and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. An RCT was identified that compared yoga and treatment as usual with treatment as usual in adults with any eating disorder. At the end the treatment, no difference was found in any of the outcomes including BMI, EDE-total or any of the EDE- sub-scales. Similar findings were found at follow up (three weeks), however there was some improvement in EDE-restraint in the yoga and treatment as usual group compared with treatment as usual. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, resource use, and service user experience. A graded body image therapy (and maintenance treatment as usual) was compared with a maintenance treatment as usual in adults with any eating disorder. No difference was found in EDE-weight concerns or EDE-shape concerns at the end of treatment or at follow up. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. An acceptance-based body image mirror exposure therapy was compared with a control therapy and showed an improvement in EDE-eating concerns, EDE-weight concerns, EDE-shape concerns, but not in EDE-restraint. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. A psychomotor therapy and support was compared with support in females with any eating disorder and showed no difference at the end of treatment on self-expression and control anger scales. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, eating disorder psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. The committee requested investigating the benefits of the Mandometer on eating disorders. A Mandometer is a device that measures how much weight is lost from a dinner plate after the person with eating disorder has finished eating. This weight is stored on a computer along with how satiated the person is after eating. The evidence on this is scarce and the sample sizes were too small (less than 10 per group) to meet our inclusion criteria as described in the protocol.
Trade-off between net health benefits and resource use	There was no evidence for the effectiveness of physical interventions in people with eating disorders. As a result, such interventions are likely to be not cost effective.
Quality of evidence	The evidence for physical interventions was mostly very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as unclear methods of randomisation, it was unclear if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. Most of the outcomes were the result of a single study with a very low number of

	participants, only binge eating disorder had more than 100 participants in total. Imprecision was detected in most outcomes because the 95% confidence interval crossed one or two minimal important differences or it did not meet the optimal information size. Also, few studies measured remission and/or compensatory behaviours relevant to that eating disorder. Some outcomes were excluded from the study because it was either unclear over what duration they measured the symptoms or it was less than the two week minimum required by the committee.
Other considerations	The committee agreed that the evidence presented was not strong enough or of sufficient quality to offer a physical intervention to people with an eating disorder. This was mostly because very few studies were identified and few participants were included in most outcomes. However, the committee decided to make a research recommendation on adding exercise to a recommended psychotherapy to determine whether it may add any benefit to those with bulimia nervosa or binge eating disorder. The committee discussed the importance of exploring what the right amount of exercise is, what is the best type of exercise and what the potential harms are. The committee suggested making a research recommendation on the effects of exercise on bulimia nervosa and binge eating disorder, as opposed to any of the other physical interventions for a number of reasons. Exercise may be useful adjunct to psychotherapy to address any co-existing weight or obesity-related issues and mood disorders, such as depression and anxiety. Exercise may also be a cost-effective and drug-free alternative to other therapeutic approaches such as transcranial magnetic stimulation or anti-depressants. See chapter 6.6

8.7 Management of long- and short-term complications

8.7.1 Review question: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all interventions that may be delivered to manage or reduce the short- or long-term physical complications of eating disorders in children, young people and adults and includes recovered as well as current service users. The interventions were categorised according to type of physical complication and intervention, the age of the participants and the type of eating disorder and were compared to the control arm as reported in the relevant studies.

Table 336: Clinical review protocol summary for the review of: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

Component	Description
Review question(s)	What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) • Recovered or current service users Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder) ○
Intervention(s)	Interventions to address the following:

Component	Description
	• Low bone mineral density (risk of fracture) • Growth (physical development) • Pubertal development • Tooth wear • Low body weight • Interventions to address the long-term physical complications may include: • GH/IGF-I • Calcium with and without Vitamin D • Bisphosphonates (age dependent and exclude pregnancy) • Exercise (low impact)/Physiotherapy • Oestrogen (patches/exogenous/pills other) • Testosterone (males/females) • Weight gain vs. Weight restoration (brain size) • Interventions to address the short-term physical complications may include • Phosphates supplementation (refeeding) • Potassium • Thiamine (refeeding) • Laxatives (for when underweight patients are constipated) • Salbutamol (reduce food intake)
Comparison	Control arm as defined by study
Critical outcomes	Primary outcome as reported by study
Important outcomes	Secondary outcome as reported by study
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs) (if no RCTs)

8.7.2 Clinical Evidence for: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

No studies were found that met the eligibility criteria for this review.

8.7.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for the management of short and long-term physical complications of binge eating disorder was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.7.4 Clinical evidence statements

No studies were found that met the eligibility criteria for this review.

8.7.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the management of short and long-term physical complications of binge eating disorder was available.

8.7.6 Recommendations and link to evidence for the review on: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

	The committee agreed no recommendation was needed for this review question on those with binge eating disorder.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of nutritional interventions for treating binge eating disorder in children, young people and adults. For this population, bingeing and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with anorexia nervosa that are of lesser importance but clearly important outcomes include, general psychopathology, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	No relevant, published, RCT or observational evidence was identified.
Trade-off between net health benefits and resource use	No relevant existing economic evidence was identified.
Quality of evidence	Not applicable
Other considerations	The committee agreed that no recommendation was needed on how to treat or manage people with binge eating disorders who have short or long-term complications.

8.8 Management of comorbidities

8.8.1 Review question: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

The review protocol summary, including the review question and the eligibility criteria used in this section of the guideline, can be found in Table 364. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers whether any intervention used to treat eating disorders in children, young people and adults needs to be modified in the presence of a common long-term health condition (i.e. comorbidity).The interventions were categorised according to their type, the type of eating disorder and comorbidity examined and the age of the participants. The comparison arm was the same intervention delivered to participants with the relevant eating disorder but without the relevant comorbidity.

Table 337: Clinical review protocol summary for the review of: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Component	Description
Review question(s)	Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Component	Description
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) and a common comorbidity (e.g. diabetes, hypothyroidism). • Mental comorbidities may include: ○ Depression ○ Anxiety ○ Social anxiety ○ Autism ○ Obsessive Compulsive Disorder ○ Personality Disorder ○ Learning disability ○ ADHD (Bulimia) ○ Self-harm ○ Substance misuse ○ Physical comorbidities may include: ○ Celiac disease ○ Diabetes (type II – relevant to obesity) ○ Irritable Bowel Disease ○ Cystic Fibrosis Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder) ○ .
Intervention(s)	Trials will be included that address the ED as primary or secondary aim to treating the comorbidity. Interventions may include: • Psychotherapy (including psychoeducation) • Pharmacological • Nutritional • Physical • Combination of any listed above
Comparison	The same intervention but delivered to people with an eating disorder without a comorbidity.
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)

Component	Description
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs) (if no RCTs)

8.8.2 Clinical Evidence for: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

8.8.2.1 Major depressive disorder and a diabetes prevention programme

One observational study was identified (n=131) that compared the effectiveness of a diabetes prevention programme in a population who had a binge eating disorder and a major depressive disorder (Pagoto 2007 (Pagoto et al., 2007)). An overview of this trial can be found in Table 338.

8.8.2.2 Diabetes

One RCT (n=34) was found that fulfilled the criteria for this review (Kenardy 2002 (Kenardy et al., 2002)). This study compared group CBT-ED with a control, non-prescriptive group therapy in adults with type II diabetes and binge eating disorder. This comparison provided insight into whether CBT-ED is effective at achieving remission in this population.

Summary of findings for how to treat a binge eating disorder in the presence of a comorbidity can be found in Table 340 and Table 341. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 338: Study information for observational study included in the review of interventions for people with binge eating disorder and a major depressive disorder.

Study ID	N	Mean age	BMI	Female	Group 1	Group 2	Intervention	Duration
Pagoto 2007	131	not reported	not reported	not reported	BED with comorbid major depressive disorder.	BED	Diabetes Prevention Program lifestyle intervention	16 sessions

Table 339: Study information for the RCT included in the review of interventions for people with binge eating disorder and type II diabetes.

Study	Eating Disorder	Age mean (SD)	BMI	N	Stage of illness	Intervention	Comparison	Duration
Kenardy 2002 Australia	Diabetes type II diabetes + binge eating disorder	51.8 (9.6) years	not reported	34	Time since diagnosis of diabetes (months)=39.1 (48.5)	Group CBT-ED	Control group Non-Prescriptive Therapy	10 weeks, 12 week FU

Table 340: Summary of findings table for healthy weight program for people with binge eating disorder and major depressive disorder versus those with binge eating disorder alone.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with BED alone group	Risk difference with BED and depression (95% CI)
Achieved Weight Loss Goal>=7%	39 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.03 (0.27 to 4)	176 per 1000	5 more per 1000 (from 129 fewer to 529 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Pagoto 2007: retrospective chart review, no control intervention and unclear length of treatment, high selection bias.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with BED alone group	Risk difference with BED and depression (95% CI)
2 CI crosses both 0.75 and 1.25.					

Table 341: Summary of findings table for group CBT-ED versus non-prescriptive control group therapy for people with binge eating disorder and type II diabetes.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with control	Risk difference with Group CBT-ED (95% CI)
Remission - Group CBT-ED v Group NPT	34 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.6 (0.66 to 3.91)	294 per 1000	176 more per 1000 (from 100 fewer to 856 more)
BMI - Group CBT-ED v Group NPT	34 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI - group cbt-ed v group not in the intervention groups was 0.63 standard deviations higher (0.06 lower to 1.32 higher)
Binge Frequency - Group CBT-ED v Group NPT	34 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency - group cbt-ed v group not in the intervention groups was 0.32 standard deviations lower (1 lower to 0.36 higher)
EDI Bulimia - Group CBT-ED v Group NPT	34 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi bulimia - group cbt-ed v group not in the intervention groups was 0.03 standard deviations lower (0.71 lower to 0.64 higher)
EDI Drive for Thinness - Group CBT-ED v Group NPT	34 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,6 due to risk of bias,		Not calculable for SMD values	The mean edi drive for thinness - group cbt-ed v group not in the intervention groups was

		indirectness, imprecision			0.17 standard deviations lower (0.84 lower to 0.5 higher)
EDI Body Dissatisfaction - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction - group cbt-ed v group not in the intervention groups was 0.06 standard deviations higher (0.61 lower to 0.73 higher)
Quality of Life - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean quality of life - group cbt-ed v group not in the intervention groups was 0 standard deviations higher (0.67 lower to 0.67 higher)
Remission FU - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,7 due to risk of bias, indirectness, imprecision	RR 3.33 (1.11 to 10.03)	176 per 1000	411 more per 1000 (from 19 more to 1000 more)
BMI FU - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean BMI fu - group cbt-ed v group not in the intervention groups was 0.64 standard deviations higher (0.06 lower to 1.33 higher)
Binge Frequency FU - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean binge frequency fu - group cbt-ed v group not in the intervention groups was 0.52 standard deviations lower (1.2 lower to 0.17 higher)
EDI Bulimia FU - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi bulimia fu - group cbt-ed v group not in the intervention groups was 0.03 standard deviations lower (0.7 lower to 0.65 higher)
EDI Drive for Thinness FU - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi drive for thinness fu - group cbt-ed v group not in the intervention groups was 0.16 standard deviations higher (0.52 lower to 0.83 higher)

EDI Body Dissatisfaction FU - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction fu - group cbt-ed v group not in the intervention groups was 0.04 standard deviations higher (0.63 lower to 0.71 higher)
Quality of Life FU - Group CBT-ED v Group NPT	34 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,6 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean quality of life fu - group cbt-ed v group not in the intervention groups was 0.17 standard deviations lower (0.84 lower to 0.51 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; NPT: Non-Prescriptive Therapy

1 Inadequate randomisation was performed and it was unclear if allocation concealment was carried out. Neither the participant or investigator was blind, nor was it clear if the assessor was blind. It was unclear how many participants completed the intervention.
2 Population included disturbed eating attitudes and behaviour based on EDI scale results.
3 95% CI crossed 2 MIDs (0.75 and 1.25)
4 95% CI crossed 1 MID (0.5)
5 95% CI crossed 1 MID (-0.5)
6 95% CI crossed 2 MIDs (-0.5 and 0.5)
7 95% CI crossed 1 MID (0.75)

1

2

3

8.8.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for binge eating disorder in the presence of common long-term conditions was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.8.4 Clinical evidence statements

8.8.4.1 Major depressive disorder

8.8.4.1.1 Diabetes prevention programme in people with a BED and major depressive disorder versus people with BED alone at the end of treatment.

Very low quality evidence from one RCT (n=39) showed no difference in the effect of a diabetes prevention program in people with binge eating disorder and comorbid major depressive disorder on the number of people achieving 7% or greater weight loss compared with people with binge eating disorder alone.

8.8.4.2 Diabetes

8.8.4.2.1 Group CBT-ED versus non-prescriptive control group therapy in people with type II diabetes and BED at the end of treatment

Very low quality evidence from one RCT (n=34) showed no difference in the effect of group CBT-ED on binge frequency, EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction, quality of life and remission compared with a control group therapy.

Very low quality evidence from one RCT (n=34) showed group CBT-ED may be less effective on decreasing BMI compared with compared with a control group therapy, but there was some uncertainty.

8.8.4.2.2 Group CBT-ED versus non-prescriptive control group therapy in people with type II diabetes and BED at the end of treatment

Very low quality evidence from one RCT (n=34) showed no difference in the effect of group CBT-ED on binge frequency, EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction and quality of life compared with a control group therapy.

Very low quality evidence from one RCT (n=34) showed group CBT-ED may be less effective on decreasing BMI compared with compared with a control group therapy, but there was some uncertainty.

Very low quality evidence from one RCT (n=34) showed group CBT-ED is more effective on remission compared with compared with a control group therapy.

8.8.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for binge eating disorder in the presence of common long-term conditions was available.

8.8.6 Recommendations and link to evidence for the review on: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Diabetes

136. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.

137. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Diabetes

138. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.

139. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.

140. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes

141. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.

142. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.

143. Offer people with an eating disorder who are misusing insulin the following treatment plan:

- a low carbohydrate diet, so that insulin can be started at a low level
- gradually increasing insulin doses to reduce blood glucose levels
- adjusted total glycaemic load and carbohydrate distribution to meet their individual needs and prevent rapid weight gain
- carbohydrate counting when adjusting their insulin dose (including via pumps)
- a diabetic educational intervention such as DAFNE
- education about the problems caused by misuse of diabetes medication.

144. For more guidance on managing diabetes, refer to the NICE

	guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem. In the case of diabetes, HbA1c levels and insulin omission days were considered critical outcomes. The other critical outcomes depended on the eating disorder included in the study. Remission is of greatest concern for any eating disorder. In addition, for those with anorexia nervosa body weight or BMI are of greatest concern. For bulimia nervosa and binge eating disorder, binge eating is a critical outcome. For any eating disorder, other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with an eating disorder that are of lesser importance but are clearly still important outcomes include general psychopathology, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	The ideal study design to answer the question of whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem would be to randomise people with an eating disorder and diabetes to two different treatment groups: one modified to address both the eating disorder and diabetes and one non-modified eating disorder treatment. **Binge eating disorder (as reviewed in chapter 8)** One study randomised adults with type II diabetes and binge eating to either group CBT-ED or a non-prescriptive control therapy (NPT). The results showed no difference in remission or binge frequency at the end of treatment. BMI showed a trend to be higher in the group CBT-ED arm, however EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction and quality of life were no different. At follow up, remission rates were higher in the CBT-ED arm, but again no difference in any of the other outcomes and BMI showed a trend to be higher in the group CBT-ED arm compared with controls. No data was available on HbA1c levels, insulin omission, all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience. An observational study compared the same diabetes prevention programme but in two populations, one with bulimia nervosa and a major depressive disorder and one with just any eating disorder. The results showed no difference in the degree of weight loss between the two populations. No data was available on HbA1c levels, insulin omission, remission, bingeing, all-cause mortality, adverse events, resource use, relapse, general functioning, eating disorder psychopathology, family functioning and service user experience. **Anorexia nervosa (as reviewed in chapter 6)** No published evidence was found on people with anorexia nervosa and diabetes, however there was a sub-group analysis from a study described below on any eating disorder that showed those with anorexia nervosa and type I diabetes are equally responsive to treatment as those with anorexia nervosa alone. No data was available on HbA1c levels, remission, all-cause mortality, adverse events, quality of life, resource use, relapse, general functioning, family functioning and service user experience. **Bulimia nervosa (as reviewed in chapter 7)** One observational study compared the effectiveness of inpatient integrated care with treatment as usual in adults with bulimia nervosa and type I diabetes. The integrated care provided CBT-ED, family based therapy and addressed control of diabetes. Whilst treatment as usual included outpatient counselling sessions on diabetes but not inpatient care or treatment for the eating disorder. This study showed better outcomes for the integrated care including, remission, general psychopathology, depression, EDI-total, the size of the binges, few compensatory behaviours but no difference in insulin omission. No data was available on HbA1c levels, all-cause mortality, adverse events, quality of life, resource use, relapse,

	general functioning, family functioning and service user experience. **Any eating disorder (as reviewed in chapter 9)** One randomised control trial compared group psychoeducation (combined with treatment as usual) with treatment as usual (diabetes treatment only) in people with type I diabetes and disturbed eating behaviours and showed no difference at the end of treatment on bingeing, EDE-restraint, EDE-shape concern, EDE-eating concern, EDE-weight concern, EDI-drive for thinness, EDI-bulimia, insulin omission days and HbA1c (%). One outcome, EDI-body dissatisfaction, favoured group psychoeducation over treatment as usual but there was some uncertainty. At follow up some benefit was found in response to group psychoeducation on bingeing but there was some uncertainty. No data was available on remission, all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience. An observational study was identified that compared the same CBT-ED intervention but in two populations, one with any eating disorder and type I diabetes, and one with just any eating disorder. Thus, this design allowed us to see whether those with a comorbidity would respond equally well to treatment as those with just an eating disorder. The results showed adults with any eating disorder and a comorbidity are less likely to recover than those with just an eating disorder. No difference was found in dropouts. No data was available on all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience.
Trade-off between net health benefits and resource use	The committee considered that providing care for eating disorders in the presence of a long-term health problems, such as diabetes, may have resource implications in terms of extra time required to provide collaborative and comprehensive care in line with the principles outlined in the recommendations 127-133. However, the committee expressed the view that if such care arrangements (that is, multidisciplinary approach, involvement of family members and carers, and the use of treatment plans) lead to better and appropriate treatment and management of health problems (including other long-term health problems such as diabetes) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating such care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence was mostly low quality from the RCT studies and very low quality from all of the observational studies. In both types of study designs the sample size was generally small and only one study was available for most outcomes, thus imprecision was often detected due to the 95% confidence interval crossing a minimal important difference or the outcome did not meet the optimal information size. **Binge eating disorder** In the RCT where they compared group CBT-ED with a control therapy in the same population (people with type I diabetes and binge eating disorder) inadequate randomisation was performed and it was unclear if allocation concealment was carried out. Neither the participant or investigator was blind, nor was it clear if the assessor was blind. It was unclear how many participants completed the intervention. The observational study identified was considered indirect evidence since it was a diabetes prevention programme and the participants had major depressive disorder in addition to binge eating disorder or binge eating disorder alone. The only outcome reported was weight loss. The committee did not consider this study helpful. **Bulimia nervosa** In the observational study where they compared inpatient integrated care with treatment as usual, the people were selected from the same recruitment site and showed no difference in their characteristics, except that binge frequency was significantly higher in the inpatient group. The duration of follow up was different for the two groups: 36 months versus 24 months in the inpatient care and treatment as usual groups, respectively. Investigators were not blind to treatment allocation and

	only 18 participants were included. **Any eating disorder (including subgroup analysis on anorexia nervosa)** In the RCT where they compared group psychoeducation and management (and treatment as usual) with treatment as usual (diabetes only programme), it was unclear if allocation concealment was performed. Neither the participant, investigator or assessor were blind and it was unclear how many completed the intervention. The population was also indirect since it included those with disturbed eating. Also, the comparison did not show whether a modified eating disorder programme is more effective at treating people with diabetes and an eating disorder compared with an eating disorder programme alone. Rather the study compared a modified diabetes programme with a regular diabetes programme. In the observational study where they compared CBT-ED in people with eating disorder alone or with a comorbidity, the authors attempted to match the groups based on age, marital status, education, catchment area and onset of diagnosis. However, it was unclear whether the two groups were followed up for the same duration and the sample size was very small. **Overall discussion** No RCT or observational study met the criteria of what would have been the ideal study design for this review (as described above). One RCT compared the effectiveness of an intervention that addressed both the eating disorder and diabetes, but the other arm addressed just the diabetes. In another RCT, one intervention was modified but it was compared with a control therapy. In the observational studies, one study compared the same intervention but in those with either an eating disorder and diabetes or just the eating disorder alone. So it only provided insight into whether one group was more responsive to treatment than the other. In the other observational study, inpatient integrated care was compared with treatment as usual, but the treatment as usual only addressed the diabetes not the eating disorder. Thus, it did not provide insight into whether a modified eating disorder treatment was needed for those with a comorbidity.
Other considerations	In summary, it was difficult for the committee to draw conclusions from these studies on whether treatment for an eating disorder needs to be modified in the presence of a comorbidity such as diabetes. The committee therefore agreed that it was best to instead provide guidance on how to manage the diabetes. Usually, the committee would refer to the diabetes NICE guideline, but because the diabetes guideline refers to this guideline, the committee needed to recommend what to do in the presence of both. The committee agreed on a series of recommendations based on their experience and knowledge on how to manage the diabetes in the presence of an eating disorder. A number of the recommendations are based on what would be considered good practice. For instance: i) establish who will monitor the physical health, ii) explain to the person that they need to monitor their diabetes during the treatment for the eating disorder, and iii) be aware of the problems caused by misuse of diabetes medication. The committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in the treatment of diabetes. They highlighted that the quality of the family environment has been shown to affect treatment compliance and metabolic control among young people with an eating disorder (Hauser 1990). Family members may also need to care for someone if they hyper or phyo (which is a case for medical emergency), so they know how to respond. There is also the possibility that eating disturbances in young girls with diabetes are associated with significantly more family dysfunction than girls with diabetes alone (i.e. 13 to 18 years of age). Specifically, they can receive less support, and have poorer communication and less trust in their relationship with their parents than diabetic girls without eating disturbances. For these reasons, the committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in their treatment. There was some indirect evidence to support the recommendation to address insulin misuse as part of any psychological treatment. One RCT (n=85) showed that a modified group psychoeducation and management programme reduced bingeing episodes at follow up compared with a programme that just addressed

	the diabetes alone. This study was considered with the reservation that it was indirect because: 1) it did not investigate the effectiveness of a modified eating disorder psychological treatment and 2) the population had a disturbed eating behaviour, not a specific eating disorder diagnosis. Nevertheless, it showed that a psychoeducation and management programme may help reduce eating disorder psychopathology in those who also have diabetes. The committee discussed the problem of a relatively high prevalence of EDNOS In young girls with diabetes. In girls who have body dissatisfaction, diabetes provides a unique but dangerous opportunity to control weight by deliberate insulin omission, which can lead to hyperglycaemia and glycosuria. It is therefore important that insulin misuse is addressed in any psychological intervention. It can be noted that the recommendations relating to diet control were contributed to by the expert opinion of a dietician on the committee, based on their experience of treating those with an eating disorder who misuse insulin. These recommendations are based upon the treatment approach of small, attainable and incremental goals. At the outset of treatment, intensive glucose management is not an appropriate goal. The first goal must be to establish medical safety for the person with diabetes by gradually increasing the doses of insulin and food intake (as described in the recommendation). Given the fear of weight gain in this population, the committee recommended that the diet is amended to prevent rapid weight gain. They also suggested an educational programme called Dose Adjustment for Normal Eating (DAFNE) that provides people with the skills necessary to estimate the carbohydrate in each meal and to inject the right dose of insulin. There was no evidence on how to treat the eating disorder in the presence of any other long-term physical health condition, such as cystic fibrosis, celiac disease, pregnancy or irritable bowel disease. Some eating disorder specialists on the committee highlighted that they would generally refer someone with an eating disorder and diabetes to a diabetologist rather than address the points raised in the recommendations on diabetes. However, the committee agreed that it should be collaborative approach for the health care professionals who treat eating disorders and diabetes. Especially for young people who may need to involve family members and carers in therapy sessions to help the person with blood glucose control. Given the lack of direct evidence to address this review question the committee agreed to make a research recommendation to ask: “Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?” See chapter 6.8

Comorbid mental health problems

	145. When deciding which order to treat an eating disorder and a comorbid mental health condition (in parallel, as part of the treatment or one after the other), take the following into account: • the severity and complexity of the eating disorder and comorbidity • the person's level of functioning • the patient's preference **146. Refer to the NICE guidelines on specific mental health problems for further guidance on treatment.**
Relative value of different	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of treating people with an eating disorder and a comorbidity. For binge eating disorder and bulimia nervosa, it was agreed binge

outcomes	eating frequency and remission are of greatest concern. For anorexia nervosa, body weight/BMI and remission are critical and for OSFED, remission and either binge eating or body weight/BMI depending on the eating disorder they most closely resemble. The other outcomes that are critical are the primary outcomes that are relevant to the physical or mental health comorbidity being treated. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	**Binge eating disorder** Very little evidence was identified on the effectiveness of treating an eating disorder in the presence of a mental health comorbidity. One observational study showed adults with binge eating disorder and major depression are equally responsive to a weight loss program (based on diabetes prevention) as those with just binge eating disorder. The only relevant outcome reported was the number who achieved weight loss greater than or equal to 7%. No other outcomes were reported. No relevant published evidence was identified in people with anorexia nervosa, bulimia nervosa or ENDOS.
Trade-off between net health benefits and resource use	The committee considered that providing care for people with eating disorders who have comorbid mental health problems may have resource implications in terms of the extra time required to provide care in line with the principles outlined in the recommendations 134-135. However, the committee expressed the view that if such care leads to better identification of health needs and this results in appropriate subsequent treatment and management of underlying health problems at an earlier stage (including eating disorder and comorbid mental health problem), before individuals require more resource intensive management, then the additional costs associated with facilitating such care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence used to generate these recommendations was very low quality. The evidence was observational therefore in GRADE (the software used to assess quality) the evidence starts at very low quality and can only be upgraded if large effect sizes are found or if a dose response is identified. Neither was the case for this review. The evidence reported only one relevant outcome that is, those who achieved weight loss greater than or equal to 7%. Remission was not reported, nor was data available at the end of treatment only at 3.5 years post treatment. The recommended therapy was a behavioural weight loss programme and it was not found to be the most effective in achieving remission in our reviews. Nevertheless, the findings suggest that people with binge eating disorder and severe depression are equally able to respond to a behavioural weight loss programme as those with just binge eating disorder.
Other considerations	The committee generated a recommendation (based on the limited data found and from their own experience) that when treating a mental health comorbidity that it could be treated either: in parallel with the eating disorder, as part of the eating disorder treatment or focus on one at a time. The order selected will depend on the severity and complexity of the comorbidity and the eating disorder, the general function of the person, preference of the family or carer (if appropriate) and the person with an eating disorder. The committee discussed the importance of taking into account the severity of the comorbidity and the eating disorder, when deciding how to sequence the treatment. Whilst some problems can be treated in parallel, there are instance where this is not always the case. For example, someone would not be offered CBT-ED for an eating disorder and obsessive compulsive disorder (OCD) at the same time but you might address the OCD as part of the eating disorder treatment

	or refer them on for subsequent treatment. Committee members also discussed another example that if substance misuse is not interfering with ability to engage in psychological therapy then the substance misuse may reduce as part of the treatment of the eating disorder. Thus in such cases, treatment of the eating disorder may be sufficient. The committee also highlighted that if treatment of a comorbidity involves different services or agencies, a multidisciplinary approach should be used because the person with an eating disorder may risk being caught between two disciplines with a different focus. Studies were excluded from this review is they randomised people with an eating disorder and mental health comorbidity to different treatments since it did not answer the question of whether the eating disorder treatment should be modified in the presence of a comorbidity or if those with a comorbidity are equally able to respond to the same treatment as those who do not have a comorbidity. The committee agreed it was important eating disorder treatment is not denied because of another mental health comorbidity such as personality disorder.

9 Treatment and management of atypical eating disorders (eating disorders not otherwise specified)

9.1 Introduction

Many people with an eating disorder do not meet the diagnostic criteria for anorexia nervosa, bulimia nervosa or binge eating disorder. There is no consensus over how to refer to these states, so they are often described as 'atypical' eating disorders – even though in some settings they are more common than the 'typical' ones. Confusing matters further, the terminology used by the DSM to refer to these conditions has changed from eating disorder not otherwise specified (EDNOS) to other specified feeding or eating disorder (OSFED). The new DSM-5 diagnosis 'avoidant/restrictive food intake disorder' (ARFID) is not classed as an OSFED (atypical eating disorder).a

In practice, these atypical states fall into two groups (Fairburn & Bohn, 2007 (Fairburn et al., 2007)). There are eating disorders that closely resemble anorexia nervosa, bulimia nervosa or binge eating disorder, but do not quite meet their diagnostic criteria. There are also 'mixed states', in which the features of anorexia nervosa, bulimia nervosa or binge eating disorder are combined in an idiosyncratic way.

A common misconception is that the atypical eating disorders are milder or less severe than the typical eating disorders. This is not the case. They are associated with the same level of distress and impairment and they are just as self-perpetuating (Fairburn 2005 (Fairburn and Bohn, 2005)). Almost all share the same over-concern about eating, shape and weight as seen in the typical eating disorders and the same tendency to engage in persistent and extreme dieting along with other forms of disordered eating (such as binge eating and purging). Body weight also tends to be low if the dietary restriction is marked.

Most people with an atypical eating disorder are female and in their 20s. Many have a history of anorexia nervosa, bulimia nervosa or both, reflecting the diagnostic migration that is common among the eating disorders (Milos, 2005 (Milos et al., 2005)). Their prevalence and incidence in the general population is uncertain, because of the difficulty in defining them and because they are ignored by some assessment instruments (Smink 2012 (Smink et al., 2012)). It seems that they are more common that the typical eating disorders.

9.1.1 Review Question: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 342. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix J.

This review considers all psychological interventions that may be delivered to children, young people and adults with an eating disorder with or without a pharmacological intervention. The interventions were categorised according to their mode of delivery, i.e. individual, group or self-help, the age of the participants and the type of eating disorder. In addition, the

a Throughout this guideline EDNOS, OFSED and atypical eating disorders will be used interchangeably.

interventions were grouped according to their type of therapy and were compared to any other intervention or to wait list controls.

Table 342: Clinical review protocol summary for the review of Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?

Component	Description
Review question(s)	Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder. • Strata: ○ children (<12), young people (13-17 years), adults ≥18 years ○ eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder and atypical eating disorder) ○ mode of delivery (i. individual ii. family iii. group iv. self-help) ○
Intervention(s)	Psychological intervention including: • Dialectical behaviour therapy (DBT) • Counselling (Nutritional/Other) • Integrative Cognitive-Affective Therapy for Binge Eating (ICAT) • Maudsley model for treatment of adults with anorexia nervosa (MANTRA) • Cognitive remediation therapy (CRT) • Specialist supportive clinical management for anorexia nervosa (SSCM) • Behavioural therapy (BT) • CBT (General or ED specific) • Dynamic (IPT, Psychodynamic General or ED specific) • Guided Self Help w therapist guidance • Pure self help • E-therapies Psychological intervention in combination with any pharmacological intervention.
Comparison	• Wait list controls • Treatment as usual • Another other intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Frequency of binge eating for bulimia nervosa and binge eating disorder; and weight/body mass index (appropriate adjustment for age) for anorexia nervosa
Important outcomes	• Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF). • Family functioning. • Service user experience

Component	Description
	• Quality of life. • All-cause mortality. • Relapse. • Adverse events • Resource use.
Study design	• Systematic reviews • RCTs

9.1.2 Clinical Evidence for: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?

One RCT (n=35) on individual therapy versus group therapy was identified in people with EDNOS (Nevonen 2005 (Nevonen and Broberg, 2005)). Three RCTs (n=396) were found that investigated the effects of self-help (guided and internet) compared with wait list controls in people with any eating disorder (Gulec 2014 (Gulec et al., 2014), Hötzel 2014 (Hotzel et al., 2014), Traviss 2011 (Traviss et al., 2011). Although the latter were on people with any eating disorder it was agreed they are best presented in the EDNOS chapter.

See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Study ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention Category	Comparison Arm Category	Sessions N	Treatment Length	Long-term FU
Eating Disorder Not Otherwise Specified EDNOS: Individual therapy										
Nevonen 2005	20.5 (18-24)	21.4	100%	4.5 years after onset of illness	35	Individual hybrid (CBT-ED-IPT)	Group hybrid (CBT-ED-IPT)	23	23 weeks	1 year and 2.5 years
Any eating disorder: Self-help										
Gulec 2014	28.2 (7.8)	21.3 (5.5)	100%	Around half of the participants (n = 46) reported a duration of illness longer than 5 years BED or EDNOS. Unclear what ratio was. The majority of the participants (n = 85) had completed outpatient treatment.	105	Guided self-help (ED) internet based	Wait list controls	16	4 months	None reported
Hötzel 2014	27.6 (8.3)	20.4 (3.6)	100%	NR AN or BN	212	Guided self-help (ED) internet based	Wait list controls	6	6 weeks	None reported
Traviss 2011	37.1 (12.8)	28.0 (7.5)	98%	NR BN (27%), BED (24%) EDNOS (24%), no diagnosis (disordered	81	Guided self-help (ED)	Wait list controls	7	12 weeks	None reported

Study ID	Mean Age (SD)	Mean BMI (SD)	Females (%)	Stage of illness: duration	N initially random-ised	Intervention Category	Comparison Arm Category	Sessions N	Treatment Length	Long-term FU
				eating) (24%)						

1

2

3 **Table 343: Summary table of findings for individual hybrid therapy versus group hybrid for adults with EDNOS.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Group hybrid	Risk difference with EDNOS Individual hybrid (95% CI)
Depression	35 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.11 standard deviations lower (0.77 lower to 0.56 higher)
General psychopathology	35 (1 study)	⨁⊖⊖⊖ VERY LOW1 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology in the intervention groups was 0.13 standard deviations lower (0.79 lower to 0.54 higher)
Dietary restraint	35 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean dietary restraint in the intervention groups was 0.08 standard deviations higher (0.58 lower to 0.74 higher)
EDI Total	35 (1 study)	⨁⨁⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total in the intervention groups was 0.29 standard deviations higher (0.38 lower to 0.96 higher)
Remission ITT	35	⨁⊖⊖⊖ VERY LOW1,4,5	RR 0.79 (0.35 to	444 per 1000	93 fewer per 1000

	(1 study)	due to risk of bias, indirectness, imprecision	1.81)		(from 289 fewer to 360 more)
Depression FU	35 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.55 standard deviations higher (0.12 lower to 1.23 higher)
General psychopathology FU	35 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean general psychopathology fu in the intervention groups was 0.33 standard deviations higher (0.33 lower to 1 higher)
Dietary restraint FU	35 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean dietary restraint fu in the intervention groups was 0.14 standard deviations higher (0.52 lower to 0.81 higher)
EDI Total FU	35 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi total fu in the intervention groups was 0.57 standard deviations higher (0.11 lower to 1.23 higher)
Remission ITT FU	35 (1 study)	⊕⊖⊖⊖ VERY LOW1,4,6 due to risk of bias, indirectness, imprecision	RR 0.81 (0.61 to 1.08)	944 per 1000	179 fewer per 1000 (from 368 fewer to 76 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Unclear methods of randomisation or if allocation concealment was performed. Participants were not blinded, unclear if investigators and assessors were blind. Considerable difference in dropout rates between individual 23% vs. group 5%,
2 95% CI crossed 2 MIDs (-0.5 and 0.5)
3 95% CI crossed 1 MID (0.5)
4 Remission was not a valid measure. It was defined as the percentage of participants who score one or more scale steps lower than their pre-treatment values for binge eating and/or purging at the RAB-R interview. However, you could move from several times each day to 5-7 days a week. Not necessarily zero times a week. Duration may be okay since it is based on DSM-IV.
5 95% CI crossed 2 MIDs (0.75 and 1.25)

6 95% CI crossed 1 MID (0.75)

Table 344: Summary table of findings for internet self-help versus wait list controls for adults with any eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Any ED Internet SH (95% CI)
EDE-Q Total score	78 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q total score in the intervention groups was 0.34 standard deviations lower (0.79 lower to 0.11 higher)
EDE-Restraint	290 (2 studies)	⨁⨀⨀⨀ VERY LOW3,4,5 due to risk of bias, inconsistency, imprecision		Not calculable for SMD values	The mean ede-restraint in the intervention groups was 0.09 standard deviations lower (0.32 lower to 0.14 higher)
EDE-Eating concern	290 (2 studies)	⨁⨀⨀⨀ VERY LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern in the intervention groups was 0.01 standard deviations lower (0.24 lower to 0.22 higher)
EDE-Weight concern	290 (2 studies)	⨁⨁⨀⨀ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern in the intervention groups was 0.13 standard deviations higher (0.1 lower to 0.37 higher)
EDE-Shape concern	290 (2 studies)	⨁⨁⨀⨀ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concern in the intervention groups was 0.09 standard deviations higher (0.14 lower to 0.32 higher)

BMI	212 (1 study)	⨁⨁⊖⊖ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.10 standard deviations higher (0.17 lower to 0.37 higher)
Depression	78 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0.31 standard deviations lower (0.76 lower to 0.14 higher)
Vomiting	212 (1 study)	⨁⨁⊖⊖ LOW3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean vomiting in the intervention groups was 0.21 standard deviations lower (0.48 lower to 0.06 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 No details were provided on how random sequence was generated and it was unclear if allocation concealment was performed. It was unclear if either the participants, investigators or assessors were blind.
2 95% CI crossed 1 MID (-0.5).
3 It was unclear if allocation concealment was performed. It was unclear if either the participants, investigators or assessors were blind. High dropouts were reported >20%.
4 Heterogeneity was detected I2 >50%.
5 For a continuous variable, there were fewer than 400 participants.

1 **Table 345: Summary table of findings for guided self-help versus wait list controls for adults with any eating disorder.**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Any ED Guided SH (ED) (95% CI)

EDE-Q Total score	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-q total score in the intervention groups was 0.68 standard deviations lower (1.13 to 0.23 lower)
EDE-Restraint	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-restraint in the intervention groups was 0.49 standard deviations lower (0.93 to 0.05 lower)
EDE-Eating concern	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-eating concern in the intervention groups was 0.6 standard deviations lower (1.05 to 0.15 lower)
EDE-Shape concern	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-shape concern in the intervention groups was 0.59 standard deviations lower (1.03 to 0.14 lower)
EDE-Weight concern	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede-weight concern in the intervention groups was 0.6 standard deviations lower (1.05 to 0.15 lower)
BMI	81 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.18 standard deviations higher (0.26 lower to 0.61 higher)
Binge eating	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean binge eating in the intervention groups was 0.07 standard deviations lower (0.5 lower to 0.37 higher)
Vomiting	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias,		Not calculable	The mean vomiting in the intervention groups was 0.12 standard deviations lower

		imprecision		for SMD values	(0.55 lower to 0.32 higher)
Laxative use	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean laxative use in the intervention groups was 0.15 standard deviations lower (0.59 lower to 0.29 higher)
Exercise frequency	81 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean exercise frequency in the intervention groups was 0.02 standard deviations higher (0.42 lower to 0.45 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 It was unclear if allocation concealment was performed. It was unclear if either the participants, investigators or assessors were blind. High dropouts were reported >20%.
2 95% CI crossed 1 MID (-0.5).
3 For a continuous outcome, there were fewer than 400 participants.

1

2

3

9.1.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for people with EDNOS was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.1.4 Clinical evidence statements

9.1.4.1 Individual therapy versus group therapy

9.1.4.1.1 Individual hybrid compared with group hybrid for people with EDNOS at the end of treatment

Very low quality evidence from one RCT (n=35) showed no difference in the effect of individual hybrid therapy on depression, general psychopathology, EDE-restraint, EDI-total and remission compared with group hybrid.

9.1.4.1.2 Individual hybrid compared with group hybrid for people with EDNOS at follow up

Very low quality evidence from one RCT (n=35) showed no difference in the effect of individual hybrid therapy on depression, general psychopathology, EDE-restraint, EDI-total and remission compared with group hybrid.

9.1.4.1.3 Internet self-help compared with wait list control for people with any eating disorder at the end of treatment

Low quality evidence from one RCT (n=78) showed no difference in the effect of internet self-help on EDE-total and depression compared with wait list controls.

Low quality evidence from one RCT (n=212) showed no difference in the effect of internet self-help on BMI compared with wait list controls.

Low quality evidence from two RCTs (n=290) showed no difference in the effect of internet EDE-restraint, EDE-eating concern, EDE-shape concern and EDE-weight concern compared with wait list controls.

9.1.4.1.4 Guided self-help compared with wait list control for people with any eating disorder at the end of treatment

Low quality evidence from one RCT (n=81) showed no difference in the effect of guided self-help on BMI, bingeing, vomiting, laxative use and exercise frequency compared with wait list controls.

Low quality evidence from one RCT (n=81) showed guided self-help is more effective on EDE-total, EDE-restraint, EDE-eating concern, EDE-shape concern and EDE-weight concern compared with wait list controls.

9.1.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for people with EDNOS was available.

9.1.6 Recommendations and link to evidence for the review on: Does any group or individual psychological intervention with or without a pharmacological intervention produce benefits/harms in people with eating disorders compared with any other intervention or controls?

Psychological treatment for OSFED (EDNOS)

	147. For people with OSFED, consider using the treatments for the eating disorder it most closely resembles.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating EDNOS (note: the recommendation replaces the term EDNOS with OSFED). For this population it was agreed that binge-eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with EDNOS that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between clinical benefits and harms	**EDNOS** Only one RCT was found on the EDNOS population. They compared an individual hybrid programme with a group hybrid programme in adult females with EDNOS. At the end of the treatment there was no difference in remission, general psychopathology, EDE-dietary restraint, depression and EDI-total. Similar results were found at follow up. No data was available on binge frequency, adverse events, all-cause mortality, quality of life, resource use, relapse, body weight, general functioning, family functioning, or service user experience. **Any eating disorder** Three studies were found on those with any eating disorder. Two studies that compared internet self-help compared with wait list controls in adults with a range of eating disorders showed no difference in EDE-subscales, BMI and depression at the end of treatment. There was a trend for a reduction in vomiting in the internet self-help group, however, there was some uncertainty. No data was available on remission, adverse events, all-cause mortality, quality of life, resource use, relapse, general functioning, family functioning, or service user experience. Another study compared guided self-help with wait list controls in adults with a range of eating disorders and showed no difference in BMI, binge eating, vomiting, laxative use and exercise frequency at the end of the treatment. However, there was a reduction in EDE-subscales in the guided self-help. No data was available on remission, adverse events, all-cause mortality, quality of life, resource use, relapse, general psychopathology, general functioning, family functioning, or service user experience.
Trade-off between net health benefits and resource use	The committee expressed the view that if something is cost effective for people with a particular eating disorder it will be for a person with EDNOS in-line with the clinical presentation of an eating disorder they most closely resemble.
Quality of evidence	The evidence was mostly low to very low quality. It was unclear how randomisation was conducted and if allocation concealment was performed. Participants were not blinded and it was unclear if investigators and assessors were blind. Outcomes were often downgraded for impression because the 95% confidence interval crossed one or two minimal important differences or it didn't meet the optimal information size.

	147. For people with OSFED, consider using the treatments for the eating disorder it most closely resembles.
	In the study on hybrid treatments, the author's definition of remission was poor and could have been excluded for this reason. However, it was kept in and downgraded for indirectness because of the lack of data on this population. The authors defined remission as the percentage of participants who score one or more scale steps lower than their pre-treatment values for binge eating and/or purging using a revised version of the Rating of Anorexia/Bulimia interview. However, this definition of remission would include people who move from the 'several times each day' category to five-seven days a week, not those who have stopped binging over a two week period. The duration over which they measured symptoms may be acceptable because it was based on DSM-IV. Results on people with any eating disorder were included in this review since people with EDNOS may have similar symptoms and behaviours to those with anorexia nervosa, bulimia nervosa and binge eating disorder. No data was available comparing the individual or group interventions with wait list controls, so it is difficult to know if they are better than no treatment alone. No data was available on children or young people with ENDOS.
Other considerations	The treatments in the hybrid study comprised CBT-ED followed by interpersonal psychotherapy in either an individual or group format. The results show either format was equally effective for this population. The guided self-help appeared to show some benefit on the EDE-subscales compared with wait list controls but not when specific compensatory behaviours were measured. Given the scarcity of data on this population, the committee agreed that it was preferable to recommend psychological treatments for a person with EDNOS (OSFED) in line with the clinical presentation of the eating disorder that their signs and symptoms most closely resembled. For example, if they presented with symptoms similar to someone with a binge eating disorder, then it was best if they follow the NICE recommendations for this population. Heterogeneity was not detected.

9.2 Carer interventions

9.2.1 Review Question: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 346. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all psychological interventions for the parents or carers of children or young people with an eating disorder. The interventions were categorised according to their mode of delivery (e.g. group, individual or self-help), the age of the people with the eating disorder and the type of eating disorder. The control arm could include wait list controls, treatment as usual or any other intervention, however results comparing an intervention with wait list controls were always presented separately.

Table 346: Clinical review protocol summary for the review of: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Component	Description
Review question(s)	Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?
Population	• Family or carers of people with an eating disorder (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder)
Intervention(s)	• Psychological interventions may include: ○ Family-based ○ Parent only (not necessarily focussed on eating disorder) ○ Parent-focused therapy (PFT) ○ Group Parent-Training (GPT) ○ Separated family therapy ○ Parents with person with ED (greater focus on eating disorder) ○ Behavioural Family Therapy (BFT) ○ Behavioural family systems therapy (BFST). ○ Family Based Treatment (FBT) ○ Family Day Workshops (FDW) ○ Family Therapy (FT) ○ Family therapy for anorexia nervosa (FT-AN) ○ Multi-Family Group Day Treatment (MFGDT) ○ Multi-Family Group Therapy (MFGT) ○ Systemic Family Therapy (SFT) ○ Systemic Family Therapy for AN (SFT-AN) ○ Multifamily therapy (MFT) is synonymous with (MFGT; MFGDT). ○ Uniting couples in the treatment of AN (UCAN ○ Conjoint family therapy
Comparison	• Wait list controls • Treatment as usual • Another intervention
Critical outcomes	• Parent's or carer's general psychopathology (including mood/depression/anxiety) • Family functioning • Quality of life • Other primary outcomes commonly reported in studies that just target the family/carer • The following outcomes will be included if the family or carer intervention includes the child or person with an eating disorder: ○ Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) ○ Binge eating for BN and BED ○ Body weight / BMI for AN
Important outcomes	• General functioning • Resource use. • Service user experience • All-cause mortality. • Adverse events • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion)

Component	Description
Study design	• Systematic Reviews • RCTs

9.2.2 Clinical Evidence for: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Two RCTs (n=204) met the eligibility criteria for this review, one of which was in young people (Spettigue 2015 (Spettigue et al., 2015)), the other which was in adults (Goddard 2011 (Goddard, 2011)). An overview of all the trials included in the review can be found in Table 347. Further information about both included and excluded studies can be found in Appendix J.

One study (n=51) examined the efficacy of psychoeducation that included a 2-hour session and bi-weekly telephone support calls in the time before formal assessment calls compared with wait list controls (Spettigue 2015). Another study (n=153) examined the efficacy of the Expert Carers Helping Others (ECHO) self-help intervention with and without guidance (Goddard 2011).

Summary of findings for interventions for carers of people with any eating disorder can be found in Table 348, Table 349 and Table 350. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 347: Study information for trials included in the review of interventions for carers of people with any eating disorder

Study ID	Mean Age of Carer (SD)	Female carers (%)	Mean Age of patient (SD)	Female patients (%)	Sample	N Random-ised	Intervention	Comparison	Number of sessions	Length of interventions	Follow up
Goddard 2011	49.6 (8.1)	89	20.9 (6.8)	95	Carers of people with an eating disorder	153 carers	Guided Self-Help (ECHO)	Self-Help Only (ECHO)	3 x 40m telephone guidance sessions	6 weeks	3 months
Spettigue 2015	Not reported	92	15.7 (1.5)	97	Carers of medically stable young people awaiting treatment from specialist ED programs	51 carers	Psychoeducation session + telephone support	Wait list control	2 hour psychoeducation session and 2 x weekly phone calls until assessment	Mean of 94 days (range 27-287) to assessment	-

Abbreviations: ECHO, Expert Carers Helping Others

Table 348: Summary table of findings for psychoeducation versus wait list control in carers of young people with any eating disorder at end of treatment.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Psychoeducation (95% CI)
Carer Self-Efficacy Parents Versus Anorexia (PVA)	31 (1 study) 260 days	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for	The mean carer self-efficacy in the intervention groups was 1.74 standard deviations higher (0.89 to 2.59 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Psychoeducation (95% CI)
				SMD values	
Carer Knowledge of ED Knowledge of Eating Disorders Scale (KEDS)	28 (1 study) 260 days	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean carer knowledge of ed in the intervention groups was 0.75 standard deviations higher (0.04 lower to 1.54 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Spettigue 2015: Randomization method unclear, allocation concealment unclear, participant and assessor not blinded, investigator blinding unclear, dropout rate for both arms >20%, available case analysis. 2 Study targeted carers of medically stable young people awaiting assessment by specialized eating disorder program. End of treatment data for wait list control was after 1 month. At time of assessment, 4 of 36 young people were not diagnosed with an eating disorder. Mean time to assessment: 94 days, range 27-287 days 3 Fewer than 400 participants. 4 CI crosses either 0.5 or -0.5 (SMD).					

1 **Table 349: Summary table of findings for psychoeducation versus wait list control in carers of young people with any eating disorder**
2 **at formal assessment**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Psychoeducation (95% CI)
Carer Self-Efficacy FU	31	⊕⊖⊖⊖		Not	The mean carer self-efficacy fu in the

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with WLC	Risk difference with Psychoeducation (95% CI)
Parents Versus Anorexia (PVA)	(1 study)	VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		calculable for SMD values	intervention groups was 0.89 standard deviations higher (0.14 to 1.64 higher)
Carer Knowledge of ED FU Knowledge of Eating Disorders Scale (KEDS)	28 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean carer knowledge of ed fu in the intervention groups was 0.99 standard deviations higher (0.18 to 1.8 higher)
Carer Burden FU Eating Disorder Symptom Impact Scale (EDSIS)	36 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean carer burden fu in the intervention groups was 0.57 standard deviations higher (0.11 lower to 1.25 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; FU: Follow up

1 Spettigue 2015: Randomization method unclear, allocation concealment unclear, participant and assessor not blinded, investigator blinding unclear, dropout rate for both arms>20%, available case analysis.
2 Study targeted carers of medically stable young people awaiting assessment by specialized eating disorder program. End of treatment data for wait list control was after 1 month. At time of assessment, 4 of 36 young people were not diagnosed with an eating disorder. Mean time to assessment: 94 days, range 27-287 days
3 CI crosses either 0.5 or -0.5 (SMD).

1 **Table 350: Summary table of findings for guided self-help versus self-help in carers of adults with any eating disorder**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Self-Help	Risk difference with Guided Self-Help (95% CI)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Self-Help	Risk difference with Guided Self-Help (95% CI)
Carer Burden ECI Negative; EDSIS	120 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer burden in the intervention groups was 0.02 standard deviations higher (0.24 lower to 0.27 higher)
Carer Quality of Life General Health Questionnaire-12 (GHQ-12)	120 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer quality of life in the intervention groups was 0.07 standard deviations lower (0.43 lower to 0.28 higher)
Carer Expressed Emotion (Family Functioning) Family Questionnaire (FQ)	120 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer expressed emotion (family functioning) in the intervention groups was 0.14 standard deviations lower (0.5 lower to 0.22 higher)
Carer Self-Efficacy Revised Scale for Caregiving Self-Efficacy (CSE)	120 (1 study) 3 months	⨁⨁⨀⨀ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer self-efficacy in the intervention groups was 0.15 standard deviations higher (0.21 lower to 0.51 higher)
Experience of Caregiving Inventory (ECI) Positive Experience of Caregiving Inventory (ECI)	120 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean experience of caregiving inventory (eci) positive in the intervention groups was 0.05 standard deviations higher (0.3 lower to 0.41 higher)
Carer Accommodation & Enabling AESED	120 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for	The mean carer accommodation & enabling in the intervention groups was 0.01 standard deviations lower (0.37 lower to 0.35 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Self-Help	Risk difference with Guided Self-Help (95% CI)
				SMD values	
Carer General Psychopathology (Distress) Hospital & Anxiety Depression Scale (HADS)	120 (1 study) 3 months	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean carer general psychopathology (distress) in the intervention groups was 0.06 standard deviations lower (0.42 lower to 0.3 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; FU: Follow up					
1 Goddard 2011: Unclear whether baseline characteristics of carers were similar. Also, dropout rate <20% and reasons not stated. 2 Fewer than 400 participants. 3 CI crosses either 0.5 or -0.5 (SMD).					

1

2

9.2.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with EDNOS was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.2.4 Clinical evidence statements

9.2.4.1 Psychoeducation versus wait list control in carers of young people with any eating disorder at end of treatment

Very low quality evidence from one RCT (n=31) showed psychoeducation is more effective on carer self-efficacy compared with wait list control.

Very low quality evidence from one RCT (n=28) showed psychoeducation may be more effective on the carer's knowledge of eating disorders compared with wait list control, although there was some uncertainty.

9.2.4.2 Psychoeducation versus wait list control in carers of young people with any eating disorder at formal assessment

Very low quality evidence from one RCT (n=31) showed psychoeducation is more effective on carer self-efficacy compared with wait list control.

Very low quality evidence from one RCT (n=28) showed psychoeducation is more effective on the carer's knowledge of eating disorders compared with wait list control.

Very low quality evidence from one RCT (n=36) showed no difference in the effect of psychoeducation on the burden of the eating disorder on the carer compared with wait list control.

9.2.4.3 Guided self-help versus self-help in carers of adults with any eating disorder

Low quality evidence from one RCT (n=120) showed no difference in the effect of guided self-help on the burden of the eating disorder on the carer, carer quality of life, carer-rated family functioning, carer self-efficacy, carer positive experience of caregiving, carer accommodation and enabling and carer general psychopathology compared with self-help only.

9.2.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the parents or carers of children or young people with EDNOS was available.

9.2.6 Recommendations and link to evidence for the review on: Does any psychological intervention produce benefits/harms in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?

Working with family members and carers

	148. Be aware that the family members or carers of a person with an eating disorder may experience severe distress. Offer them an assessment of their own needs, including

	• what impact the eating disorder has on them • what support they need, including practical support and emergency plans for increasing medical or psychiatric risk. **149. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes, when assessing whether any interventions help the parents and carers of children and young people with an eating disorder. The critical outcomes for the parents and carers were: general psychopathology, family functioning, quality of life, and other primary outcomes reported by the study. Other outcomes that are critical for the child or young person with the eating disorder include remission and bingeing or body weight, depending on the eating disorder. Other outcomes that are of lesser importance but clearly important outcomes include, general functioning, service user experience, all-cause mortality, adverse events and eating disorder psychopathology.
Trade-off between clinical benefits and harms	**Any eating disorder (evidence found in this chapter)** Randomised control trials investigating interventions for the carers of young people with any eating disorder failed to show many favourable outcomes. Psychoeducation compared with waitlist control showed a positive effect on carer self-efficacy and a trend to improve carer knowledge of eating disorders at the end of treatment. Long-term follow up showed favourable results in both but carer burden (only measured at follow up) was not different compared with wait list controls. No evidence was found on the critical outcomes of carer general psychopathology, family functioning, and quality of life, nor on the other important outcomes. Comparing guided self-help with self-help showed no difference in any of the carer-related outcomes at the end of treatment. No evidence was found on the other important outcomes. **Anorexia nervosa (evidence in chapter 6)** One randomised controlled trial (RCT), aimed at carers of young people with anorexia nervosa, and compared the effectiveness of guided self-help or self-help (and treatment as usual) with treatment as usual alone. After 12 months there was no difference in carer general psychopathology. No evidence was found on the critical outcomes of carer general psychopathology, carer family functioning, carer quality of life, nor the important outcomes of eating psychopathology, carer general functioning, service user experience, resource use, adverse events, and all-cause mortality. Another study compared self-help (with treatment as usual) with treatment as usual and showed no difference in the carer's general psychopathology or carer skills after 6 to 12 months but a trend for poorer results for family functioning but there was some uncertainty. In the young people with anorexia nervosa whom they care for, there was no difference in BMI, weight, severity index (SEED), general psychopathology, clinical improvement, peer related problems between the two treatment arms. However, there was a trend for poorer outcomes in prosocial behaviour in the self-help group but there was some uncertainty. No evidence was found on the critical outcomes of remission, carer general psychopathology, nor the important outcomes of service user experience, resource use, adverse events, and all-cause mortality. Comparing guided self-help (and treatment as usual) with treatment as usual showed at 12 months a trend for positive outcomes in the combined treatment group on carer burden and quality of life, but no difference in family functioning, carer skills or carer psychopathology. There was a trend for poorer outcomes in

	carer accommodation and enabling. At 24 months, there was a trend for a positive result on carer burden, quality of life, carer accommodation and enabling, and carer psychopathology. In addition, a trend for poorer outcomes in family functioning and time spent caring. No evidence was found on the critical outcome of carer general psychopathology, nor the important outcomes of service user experience and resource use. In the same intervention, the guided self-help for the carers did not translate to many benefits in the young people with anorexia nervosa whom they care for. At 12 months, no differences were found in any of the outcomes including mortality, admission to hospital, patient relapse, BMI, EDE-global, severity index (SEED), general psychopathology, clinical improvement. However there was improvement in peer problems but a trend for a negative result in prosocial behaviour. At 24 months, there was a trend for positive increase in BMI and EDE-global, no difference in general psychopathology, and a trend for a negative result in quality of life. No evidence was found on the critical outcome of remission, nor the important outcomes of adverse events, and all-cause mortality. Comparing 2 active treatments generally showed no difference in effectiveness in the carer-related outcomes. Guided self-help compared with self-help were equally effective on all outcomes 6 to 12 months after the young people with anorexia nervosa had been discharged from inpatient care, except there was a trend for carer accommodation to favour guided self-help. No evidence was found on the critical outcome of remission, nor the important outcomes of adverse events, and all-cause mortality. In the young people with anorexia nervosa, there was a trend for poorer results on BMI and peer problems in the guided self-help group compared with self-help. No difference was found in clinical severity (SEED), general psychopathology, clinical improvement, prosocial behaviour but there was a trend for better results in peer problems. No evidence was found on the critical outcome of remission. Web-based guided self-help also failed to show convincing benefits for the carers of young people with anorexia nervosa compared with treatment as usual. At the end of treatment, a poorer outcome in distress was found but there was some uncertainty. The other outcomes, such as carer accommodation and enabling, family functioning, carer burden and caregiving experience showed no difference. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. Web-based guided self-help compared with web-based self-help showed no difference in the outcomes for carers at the end of treatment. At follow up, favourable results were found on family functioning in the guided web-based self-help group, but no difference in carer experience, quality of life and general psychopathology. There was a trend for poorer results in carer burden, but there was some uncertainty. No evidence was found on the critical outcome of quality of life nor on the important outcome of resource use. **Bulimia nervosa or binge eating disorder (chapter 7 and 8)** No relevant published evidence was found on parents or carers of children and young people with bulimia nervosa or binge eating disorder. Adverse events or all-cause mortality were not reported in any of the studies.
Trade-off between net health benefits and resource use	The committee expressed the view that offering family members and carers an assessment of their own needs may incur additional healthcare resources (that is, time required to perform such assessment). However, the committee considered the cost of providing such assessment to be small, taking into account the potential reduction in family and carers' burden, potential depression and other health vulnerabilities which may be costly to other parts of the healthcare system, especially considering that the burden on family and carers can last for many years and increase their morbidity and stress. Consequently, the committee judged that assessment that aims to improve family and carers' experience are likely to represent good value for money.
Quality of evidence	The quality of the evidence was mostly very low. The outcomes were downgraded because it was unclear how they randomised or if allocation concealment was performed, if participants or investigators were blinded. In some, not all, assessors

	were blinded. High dropout rates were also detected in some groups >20%. Imprecision was detected in most outcomes due to the 95% confidence interval crossing one or two minimal important differences or because it did not meet the optimal information size. Outcomes were not always measured at the end of treatment or at follow up. It is not known if any improvements in the carer's general psychopathology also translated to benefits in the children with the eating disorder.
Other considerations	Given the very low quality of the data with very few positive findings favouring one arm over the other, the committee came to the consensus that there was not enough evidence to support a recommendation on any specific treatment for parents or carers of people with an eating disorder. Nevertheless, the committee acknowledged the stress and burden that a person with an eating disorder, in particular anorexia nervosa, can have on family members over a long period of time. Therefore, they agreed that offering family members and carers an assessment of their own needs, including: personal, social and emotional support available to them, the need for support in the caring role for example if the child should need urgent care and there are other children to take care of and to offer advice on where they can get some practical support. The extent to which the family need to be involved in treatment depends on the age and developmental needs of the person with the eating disorder, the severity of the illness, the risk from harm and the person receiving treatment's wishes. In general, parents and other family members will want to be involved in the treatment. If a parent or carer cannot attend a meeting the healthcare professional should provide written information on the outcome of an assessment or treatment where appropriate. The committee acknowledged the importance of consent and confidentiality and their discussion can be found in the LETRs relating to this. They also discussed that although the evidence found was for carers and parents of people with anorexia nervosa or any eating disorder, the recommendation is relevant for parents and carers of people with bulimia nervosa and binge eating disorder. This is mostly because no specific intervention was recommended, rather to offer an assessment of their needs and to help them find the necessary support. In absence of good evidence, the committee agreed to generate a research recommendation to address the question "What is the effectiveness of a carer-focused psychological intervention in the parents or carers of children or young people with an eating disorder compared with any other intervention or controls?" See chapter 6.2.

9.3 Pharmacological interventions

9.3.1 Review Question: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 351. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all pharmacological interventions that may be delivered to children, young people and adults with an eating disorder with or without a psychological intervention. The interventions were categorised according to the type of pharmacological intervention, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to placebo, wait list controls or any other intervention

Table 351: Clinical review protocol summary for the review of: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

Component	Description
Review question(s)	Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). • Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Pharmacological interventions may include: • Antidepressants, e.g. SSRIs, fluoxetine (Prozac) • Anxiolytics (antianxiety) • Antipsychotics • Antiemetic medication, e.g. ondansetron • Antiepileptic/anticonvulsant, e.g. topiramate (Topomax) • Appetite suppressant, e.g. lisdexamfetamine dimesylate • Pharmacological in combination with any psychological intervention
Comparison	• Placebo • Wait list control • Treatment as usual • Another intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN • Adverse events
Important outcomes	• All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Relapse • Resource use • Quality of life • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs

9.3.2 Clinical Evidence for: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

No studies were identified that met the eligibility criteria for this review. Further information about excluded studies can be found in Appendix J. See also the study selection flow chart in Appendix K.

9.3.3 Economic Evidence

No economic evidence on the cost effectiveness of pharmacological interventions for the treatment of EDNOS was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.3.4 Clinical evidence statements

No studies were identified that met the eligibility criteria for this review.

9.3.5 Economic Evidence statements

No economic evidence on the cost effectiveness of pharmacological interventions for people with EDNOS was available.

9.3.6 Recommendations and link to evidence for the review on: Does any pharmacological intervention produce benefits/harms on specified outcomes in people with eating disorders?

	The committee agreed that people with OFSED should be treated in-line with the eating disorder their symptoms most closely resemble
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating OSFEDs. For this population it was agreed that binge-eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making.
Trade-off between clinical benefits and harms	No RCT evidence was identified on pharmacological interventions for people with EDNOS. The committee expressed the view that if something is cost effective for people with a particular eating disorder it will be for a person with EDNOS in-line with the clinical presentation of an eating disorder they most closely resemble.
Trade-off between net health benefits and resource use	The network meta-analysis found no evidence for the effectiveness of pharmacological interventions for the management of people with BN and BED. As a result the committee expressed the view that such treatments are unlikely to be effective nor cost effective in people with EDNOS.
Quality of evidence	Not applicable
Other considerations	The committee agreed that people with EDNOS should be treated in-line with the eating disorder their symptoms most closely resemble.

9.4 Nutritional interventions

9.4.1 Review Question: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 352. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all nutritional interventions that may be delivered to children, young people and adults with an eating disorder with or without a pharmacological intervention. The interventions were categorised according to type of nutritional intervention, the age of the participants and the type of eating disorder. In addition, the interventions were grouped according to their type of therapy and were compared to wait list controls, placebo, or any other intervention.

Table 352: Clinical review protocol summary for the review of: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

Component	Description
Review question(s)	Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). • Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Nutritional intervention • Method of feeding • Nutritional in combination with any pharmacological intervention • Examples of nutritional interventions are nutritional counselling (with or without educational and supportive groups) and supplements (e.g. zinc)
Comparison	• Placebo • Wait list control • Treatment as usual • Another intervention (psychological, pharmacological, nutritional, physical)
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global

Component	Description
	Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs

9.4.2 Clinical Evidence for: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

No studies were identified that met the eligibility criteria for this review. Further information about excluded studies can be found in Appendix J. See also the study selection flow chart in Appendix K.

9.4.3 Economic Evidence

No economic evidence on the cost effectiveness of nutritional interventions for people with EDNOS was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.4.4 Clinical evidence statements

No studies were identified that met the eligibility criteria for this review

9.4.5 Economic Evidence statements

No economic evidence on the cost effectiveness of nutritional interventions for people with EDNOS was available.

9.4.6 Recommendations and link to evidence for the review on: Does any nutritional intervention produce benefits/harms on specified outcomes in people with eating disorders?

	The committee agreed that people with OFSED should be treated in-line with the eating disorder their symptoms most closely resemble
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating OSFEDs. For this population it was agreed that binge-eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making.
Trade-off between clinical benefits and harms	No relevant published evidence was identified.
Trade-off between net health benefits and resource use	No relevant published economic evidence was identified.

Quality of evidence	Not applicable
Other considerations	The committee agreed that people with EDNOS should be treated in-line with the eating disorder their symptoms most closely resemble.

9.5 Physical interventions

9.5.1 Review Question: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 353. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all physical interventions that may be delivered to children, young people and adults with an eating disorder. The interventions were categorised according to type of physical intervention, the age of the participants and the type of eating disorder and were compared to wait list controls, placebo, treatment as usual or any other intervention.

Table 353: Clinical review protocol summary for the review of Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?

Component	Description
Review question(s)	Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder). • Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Physical interventions may include: ○ transcranial magnetic stimulation ○ deep brain stimulation ○ physiotherapy ○ yoga ○ physical exercise ○ acupuncture ○ mandometer ○ massage
Comparison	• Placebo • Wait list control • Treatment as usual • Another intervention
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events

Component	Description
	• All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety) • General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs

9.5.2 Clinical Evidence for: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders

Five RCTs (n=425) met the eligibility criteria for this review (Bloomgarden 2008 (Bloomgarden and Calogero, 2008), Boerhout 2016 (Boerhout et al., 2016), Carei 2010 (Carei et al., 2010), Hildebrandt 2012 (Hildebrandt et al., 2012), Trottier 2015). The majority of studies were in an outpatient setting and the majority of participants were adult females (no young people). One study was conducted in an inpatient setting (Bloomgarden 2008), whilst one study was conducted after participants had received treatment in an intensive day hospital setting (Trottier 2015). Further information about both included and excluded studies can be found in Appendix J.

Summary of findings for those on any eating disorder can be found in Table 355, Table 356, Table 357, Table 358, Table 359, Table 360, Table 361 and Table 362. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 354: Study information for trials included in the analysis of physical interventions versus any other intervention or wait list control for people with any eating disorder.

Study ID	N Random-ized	Female (%)	Mean BMI (SD) (kg/m2) or other	Sample	Intervention Age at onset and/or duration of illness (years)	Comparison(s) Age at onset and/or duration of illness (years)	Duration
Boerhout 2016	40	100	Not reported	Adult AN=9 BN=16 BED=4	Psychomotor therapy + supportive contact	Supportive contact	Not reported, Six 1 hour sessions of psychomotor therapy
Bloomgarden 2008	86	100	20.1 (5.3)	Adult inpatients AN-R=27 BN=23 EDNOS=36	Eye movement desensitization and reprocessing therapy + treatment as usual	Treatment as usual	18 months; 12-mo FU
Carei 2010	53	92	19.2 (2.7)	Young people & Adult outpatients AN=29; BN=9; EDNOS=15	Yoga + treatment as usual Length of illness for whole sample: 1.2 (1.4)	Treatment as usual Length of illness for whole sample: 1.2 (1.4)	9 weeks; 3 weeks FU
Hildebrandt 2012	33	88	21.9 (2.3)	Adult AN in partial remission=3 BN=2 EDNOS=26 BED=2	Acceptance-based mirror exposure therapy + treatment as usual	Non-Directive Body Image Therapy + treatment as usual	5 x 50 min sessions; 1-mo FU
Trottier 2015	45	100	22.6 (3.1)	Adults in partial remission AN=1; BN=29; EDNOS=4	Graded body image exposure + maintenance treatment as usual	Maintenance treatment as usual	4 weeks + 4-16 weeks of maintenance treatment as usual

*Notes: *, participants randomised to interventions after intensive day hospital treatment. Abbreviations: FU: follow up*

Table 355: Summary table of findings for eye movement desensitization and reprocessing therapy versus treatment as usual in adult inpatients with any eating disorder at end of treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Eye Movement Desensitization & Reprocessing Therapy (95% CI)
Body Image Memory Questionnaire - Earliest Memory	86 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean body image memory questionnaire - earliest memory in the intervention groups was 0.63 standard deviations lower (1.06 to 0.19 lower)
Body Image Memory Questionnaire - Worst Memory	86 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean body image memory questionnaire - worst memory in the intervention groups was 0.77 standard deviations lower (1.21 to 0.33 lower)
Body Image Memory Questionnaire - Most Recent Memory	86 (1 study) 12 months	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean body image memory questionnaire - most recent memory in the intervention groups was 0.38 standard deviations lower (0.81 lower to 0.04 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Bloomgarden 2008: No participant blinding, Investigator and assessor blinding unclear. Sample consisted of 29 AN-R, 23 BN, and 36 EDNOS.
2 CI crosses -0.5.

Table 356: Summary table of findings for eye movement desensitization and reprocessing therapy versus treatment as usual in adult inpatients with any eating disorder at follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Eye Movement Desensitization & Reprocessing Therapy (95% CI)
Body Image Memory Questionnaire - Earliest Memory FU	66 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean body image memory questionnaire - earliest memory fu in the intervention groups was 0.22 standard deviations lower (0.71 lower to 0.26 higher)
Body Image Memory Questionnaire - Worst Memory FU	66 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean body image memory questionnaire - worst memory fu in the intervention groups was 0.7 standard deviations lower (1.2 to 0.21 lower)
Body Image Memory Questionnaire - Most Recent Memory FU	66 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean body image memory questionnaire - most recent memory fu in the intervention groups was 0.08 standard deviations lower (0.56 lower to 0.4 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; FU: Follow up

1 Bloomgarden 2008: No participant blinding, Investigator and assessor blinding unclear. Sample consisted of 29 AN-R, 23 BN, and 36 EDNOS.
2 CI crosses -0.5.

Table 357: Summary table of findings for yoga and treatment as usual versus treatment as usual in young people with any eating disorder at end of treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Yoga+ treatment as usual (95% CI)
BMI	53 (1 study) 3 weeks	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI in the intervention groups was 0.22 standard deviations higher (0.32 lower to 0.76 higher)
EDE Global	53 (1 study) 3 weeks	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede global in the intervention groups was 0.05 standard deviations higher (0.49 lower to 0.59 higher)
EDE Restraint	53 (1 study) 3 weeks	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint in the intervention groups was 0.22 standard deviations lower (0.76 lower to 0.32 higher)
EDE Weight Concern	53 (1 study) 3 weeks	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern in the intervention groups was 0.14 standard deviations higher (0.4 lower to 0.68 higher)
EDE Shape Concern	53 (1 study) 3 days	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern in the intervention groups was 0.14 standard deviations higher (0.4 lower to 0.68 higher)
EDE Eating Concern	53 (1 study) 3 weeks	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concern in the intervention groups was 0.09 standard deviations higher (0.45 lower to 0.62 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Yoga+ treatment as usual (95% CI)
Depression BDI-2	53 (1 study) 3 weeks	⨁⨀⨀⨀ VERY LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression in the intervention groups was 0 standard deviations higher (0.54 lower to 0.54 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Carei 2010: Unclear randomization method (stratified, permuted block scheme after baseline measures. No participant blinding; unclear investigator and assessor blinding. Sample consisted of 29 AN, 9 BN, and 15 EDNOS. 2 CI crosses either 0.5 or -0.5 (SMD). 3 CI crosses both 0.5 and 0.5 (SMD).					

Table 358: Summary table of findings for yoga and treatment as usual versus treatment as usual in young people with any eating disorder at follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Yoga + treatment as usual (95% CI)
BMI FU	53 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI fu in the intervention groups was 0.21 standard deviations higher (0.33 lower to 0.75 higher)
EDE Global FU	53 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD	The mean ede global fu in the intervention groups was 0.38 standard deviations lower (0.92 lower to 0.17 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Yoga + treatment as usual (95% CI)
				values	
EDE Restraint FU	53 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint fu in the intervention groups was 0.65 standard deviations lower (1.2 to 0.09 lower)
EDE Weight Concern FU	53 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concern fu in the intervention groups was 0.09 standard deviations lower (0.63 lower to 0.45 higher)
EDE Shape Concern FU	53 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concern fu in the intervention groups was 0.36 standard deviations lower (0.9 lower to 0.19 higher)
EDE Eating Concern FU	53 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concern fu in the intervention groups was 0.28 standard deviations lower (0.82 lower to 0.27 higher)
Depression FU BDI-2	53 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression fu in the intervention groups was 0.09 standard deviations lower (0.63 lower to 0.45 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual	Risk difference with Yoga + treatment as usual (95% CI)
CI: Confidence interval; FU: Follow up					
1 Carei 2010: Unclear randomization method (stratified, permuted block scheme after baseline measures. No participant blinding; unclear investigator and assessor blinding. Sample consisted of 29 AN, 9 BN, and 15 EDNOS. 2 CI crosses either 0.5 or -0.5 (SMD).					

Table 359: Summary table of findings for body image therapy and maintenance treatment as usual versus maintenance treatment as usual in adults with any eating disorder at end of treatment

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with MTAU for adult ED	Risk difference with Body image therapy+MTAU (95% CI)
EDE weight concerns	45 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concerns in the intervention groups was 0.11 standard deviations lower (0.7 lower to 0.47 higher)
EDE shape concerns	45 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede shape concerns in the intervention groups was 0.24 standard deviations higher (0.35 lower to 0.82 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; FU: Follow up					
1 Trottier 2015: Randomization method not specified, unclear allocation concealment; no participant nor investigator blinding, unclear assessor blinding. Dropout both groups>20%. **2 Participants received interventions after intensive day hospital treatment involving group cognitive behavioural program.** **3 CI crosses either 0.5 or -0.5 (SMD).**					

Table 360: Summary table of findings for body image therapy and maintenance treatment as usual versus maintenance treatment as usual in adults with any eating disorder at follow up

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with MTAU for adult ED 6-mo FU	Risk difference with Body image therapy+MTAU (95% CI)
EDE weight concerns	45 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede weight concerns in the intervention groups was 0.2 standard deviations higher (0.39 lower to 0.79 higher)
EDE shape concerns	45 (1 study) 6 months	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede shape concerns in the intervention groups was 0.03 standard deviations lower (0.61 lower to 0.56 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Trottier 2015: Randomization method not specified, unclear allocation concealment; no participant nor investigator blinding, unclear assessor blinding. Dropout both groups>20%.
2 Participants received interventions after intensive day hospital treatment involving group cognitive behavioural program.
3 CI crosses either 0.5 or -0.5 (SMD).
4 CI crosses both 0.5 and -0.5 (SMD).

Table 361: Summary table of outcomes for body image therapy 1 (acceptance-based mirror exposure therapy and treatment as usual) versus body image therapy 2 (non-directive body image therapy and TAU) in adults with any eating disorder

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Body Image Therapy-2	Risk difference with Body Image Therapy-1 (95% CI)
EDE-Q Restraint	33 (1 study) 1 months	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede-q restraint in the intervention groups was 0.11 standard deviations lower (0.35 lower to 0.13 higher)
EDE-Q Eating Concern	33 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,4		Not calculable for	The mean ede-q eating concern in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Body Image Therapy-2	Risk difference with Body Image Therapy-1 (95% CI)
	1 months	due to risk of bias, indirectness, imprecision		SMD values	0.33 standard deviations lower (0.57 to 0.09 lower)
EDE-Q Shape Concern	33 (1 study) 1 months	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede-q shape concern in the intervention groups was 0.68 standard deviations lower (0.94 to 0.43 lower)
EDE-Q Weight Concern	33 (1 study) 1 months	⨁⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean ede-q weight concern in the intervention groups was 0.73 standard deviations lower (0.99 to 0.48 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Hildebrandt 2012: Unclear randomization and allocation concealment. No assessor blinding. Control group dropout rate>20%.
2 Inclusion criteria included participation in concurrent psychotherapy. 18 of the 31 participants were receiving either CBT or Family Therapy.
3 <400 participants.
4 CI crosses either 0.5 or -0.5 (SMD).

1 **Table 362: Summary table of findings for psychomotor therapy and support versus support in adult females with any eating disorder**

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Support	Risk difference with Psychomotor Therapy + Support (95% CI)
Self-Expression & Control Scale - Anger In	29 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision		Not calculable for SMD values	The mean self-expression & control scale - anger in in the intervention groups was 0.49 standard deviations lower (1.24 lower to 0.26 higher)
Self-Expression & Control Scale - Anger Out	29 (1 study)	⨁⊖⊖⊖ VERY LOW1,2 due to risk of bias,		Not calculable	The mean self-expression & control scale - anger out in the intervention groups was 0.28 standard deviations lower

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Support	Risk difference with Psychomotor Therapy + Support (95% CI)
		indirectness, imprecision		for SMD values	(1.02 lower to 0.47 higher)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

1 Boerhout 2016: unclear randomisation method; no participant nor investigator blinding. Dropout rate of both groups >20%. Supportive contact included consultation with hospital staff once every one or two weeks, prescription of medication, psychoeducation, and diet management. Sample consisted of 9 AN, 16 BN and 4 BED participants.
2 CI crosses either 0.5 or -0.5.

1

2

3

9.5.3 Economic Evidence

No economic evidence on the cost effectiveness of physical interventions for people with EDNOS was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.5.4 Clinical evidence statements

9.5.4.1 Eye movement desensitization and reprocessing therapy versus treatment as usual in adult inpatients with any eating disorder at end of treatment

Low quality evidence from one RCT (n=86) showed eye movement desensitization and reprocessing therapy is more effective on Body Image Memory Questionnaire-earliest memory and Body Image Memory Questionnaire-worst memory compared with treatment as usual.

Low quality evidence from one RCT (n=86) showed eye movement desensitization and reprocessing therapy may be more effective on Body Image Memory Questionnaire-recent memory compared with treatment as usual, though there was some uncertainty.

9.5.4.2 Eye movement desensitization and reprocessing therapy versus treatment as usual in adult inpatients with any eating disorder at follow up

Low quality evidence from one RCT (n=66) showed eye movement desensitization and reprocessing therapy is more effective on Body Image Memory Questionnaire-worst memory compared with treatment as usual.

Low quality evidence from one RCT (n=66) showed no difference in the effect of eye movement desensitization and reprocessing therapy on Body Image Memory Questionnaire-earliest memory and Body Image Memory Questionnaire-most recent memory compared with treatment as usual.

9.5.4.3 Yoga and treatment as usual versus treatment as usual in adults with any eating disorder at end of treatment

Low quality evidence from one RCT (n=53) showed no difference in the effect of yoga and treatment as usual on BMI, EDE-global, EDE-dietary restraint, EDE-weight concern, EDE-shape concern, EDE-eating concern and depression compared with treatment as usual.

9.5.4.4 Yoga and treatment as usual versus treatment as usual in adults with eating disorder at follow up

Low quality evidence from one RCT (n=53) showed no difference in the effect of yoga and treatment as usual on BMI, EDE-global, EDE-weight concern, EDE-shape concern, EDE-eating concern and depression compared with treatment as usual.

Low quality evidence from one RCT (n=53) showed yoga and treatment as usual is more effective on EDE-dietary restraint compared with treatment as usual.

9.5.4.5 Graded body image therapy and maintenance treatment as usual versus maintenance treatment as usual in adults with any eating disorder at end of treatment

Very low quality evidence from one RCT (n=45) showed no difference in the effect of graded body image exposure therapy and maintenance treatment as usual on EDE-weight concern and EDE-shape concern compared with maintenance treatment as usual.

9.5.4.6 Graded body image therapy and maintenance treatment as usual versus maintenance treatment as usual in adults with any eating disorder at follow up

Very low quality evidence from one RCT (n=45) showed no difference in the effect of graded body image exposure therapy and maintenance treatment as usual on EDE-weight concern and EDE-shape concerns compared with maintenance treatment as usual.

9.5.4.7 Acceptance-based mirror exposure therapy and treatment as usual versus non-directive body image therapy and treatment as usual in adults with any eating disorder at end of treatment

Very low quality evidence from one RCT (n=33) showed acceptance-based mirror exposure therapy is more effective on EDE-Q-eating concern, EDE-Q-shape concern and EDE-Q-weight concern compared with non-directive body image therapy.

Very low quality evidence from one RCT (n=33) showed no difference in the effect of acceptance-based mirror exposure therapy on EDE-Q-dietary restraint compared with non-directive body image therapy.

9.5.4.8 Psychomotor therapy and supportive contact versus supportive contact in adults with any eating disorder

Very low quality evidence from one RCT (n=29) showed no difference in the effect of psychomotor therapy and support on Self-Expression and Control Scale-anger in and Self-Expression and Control Scale-anger out compared with support only.

9.5.5 Economic Evidence statements

No economic evidence on the cost effectiveness of physical interventions for people with EDNOS was available.

9.5.6 Recommendations and link to evidence for the review on: Do physical interventions, such as transcranial magnetic stimulation or physiotherapy, produce benefits/harms in people with eating disorders

Physical therapy

	150. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitization, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes for the review on the effectiveness of physical interventions, such as transcranial magnetic stimulation or physiotherapy in people with eating disorders and it was agreed that for any eating disorder remission is of greatest concern. The other critical outcomes for anorexia nervosa are body weight and BMI and for binge eating disorder and bulimia nervosa it is bingeing. Other outcomes that are important but are considered rare events or rarely measured in randomised controlled trials for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse, thus they were extracted where possible, but did not factor strongly in the decision making. Other outcomes of concern for people with an eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body general functioning, family functioning and service user experience.
Trade-off between	**Any eating disorder (as reviewed in this chapter 9)** One study compared eye movement desensitization and reprocessing therapy with

clinical benefits and harms	treatment as usual in adults with any eating disorder. The results showed some improvement in the outcomes reported by the body image memory questionnaire, including the earliest memory and worst memory on body image, and only a trend for the most recent memory. At 12 months follow up the worst memory on body image was still better but not the earliest or most recent. No evidence was found on the critical outcomes of remission, bingeing and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. An RCT was identified that compared yoga and treatment as usual with treatment as usual in adults with any eating disorder. At the end the treatment, no difference was found in any of the outcomes including BMI, EDE-total or any of the EDE- sub-scales. Similar findings were found at follow up (three weeks), however there was some improvement in EDE-restraint in the yoga and treatment as usual group compared with treatment as usual. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, relapse, general functioning, family functioning, resource use, and service user experience. A graded body image therapy (and maintenance treatment as usual) was compared with a maintenance treatment as usual in adults with any eating disorder. No difference was found in EDE-weight concerns or EDE-shape concerns at the end of treatment or at follow up. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. An acceptance-based body image mirror exposure therapy was compared with a control therapy and showed an improvement in EDE-eating concerns, EDE-weight concerns, EDE-shape concerns, but not in EDE-restraint. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. A psychomotor therapy and support was compared with support in females with any eating disorder and showed no difference at the end of treatment on self-expression and control anger scales. No evidence was found on the critical outcomes of remission, weight and bingeing, nor on the important outcomes of quality of life, all-cause mortality, eating disorder psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. **Young people with anorexia nervosa (chapter 6)** For young people with anorexia nervosa, bright light treatment and CBT showed benefits on depression compared with any other intervention. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. Video feedback and nutritional counselling compared with nutritional counselling alone showed no additional benefit of the video feedback on BMI. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. Resistance training and treatment as usual showed no difference on BMI and quality of life in young people with anorexia nervosa compared with treatment as usual. At 4 weeks follow up, resistance training and treatment as usual appeared to be less effective on BMI compared with treatment as usual. No evidence was found on the critical outcomes of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, family functioning, resource use, and service user experience. **Adults with anorexia nervosa (chapter 6)** Repetitive transcranial magnetic stimulation versus sham showed no difference in anorexia nervosa symptoms (urge to restrict, feeling full and feeling fat), urge to

binge or side-effects from treatment. However, at one day follow up some benefits were detected on anorexia nervosa symptoms including feeling full and feeling flat compared with sham, but no difference in the symptom of urge to restrict. No evidence was found on the critical outcomes of remission and weight, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Warming therapy on top of refeeding had no effect on change in BMI compared with refeeding alone in adults with anorexia nervosa. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Acupuncture and treatment as usual compared with acupressure, massage and treatment as usual showed acupuncture is more effective on EDE-shape concerns but no other outcome was different between the two groups including EDI-subscales, EDE-subscales, depression, general psychopathology and weight or BMI. No evidence was found on the critical outcome of remission, nor on the important outcomes of all-cause mortality, relapse, general functioning, family functioning, resource use, and service user experience.

Adults with bulimia nervosa (chapter 7)

Repetitive transcranial magnetic stimulation versus sham showed no difference in the effect on bingeing and food cravings within 24 hours of treatment, nor on the urge to eat or the number who withdrew due to adverse events. There was a trend for hunger and the number of those who binged to be reduced but there was some uncertainty. No evidence was found on the critical outcome of remission, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Aerobic exercise appeared to be less effective on EDI-drive for thinness. No difference was found on the number of people who recovered from bulimia nervosa nor who satisfied the EDNOS criteria.
Compared with wait list control, aerobic exercise was less effective on the number who had recovered (unclear definition) from bulimia nervosa but showed no difference on the number who satisfied the criteria for EDNOS. No evidence was found on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Adults with binge eating disorder (chapter 8)

Yoga appears to be effective at reducing scores on the binge eating scale compared with wait list controls. However, this did not translate to a benefit in BMI. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

Aerobic exercise and group CBT-ED appeared to be more effective at reducing BMI compared with group CBT-ED alone in adults with binge eating disorder. No difference was found in depression scores. Similar results were found at follow up. When a maintenance component (12 biweekly meetings over six months) was added to both arms to make this part of the intervention more comparable with the aerobic exercise group (because they continued to meet up), there was a trend for a reduced BMI and depression in the aerobic exercise, group CBT-ED and maintenance group compared with the group CBT-ED and maintenance group at the end of treatment and for the trend in the benefit on BMI to be maintained at follow up but not depression. No evidence was found on the critical outcomes of remission and bingeing, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, relapse, general functioning, family functioning, resource use, and service user experience.

The committee requested investigating the benefits of the Mandometer on eating disorders. A Mandometer is a device that measures how much weight is lost from a

	dinner plate after the person with eating disorder has finished eating. This weight is stored on a computer along with how satiated the person is after eating. The evidence on this is scarce and the sample sizes were too small (less than 10 per group) to meet our inclusion criteria as described in the protocol.
Trade-off between net health benefits and resource use	There was no evidence for the effectiveness of physical interventions in people with eating disorders. As a result, such interventions are likely to be not cost effective.
Quality of evidence	The evidence for physical interventions was mostly very low quality. The evidence was downgraded for imprecision and risk of bias for reasons such as unclear methods of randomisation, it was unclear if allocation concealment was performed, if either or all of the participants, investigators or assessors were blinded. High dropout rates were also detected, with more than 20% dropping out in each arm. Most of the outcomes were the result of a single study with a very low number of participants, only binge eating disorder had more than 100 participants in total. Imprecision was detected in most outcomes because the 95% confidence interval crossed one or two minimal important differences or it did not meet the optimal information size. Also, few studies measured remission and/or compensatory behaviours relevant to that eating disorder. Some outcomes were excluded from the study because it was either unclear over what duration they measured the symptoms or it was less than the two week minimum required by the committee.
Other considerations	The committee agreed that the evidence presented was not strong enough or of sufficient quality to offer a physical intervention to people with an eating disorder. This was mostly because very few studies were identified and few participants were included in most outcomes. However, the committee decided to make a research recommendation on adding exercise to a recommended psychotherapy to determine whether it may add any benefit to those with bulimia nervosa or binge eating disorder. The committee discussed the importance of exploring what the right amount of exercise is, what is the best type of exercise and what the potential harms are. The committee suggested making a research recommendation on the effects of exercise on bulimia nervosa and binge eating disorder, as opposed to any of the other physical interventions for a number of reasons. Exercise may be useful adjunct to psychotherapy to address any co-existing weight or obesity-related issues and mood disorders, such as depression and anxiety. Exercise may also be a cost-effective and drug-free alternative to other therapeutic approaches such as transcranial magnetic stimulation or anti-depressants. See chapter 6.6

9.6 Management of long- and short-term complications

9.6.1 Review Question: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 363. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers all interventions that may be delivered to manage or reduce the short- or long-term physical complications of eating disorders in children, young people and adults and includes recovered as well as current service users. The interventions were categorised according to type of physical complication and intervention, the age of the participants and the type of eating disorder. The control arm varied depending on the study.

Table 363: Clinical review protocol summary for the review of: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

Component	Description
Review question(s)	What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) • Recovered or current service users • Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	• Interventions to address the following: ○ Low bone mineral density (risk of fracture) ○ Growth (physical development) ○ Pubertal development ○ Tooth wear ○ Low body weight ○ Interventions to address the long-term physical complications may include: ○ GH/IGF-I ○ Calcium with and without Vitamin D ○ Bisphosphonates (age dependent and exclude pregnancy) ○ Exercise (low impact)/Physiotherapy ○ Oestrogen (patches/exogenous/pills other) ○ Testosterone (males/females) ○ Weight gain vs. Weight restoration (brain size) ○ Interventions to address the short-term physical complications may include ○ Phosphates supplementation (refeeding) ○ Potassium ○ Thiamine (refeeding) ○ Laxatives (for when underweight patients are constipated) ○ Salbutamol (reduce food intake)
Comparison	• Control arm as defined by study
Critical outcomes	• Primary outcome as reported by study
Important outcomes	• Secondary outcome as reported by study
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs) (if no RCTs)

9.6.2 Clinical Evidence

No studies were identified that met the eligibility criteria for this review.

9.6.3 Economic Evidence

No economic evidence on the cost effectiveness of interventions for managing short and long-term physical complications for people with EDNOS was identified by the systematic

search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.6.4 Clinical evidence statements

No studies were identified that met the eligibility criteria for this review.

9.6.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for managing short and long-term physical complications for people with EDNOS was available.

9.6.6 Recommendations and link to evidence for the review on: What interventions are effective at managing or reducing short and long-term physical complications of eating disorders?

	The committee agreed that people with OFSED who have short or long-term physical complications associated with the eating disorder, should be treated in-line with the eating disorder their symptoms most closely resemble
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of psychotherapies for treating OSFEDs. For this population it was agreed that binge-eating frequency and remission are of greatest concern. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making.
Trade-off between clinical benefits and harms	No relevant clinical evidence was identified.
Trade-off between net health benefits and resource use	No relevant existing economic evidence was identified. The committee expressed the view that if something is potentially cost effective and represents value for money for people with BN, BED or AN it will also do so for people with EDNOS.
Quality of evidence	Not applicable
Other considerations	The committee agreed that people with EDNOS who have short or long-term physical complications associated with the eating disorder, should be treated in-line with the eating disorder their symptoms most closely resemble.

9.7 Management of comorbidities

9.7.1 Review question: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

The review protocol summary, including the review question and the eligibility criteria used in this section of the guideline, can be found in. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers whether any intervention used to treat eating disorders in children, young people and adults needs to be modified in the presence of a common long-term health condition (i.e. comorbidity).The interventions were categorised according to their type, the type of eating disorder and comorbidity examined and the age of the participants. The

comparison arm was the same intervention delivered to participants with the relevant eating disorder but without the relevant comorbidity.

Table 364: Clinical review protocol summary for the review of: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Component	Description
Review question(s)	Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) and a common comorbidity (e.g. diabetes, hypothyroidism). • Mental comorbidities may include: ○ Depression ○ Anxiety ○ Social anxiety ○ Autism ○ Obsessive Compulsive Disorder ○ Personality Disorder ○ Learning disability ○ ADHD (Bulimia) ○ Self-harm ○ Substance misuse ○ Physical comorbidities may include: ○ Celiac disease ○ Diabetes (type II – relevant to obesity) ○ Irritable Bowel Disease ○ Cystic Fibrosis • Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Intervention(s)	Trials will be included that address the ED as primary or secondary aim to treating the comorbidity. Interventions may include: • Psychotherapy (including psychoeducation) • Pharmacological • Nutritional • Physical • Combination of any listed above
Comparison	• The same intervention but delivered to people with an eating disorder without a comorbidity.
Critical outcomes	• Remission and long-term recovery (if symptoms were measured over a minimum 2 week period) • Binge eating for BN and BED • Body weight / BMI for AN
Important outcomes	• Adverse events • All-cause mortality • Cost effectiveness • Eating disorders psychopathology (cognitive distortion/eating behaviours/body image distortion) • Family functioning • General psychopathology (including mood/depression/anxiety)

Component	Description
	• General functioning, measured by return to normal activities, or by general mental health functioning measures such as Global Assessment of Functioning (GAF) • Quality of life • Relapse • Resource use • Service user experience (in patient vs. community)
Study design	• Systematic Reviews • RCTs • Observational studies: including prospective or retrospective cohort studies, (if no RCTs) (if no RCTs)

9.7.2 Clinical Evidence for: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

9.7.2.1 Diabetes

One RCT (n=85) was identified that compared the effectiveness of a parental and patient group psychoeducation programme with treatment as usual for reducing symptoms in young people with type 1 diabetes and disturbed eating attitudes (Olmsted 2002 (Olmsted et al., 2002)).

Two observational studies (n=878) were found that met the inclusion criteria (Colton 2015 (Colton et al., 2015), Custal 2014 (Custal et al., 2014). Both studies used two different populations, one with any eating disorder and type I diabetes and compared them with one that just had any eating disorder. The two groups were compared after either receiving the same treatment (CBT-ED) or different treatments (group CBT-ED with additional care by a multidisciplinary team for those with type I diabetes or just group CBT-ED alone). These comparisons allowed us to see if those with diabetes can respond equally well to treatment as those with just an eating disorder.

9.7.2.2 High alcohol misuse

One observational study (n=149) was found that addressed the comorbidity of alcohol misuse in people with an eating disorder (Karacic 2011 (Karacic et al., 2011)). The study examined the effect of transdiagnostic CBT-enhanced for eating disorders in adults with bulimia nervosa and other eating disorders not otherwise specified (EDNOS).

Although this review question includes people with any eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder, EDNOS), the committee wanted to firstly consider the evidence for individual eating disorders to see if specific recommendations could be made. If none was available, or it was deemed insufficient, then they agreed to make a general recommendation for treating people with any eating disorder and a common long-term health condition.

Table 365: Study information of the RCTs included in the review of interventions for young people with disturbed eating and type I diabetes.

Study	Eating Disorder	Age mean (SD)	BMI	N	Stage of illness	Intervention	Comparison	Duration
Olmsted 2002	Diabetes Type 1 + Disturbed eating attitudes	16 (2)	23.4 (3.5)	85	Age of diabetes onset=9.1 (3.6) years Duration of DM=7 (3.4) years	Group Psychoeducation + treatment as usual Carers and patients	Treatment as usual. Quarterly visits and diabetes management conducted in multidisciplinary treatment setting	6 weeks, assessment at 10 weeks and 6-mo FU

Abbreviations: FU, follow up; DM, diabetes mellitus

Table 366: Study information of the observational studies included in the review that compared outcomes in those with any eating disorder and type I diabetes versus eating disorder alone.

Study	Eating Disorder	Age mean (SD)	BMI	N	Stage of illness	Intervention	Comparison	Duration
Colton 2015 Canada	AN/BN/EDNOS with or without type I diabetes (T1DM)	25.6 (6.4)	21.8 (4.5)	838	Age at diabetes diagnosis: 14.3 (8.2) years	Group CBT-ED + care by multidisciplinary team ED + T1DM	Group CBT-ED ED only	BN/EDNOS 6-8 weeks; AN 10-14 weeks Mean: 6.6 weeks (2.8)
Custal 2014 Spain	AN/BN/EDNOS with or without type I diabetes (T1DM)	25.3 (8)	23.34 (6.35)	40	Age of ED onset=19.5 (7.4) Duration of T1DM=10.3 (8.2) years	CBT-ED ED + T1DM	CBT-ED ED only	3-4 months

Abbreviations: AN, anorexia nervosa; BN, bulimia nervosa; EDNOS eating disorder not otherwise stated; ED, eating disorder; T1DM, type I diabetes mellitus;

Table 367: Study information of the trials included in the review of interventions with people with an eating disorder and a comorbidity of alcohol misuse.

Study ID	N	Mean age	BMI (kg/m2) or Weight (kg)	Female %	Sample characteristics	Group 1	Group 2	Duration
Karacic 2011	149	26.2 (7.1)	23.0 (4.2)	95	DSM-IV BN, or other EDNOS	Eating disorder and high-alcohol use	Eating disorder and low alcohol use	20 weeks + 20/40/60 week FU

*Note: *, High-alcohol use was defined as consuming ≥14 units/week in women and ≥21 units/week in men. Abbreviations: DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; EDNOS eating disorder not otherwise stated; FU, follow up*

Table 368: Summary table of findings for group psychoeducation versus treatment as usual for young people with disturbed eating and type I diabetes.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with treatment as usual for Disturbed eating + Diabetes TI - Young people	Risk difference with RCT: Psychoeducation (95% CI)
EDE Objective Binge Episodes - Group Psychoeducation-ED	85 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede objective binge episodes - group psychoeducation-ed in the intervention groups was 0.13 standard deviations lower (0.56 lower to 0.31 higher)
EDE Restraint - Group Psychoeducation-ED	85 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint - group psychoeducation-ed in the intervention groups was 0.33 standard deviations lower (0.77 lower to 0.1 higher)
EDE Eating Concerns - Group Psychoeducation-ED	85 (1 study)	⨁⨁⨀⨀ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concerns - group psychoeducation-ed in the intervention groups was 0.32 standard deviations lower (0.75 lower to 0.12 higher)

EDE Shape Concerns - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concerns - group psychoeducation-ed in the intervention groups was 0.07 standard deviations lower (0.5 lower to 0.36 higher)
EDE Weight Concerns - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concerns - group psychoeducation-ed in the intervention groups was 0.15 standard deviations lower (0.58 lower to 0.28 higher)
EDI Drive for Thinness - Group Psychoeducation-ED	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness - group psychoeducation-ed in the intervention groups was 0.28 standard deviations lower (0.73 lower to 0.17 higher)
EDI Bulimia - Group Psychoeducation-ED	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia - group psychoeducation-ed in the intervention groups was 0.35 standard deviations lower (0.8 lower to 0.1 higher)
EDI Body Dissatisfaction - Group Psychoeducation-ED	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction - group psychoeducation-ed in the intervention groups was 0.38 standard deviations lower (0.83 lower to 0.07 higher)
Insulin Omission Days - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean insulin omission days - group psychoeducation-ed in the intervention groups was 0.17 standard deviations higher (0.26 lower to 0.6 higher)
HbA1c Level (%) - Group Psychoeducation-ED	82 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean hba1c level (%) - group psychoeducation-ed in the intervention groups was 0 standard deviations higher (0.44 lower to 0.44 higher)
EDE Objective Binge Episodes FU - Group Psychoeducation-	85 (1 study)	⊕⊕⊖⊖ LOW1,2		Not calculable for SMD values	The mean ede objective binge episodes fu - group psychoeducation-ed in the

ED		due to risk of bias, imprecision			intervention groups was 0.34 standard deviations lower (0.78 lower to 0.09 higher)
EDE Restraint FU - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede restraint fu - group psychoeducation-ed in the intervention groups was 0 standard deviations higher (0.43 lower to 0.43 higher)
EDE Overeating FU - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede overeating fu - group psychoeducation-ed in the intervention groups was 0.22 standard deviations lower (0.66 lower to 0.21 higher)
EDE Eating Concerns FU - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede eating concerns fu - group psychoeducation-ed in the intervention groups was 0.25 standard deviations lower (0.69 lower to 0.18 higher)
EDE Shape Concerns FU - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede shape concerns fu - group psychoeducation-ed in the intervention groups was 0.07 standard deviations lower (0.5 lower to 0.36 higher)
EDE Weight Concerns FU - Group Psychoeducation-ED	85 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean ede weight concerns fu - group psychoeducation-ed in the intervention groups was 0.08 standard deviations lower (0.51 lower to 0.36 higher)
EDI Drive for Thinness FU - Group Psychoeducation-ED	81 (1 study)	⊕⊕⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi drive for thinness fu - group psychoeducation-ed in the intervention groups was 0.03 standard deviations lower (0.48 lower to 0.41 higher)
EDI Bulimia FU - Group Psychoeducation-ED	81 (1 study)	⊕⊕⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi bulimia fu - group psychoeducation-ed in the intervention groups was 0.34 standard deviations lower (0.79 lower to 0.11 higher)

EDI Body Dissatisfaction FU - Group Psychoeducation-ED	81 (1 study)	⨁⨁⊖⊖ LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean edi body dissatisfaction fu - group psychoeducation-ed in the intervention groups was 0.13 standard deviations lower (0.58 lower to 0.31 higher)
Insulin Omission Days FU - Group Psychoeducation-ED	85 (1 study)	⨁⨁⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean insulin omission days fu - group psychoeducation-ed in the intervention groups was 0.04 standard deviations higher (0.4 lower to 0.47 higher)
HbA1c Level (%) FU - Group Psychoeducation-ED	82 (1 study)	⨁⨁⊖⊖ LOW1,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean hba1c level (%) fu - group psychoeducation-ed in the intervention groups was 0 standard deviations higher (0.44 lower to 0.44 higher)
*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval;					
1 Unclear if allocation concealment was performed. Neither the participant, investigator nor assessor were blind. Unclear how many completed the intervention. 2 95% CI crossed 1 MID (-0.5) 3 95% CI crossed 1 MID (0.5) 4 For a continuous outcome, there were fewer than 400 participants.					

Table 369: Summary table of findings for outcomes from observational studies in people with any eating disorder and type 1 diabetes versus any eating disorder.

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Any ED only	Risk difference with Any ED+Diabetes TI (95% CI)
Dropouts	40 (1 study)	⨁⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, indirectness, imprecision	RR 1.45 (0.9 to 2.34)	500 per 1000	225 more per 1000 (from 50 fewer to 670 more)
Dropouts - Anorexia Nervosa	4	⨁⊖⊖⊖	RR 1.00	See comment**	-

	(1 study)	VERY LOW1,2,8 due to risk of bias, indirectness, imprecision	(0.49 to 2.05)		
Dropouts - Bulimia Nervosa	10 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,4 due to risk of bias, indirectness, imprecision	RR 1.57 (0.77 to 3.22)	400 per 1000	228 more per 1000 (from 92 fewer to 888 more)
Dropouts - EDNOS	22 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision	RR 1.75 (0.71 to 4.31)	636 per 1000	477 more per 1000 (from 185 fewer to 1000 more)
Dropouts - Binge Eating Disorder	4 (1 study)	⊕⊖⊖⊖ VERY LOW1,2,5 due to risk of bias, indirectness, imprecision	RR 1 (0.14 to 7.1)	500 per 1000	0 fewer per 1000 (from 430 fewer to 1000 more)
Full or Partial Remission	873 (2 studies)	⊕⊖⊖⊖ VERY LOW2,6,7 due to risk of bias, indirectness, imprecision	RR 0.52 (0.33 to 0.81)	469 per 1000	225 fewer per 1000 (from 89 fewer to 314 fewer)
Full or Partial Remission - Anorexia Nervosa	276 (2 studies)	⊕⊖⊖⊖ VERY LOW 2,5,6 due to risk of bias, indirectness, imprecision	RR 0.44 (0.13 to 1.48)	465 per 1000	260 fewer per 1000 (from 404 fewer to 223 more)
Full or Partial Remission - Bulimia Nervosa	293 (2 studies)	⊕⊖⊖⊖ VERY LOW 2,6,7 due to risk of bias, indirectness, imprecision	RR 0.47 (0.23 to 0.97)	730 per 1000	387 fewer per 1000 (from 22 fewer to 562 fewer)
Full or Partial Remission - EDNOS	300 (2 studies)	⊕⊖⊖⊖ VERY LOW2,6,7 due to risk of bias, indirectness, imprecision	RR 0.58 (0.29 to 1.15)	471 per 1000	198 fewer per 1000 (from 335 fewer to 71 more)
Full or Partial Remission - Binge Eating Disorder	4 (1 study)	⊕⊖⊖⊖ VERY LOW2,5,6 due to risk of bias, indirectness, imprecision	RR 1 (0.14 to 7.1)	500 per 1000	0 fewer per 1000 (from 430 fewer to 1000 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

** Absolute effects could not be calculated because zero events were included in the outcome.

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 The authors attempted to match the groups based on age, marital status, education, catchment area, onset of diagnosis. It was unclear whether the two groups were followed up for the same duration. The sample size was very small.
2 They compared two different therapies for two different populations. The patients with an ED and T1DM were treated for both conditions, whilst the comparison group was an ED only group and were treated for just their ED.
3 95% CI crossed 1 MID (1.25).
4 95% CI crossed 1 MID (1.25).
5 95% CI crossed 2 MIDs (0.75 and 1.25).
6 In Custal 2014 the authors attempted to match the groups based on age, marital status, education, catchment area, onset of diagnosis. It was unclear whether the two groups were followed up for the same duration. The sample size was very small. In Cotton 2015, the authors did not attempt to match the groups, nor adjust for potential confounders. The control group data was selected from a different study/data base. It was unclear what the duration of follow up was for both groups. The investigators were not blind to participant's exposure to treatment.
7 95% CI crossed 1 MID (0.75).
8. Fewer than 300 events.

1 Alcohol misuse

2 Table 370: Summary table of findings for CBT-Enhanced for eating disorders in people with an eating disorder and high-alcohol use
3 versus people with an eating disorder and low-alcohol use

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with CBT-E for ED (95% CI)
EDE >1 SD above community norm	119 (1 study) 60 weeks	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, indirectness, imprecision	RR 1.16 (0.68 to 1.97)	321 per 1000	51 more per 1000 (from 103 fewer to 312 more)
Excessive Drinking	119 (1 study) 60 weeks	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, indirectness, imprecision	RR 4.08 (2.08 to 8.01)	119 per 1000	367 more per 1000 (from 129 more to 835 more)
EDE Global 60 week FU	104 (1 study)	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, indirectness, imprecision		Not calculable for SMD	The mean ede global 60 week fu in the intervention groups was 0.23 standard deviations lower (0.66 lower to 0.2 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with CBT-E for ED (95% CI)
				values	

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Karacic 2011: attrition bias (dropout for low alcohol group >20 %); sample did not have current alcohol use disorder comorbidity; group allocated on basis of self-reported alcohol use. Sample consisted of 67 BN, 10 BED and 72 EDNOS. Participants with anorexia nervosa were excluded.
2 CI crosses both 0.75 and 1.25 (Risk Ratio) or 0.5 and -0.5 (SMD).
3 <300 events.
4 CI crosses 0.5 or -0.5 (SMD).

1

2

9.7.3 Economic Evidence

- No economic evidence on the cost effectiveness of interventions for the management of comorbidities of EDNOS was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

9.7.4 Clinical evidence statements

Diabetes

9.7.4.1 Group psychoeducation versus treatment as usual in people with type I diabetes and disturbed eating disorders at end of treatment

Low quality evidence from one RCT (n=85) showed no difference in the effect of group psychoeducation on binges, EDE-restraint, EDE-shape concern, EDE-eating concern, EDE-weight concern, EDI-drive for thinness, EDI-bulimia, insulin omission days and HbA1c (%) compared with treatment as usual.

Low quality evidence from one RCT (n=85) showed group psychoeducation is more effective on EDI-body dissatisfaction compared with treatment as usual but there was some uncertainty.

9.7.4.2 Group psychoeducation versus treatment as in people with type I diabetes and usual disturbed eating disorders at follow up

Low quality evidence from one RCT (n=85) showed no difference in the effect of group psychoeducation on EDE-restraint, EDE-shape concern, EDE-eating concern, EDE-weight concern, EDI-drive for thinness, EDI-bulimia, EDI-body dissatisfaction, insulin omission days and HbA1c (%) compared with treatment as usual.

Low quality evidence from one RCT (n=85) showed group psychoeducation is more beneficial on frequency of binges compared with treatment as usual but there was some uncertainty.

9.7.4.3 Any eating disorder and type I diabetes versus any eating disorder at end of treatment. Observational study.

Very low quality evidence from one observational study (n=40) showed no difference in the number of dropouts in those with any eating disorder and type I diabetes compared with an eating disorder alone. This trend was apparent in all types of eating disorders: anorexia nervosa, bulimia nervosa, binge eating disorder and EDNOS.

Very low quality evidence from two observational studies (n=873) showed lower rates of remission in those with any eating disorder and type I diabetes compared with an eating disorder alone. This trend was apparent in those with bulimia nervosa, but no difference was found in those with anorexia nervosa, binge eating disorder and EDNOS

Alcohol misuse

9.7.4.4 CBT-ED for people with an eating disorder and high or low alcohol misuse at end of treatment

Very low quality evidence from one observational study (n=119) showed no difference in the effect of CBT-ED on the number of people who were more than 1 standard deviation above

EDE community norms in people with an eating disorder whose alcohol use was high compared with those whose alcohol use was low.

Very low quality evidence from one observational study (n=119) showed that CBT-ED is less effective on the number of people who were engaging in excessive drinking in people with an eating disorder whose alcohol use was high compared with people with those whose alcohol use was low.

9.7.4.5 CBT-ED for people with an eating disorder and high or low alcohol misuse at follow up

Very low quality evidence from one observational study (n=104) showed no difference in the effect of CBT-ED for eating disorders on EDE-global in people with an eating disorder whose alcohol use was high compared with people with an eating disorder whose alcohol use was low.

9.7.5 Economic Evidence statements

No economic evidence on the cost effectiveness of interventions for the management of comorbidities of EDNOS was available.

9.7.6 Recommendations and link to evidence for the review on: Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?

Diabetes

	151. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder. **152. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.** **Diabetes** **153. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.** **154. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.** **155. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes.** **156. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.** **157. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.**

	158. Offer people with an eating disorder who are misusing insulin the following treatment plan: • a low carbohydrate diet, so that insulin can be started at a low level • gradually increasing insulin doses to reduce blood glucose levels • adjusted total glycaemic load and carbohydrate distribution to meet their individual needs and prevent rapid weight gain • carbohydrate counting when adjusting their insulin dose (including via pumps) • a diabetic educational intervention such as DAFNE • education about the problems caused by misuse of diabetes medication. **159. For more guidance on managing diabetes, refer to the NICE guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem. In the case of diabetes, HbA1c levels and insulin omission days were considered critical outcomes. The other critical outcomes depended on the eating disorder included in the study. Remission is of greatest concern for any eating disorder. In addition, for those with anorexia nervosa body weight or BMI are of greatest concern. For bulimia nervosa and binge eating disorder, binge eating is a critical outcome. For any eating disorder, other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with an eating disorder that are of lesser importance, but are clearly still important outcomes, include general psychopathology, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	The ideal study design to answer the question of whether a treatment for eating disorders needs to be modified in the presence of a long-term health problem would be to randomise people with an eating disorder and diabetes to two different treatment groups: one modified to address both the eating disorder and diabetes and one non-modified eating disorder treatment. **Any eating disorder (as reviewed in this chapter)** One randomised control trial compared group psychoeducation (combined with treatment as usual) with treatment as usual (diabetes treatment only) in people with type I diabetes and disturbed eating behaviours and showed no difference at the end of treatment on bingeing, EDE-restraint, EDE-shape concern, EDE-eating concern, EDE-weight concern, EDI-drive for thinness, EDI-bulimia, insulin omission days and HbA1c (%). One outcome, EDI-body dissatisfaction, favoured group psychoeducation over treatment as usual but there was some uncertainty. At follow up some benefit was found in response to group psychoeducation on bingeing but there was some uncertainty. No data was available on remission, all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience. An observational study was identified that compared the same CBT-ED intervention but in two populations, one with any eating disorder and type I diabetes, and one with just any eating disorder. Thus, this design allowed us to see

	whether those with a comorbidity would respond equally well to treatment as those with just an eating disorder. The results showed adults with any eating disorder and a comorbidity are less likely to recover than those with just an eating disorder. No difference was found in dropouts. No data was available on all-cause mortality, resource use, relapse, general functioning, family functioning, or service user experience **Anorexia nervosa (as reviewed in chapter 6)** No published evidence was found on people with anorexia nervosa and diabetes, however there was a sub-group analysis from a study described below on any eating disorder that showed those with anorexia nervosa and type I diabetes are equally responsive to treatment as those with anorexia nervosa alone. No data was available on HbA1c levels, remission, all-cause mortality, adverse events, quality of life, resource use, relapse, general functioning, family functioning and service user experience. **Bulimia nervosa (as reviewed in chapter 7)** One observational study compared the effectiveness of inpatient integrated care with treatment as usual in adults with bulimia nervosa and type I diabetes. The integrated care provided CBT-ED, family based therapy and addressed control of diabetes. Whilst treatment as usual included outpatient counselling sessions on diabetes but not inpatient care or treatment for the eating disorder. This study showed better outcomes for the integrated care including, remission, general psychopathology, depression, EDI-total, the size of the binges, few compensatory behaviours but no difference in insulin omission. No data was available on HbA1c levels, all-cause mortality, adverse events, quality of life, resource use, relapse, general functioning, family functioning and service user experience. **Binge eating disorder (as reviewed in chapter 8)** One study randomised adults with type II diabetes and binge eating to either group CBT-ED or a non-prescriptive control therapy (NPT). The results showed no difference in remission or binge frequency at the end of treatment. BMI showed a trend to be higher in the group CBT-ED arm, however EDI-bulimia, EDI-drive for thinness, EDI-body dissatisfaction and quality of life were no different. At follow up, remission rates were higher in the CBT-ED arm, but again no difference in any of the other outcomes and BMI showed a trend to be higher in the group CBT-ED arm compared with controls. No data was available on HbA1c levels, insulin omission, all-cause mortality, adverse events, resource use, relapse, general functioning, family functioning and service user experience. An observational study compared the same diabetes prevention programme but in two populations, one with bulimia nervosa and a major depressive disorder and one with just any eating disorder. The results showed no difference in the degree of weight loss between the two populations. No data was available on HbA1c levels, insulin omission, remission, bingeing, all-cause mortality, adverse events, resource use, relapse, general functioning, eating disorder psychopathology, family functioning and service user experience.
Trade-off between net health benefits and resource use	The committee considered that providing care for eating disorders in the presence of a long-term health problems, such as diabetes, may have resource implications in terms of extra time required to provide collaborative and comprehensive care in line with the principles outlined in the recommendations 141-147. However, the committee expressed the view that if such care arrangements (that is, multidisciplinary approach, involvement of family members and carers, and the use of treatment plans) lead to better and appropriate treatment and management of health problems (including other long-term health problems such as diabetes) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating such care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence was mostly low quality from the RCT studies and very low quality from all of the observational studies. In both types of study designs the sample size was generally small and only one study was available for most outcomes, thus imprecision was often detected due to the 95% confidence interval crossing a minimal important difference or the outcome did not meet the optimal information

	size. **Any eating disorder (including subgroup analysis on anorexia nervosa)** In the RCT where they compared group psychoeducation and management (and treatment as usual) with treatment as usual (diabetes only programme), it was unclear if allocation concealment was performed. Neither the participant, investigator or assessor were blind and it was unclear how many completed the intervention. The population was also indirect since it included those with disturbed eating. Also the comparison did not show whether a modified eating disorder programme is more effective at treating people with diabetes and an eating disorder compared with an eating disorder programme alone. Rather the study compared a modified diabetes programme with a regular diabetes programme. In the observational study where they compared CBT-ED in people with eating disorder alone or with a comorbidity, the authors attempted to match the groups based on age, marital status, education, catchment area and onset of diagnosis. However, it was unclear whether the two groups were followed up for the same duration and the sample size was very small. **Bulimia nervosa** In the observational study where they compared inpatient integrated care with treatment as usual, the people were selected from the same recruitment site and showed no difference in their characteristics, except that binge frequency was significantly higher in the inpatient group. The duration of follow up was different for the two groups: 36 months versus 24 months in the inpatient care and treatment as usual groups, respectively. Investigators were not blind to treatment allocation and only 18 participants were included. **Binge eating disorder** In the RCT where they compared group CBT-ED with a control therapy in the same population (people with type I diabetes and binge eating disorder) inadequate randomisation was performed and it was unclear if allocation concealment was carried out. Neither the participant or investigator was blind, nor was it clear if the assessor was blind. It was unclear how many participants completed the intervention. The observational study identified was considered indirect evidence since it was a diabetes prevention programme and the participants had major depressive disorder in addition to binge eating disorder or binge eating disorder alone. The only outcome reported was weight loss. The committee did not consider this study helpful. **Overall discussion** No RCT or observational study met the criteria of what would have been the ideal study design for this review (as described above). One RCT compared the effectiveness of an intervention that addressed both the eating disorder and diabetes, but the other arm addressed just the diabetes. In another RCT, one intervention was modified but it was compared with a control therapy. In the observational studies, one study compared the same intervention but in those with either an eating disorder and diabetes or just the eating disorder alone. Therefore, it only provided insight into whether one group was more responsive to treatment than the other. In the other observational study, inpatient integrated care was compared with treatment as usual, but the treatment as usual only addressed the diabetes not the eating disorder. Thus, it did not provide insight into whether a modified eating disorder treatment was needed for those with a comorbidity.
Other considerations	In summary, it was difficult for the committee to draw conclusions from these studies on whether treatment for an eating disorder needs to be modified in the presence of a comorbidity such as diabetes. The committee therefore agreed that it was best to instead provide guidance on how to manage the diabetes. Usually, the committee would refer to the diabetes NICE guideline, but because the diabetes guideline refers to this guideline, the committee needed to recommend what to do in the presence of both. The committee agreed on a series of recommendations based on their experience and knowledge on how to manage the diabetes in the presence of an eating disorder. A number of the recommendations are based on what would be

considered good practice. For instance: i) establish who will monitor the physical health, ii) explain to the person that they need to monitor their diabetes during the treatment for the eating disorder, and iii) be aware of the problems caused by misuse of diabetes medication.

The committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in the treatment of diabetes. They highlighted that the quality of the family environment has been shown to affect treatment compliance and metabolic control among young people with an eating disorder (Hauser 1990). Family members may also need to care for someone if they hyper or phyo (which is a case for medical emergency), so they know how to respond. There is also the possibility that eating disturbances in young girls with diabetes are associated with significantly more family dysfunction than girls with diabetes alone (i.e. 13 to 18 years of age). Specifically, they can receive less support, and have poorer communication and less trust in their relationship with their parents than diabetic girls without eating disturbances. For these reasons, the committee agreed that healthcare professionals should consider involving family members and carers (as appropriate) in their treatment.

There was some indirect evidence to support the recommendation to address insulin misuse as part of any psychological treatment. One 1 RCT (n=85) showed that a modified group psychoeducation and management programme reduced bingeing episodes at follow up compared with a programme that just addressed the diabetes alone. This study was considered with the reservation that it was indirect because: 1) it did not investigate the effectiveness of a modified eating disorder psychological treatment and 2) the population had a disturbed eating behaviour, not a specific eating disorder diagnosis. Nevertheless, it showed that a psychoeducation and management programme may help reduce eating disorder psychopathology in those who also have diabetes.

The committee discussed the problem of a relatively high prevalence of EDNOS In young girls with diabetes. In girls who have body dissatisfaction, diabetes provides a unique but dangerous opportunity to control weight by deliberate insulin omission, which can lead to hyperglycaemia and glycosuria. It is therefore important that insulin misuse is addressed in any psychological intervention.

The recommendations relating to diet control were contributed to by the expert opinion of a dietician on the committee, based on their experience of treating those with an eating disorder who misuse insulin. These recommendations are based upon the treatment approach of small, attainable and incremental goals. At the outset of treatment, intensive glucose management is not an appropriate goal. The first goal must be to establish medical safety for the person with diabetes by gradually increasing the doses of insulin and food intake (as described in the recommendation). Given the fear of weight gain in this population, the committee recommended that the diet is amended to prevent rapid weight gain. They also suggested an educational programme called Dose Adjustment for Normal Eating (DAFNE) that provides people with the skills necessary to estimate the carbohydrate in each meal and to inject the right dose of insulin.

There was no evidence on how to treat the eating disorder in the presence of any other long-term physical health condition, such as cystic fibrosis, celiac disease, pregnancy or irritable bowel disease.

Some eating disorder specialists on the committee highlighted that they would generally refer someone with an eating disorder and diabetes to a diabetologist rather than address the points raised in the recommendations on diabetes. However, the committee agreed that it should be collaborative approach for the health care professionals who treat eating disorders and diabetes. Especially for young people who may need to involve family members and carers in therapy sessions to help the person with blood glucose control.

Given the lack of direct evidence to address this review question the committee agreed to make a research recommendation to ask: “Does any intervention for an eating disorder need to be modified in the presence of common long-term health conditions?” See chapter 6.8

1 **Substance and medication misuse**

	160. For people with an eating disorder who are misusing substances, or over the counter or prescribed medication, provide treatment for the eating disorder unless the substance misuse is interfering with this treatment. **161. If substance misuse or medication is interfering with treatment, consider a multi-disciplinary approach with substance misuse services.**
Relative value of different outcomes	The committee discussed the importance and relevance of various outcomes when assessing the effectiveness of treating people with an eating disorder and a comorbidity. For binge eating disorder and bulimia nervosa, it was agreed binge eating frequency and remission are of greatest concern. For anorexia nervosa, body weight/BMI and remission are critical and for ENDOS, remission and either binge eating or body weight/BMI depending on the eating disorder they most closely resemble. Other critical outcomes will include the primary outcomes relevant to the physical or mental health comorbidity being treated. Other outcomes that are important but are considered rare events or rarely measured in RCTs for eating disorders include all-cause mortality, adverse events, quality of life, resource use and relapse. They were therefore extracted where possible, but did not factor strongly in the decision-making. Other outcomes of concern for people with binge eating disorder that are of lesser importance but clearly important outcomes include, general psychopathology, body weight, general functioning, family functioning and service user experience.
Trade-off between clinical benefits and harms	**Bulimia nervosa and EDNOS** An observational study was identified were they extracted data from a randomised control trial and compared the outcomes in those with bulimia nervosa and EDNOS who had a low or high alcohol intake. The participants were treated with either broad or focused CBT-ED. At the end of 20 weeks of treatment, there was no difference in the number who had EDE scores one standard deviation above the community norms (i.e., relatively abnormal eating psychopathology) in those with a low or high alcohol intake. However, the number who continued to have excessive alcohol intake (defined as >21 units or >14 units/week for males and females respectively) was higher in those whose alcohol intake was high compared with those whose intake was low. At 60 weeks follow up, there continued to be no difference in EDE scores between those who had low versus high alcohol intake. No evidence was found on the critical outcomes of remission and binge eating, nor on the important outcomes of quality of life, all-cause mortality, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. **Bulimia nervosa (data presented in chapter 7)** An observational study compared the long-term outcomes (two to five years) of women with bulimia nervosa who had a history of substance abuse with those who no history of substance abuse. Both groups had received outpatient group cognitive behavioural psychotherapy and showed no different in long-term remission rates or being hospitalised for substance abuse. No evidence was found on the critical outcome binge eating, nor on the important outcomes of quality of life, all-cause mortality, eating disorders psychopathology, general psychopathology, relapse, general functioning, family functioning, resource use, and service user experience. No randomised control trial evidence was found. No relevant published evidence was found in those with anorexia nervosa or binge eating disorder.
Trade-off between net health	The committee considered that providing care for people with an eating disorder who are misusing substances or medication may have resource implications in terms of the extra time required to facilitate care for such people (in particular the

benefits and resource use	use of a multi-disciplinary approach). However, the committee expressed the view that if such care leads to better identification of health needs and this results in appropriate subsequent treatment and management of health problems (including eating disorder and substance and medication misuse) at an earlier stage, before individuals require more resource intensive management, then the additional costs associated with facilitating a multi-disciplinary care is expected to result in improved health outcomes in the longer term and potential future cost savings to the healthcare system.
Quality of evidence	The evidence used to generate these recommendations was very low quality. The evidence was observational therefore in the GRADE software used to assess quality, the evidence starts at very low quality and can only be upgraded if large effect sizes are found or if a dose response is identified. Neither was the case for this review. In the absence of RCT evidence the committee still considered this evidence useful. In one study a reasonable number of participants were included (n=119) and they underwent a currently recommended CBT-ED programme. However, there were few outcomes reported and no remission data. In the other study, again there was a reasonable number of participants included (n=81), but there was no data at the end of treatment (only at follow up) and again few outcomes were reported. They did however measure remission.
Other considerations	Limited published evidence was found on individual eating disorders, so the committee generated a recommendation incorporating the evidence from people with BN and EDNOS and made it relevant for treating people with any eating disorder and a substance misuse problem. The observational evidence suggested that people with BN or EDNOS, with a low or high alcohol intake, may be equally responsive to an eating disorder treatment. And for people with BN alone, a positive long-term response to treatment may be equally found in those with a history of substance misuse as those with no history. Thus, the committee recommended that for people with an eating disorder who are misusing substances, offer treatment for the eating disorder unless the substance misuse is demonstrably interfering with this treatment. The committee discussed the scenario of when a clinician is faced with someone who has an eating disorder and a drinking problem, they need to consider two important questions: firstly, do patients with an eating disorder and concurrent high alcohol consumption do less well in treatment? And secondly, do they require treatment that is different and modified to focus upon both drinking and eating problems? The evidence presented in this review answered both question with no; people with a high or low alcohol intake or a history or no history of substance have almost identical responses to CBT-ED, therefore an amended programme may not be needed. The observational evidence also suggested that treatment for an eating disorder may also reduce alcohol intake. In the study by Karacic 2011 over half the high alcohol intake group were no longer drinking excessively (52.8%, n=19) at the end of treatment, however, 12.5% (n=10) of the low alcohol intake group were now drinking above the safe limit (this data was not extracted because change scores were not presented). Another important finding was that mean intake for the high alcohol intake group changed from a risky drinking pattern to one closer to recommended guidelines (again this data could not be extracted because no error estimates around the mean were provided). These changes happened despite the programme not specifically addressing alcohol consumption. For these reasons, the committee were confident recommending the person undergoes treatment for the eating disorder unless the substance misuse is interfering with the treatment. In such cases, a multidisciplinary approach may be needed Although the evidence for this recommendation was found in those with bulimia nervosa and ENDOS, the committee were confident that the findings would translate to those with any eating disorder. For this reason, they did not specify the type of eating disorder. It was discussed in the committee meeting that comorbid alcoholism has been associated with an increased risk of mortality in people with an eating disorder.

10 Coordinating care and compulsory treatment

10.1 Introduction

Anorexia nervosa is a mental disorder and therefore the question of whether the Mental Health Acts (1983, substantially amended 2007) can be used has been answered in the affirmative. The provision of nutrition, including by artificial means, is considered to be treatment for the mental disorder, along with other components. The Care Quality Commission (CQC) has issued guidance, which says, 'some patients with anorexia nervosa – who might have the intellectual capacity to understand the nature, purpose and likely effect of treatment – may be unable to give valid consent, perhaps because their capacity to consent is compromised by fears of obesity or by denial of the consequences of their actions.' They also say, 'Consideration of whether the treatment environment constitutes a deprivation of liberty might be an additional reason for considering compulsory treatment under the MHA may be required.'

Compulsory treatment in adults

The CQC's 'Guidance on the treatment of anorexia nervosa under the Mental Health Act 1983' (2008) clarifies that:

- the Mental Health Act (1983) can be used for patients with anorexia nervosa, because it is a mental disorder
- 'there are likely to be particularly strong reasons for any application … to be made by the Approved Mental Health Professional (AMHP) rather than by the Nearest Relative'
- medical treatment may include feeding by 'nasogastric tube or other means'
- some patients may have the intellectual capacity to understand the treatment, but 'be unable to give valid consent to it … because their capacity to consent is compromised by fears of obesity'
- diagnostic and monitoring procedures, including blood tests, may be necessary as part of the medical treatment for the eating disorder
- 'consideration of whether the treatment environment constitutes a deprivation of liberty might be an additional reason for considering compulsory treatment' under the Mental Health Act (1983).

There appears to be an anomaly in that feeding by nasogastric tube or other 'artificial means' is considered to be 'medical treatment', but not a medical treatment for which consent or a second opinion is required after the first 3 months of compulsory treatment.

The Mental Capacity Act 2005 Code of Practice also specifically mentions that people with anorexia nervosa may have impaired decisional capacity despite having a good understanding of risk. It is possible to treat people under the Mental Capacity Act but, due to the complexities of assessing capacity and the extent of any deprivation of liberty, the Mental Health Act is the preferred instrument. The CQC guidance has not been updated since the introduction of the Deprivation of Liberty Safeguards in 2009. However, the Court of Protection recognised and upheld what it called the 'primacy' of the Mental Health Act (1983) in situations where patients met the criteria for detention under its powers and were objecting to treatment or admission. The Deprivation of Liberty Safeguards was intended to fill a gap left between the Mental Health Act (1983) and common law, not to provide an alternative to detention under that Act.

Compulsory treatment for atypical presentations of anorexia nervosa would be approached in the same way. Consideration of the use of the Mental Health Act (1983) for treatment of

bulimia nervosa is rare, reserved for cases where risks to health or life are severe or there is significant psychiatric comorbidity that may constitute the grounds for detention.

Compulsory treatment will only be considered when less restrictive options, including informal inpatient treatment, have been unsuccessful or are not appropriate. It is therefore not surprising to find that detained patients have higher rates of comorbidity, take longer to gain weight and have a subsequent higher mortality rate compared with non-detained patients (Ramsay et al., 1999).

Studies have compared the outcomes for patients treated with and without their consent, but the salient issues for patients, professionals and families relate to when compulsory powers can be used, what can be done under them and how to decide whether the patient has capacity to make particular treatment decisions.

Children and Young People

The CQC Guidance also applies to children and young people but issues of capacity and consent are potentially more complex in that, below the age of 16, there is no assumption of capacity, and even for those over 16 up to age 18, there are caveats to the assumption of capacity. For example, advance directives about care cannot be given by those age 16-18, and the safeguards provided by the Children Act 1989 apply up to the age of 18.

For young people under age 16 the onus is on demonstrating competence to make decisions ("Gillick competence"). Typically these would be decisions involving assent (agreement) rather than refusal. If a young person refuses a recommended treatment, the concept of the 'zone of parental control' is evoked. Sometimes criticised as a vague concept, the 'zone of parental control' describes the changing level of parental authority appropriate to a child's development stage. For example, it would be normal for the parents of a child or younger young people to decide what he or she ate, or whether they needed to see a doctor, but decisions such as the need for restraint or feeding against consent lie outside the normal range of parental decisions.

The Department of Health provides the following guidance on consent in minors:

"Where a child under the age of 16 lacks capacity to consent (i.e. is not Gillick competent), consent can be given on their behalf by any one person with parental responsibility (if the matter is within the 'zone of parental control') or by the court. As is the case where patients are giving consent for themselves, those giving consent on behalf of child patients must have the capacity to consent to the intervention in question, be acting voluntarily and be appropriately informed. The power to consent must be exercised according to the 'welfare principle': that the child's 'welfare' or 'best interests' must be paramount. Even where a child lacks capacity to consent on their own behalf, it is good practice to involve the child as much as possible in the decision-making process."

In recent years there has been a preference in young people for use of the MHA over parental consent or the Children Act, although all three are potentially applicable in young people who do not consent to treatment.

Although a young person may not have given consent, the term compulsory treatment is only utilised when treatment is given under the legal framework of the Mental Health Act or Children Act. The limited evidence base indicates that the outcome of compulsory treatment may be different in young people compared to adult patients. Ayton et al (2009) found that detained patients had more severe comorbidity and greater risk on admission but did no worse than informal patients on follow up.

10.2 Compulsory treatment

10.2.1 Review Question: What factors/indicators should be considered when assessing whether a person with an eating disorder should be admitted for compulsory treatment (including any form of restrictive interventions usually implemented in refeeding)?

The review protocol summary, including the review question and the eligibility criteria used for this section of the guideline, can be found in Table 371. Further information about the search strategy can be found in Appendix H; the full review protocols can be found in Appendix F.

This review considers what factor or indicators should be considered when assessing whether children, young people or adults with an eating disorder should be admitted for compulsory treatment. The studies were categorised according to the age of the participants and whether they compared compulsory treatment under the legal auspices of the relevant country with voluntary treatment or whether they conducted a regression analysis

Table 371: Clinical review protocol summary for the review of: What factors/indicators should be considered when assessing whether a person with an eating disorder should be admitted for compulsory treatment (including any form of restrictive interventions usually implemented in refeeding)?

Component	Description
Review question(s)	What factors/indicators should be considered when assessing whether a person with an eating disorder should be admitted for compulsory treatment (including any form of restrictive interventions usually implemented in refeeding)?
Population	• Children, young people and adults with eating disorders (anorexia nervosa, bulimia nervosa, binge eating, atypical eating disorder) who need to be admitted for compulsory treatment Strata: ○ children (≤12), young people (13-≤17 years), adults ≥18 years ○ eating disorder (i. anorexia nervosa, ii. bulimia nervosa, iii. binge eating, iv. typical eating disorder)
Factors/indicators	The following factors may be considered when admitting for compulsory treatment: • body weight • consent • family functioning • general functioning or general mental health functioning measures such as Global Assessment of Functioning (GAF). • other medical indicators (i.e. low potassium) • MARSIPAN check list
Comparison	• Relevant comparison as reported by the papers
Critical outcomes	• Primary outcomes as reported by the authors (may include ANOVA, or multiple regression analysis showing what factors are associated with a higher likelihood of compulsory treatment)
Important outcomes	• Secondary outcomes as reported by the papers
Study design	• Individual patient data meta-analysis • Systematic reviews • Observational non-RCT studies (prospective, retrospective or cross-sectional studies) • RCTs will be included if they provided a multiple regression analysis

Component	Description
	looking at predictors of any relevant outcomes • In the absence of any direct evidence, observational or RCTs studies will be included if they provide insight into the relative success of compulsory versus voluntary inpatient care or which factors are associated with a higher likelihood of being admitted.

10.2.2 Clinical Evidence for: What factors/indicators should be considered when assessing whether a person with an eating disorder should be admitted for compulsory treatment (including any form of restrictive interventions usually implemented in refeeding)?

Overall 6 observational studies (N = 1144) provided indirect evidence for this review question, all of which were on young people (Ayton 2009 (Ayton et al., 2009), Carney 2004 (Carney et al., 2004)/Carney 2006 (Carney et al., 2006), Carney 2008 (Carney et al., 2008), Griffiths 1987, Ramsay 1999 (Ramsay et al., 1999), Ward 2015 (Ward et al., 2015), Watson 2000 (Watson et al., 2000). Two studies (n=441) were relevant to those with any eating disorder (Ayton 2009, Watson 2000).

Three studies (n=325) compared compulsory treatment with voluntary treatment in people with anorexia nervosa, all of which were on adults (Carney 2004/2006/2008, Griffiths 1997, Ramsay 1999/Ward 2015) and two were conducted in people with any eating disorder, one of which was in young people (Ayton 2009) and one in adults (Watson 2000). The results from these studies showed the relative success of inpatient treatment if the person was admitted under compulsory conditions versus voluntarily. Although the results from these studies did not provide indicators or what factors should be considered when assessing whether a person with an eating disorder should be admitted for compulsory treatment, the results were extracted and presented to the committee in case they helped them decide whether compulsory treatment is a viable treatment option compared with patients who are treated voluntarily.

Two studies (n=279) conducted a regression analysis of the predictors or factors associated with compulsory treatment in people with anorexia nervosa, both of which were on adults (Carney 2008, Schreyer 2015 (Schreyer et al., 2015)). Another study conducted by Vandereycken 2009 (Vandereycken and Vansteenkiste, 2009) (n=174) conducted a regression analysis to assess the impact of compulsory treatment on the likelihood of patients dropping out of treatment.

None of the 3 studies that conducted a regression analysis directly answered the review question. They investigated whether: a patient was more likely to drop out from inpatient care; the factors associated with patients who undergo compulsory treatment (not necessarily what factors they should be admitted for); and the impact of compulsory treatment on outcomes from hospital. Nevertheless, they were presented in the hope they may provide some insight or context for developing a recommendation.

An overview of the trials included in the narrative synthesis can be found in Table 372, Table 373 and Table 374. Further information about both included and excluded studies can be found in Appendix J.

Summary of findings for those on anorexia nervosa can be found in Table 375, Table 376, Table 377 and Table 378. See also the study selection flow chart in Appendix K, sensitivity and specificity forest plots in Appendix M, study evidence tables in Appendix L and exclusion list in Appendix J.

Table 372: Study information for studies included in the analysis of compulsory treatment versus voluntary treatment in people with any eating disorder

Study ID	Number	Mean Age (SD)	Female (%)	Mean BMI (SD), kg/m2	Compulsory treatment Country, legal statute	Mean length of hospitalisation (days)	Comparison	Mean length of hospitalisation (days)
Ayton 2009	50	16.2 (1.3)	94	15.0 (2.4)	UK, Section 3, MHA 1983/2007 Age at ED onset: 12.5 (1.9) years Duration of ED: 3.8 (2.1) years	243	Parental consent Age at ED onset: 14.3 (1.8) years Duration of ED: 1.9 (1.5) years	423
Watson 2000	391	24.8 (8.6)	88	17.4 (4.7)	USA, legally committed for involuntary treatment, Iowa Duration of ED: 1.6 (1.6) years	41 (36)	Voluntary treatment Duration of ED: 1.9 (1.5)	58 (47)

Abbreviations: ED, eating disorder; MHA, Mental Health Act.

Table 373: Study information for studies included in analysis of compulsory treatment versus voluntary treatment in adults with anorexia nervosa

Study ID	Number	Female (%)	Mean BMI (SD), kg/m2	Compulsory Treatment Country, legal statute	Mean length of hospitalisation (days)	Comparison	Mean length of hospitalisation (days)	Follow up
Carney 2004/2006/ 2008	75*	96	13.2 (1.7)	Australia, Guardianship Order, Guardianship Act 1987 Duration of AN: 8.1 (7.6) years	52 (47)	Informal admission Duration of AN: 6.3 (6.5) years	47 (53)	na
Griffiths 1997	88	Not reported	14.1 (2.2)	Australia, Guardianship Order, Guardianship Act 1987	105 (76)	Voluntary treatment	62 (42)	1 year
Ramsay 1999/Ward 2015	162	96	14.3 (2.5)	UK, Sections 2, 3, 4 or 5, MHA 1983/2007 Duration of AN: 8.2 (6.1)	113 (90)	Voluntary admission Duration of AN: 7.6 (6.4)	88 (53)	~5 and ~20 years

*Notes: *, data is for 96 admissions. Abbreviations: AN, anorexia nervosa; MHA, Mental Health Act; na, not applicable.*

Table 374: Study information for studies that explore factors associated with compulsory treatment

Study ID	Number of participants	Female (%)	Mean BMI (SD), kg/m2	Groups	Type of regression: outcome investigated	Predictors of outcomes
Carney 2008	75*	96	13.8 (1.8)	Compulsory treatment Duration of AN: 8.1 (7.6) years vs. Voluntary treatment Duration of AN: 6.3 (6.5)	Multiple logistic regression: • likelihood of compulsory treatment	• age • aged 20-29 years • previous admissions • type of ED • psychiatric comorbidity • admission BMI • refeeding syndrome • length of treatment • tube feeding • locked ward
Schreyer 2015	204	95	16.3 (2.1)	Inpatients in specialty ED program completed questionnaire about perceived coercion	• Linear multiple regressions: • inpatient length of stay • discharge BMI • Binary logistic regression: • transition to partial hospital vs. early drop out • achieved target weight	• perceived coercion • extraversion • admission BMI • binary logistic regression: • perceived coercion • EDI-2 Drive for Thinness • admission BMI
Vandereycken 2009	174	100	Not reported	Treatment in ED unit within psychiatric hospital.	• Logistic regression • dropout from treatment	• compulsory treatment • voluntary treatment • duration of illness • inpatient treatment • outpatient treatment

Table 375: Summary table of findings for compulsory treatment versus voluntary treatment in young people or adults with any eating disorder at discharge from hospital

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with Voluntary Treatment	Risk difference with Any ED: Compulsory Treatment (95% CI)
BMI at discharge - young people	47 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI at discharge - young people in the intervention groups was 0.69 standard deviations higher (0.06 to 1.32 higher)
BMI at discharge - adults	397 (1 study)	⊕⊖⊖⊖ VERY LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean bmi at discharge - adults in the intervention groups was 0.05 standard deviations lower (0.32 lower to 0.21 higher)
Weight Gain (lbs) - adults	397 (1 study)	⊕⊖⊖⊖ VERY LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight gain (lbs) - adults in the intervention groups was 0.33 standard deviations higher (0.07 to 0.6 higher)
Rate of Weight Gain (lbs/week) - adults	397 (1 study)	⊕⊖⊖⊖ VERY LOW3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean rate of weight gain (lbs/week) - adults in the intervention groups was 0.18 standard deviations higher (0.09 lower to 0.44 higher)
# achieving >85% ABW or BMI>18 - adults	397 (1 study)	⊕⊖⊖⊖ VERY LOW3,4 due to risk of bias, imprecision	RR 0.98 (0.85 to 1.12)	807 per 1000	16 fewer per 1000 (from 121 fewer to 97 more)
# AN patients achieving >85% ABW - adults	178 (1 study)	⊕⊖⊖⊖ VERY LOW2,3 due to risk of bias, imprecision	RR 1.03 (0.82 to 1.31)	727 per 1000	22 more per 1000 (from 131 fewer to 225 more)
Morgan-Russell Outcome (change scores) - young people	47 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean morgan-russell outcome (change scores) - young people in the intervention groups was 0.53 standard deviations higher (0.09 lower to 1.16 higher)
Length of Hospital Stay (days) - adults	397 (1 study)	⊕⊖⊖⊖ VERY LOW2,3 due to risk of bias, imprecision		Not calculable for SMD values	The mean length of hospital stay (days) - adults in the intervention groups was 0.45 standard deviations higher (0.19 to 0.72 higher)
Regular Menstruation - young	47	⊕⊖⊖⊖	RR 4.27	156 per 1000	511 more per 1000

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Any ED: Compulsory Treatment (95% CI)
people	(1 study) 12 months	VERY LOW1,4 due to risk of bias, imprecision	(1.77 to 10.3)		(from 120 more to 1000 more)
Disengaged from Family Therapy - young people	50 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision	RR 0.57 (0.22 to 1.44)	441 per 1000	190 fewer per 1000 (from 344 fewer to 194 more)
Required Nasogastric Feeding - young people	50 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 5.84 (2.2 to 15.54)	118 per 1000	569 more per 1000 (from 141 more to 1000 more)
Prematurely Discharged - young people	50 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision	RR 0.35 (0.09 to 1.4)	353 per 1000	229 fewer per 1000 (from 321 fewer to 141 more)
General Functioning - young people	47 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean general functioning - young people in the intervention groups was 0.91 standard deviations lower (1.36 to 0.45 lower)
Depression - young people	47 (1 study) 12 months	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision		Not calculable for SMD values	The mean depression - young people in the intervention groups was 0.77 standard deviations lower (1.41 to 0.14 lower)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Ayton 2009: high selection bias (group allocation likely to affect outcome, no attempt to balance design, baseline not comparable); high performance bias (compulsory group treated significantly longer than voluntary group, sig more in compulsory group required nasogastric feeding).
2 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).
3 Watson 2000: low selection bias (group allocation likely to affect outcome); high performance bias (no participant nor investigator blinding).
4 <300 events (dichotomous outcome) or <400 participants (continuous outcome).

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Any ED: Compulsory Treatment (95% CI)
5 CI crosses both 0.75 and 1.25 (Risk Ratio).					

Table 376: Summary table of findings for compulsory treatment versus voluntary treatment in young people with any eating disorder at 12 months after discharge from hospital

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Compulsory Treatment (95% CI)
>90% Weight for Height 12-mo after discharge - young people	41 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.32 (0.63 to 2.74)	379 per 1000	121 more per 1000 (from 140 fewer to 660 more)
Intermediate Outcome 12-mo after discharge - young people Clinically underweight and either receiving ongoing OP treatment or prematurely disengaged with services	41 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 1.61 (0.55 to 4.7)	207 per 1000	126 more per 1000 (from 93 fewer to 766 more)
Patients alive 12-mo after discharge - young people	41 (1 study)	⊕⊖⊖⊖ VERY LOW1,3 due to risk of bias, imprecision	RR 1.05 (0.9 to 1.22)	931 per 1000	47 more per 1000 (from 93 fewer to 205 more)
Readmitted to Hospital 12-mo after discharge - young people	41 (1 study)	⊕⊖⊖⊖ VERY LOW1,2 due to risk of bias, imprecision	RR 0.46 (0.02 to 8.96)	69 per 1000	37 fewer per 1000 (from 68 fewer to 549 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Ayton 2009: high selection bias (group allocation likely to affect outcome, no attempt to balance design, baseline not comparable); high performance bias (compulsory group treated significantly longer than voluntary group, sig more in compulsory group required nasogastric

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Compulsory Treatment (95% CI)
feeding). 2 CI crosses both 0.75 and 1.25 (Risk Ratio). 3 <300 events (dichotomous outcome) or <400 participants (continuous outcome).					

Table 377: Summary table of findings for compulsory treatment versus voluntary treatment in adults with anorexia nervosa at discharge

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Compulsory Treatment (95% CI)
BMI at discharge	346 (3 studies) 5.7 years	⨁⊖⊖⊖ VERY LOW1,2,3,4 due to risk of bias, imprecision		Not calculable for SMD values	The mean BMI at discharge in the intervention groups was 0.04 standard deviations higher (0.19 lower to 0.27 higher)
Weight Gain	96 (1 study)	⨁⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean weight gain in the intervention groups was 0.23 standard deviations higher (0.22 lower to 0.68 higher)
Duration of hospital stay	250 (2 studies)	⨁⊖⊖⊖ VERY LOW2,3,5 due to risk of bias, imprecision		Not calculable for SMD values	The mean duration of hospital stay in the intervention groups was 0.46 standard deviations higher (0.18 to 0.73 higher)
Refeeding Syndrome	96 (1 study)	⨁⊖⊖⊖ VERY LOW1,5 due to risk of bias, imprecision	RR 2.24 (1.1 to 4.56)	171 per 1000	213 more per 1000 (from 17 more to 610 more)
Locked Ward	96 (1 study)	⨁⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 29.62 (4.02 to 218.18)	14 per 1000	409 more per 1000 (from 43 more to 1000 more)
Required Tube Feeding	96 (1 study)	⨁⊖⊖⊖ VERY LOW1,4	RR 2.94 (1.48 to	157 per 1000	305 more per 1000 (from 75 more to 757 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Compulsory Treatment (95% CI)
		due to risk of bias, imprecision	5.82)		
Achieved Target Weight	88 (1 study)	⊕⊖⊖⊖ VERY LOW3,5 due to risk of bias, imprecision	RR 0.65 (0.27 to 1.57)	411 per 1000	144 fewer per 1000 (from 300 fewer to 234 more)
Required >1 Specialist Medical Consultation	88 (1 study)	⊕⊖⊖⊖ VERY LOW3,5 due to risk of bias, imprecision	RR 1.29 (1.06 to 1.56)	726 per 1000	211 more per 1000 (from 44 more to 407 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Carney 2006: high selection bias (group allocation likely to affect study outcome, no attempt made to balance design, groups not comparable at baseline); high performance bias (Voluntary group not likely to be on locked ward nor subject to tube feeding).
2 Ramsay 1999/Ward 2015: high selection bias (allocation to group likely to affect study outcome, no attempt to balance design, groups not comparable at baseline).
3 Griffiths 1997: high selection bias (group allocation likely to affect study outcome, no attempt made to balance design, socioeconomic status of compulsory group significantly higher than voluntary group); low performance bias (compulsory group had significantly longer treatment).
4 <300 events (dichotomous outcome) or <400 participants (continuous outcome).
5 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

Table 378: Summary table of findings for compulsory treatment versus voluntary treatment in adults with anorexia nervosa at follow up from discharge

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Compulsory Treatment (95% CI)
Patient Deaths FU	245 (2 studies)	⊕⊖⊖⊖ VERY LOW1,2,3 due to risk of bias, imprecision	RR 5.66 (1.49 to 21.54)	13 per 1000	62 more per 1000 (from 6 more to 272 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Voluntary Treatment	Risk difference with Compulsory Treatment (95% CI)
Patient Deaths 20-yr FU	157 (1 study)	⨁⊖⊖⊖ VERY LOW1,4 due to risk of bias, imprecision	RR 1.68 (0.82 to 3.43)	128 per 1000	87 more per 1000 (from 23 fewer to 312 more)

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio; ITT: Intention to treat; FU: Follow up

1 Ramsay 1999/Ward 2015: high selection bias (allocation to group likely to affect study outcome, no attempt to balance design, groups not comparable at baseline).
2 Griffiths 1997: high selection bias (group allocation likely to affect study outcome, no attempt made to balance design, socioeconomic status of compulsory group significantly higher than voluntary group); low performance bias (compulsory group had significantly longer treatment).
3 <300 events (dichotomous outcome) or <400 participants (continuous outcome).
4 CI crosses either 0.75 or 1.25 (Risk Ratio), or either 0.5 or -0.5 (SMD).

1 **Table 379: Logistic regression predicting dropout from inpatient care (Vandereycken 2009)**

	Total Period	Drop out <1 month	Drop out <2 months	Drop out <3 months	Drop out <4 months
Compulsory versus voluntary	0.19	5.73*	0.76	0.41	0.78
Duration of illness	0.68	0.29	1.18	0.22	0.36
Inpatient treatments	0.1	0.78	0.13	0.72	0.78
Outpatient treatments	1.93	4.31*	0.97	0.1	0.11

2 **Table 380: Regression analysis on factors predicting likelihood of compulsory treatment (Carney 2008)**

	Multiple regression OR	Significance
Age		NS
Age 20-29		NS
Previous admissions	1.29	0.0225
Type of eating disorder		NS

	Multiple regression OR	Significance
Psychiatric co-morbidities	1.75	0.0315
BMI on admission		NS
Re-feeding syndrome		NS
Days in treatment		NS
Tube feeding		NS
Locked ward	23.45	0.0235

Abbreviations: NS, not significant; OR, odds ratio

Table 381: Regression analysis on the effect of perceived coercion on outcome in hospital (Schreyer 2015)

Predictors	B co-efficient	OR	Hospital outcome
Perceived coercion	NS		Inpatient length of stay
Extraversion	NS		Inpatient length of stay
Admission BMI	NS		Inpatient length of stay
Perceived coercion	NS		Inpatient rate of weight gain
Perceived coercion	NS		Final discharge BMI
Admission BMI	NS		Final discharge BMI
Perceived coercion	β=4.20*	0.92	Successful transition to partial hospital vs early drop out
Drive for thinness	β =8.79**	1.08	Successful transition to partial hospital vs early drop out
Admission BMI	β =7.37**	1.28	Successful transition to partial hospital vs early drop out
Perceived coercion	NS	0.99	Achieved target weight
Admission BMI	β =23.1**	1.48	Achieved target weight

Abbreviations: β = beta coefficient; NS = not significant

Table 382: Summary on the quality of the studies that conducted a regression analysis on compulsory treatment

	1.1 Is the source population or source area well described?	1.2 Is the eligible population or area representative of the source population or area?	1.3 Do the selected participants or areas represent the eligible population or area?	2.1 Selection of exposure (and comparison) group. How was selection bias minimised?	2.2 Was the selection of explanatory variables based on a sound theoretical basis?	2.3 Was the contamination acceptably low?	2.4 How well were likely confounding factors identified and controlled?	2.5 Is the setting applicable to the UK?	3.1 Were the outcome measures and procedures reliable?	3.2 Were the outcome measurements complete?	3.3 Were all the important outcomes assessed?	3.4 Was there a similar follow up time in exposure and comparison groups?	3.5 Was follow up time meaningful?	4.1 Was the study sufficiently powered to detect an intervention effect (if one exists)?	4.2 Were multiple explanatory variables considered in the analyses?	4.3 Were the analytical methods appropriate?	4.6 Was the precision of association given or calculable? Is association meaningful?	5.1 Are the study results internally valid (i.e. unbiased)?	5.2 Are the findings generalizable to the source population (i.e. externally valid)?	Overall. Average
Carney 2008	(+)	(+)	(+)	NR	(++)	NR	(+)	(+)	(++)	(++)	(-)	(++)	(++)	(+)	(-)	(++)	(++)	(+)	(+)	(+) LOW
Schreyer 2015	(+)	(+)	(+)	NR	(++)	NR	(+)	(-)	(++)	(++)	(-)	(++)	(++)	(+)	(+)	(++)	(++)	(+)	(+)	(+) LOW
Vandereycken 2009	(+)	(+)	(+)	NR	(++)	NR	(-)	(+)	(++)	(++)	(-)	(++)	(++)	(+)	(+)	(++)	(++)	(+)	(+)	(+) LOW

Abbreviations: ++ Some of the checklist criteria have been fulfilled, where they have not been fulfilled, or not adequately described, the conclusions are unlikely to alter; +Some of the checklist criteria have been fulfilled, where they have not been fulfilled, or not adequately described, the conclusions may be likely to alter; -, (-) Few or no checklist criteria have been fulfilled and the conclusions are likely or very likely to alter; NR, not reported

10.2.3 Economic Evidence

No economic evidence on the cost effectiveness of compulsory treatment was identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

10.2.4 Clinical evidence statements

10.2.4.1 Compulsory treatment versus voluntary treatment in young people with any eating disorder at discharge

Very low quality evidence from one observational study (n=47) showed compulsory treatment is more effective on BMI, the number of people achieving regular menstruation and depression compared with voluntary treatment.

Very low quality evidence from one observational study (n=47) showed compulsory treatment is more effective on change in Morgan-Russell-scores compared with voluntary treatment but there was some uncertainty.

Very low quality evidence from one observational study (n=50) showed no difference in the effect of compulsory treatment on the number of people who disengaged from family therapy during treatment and the number of people who were prematurely discharged compared with voluntary treatment.

Very low quality evidence from one observational study (n=50) showed compulsory treatment is less effective on general functioning and the number of people who required nasogastric feeding compared with voluntary treatment.

10.2.4.2 Compulsory treatment versus voluntary treatment in young people with any eating disorder at follow up

Very low quality evidence from one observational study (n=41) showed no difference in the effect of compulsory treatment on the number of people who achieve greater than 90% weight for height compared with voluntary treatment.

Very low quality evidence from one observational study (n=41) showed compulsory treatment may be more effective on the number of people who are still alive, although there was some uncertainty.

Very low quality evidence from one observational study (n=41) showed compulsory treatment may be less effective on the number of people who are either clinically underweight or are prematurely discharged, and on the number of people who are readmitted to hospital compared with voluntary treatment, although there was some uncertainty.

10.2.4.3 Compulsory treatment versus voluntary treatment in adults with any eating disorder at discharge

Very low quality evidence from one observational study (n=397) showed no difference in the effect of compulsory treatment on BMI and the number of people who achieving greater than 85% average body weight compared with voluntary treatment.

Very low quality evidence from one observational study (n=397) showed compulsory treatment is more effective on weight gain compared with voluntary treatment.

Very low quality evidence from one observational study (n=397) showed compulsory treatment may be more effective on weight gain compared with voluntary treatment, although there was some uncertainty.

Very low quality evidence from one observational study (n=397) showed compulsory treatment may be less effective on the number of people achieving either more than 85% average body weight or a BMI of more than 18 kg/m2 compared with voluntary treatment, although there was some uncertainty.

Very low quality evidence from one observational study (n=397) showed compulsory treatment is less effective on length of treatment compared with voluntary treatment.

10.2.4.4 Compulsory treatment versus voluntary treatment in adults with anorexia nervosa at discharge

Very low quality evidence from three observational studies (n=346) showed no difference in the effect of compulsory treatment on BMI compared with voluntary treatment.

Very low quality evidence from one observational study (n=96) showed no difference in the effect of compulsory treatment on weight gain compared with voluntary treatment.

Very low quality evidence from two observational studies (n=250) showed compulsory treatment is less effective on duration of hospital stay compared with voluntary treatment.

Very low quality evidence from one observational study (n=96) showed compulsory treatment is less effective on the number of people experiencing refeeding syndrome, the number of people who were put on a locked ward and the number of people who required tube feeding compared with voluntary treatment.

Very low quality evidence from one observational study (n=88) showed no difference in the effect of compulsory treatment on the number of people who achieved their target weight compared with voluntary treatment.

Very low quality evidence from one observational study (n=88) showed compulsory treatment is less effective on the number of people required more than one specialist medical consultation compared with voluntary treatment.

10.2.4.5 Compulsory treatment versus voluntary treatment in adults with anorexia nervosa at follow up

Very low quality of evidence from two observational studies (n=245) showed that there are more patient deaths at 5 years follow up compared with voluntary treatment.

Very low quality of evidence from one observational study (n=157) showed compulsory treatment may be less effective on the number of people who died at long-term follow up (mean of 19.3 years [range: 14.4-26.3] from admission date) compared with voluntary treatment, although there was some uncertainty.

10.2.4.6 Regression analysis predicting dropouts from treatment if compulsory versus voluntary treatment

Low quality evidence from one observational study (n=174) investigated the hypothesis that the provision of choice will result in more self-chosen and fewer rebellious dropouts. The results showed fewer dropouts within one month in those who had voluntary treatment compared with compulsory treatment after adjusting for duration of illness, previous inpatient and outpatient treatments.

10.2.4.7 Regression analysis on the factors associated with the likelihood of compulsory treatment

Low quality evidence from one observational study (n=96) aimed to explore the circumstances that lead clinicians to use legal coercion in the management of patients with severe AN. The results showed patients are more likely to be admitted for compulsory treatment if they had previous admissions and psychiatric comorbidities after controlling for age, type of eating disorder, BMI at admission, re-feeding syndrome, days in treatment, tube feeding and locked ward. None of the latter variables were associated with being more likely to have compulsory treatment.

10.2.4.8 Regression analysis on the impact of perceived coercion on outcomes from hospital.

Low quality evidence from one observational study (n=204) investigated the hypothesis that a perceived coercion at admission is associated with poorer hospital outcomes. A regression analysis partially supported the hypothesis. The results showed an increased likelihood of dropout prior to successful transition to partial hospitalisation after controlling for EDI-drive for thinness and BMI.

Low quality evidence from one observational study (n=204) showed higher perceived coercion was not associated with inpatient length of stay after adjusting for admission BMI or extraversion.

Low quality evidence from one observational study (n=204) showed higher perceived coercion was not associated with achieving target weight after adjusting for admission BMI

Low quality evidence from one observational study (n=204) showed higher perceived coercion was not associated with discharge BMI after adjusting for admission BMI.

10.2.5 Economic Evidence statements

No economic evidence on the cost effectiveness of compulsory treatment was available.

10.2.6 Recommendations and link to evidence for the review on: What factors/indicators should be considered when assessing whether a person with an eating disorder should be admitted for compulsory treatment (including any form of restrictive interventions usually implemented in refeeding)?

Using the Mental Health Act and compulsory treatment

	162. If a person's physical health is at serious risk due to their eating disorder, they do not consent to treatment, and they can only be treated safely in an inpatient setting, use an appropriate legal framework for compulsory treatment (for example the Mental Health Act 1983) **163. If a child or young person lacks capacity, their physical health is at serious risk and they do not consent to treatment, ask their parents or carers to consent on their behalf and if necessary, use an appropriate legal framework for compulsory treatment (such as the Mental Health Act 1983 or the Children Act 1989).** **164. Feeding people without their consent should only be done by multidisciplinary teams who are competent in doing so.**
Relative value of	For the review: “what factors/indicators should be considered when assessing whether a person with an eating disorder should be admitted for compulsory

different outcomes	treatment (including any form of restrictive interventions usually implemented in refeeding", the following were considered critical outcomes: body weight, consent, family functioning, general functioning, other medical indicators (for example low potassium) and MARSIPAN check list. Important outcomes were the secondary outcomes reported by the authors. It was not possible to be specific about these given the variability in the studies that could be included in this review.
Trade-off between clinical benefits and harms	No direct evidence was identified on what factors/indicators should be considered when assessing whether a person with an eating disorder should be admitted for compulsory treatment. As such indirect evidence was included in this review in case it helped the committee generate a recommendation. **Observational studies** An observational study explored how successful inpatient care was for people with any eating disorder who were admitted for compulsory versus voluntary treatment. In adults, no difference was found at discharge for any outcome including weight and the number who achieved a target body weight between compulsory and voluntary treatment. However, the rate of weight gain was greater in those who were admitted for compulsory treatment (though there was some uncertainty), but they also stayed longer in hospital. In young people, favourable results were found for weight at discharge, Morgan-Russell scores (although there was some uncertainty), resumption of menses, depression and general functioning in those who were admitted for compulsory treatment compared with voluntary admission. However, more patients required nasogastric feeding. No difference was found in the number of patients who disengaged from family therapy nor in the number of those who were prematurely discharged. At long-term follow up (12 months after discharge from an inpatient unit), more people who were treated compulsorily compared with voluntarily were alive, more people were either clinically underweight and receiving ongoing outpatient treatment or were prematurely discharged from services, and more people were readmitted to hospital (although there was some uncertainty). However, mode of treatment (compulsory or voluntary) made no difference in the number of people who achieved a target body weight. In adults with anorexia nervosa, the results were similar between those who were admitted under compulsory conditions versus those who were voluntary for weight gain and achieving target body weight. However, patients admitted under compulsory conditions had a longer hospital stay, required nasogastric feeding, a higher incidence of refeeding syndrome, required specialist care and were kept on a locked ward. At five years follow up, the number of patients alive was less in the compulsory treated group, however, after 20 years the number of patient deaths was no different (although there was some uncertainty). **Regression analysis** A regression analysis showed dropout from inpatient treatment within one month is significantly lower if patients are admitted under voluntary conditions compared with compulsory treatment after controlling for duration of illness, inpatient and outpatient treatments. One study explored the circumstances that lead clinicians to use legal coercion in the management of patients with severe AN. The results showed patients are more likely to be admitted for compulsory treatment if they had previous admissions and psychiatric comorbidities after controlling for age, type of eating disorder, BMI at admission, re-feeding syndrome, days in treatment, tube feeding and locked ward. None of the latter variables were associated with admission for compulsory treatment. Another regression analysis on people admitted for compulsory care showed a higher perceived coercion was associated with a decreased likelihood of successful transition to partial hospitalisation after adjusting for EDI-drive for thinness and BMI. The analysis also showed a higher perceived coercion was not associated with length of inpatient stay after adjusting for admission BMI or extraversion. Nor was it associated with achieving target weight after adjusting for admission BMI, or discharge BMI after adjusting for admission BMI.

Trade-off between net health benefits and resource use	The committee expressed the view that this is a clinical question (that is, a safety issue), and therefore is not relevant for economic analysis.
Quality of evidence	No evidence was identified that explored what factors predicted the need for compulsory treatment. The studies found that conducted a regression analysis on compulsory treatment did not directly answer the question. They investigated whether: a patient was more likely to drop out from inpatient care; factors associated with patients who undergo compulsory treatment (not necessarily what factors they should be admitted for); and the impact of compulsory treatment on outcomes from hospital. The evidence from these studies was low quality for the reasons that they either: did not fully explore all explanatory variables, the size of the studies was relatively small ranging from 96 to 204 participants, provided indirect evidence; they were not from the UK. Observational studies were included in this review and presented to the committee in case they proved useful. All of the outcomes were graded very low quality since they were observational and there were no reasons identified to justify upgrading the quality (i.e. large effect size or dose response). The committee agreed that although these studies were interesting, they did not answer the question. They also questioned their relevance because they were cohort studies and the patients admitted for compulsory care were more likely to be unwell compared with those admitted voluntarily. Thus the committee decided to generate the recommendations based on their knowledge and expertise, in addition to what is outlined in UK legal frameworks. They agreed on the wording via informal consensus.
Other considerations	No direct evidence was identified on what factors need to be considered before admitting a patient for compulsory admission. Therefore, the committee developed this recommendation using their expertise and used the Mental Health Act 1983 and Children Act 1989 as guidance. The committee considered the Mental Health Act 1983 in making these recommendations since it provides the legislation by which people diagnosed with a mental disorder can be detained in hospital and have their disorder assessed or treated against their wishes, unofficially known as "sectioning". For children and young people, the Children Act 1989 means that duty can be allocated to local authorities, courts, parents and other agencies in the UK to ensure children are safeguarded and their welfare is promoted. It makes provisions for instances when parents and families do not co-operate with statutory bodies. Compulsory treatment is only relevant to a few cases. Reconciling respect for the wishes of people receiving treatment and their right to receive good treatment can be difficult, and compulsory treatment, though legally permissible, should never be undertaken lightly. In the case of eating disorders the compulsory treatment would mean admitting a critically ill person who is refusing treatment to an inpatient unit for refeeding. This applied to both adults and children. Although with children, if they are especially young and lack capacity, and they refuse treatment then the parents and carers should be asked to consent on their behalf. If not, then the Mental Health Act can be used to ensure they get the necessary care. Legal advice can be sought where necessary but again compulsory treatment is legal under the Mental Health Act. The committee agreed it is important that feeding people without their consent should be only be done by teams competent in doing so. The person may be resistant to the treatment and there is a risk of injuring the person or they may injure themselves. The committee noted that for nasogastric feeding to be done, it will almost always be necessary for the person to be detained under the Mental Health Act, although occasionally someone may request this treatment. If the person is detained under Section 2 or Section 3, nasogastric feeding does not require a second opinion, although there will be circumstances in which it will be good practice to obtain one.

11 References

AGH, T., PAWASKAR, M., NAGY, B., LACHAINE, J. & VOKO, Z. 2016. The Cost Effectiveness of Lisdexamfetamine Dimesylate for the Treatment of Binge Eating Disorder in the USA. *Clinical Drug Investigation,* 36**,** 305-312.

AGRAS, W. S., LOCK, J., BRANDT, H., BRYSON, S. W., DODGE, E., HALMI, K. A., JO, B., JOHNSON, C., KAYE, W., WILFLEY, D. & WOODSIDE, B. 2014. Comparison of 2 family therapies for adolescent anorexia nervosa: A randomized parallel trial. *JAMA Psychiatry,* 71**,** 1279-1286.

AGRAS, W. S., ROSSITER, E. M., ARNOW, B., SCHNEIDER, J. A., TELCH, C. F., RAEBURN, S. D., BRUCE, B., PERL, M. & KORAN, L. M. 1992. Pharmacologic and cognitive-behavioral treatment for bulimia nervosa: a controlled comparison. *American journal of psychiatry* [Online], 149. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/401/CN-00080401/frame.html.

AGRAS, W. S., ROSSITER, E. M., ARNOW, B., TELCH, C. F., RAEBURN, S. D., BRUCE, B. & KORAN, L. M. 1994a. One-year follow-up of psychosocial and pharmacologic treatments for bulimia nervosa. *Journal of clinical psychiatry* [Online], 55. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/950/CN-00103950/frame.html.

AGRAS, W. S., TELCH, C. F., ARNOW, B., ELDREDGE, K., WILFLEY, D. E., RAEBURN, S. D., HENDERSON, J. & MARNELL, M. 1994b. Weight loss, cognitive-behavioral, and desipramine treatments in binge eating disorder. An additive design. *Behavior therapy* [Online], 25. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/603/CN-00168603/frame.html.

ALFONSSON, S., PARLING, T. & GHADERI, A. 2015. Group behavioral activation for patients with severe obesity and binge eating disorder: A randomized controlled trial. *Behavior Modification,* 39**,** 270-294.

ALLEN, K. L., FURSLAND, A., WATSON, H. & BYRNE, S. M. 2011. Eating disorder diagnoses in general practice settings: Comparison with structured clinical interview and self-report questionnaires. *Journal of Mental Health,* 20**,** 270-280.

ALVAREZ-RAYON, G., MANCILLA-DIAZ, J. M., VAZQUEZ-AREVALO, R., UNIKEL-SANTONCINI, C., CABALLERO-ROMO, A. & MERCADO-CORONA, D. 2004. Validity of the Eating Attitudes Test: a study of Mexican eating disorders patients. *Eating and weight disorders : EWD,* 9**,** 243-248.

ANDRIES, A., FRYSTYK, J., FLYVBJERG, A. & STOVING, R. K. 2014. Dronabinol in severe, enduring anorexia nervosa: A randomized controlled trial. *International Journal of Eating Disorders,* 47**,** 18-23.

ANDRIES, A., GRAM, B. & STOVING, R. K. 2015. Effect of dronabinol therapy on physical activity in anorexia nervosa: a randomised, controlled trial. *Eating and Weight Disorders,* 20**,** 13-21.

AOUN, A., AZZAM, J., EL JABBOUR, F., HLAIS, S., DAHAM, D., EL AMM, C., HONEIN, K. & DECHELOTTE, P. 2015. Validation of the Arabic version of the SCOFF questionnaire for the screening of eating disorders. *Eastern Mediterranean Health Journal,* 21**,** 326-331.

APA 1994. *Diagnostic and statistical manual of mental disorders: DSM-IV,* Washington, DC, American Psychiatric Association.

APA 2013. *Diagnostic and statistical manual of mental disorders: DSM-5,* Arlington, VA, American Psychiatric Association.

ARCELUS, J., BOUMAN, W. P. & MORGAN, J. F. 2008. Treating young people with eating disorders: Transition from child mental health to specialist adult eating disorder services. *European Eating Disorders Review,* 16**,** 30-36.

ARCELUS, J., MITCHELL, A. J., WALES, J. & NIELSEN, S. 2011. Mortality rates in patients with anorexia nervosa and other eating disorders. A meta-analysis of 36 studies. *Arch Gen Psychiatry,* 68**,** 724-31.

ARNOLD, L. M., MCELROY, S. L., HUDSON, J. I., WELGE, J. A., BENNETT, A. J. & KECK, P. E. 2002. A placebo-controlled, randomized trial of fluoxetine in the treatment of binge-eating disorder. *Journal of clinical psychiatry* [Online], 63. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/677/CN-00411677/frame.html.

ATTIA, E., HAIMAN, C., WALSH, B. T. & FLATER, S. R. 1998. Does fluoxetine augment the inpatient treatment of anorexia nervosa? *American journal of psychiatry* [Online], 155. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/947/CN-00686947/frame.html.

ATTIA, E., KAPLAN, A. S., WALSH, B. T., GERSHKOVICH, M., YILMAZ, Z., MUSANTE, D. & WANG, Y. 2011. Olanzapine versus placebo for out-patients with anorexia nervosa. Psychological medicine, 41, 2177-2182.

AYTON, A., KEEN, C. & LASK, B. 2009. Pros and cons of using the mental health act for severe eating disorders in adolescents. European Eating Disorders Review, 17, 14-23.

BAILER, U., ZWAAN, M., LEISCH, F., STRNAD, A., LENNKH-WOLFSBERG, C., EL-GIAMAL, N., HORNIK, K. & KASPER, S. 2004. Guided self-help versus cognitive-behavioral group therapy in the treatment of bulimia nervosa. International journal of eating disorders [Online], 35. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/389/CN-00482389/frame.html.

BANASIAK, S. J., PAXTON, S. J. & HAY, P. 2005. Guided self-help for bulimia nervosa in primary care: a randomized controlled trial. Psychol Med, 35, 1283-94.

BARNES, R. D., WHITE, M. A., MARTINO, S. & GRILO, C. M. 2014. A randomized controlled trial comparing scalable weight loss treatments in primary care. Obesity, 22, 2508-2516.

BAUDET, M. L., MONTASTIER, E., MESTHE, P., OUSTRIC, S., LEPAGE, B. & RITZ, P. 2013. The SCOFF score: A screening tool for eating disorders in family practice. e-SPEN Journal, 8, e86-e89.

BAUER, S., OKON, E., MEERMANN, R. & KORDY, H. 2012. Technology-enhanced maintenance of treatment gains in eating disorders: efficacy of an intervention delivered via text messaging. Journal of consulting and clinical psychology, 80, 700-706.

BEAT 2015. The Costs of Eating Disorders. Social, Health and Economic Impacts. .

BELLOWS, B. K., DUVALL, S. L., KAMAUU, A. W. C., SUPINA, D., BABCOCK, T. & LAFLEUR, J. 2015. Healthcare costs and resource utilization of patients with binge-eating disorder and eating disorder not otherwise specified in the Department of Veterans Affairs. International Journal of Eating Disorders, 48, 1082-1091.

BERG, K. C., PETERSON, C. B., FRAZIER, P. & CROW, S. J. 2012. Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: A systematic review of the literature. International Journal of Eating Disorders, 45, 428-438.

BERLIN, J. A. 2001. Does blinding of readers affect the results of meta-analyses? The Lancet, 350, 185-86.

BEUMONT, P. J., RUSSELL, J. D., TOUYZ, S. W., BUCKLEY, C., LOWINGER, K., TALBOT, P. & JOHNSON, G. F. 1997. Intensive nutritional counseling in bulimia nervosa: A role for supplementation with fluoxetine? Australian and New Zealand Journal of Psychiatry, 31, 514-524.

BIRCHALL, H., PALMER, R. L., WAINE, J., GADSBY, K. & GATWARD, N. 2002. Intensive day programme treatment for severe anorexia nervosa - The Leicester experience. Psychiatric Bulletin, 26, 334-336.

BIRMINGHAM, C. L., GUTIERREZ, E., JONAT, L. & BEUMONT, P. 2004a. Randomized controlled trial of warming in anorexia nervosa. International journal of eating disorders [Online], 35. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/337/CN-00469337/frame.html.

BIRMINGHAM, C. L., GUTIERREZ, E., JONAT, L. & BEUMONT, P. 2004b. Randomized controlled trial of warming in anorexia nervosa. Int J Eat Disord, 35, 234-8.

BISSADA, H., TASCA, G. A., BARBER, A. M. & BRADWEJN, J. 2008. Olanzapine in the treatment of low body weight and obsessive thinking in women with anorexia nervosa: a randomized, double-blind, placebo-controlled trial. American Journal of Psychiatry, 165, 1281-8.

BLOOMGARDEN, A. & CALOGERO, R. M. 2008. A randomized experimental test of the efficacy of EMDR treatment on negative body image in eating disorder inpatients. Brunner-Mazel Eating Disorders Monograph Series, 16, 418-27.

BOERHOUT, C., SWART, M., VAN BUSSCHBACH, J. T. & HOEK, H. W. 2016. Effect of Aggression Regulation on Eating Disorder Pathology: RCT of a Brief Body and Movement Oriented Intervention. European Eating Disorders Review, 24, 114-121.

BORN, C., DE LA FONTAINE, L., WINTER, B., MULLER, N., SCHAUB, A., FRUSTUCK, C., SCHULE, C., VODERHOLZER, U., CUNTZ, U., FALKAI, P. & MEISENZAHL, E. 2015. First results of a refeeding program in a psychiatric intensive care unit for patients with extreme anorexia nervosa. BMC Psychiatry, 15, 57.

BOSSUYT, P. M., DAVENPORT, C., DEEKS, J. J., HYDE, C., LEEFLANG, M. & SCHOLTEN, R. 2013. Interpreting results and drawing conclusions. In: DEEKS, J. J., BOSSUYT, P. M. & GATSONIS, C. (eds.) Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, Version 9.0. Version 9 ed.: The Cochrane Collaboration.

BRAMBILLA, F., GARCIA, C. S., FASSINO, S., DAGA, G. A., FAVARO, A., SANTONASTASO, P., RAMACIOTTI, C., BONDI, E., MELLADO, C., BORRIELLO, R. & MONTELEONE, P. 2007. Olanzapine therapy in anorexia nervosa: Psychobiological effects. International Clinical Psychopharmacology, 22, 197-204.

BRAMBILLA, F., SAMEK, L., COMPANY, M., LOVO, F., CIONI, L. & MELLADO, C. 2009. Multivariate therapeutic approach to binge-eating disorder: combined nutritional, psychological and pharmacological treatment. International Clinical Psychopharmacology, 24, 312-7.

BROWN, A., MOUNTFORD, V. & WALLER, G. 2013. Therapeutic alliance and weight gain during cognitive behavioural therapy for anorexia nervosa. Behav Res Ther, 51, 216-20.

BURTON, E. & STICE, E. 2006. Evaluation of a healthy-weight treatment program for bulimia nervosa: a preliminary randomized trial. Behaviour Research & Therapy, 44, 1727-38.

BYFORD, S., BARRETT, B., ROBERTS, C., CLARK, A., EDWARDS, V., SMETHURST, N. & GOWERS, S. G. 2007. Economic evaluation of a randomised controlled trial for anorexia nervosa in adolescents. British Journal of Psychiatry, 191, 436-40.

BYRNE, S. M., FURSLAND, A., ALLEN, K. L. & WATSON, H. 2011. The effectiveness of enhanced cognitive behavioural therapy for eating disorders: An open trial. Behaviour Research and Therapy, 49, 219-226.

CAREI, T. R., FYFE-JOHNSON, A. L., BREUNER, C. C. & BROWN, M. A. 2010. Randomized controlled clinical trial of yoga in the treatment of eating disorders. Journal of Adolescent Health, 46, 346-51.

CARMEN, D., SHAH, S., GILBER, S. & RUSSELL, F. 2007. Improving attendance at an eating disorders unit by means of an 'opt-in' letter. Clinical Psychology Forum, 22-25.

CARNEY, T., SAUNDERS, D., TAIT, D., TOUYZ, S. & INGVARSON, M. 2004. Therapeutic Pathways in Treatment of Severe Anorexia Nervosa. Contemporary Issues in Law, 7, 100-118.

CARNEY, T., TAIT, D., RICHARDSON, A. & TOUYZ, S. 2008. Why (and when) clinicians compel treatment of anorexia nervosa patients. European Eating Disorders Review, 16, 199-206.

CARNEY, T., WAKEFIELD, A., TAIT, D. & TOUYZ, S. 2006. Reflections on coercion in the treatment of severe anorexia nervosa. Israel Journal of Psychiatry and Related Sciences, 43, 159-165.

CARRARD, I., CREPIN, C., ROUGET, P., LAM, T., GOLAY, A. & VAN DER LINDEN, M. 2011. Randomised controlled trial of a guided self-help treatment on the Internet for binge eating disorder. Behaviour Research and Therapy, 49, 482-491.

CARRUBA, M. O., CUZZOLARO, M., RIVA, L., BOSELLO, O., LIBERTI, S., CASTRA, R., DALLE GRAVE, R., SANTONASTASO, P., GAROSI, V. & NISOLI, E. 2001. Efficacy and tolerability of moclobemide in bulimia nervosa: a placebo-controlled trial. International clinical psychopharmacology [Online], 16. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/309/CN-00328309/frame.html.

CARTER, F. A., JORDAN, J., MCINTOSH, V. V. W., LUTY, S. E., MCKENZIE, J. M., FRAMPTON, C. M. A., BULIK, C. M. & JOYCE, P. R. 2011. The long-term efficacy of three psychotherapies for anorexia nervosa: A randomized, controlled trial. International Journal of Eating Disorders, 44, 647-654.

CARTER, F. A., MCINTOSH, V. V., JOYCE, P. R., SULLIVAN, P. F. & BULIK, C. M. 2003. Role of exposure with response prevention in cognitive-behavioral therapy for bulimia nervosa: three-year follow-up results. International journal of eating disorders [Online], 33. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/641/CN-00435641/frame.html.

CARTER, J. C. & FAIRBURN, C. G. 1998. Cognitive-behavioral self-help for binge eating disorder: a controlled effectiveness study. Journal of consulting and clinical psychology [Online], 66. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/706/CN-00154706/frame.html.

CASSIN, S. E., VON RANSON, K. M., HENG, K., BRAR, J. & WOJTOWICZ, A. E. 2008. Adapted motivational interviewing for women with binge eating disorder: a randomized controlled trial. Psychology of Addictive Behaviors, 22, 417-25.

CASTELNUOVO, G., MANZONI, G. M., VILLA, V., CESA, G. L. & MOLINARI, E. 2011. Brief strategic therapy vs cognitive behavioral therapy for the inpatient and telephone-based outpatient treatment of binge eating disorder: The STRATOB randomized controlled clinical trial. Clinical Practice and Epidemiology in Mental Health, 7, 29-37.

CHEN, E., TOUYZ, S. W., BEUMONT, P. J., FAIRBURN, C. G., GRIFFITHS, R., BUTOW, P., RUSSELL, J., SCHOTTE, D. E., GERTLER, R. & BASTEN, C. 2003. Comparison of

group and individual cognitive-behavioral therapy for patients with bulimia nervosa. International journal of eating disorders [Online], 33. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/062/CN-00436062/frame.html.

CLAUDINO, A. M., DE OLIVEIRA, I. R., APPOLINARIO, J. C., CORDAS, T. A., DUCHESNE, M., SICHIERI, R. & BACALTCHUK, J. 2007. Double-blind, randomized, placebo-controlled trial of topiramate plus cognitive-behavior therapy in binge-eating disorder. Journal of Clinical Psychiatry, 68, 1324-32.

COLLABORATION, A. 2003. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care, 12, 18-23.

COLLABORATION, C. 2014. Review Manager. Version 5.3.5. ed. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration.

COLTON, P. A., OLMSTED, M. P., WONG, H. & RODIN, G. M. 2015. Eating disorders in individuals with type 1 diabetes: case series and day hospital treatment outcome. Eur Eat Disord Rev, 23, 312-7.

COURT, A., MULDER, C., KERR, M., YUEN, H. P., BOASMAN, M., GOLDSTONE, S., FLEMING, J., WEIGALL, S., DERHAM, H., HUANG, C., MCGORRY, P. & BERGER, G. 2010. Investigating the effectiveness, safety and tolerability of quetiapine in the treatment of anorexia nervosa in young people: a pilot study. J Psychiatr Res, 44, 1027-34.

COUTURIER, J., ISSERLIN, L. & LOCK, J. 2010. Family-based treatment for adolescents with anorexia nervosa: a dissemination study. Eat Disord, 18, 199-209.

CRISP, A. H., NORTON, K., GOWERS, S., HALEK, C., BOWYER, C., YELDHAM, D., LEVETT, G. & BHAT, A. 1991. A controlled study of the effect of therapies aimed at adolescent and family psychopathology in anorexia nervosa. British journal of psychiatry [Online], 159. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/799/CN-00079799/frame.html.

CRISTINA, L., FRANCESCA, C., ROSALBA, L., ANTONIO, B. & VALENTINA, M. 2014. Effectiveness of cognitive behavioural psychotherapy alone and combined with pharmacotherapy in binge eating disorder: A differential research. British Journal of Medical Practitioners, 7.

CROW, S. J., AGRAS, W. S., HALMI, K. A., FAIRBURN, C. G., MITCHELL, J. E. & NYMAN, J. A. 2013. A cost effectiveness analysis of stepped care treatment for bulimia nervosa. International Journal of Eating Disorders, 46, 302-307.

CROW, S. J., MITCHELL, J. E., CROSBY, R. D., SWANSON, S. A., WONDERLICH, S. & LANCANSTER, K. 2009. The cost effectiveness of cognitive behavioral therapy for bulimia nervosa delivered via telemedicine versus face-to-face. Behaviour Research & Therapy, 47, 451-3.

CROW, S. J. & NYMAN, J. A. 2004. The Cost-Effectiveness of Anorexia Nervosa Treatment. International Journal of Eating Disorders, 35, 155-160.

CURRIN, L., SCHMIDT, U., TREASURE, J. & JICK, H. 2005. Time trends in eating disorder incidence. Br J Psychiatry, 186, 132-5.

CUSTAL, N., ARCELUS, J., AGUERA, Z., BOVE, F. I., WALES, J., GRANERO, R., JIMENEZ-MURCIA, S., SANCHEZ, I., RIESCO, N., ALONSO, P., CRESPO, J. M., VIRGILI, N., MENCHON, J. M. & FERNANDEZ-ARANDA, F. 2014. Treatment outcome of patients with comorbid type 1 diabetes and eating disorders. BMC Psychiatry, 14, 140.

DALLE GRAVE, R., CALUGI, S., MARCHESINI, G., BECK-PECCOZ, P., BOSELLO, O., COMPARE, A., CUZZOLARO, M., GROSSI, E., MANNUCCI, E., MOLINARI, E., TOMASI, F. & MELCHIONDA, N. 2013. Personality features of obese women in relation to binge eating and night eating. Psychiatry Research, 207, 86-91.

DARE, C., EISLER, I., RUSSELL, G., TREASURE, J. & DODGE, L. 2001. Psychological therapies for adults with anorexia nervosa: randomised controlled trial of out-patient treatments. British journal of psychiatry [Online], 178. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/433/CN-00327433/frame.html.

DAVIS, R., MCVEY, G., HEINMAA, M., ROCKERT, W. & KENNEDY, S. 1999. Sequencing of cognitive-behavioral treatments for bulimia nervosa. International Journal of Eating Disorders, 25, 361-374.

DEBAR, L. L., STRIEGEL-MOORE, R. H., WILSON, G. T., PERRIN, N., YARBOROUGH, B. J., DICKERSON, J., LYNCH, F., ROSSELLI, F. & KRAEMER, H. C. 2011. Guided self-help treatment for recurrent binge eating: Replication and extension. Psychiatric Services, 62, 367-373.

DEBAR, L. L., WILSON, G. T., YARBOROUGH, B. J., BURNS, B., OYLER, B., HILDEBRANDT, T., CLARKE, G. N., DICKERSON, J. & STRIEGEL, R. H. 2013. Cognitive Behavioral Treatment for Recurrent Binge Eating in Adolescent Girls: A Pilot Trial. Cognitive and Behavioral Practice, 20, 147-161.

DEL VALLE, M. F., LARUMBE-ZABALA, E., VILLASENOR-MONTARROSO, A., CARDONA GONZALEZ, C., DIEZ-VEGA, I., LOPEZ MOJARES, L. M. & PEREZ RUIZ, M. 2014. Resistance training enhances muscular performance in patients with anorexia nervosa: A randomized controlled trial. International Journal of Eating Disorders, 47, 601-609.

DEL VALLE, M. F., PEREZ, M., SANTANA-SOSA, E., FIUZA-LUCES, C., BUSTAMANTE-ARA, N., GALLARDO, C., VILLASENOR, A., GRAELL, M., MORANDE, G., ROMO, G. R., LOPEZ-MOJARES, L. M., RUIZ, J. R. & LUCIA, A. 2010. Does resistance training improve the functional capacity and well being of very young anorexic patients? A randomized controlled trial. Journal of Adolescent Health, 46, 352-8.

DEVLIN, M. J., GOLDFEIN, J. A., PETKOVA, E., JIANG, H., RAIZMAN, P. S., WOLK, S., MAYER, L., CARINO, J., BELLACE, D., KAMENETZ, C., DOBROW, I. & WALSH, B. T. 2005. Cognitive behavioral therapy and fluoxetine as adjuncts to group behavioral therapy for binge eating disorder. Obesity Research, 13, 1077-88.

DIAMANTI, A., BASSO, M. S., CASTRO, M., BIANCO, G., CIACCO, E., CALCE, A., CARAMADRE, A. M., NOTO, C. & GAMBARARA, M. 2008. Clinical efficacy and safety of parenteral nutrition in adolescent girls with anorexia nervosa. J Adolesc Health, 42, 111-8.

DUNN, E. C., NEIGHBORS, C. & LARIMER, M. E. 2006. Motivational enhancement therapy and self-help treatment for binge eaters. Psychology of Addictive Behaviors, 20, 44-52.

DURAND, M. A. & KING, M. 2003. Specialist treatment versus self-help for bulimia nervosa: A randomised controlled trial in general practice. British Journal of General Practice, 53, 371-377.

EISLER, I., DARE, C., HODES, M., RUSSELL, G., DODGE, E. & GRANGE, D. 2000. Family therapy for adolescent anorexia nervosa: the results of a controlled comparison of two family interventions. Journal of child psychology and psychiatry, and allied disciplines [Online], 41. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/756/CN-00315756/frame.html.

EISLER, I., DARE, C., RUSSELL, G. F., SZMUKLER, G., LE GRANGE, D. & DODGE, E. 1997. Family and individual therapy in anorexia nervosa: A 5-year follow-up. Archives of General Psychiatry, 54, 1025-1030.

EISLER, I., SIMIC, M., HODSOLL, J., ASEN, E., BERELOWITZ, M., CONNAN, F., ELLIS, G., HUGO, P., SCHMIDT, U., TREASURE, J., YI, I. & LANDAU, S. 2016. A pragmatic randomised multi-centre trial of multifamily and single family therapy for adolescent anorexia nervosa.

FAIRBURN, C. G. & BOHN, K. 2005. Eating disorder NOS (EDNOS): an example of the troublesome "not otherwise specified" (NOS) category in DSM-IV. Behav Res Ther, 43, 691-701.

FAIRBURN, C. G. & COOPER, Z. 2011. Therapist competence, therapy quality, and therapist training. Behav Res Ther, 49, 373-8.

FAIRBURN, C. G., COOPER, Z., BOHN, K., O'CONNOR, M. E., DOLL, H. A. & PALMER, R. L. 2007. The severity and status of eating disorder NOS: implications for DSM-V. Behav Res Ther, 45, 1705-15.

FAIRBURN, C. G. & HARRISON, P. J. 2003. Eating disorders. Lancet, 361, 407-16.

FARIS, P. L., KIM, S. W., MELLER, W. H., GOODALE, R. L., OAKMAN, S. A., HOFBAUER, R. D., MARSHALL, A. M., DAUGHTERS, R. S., BANERJEE-STEVENS, D., ECKERT, E. D. & HARTMAN, B. K. 2000. Effect of decreasing afferent vagal activity with ondansetron on symptoms of bulimia nervosa: a randomised, double-blind trial. Lancet [Online], 355. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/052/CN-00276052/frame.html.

FASSINO, S., ABBATE-DAGA, G., AMIANTO, F., LEOMBRUNI, P., BOGGIO, S. & ROVERA, G. G. 2002. Temperament and character profile of eating disorders: a controlled study with the Temperament and Character Inventory. International Journal of Eating Disorders, 32, 412-25.

FAZELI, P. K., LAWSON, E. A., PRABHAKARAN, R., MILLER, K. K., DONOHO, D. A., CLEMMONS, D. R., HERZOG, D. B., MISRA, M. & KLIBANSKI, A. 2010. Effects of recombinant human growth hormone in anorexia nervosa: a randomized, placebo-controlled study. Journal of Clinical Endocrinology & Metabolism, 95, 4889-97.

FAZELI, P. K., WANG, I. S., MILLER, K. K., HERZOG, D. B., MISRA, M., LEE, H., FINKELSTEIN, J. S., BOUXSEIN, M. L. & KLIBANSKI, A. 2014. Teriparatide increases bone formation and bone mineral density in adult women with anorexia nervosa. Journal of Clinical Endocrinology and Metabolism, 99, 1322-1329.

FBNCSG 1992. Fluoxetine in the treatment of bulimia nervosa. A multicenter, placebo-controlled, double-blind trial. Arch Gen Psychiatry, 49, 139-47.

FICHTER, M. & QUADFLIEG, N. 2001. The structured interview (SIAB-EX) and questionnaire (SIAB-S) for Anorexic and Bulimic eating disorders for DSM-IV and ICD-10 [German]. Verhaltenstherapie, 11, 314-325.

FICHTER, M. M. & QUADFLIEG, N. 2000. Comparing self- and expert rating: A self-report screening version (SIAB-S) of the structured interview for anorexic and bulimic syndromes for DSM-IV and ICD-10 (SIAB-EX). European Archives of Psychiatry and Clinical Neuroscience, 250, 175-185.

FICHTER, M. M., QUADFLIEG, N., GIERK, B., VODERHOLZER, U. & HEUSER, J. 2015. The Munich eating and feeding disorder questionnaire (Munich ED-Quest) DSM-5/ICD-10: validity, reliability, sensitivity to change and norms. European eating disorders review : the journal of the Eating Disorders Association, 23, 229-240.

FICHTER, M. M., QUADFLIEG, N., NISSLMULLER, K., LINDNER, S., OSEN, B., HUBER, T. & WUNSCH-LEITERITZ, W. 2012. Does internet-based prevention reduce the risk of relapse for anorexia nervosa? Behaviour Research and Therapy, 50, 180-190.

FISCHER, S., MEYER, A. H., DREMMEL, D., SCHLUP, B. & MUNSCH, S. 2014. Short-term cognitive-behavioral therapy for binge eating disorder: long-term efficacy and predictors of long-term treatment success. Behaviour research and therapy, 58, 36-42.

FRANZONI, F., MATALONI, E., FEMIA, R. & GALETTA, F. 2002. Effect of oral potassium supplementation on QT dispersion in anorexia nervosa. Acta Paediatrica, 91, 653-6.

FRISCH, M. J., HERZOG, D. B. & FRANKO, D. L. 2006. Residential treatment for eating disorders. International Journal of Eating Disorders, 39, 434-442.

GARCIA-CAMPAYO, J., SANZ-CARRILLO, C., IBANEZ, J. A., LOU, S., SOLANO, V. & ALDA, M. 2005. Validation of the Spanish version of the SCOFF questionnaire for the screening of eating disorders in primary care. Journal of Psychosomatic Research, 59, 51-55.

GEIST, R., HEINMAA, M., STEPHENS, D., DAVIS, R. & KATZMAN, D. K. 2000. Comparison of family therapy and family group psychoeducation in adolescents with anorexia nervosa. Canadian journal of psychiatry. Revue canadienne de psychiatrie [Online], 45. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/474/CN-00292474/frame.html.

GHADERI, A. & SCOTT, B. 2003. Pure and guided self-help for full and sub-threshold bulimia nervosa and binge eating disorder. British journal of clinical psychology [Online], 42. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/016/CN-00463016/frame.html.

GODART, N., BERTHOZ, S., CURT, F., PERDEREAU, F., REIN, Z., WALLIER, J., HORREARD, A. S., KAGANSKI, I., LUCET, R., ATGER, F., CORCOS, M., FERMANIAN, J., FALISSARD, B., FLAMENT, M., EISLER, I. & JEAMMET, P. 2012. A randomized controlled trial of adjunctive family therapy and treatment as usual following inpatient treatment for anorexia nervosa adolescents. PLoS ONE, 7.

GODDARD, E. 2011. Does a proven intervention to improve functioning of carers also benefit the anorexia nervosa sufferer for whom they care? A pilot study of our Expert Carer Helping Others (ECHO) intervention. Http://www.controlled-trials.com/ISRCTN83003225 [Online]. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/568/CN-00974568/frame.html.

GODDARD, E., HIBBS, R., RAENKER, S., SALERNO, L., ARCELUS, J., BOUGHTON, N., CONNAN, F., GOSS, K., LASZLO, B., MORGAN, J., MOORE, K., ROBERTSON, D., S, S., SCHREIBER-KOUNINE, C., SHARMA, S., WHITEHEAD, L., SCHMIDT, U. & TREASURE, J. 2013. A multi-centre cohort study of short term outcomes of hospital treatment for anorexia nervosa in the UK. BMC psychiatry, 13, 287.

GOLAN, M. & HEYMAN, N. 2005. Managing eating disorders in the natural environment: Intensive vs. limited programs. Israel Journal of Psychiatry and Related Sciences, 42, 163-171.

GOLDBLOOM, D. S., OLMSTED, M., DAVIS, R., CLEWES, J., HEINMAA, M., ROCKERT, W. & SHAW, B. 1997. A randomized controlled trial of fluoxetine and cognitive behavioral therapy for bulimia nervosa: short-term outcome. Behaviour research and therapy [Online], 35. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/593/CN-00143593/frame.html.

GOLDEN, N. H., IGLESIAS, E. A., JACOBSON, M. S., CAREY, D., MEYER, W., SCHEBENDACH, J., HERTZ, S. & SHENKER, I. R. 2005. Alendronate for the treatment of

osteopenia in anorexia nervosa: A randomized, double-blind, placebo-controlled trial. Journal of Clinical Endocrinology and Metabolism, 90, 3179-3185.

GOODRICK, G. K., POSTON, W. S., KIMBALL, K. T., REEVES, R. S. & FOREYT, J. P. 1998. Nondieting Versus Dieting Treatment for Overweight Binge-Eating Women. Journal of consulting and clinical psychology [Online], 66. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/060/CN-00684060/frame.html.

GORDON, C. M., GRACE, E., EMANS, S. J., FELDMAN, H. A., GOODMAN, E., BECKER, K. A., ROSEN, C. J., GUNDBERG, C. M. & LEBOFF, M. S. 2002. Effects of oral dehydroepiandrosterone on bone density in young women with anorexia nervosa: a randomized trial. Journal of clinical endocrinology and metabolism [Online], 87. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/101/CN-00411101/frame.html.

GORIN, A. A., GRANGE, D. & STONE, A. A. 2003. Effectiveness of spouse involvement in cognitive behavioral therapy for binge eating disorder. International journal of eating disorders [Online], 33. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/106/CN-00436106/frame.html.

GOWERS, S., CLARK, A., ROBERTS, C., BYFORD, S. & BRARRETT, B. 2010. A randomised controlled multi-centre treatment trial of adolescent anorexia nervosa, including assessment of cost effectiveness and patient acceptability - The TOuCAN trial. Health Technology Assessment.

GOWERS, S., NORTON, K., HALEK, C. & CRISP, A. H. 1994. Outcome of outpatient psychotherapy in a random allocation treatment study of anorexia nervosa. International Journal of Eating Disorders, 15, 165-177.

GOWERS, S. G., CLARK, A., ROBERTS, C., GRIFFITHS, A., EDWARDS, V., BRYAN, C., SMETHURST, N., BYFORD, S. & BARRETT, B. 2007. Clinical effectiveness of treatments for anorexia nervosa in adolescents: randomised controlled trial. British Journal of Psychiatry, 191, 427-35.

GRAVES, T. A., TABRI, N., THOMPSON-BRENNER, H., FRANKO, D. L., EDDY, K. T., BOURION-BEDES, S., BROWN, A., CONSTANTINO, M. J., FLÜCKIGER, C., FORSBERG, S., HILDEBRANDT, T., ISSERLIN, L., COUTURIER, J., PAULSSON KARLSSON, G., MANDER, J., TEUFEL, M., MITCHELL, J. E., CROSBY, R. D., PRESTANO, C., SATIR, D. A., SIMPSON, S., SLY, R., LACEY, J. H., STILES-SHIELDS, C., TASCA, G. A., WALLER, G., ZAITSOFF, S. L., RIENECKE, R., LE GRANGE, D. & THOMAS, J. J. 2016. A meta-analysis of the relation between therapeutic alliance and treatment outcome in eating disorders. Journal of Consulting & Clinical Psychology.

GRENON, R., TASCA, G. A., CWINN, E., COYLE, D., SUMNER, A., GICK, M. & BISSADA, H. 2010. Depressive symptoms are associated with medication use and lower health-related quality of life in overweight women with binge eating disorder. Women's Health Issues, 20, 435-440.

GRILO, C., M., CROSBY, R., D., WILSON, G. T. & MASHEB, R., M. 2012. 12-month follow-up of fluoxetine and cognitive behavioral therapy for binge eating disorder. [References]. Journal of consulting and clinical psychology [Online], 80. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/849/CN-00852849/frame.html.

GRILO, C. M. & MASHEB, R. M. 2005. A randomized controlled comparison of guided self-help cognitive behavioral therapy and behavioral weight loss for binge eating disorder. Behaviour Research & Therapy, 43, 1509-25.

GRILO, C. M., MASHEB, R. M. & WILSON, G. T. 2005. Efficacy of cognitive behavioral therapy and fluoxetine for the treatment of binge eating disorder: a randomized double-blind placebo-controlled comparison. Biological Psychiatry, 57, 301-9.

GRILO, C. M., MASHEB, R. M., WILSON, G. T., GUEORGUIEVA, R. & WHITE, M. A. 2011. Cognitive-behavioral therapy, behavioral weight loss, and sequential treatment for obese patients with binge-eating disorder: A randomized controlled trial. Journal of Consulting and Clinical Psychology, 79, 675-685.

GRILO, C. M., WHITE, M. A., GUEORGUIEVA, R., BARNES, R. D. & MASHEB, R. M. 2013. Self-help for binge eating disorder in primary care: A randomized controlled trial with ethnically and racially diverse obese patients. Behaviour Research and Therapy, 51, 855-861.

GRINSPOON, S., MILLER, K., HERZOG, D., CLEMMONS, D. & KLIBANSKI, A. 2003. Effects of recombinant human insulin-like growth factor (IGF)-I and estrogen administration on IGF-I, IGF binding protein (IGFBP)-2, and IGFBP-3 in anorexia nervosa: a randomized-controlled study. Journal of clinical endocrinology and metabolism [Online], 88. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/463/CN-00413463/frame.html.

GRINSPOON, S., THOMAS, L., MILLER, K., HERZOG, D. & KLIBANSKI, A. 2002. Effects of recombinant human IGF-I and oral contraceptive administration on bone density in anorexia nervosa. The Journal of clinical endocrinology and metabolism [Online], 87. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/454/CN-00389454/frame.html.

GROVER, M., NAUMANN, U., MOHAMMAD-DAR, L., GLENNON, D., RINGWOOD, S., EISLER, I., WILLIAMS, C., TREASURE, J. & SCHMIDT, U. 2011. A randomized controlled trial of an Internet-based cognitive-behavioural skills package for carers of people with anorexia nervosa. Psychological medicine, 41, 2581-2591.

GUERDJIKOVA, A. I., MCELROY, S. L., KOTWAL, R., WELGE, J. A., NELSON, E., LAKE, K., ALESSIO, D. D., KECK, P. E., JR. & HUDSON, J. I. 2008. High-dose escitalopram in the treatment of binge-eating disorder with obesity: a placebo-controlled monotherapy trial. Human Psychopharmacology, 23, 1-11.

GUERDJIKOVA, A. I., MCELROY, S. L., WELGE, J. A., NELSON, E., KECK, P. E. & HUDSON, J. I. 2009. Lamotrigine in the treatment of binge-eating disorder with obesity: a randomized, placebo-controlled monotherapy trial. International Clinical Psychopharmacology, 24, 150-8.

GUERDJIKOVA, A. I., MCELROY, S. L., WINSTANLEY, E. L., NELSON, E. B., MORI, N., MCCOY, J., KECK, P. E. & HUDSON, J. I. 2012. Duloxetine in the treatment of binge eating disorder with depressive disorders: A placebo-controlled trial. International Journal of Eating Disorders, 45, 281-289.

GULEC, H., MOESSNER, M., TURY, F., FIEDLER, P., MEZEI, A. & BAUER, S. 2014. A randomized controlled trial of an Internet-based posttreatment care for patients with eating disorders. Telemedicine and e-Health, 20, 916-922.

GUYATT, G., OXMAN, A. D., AKL, E. A., KUNZ, R., VIST, G., BROZEK, J., NORRIS, S., FALCK-YTTER, Y., GLASZIOU, P., DEBEER, H., JAESCHKE, R., RIND, D., MEERPOHL, J., DAHM, P. & SCHUNEMANN, H. J. 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol, 64, 383-94.

HAAS, L., STARGARDT, T., SCHREYOEGG, J., SCHLOSSER, R., DANZER, G. & KLAPP, B. F. 2012. Inpatient costs and predictors of costs in the psychosomatic treatment of anorexia nervosa. International Journal of Eating Disorders, 45, 214-221.

HAGMAN, J., GRALLA, J., SIGEL, E., ELLERT, S., DODGE, M., GARDNER, R., O'LONERGAN, T., FRANK, G. & WAMBOLDT, M. Z. 2011. A double-blind, placebo-controlled study of risperidone for the treatment of adolescents and young adults with anorexia nervosa: A pilot study. Journal of the American Academy of Child and Adolescent Psychiatry, 50, 915-924.

HALL, A. & CRISP, A. H. 1987. Brief psychotherapy in the treatment of anorexia nervosa. Outcome at one year. British journal of psychiatry [Online], 151. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/323/CN-00051323/frame.html.

HALMI, K. A., ECKERT, E., LADU, T. J. & COHEN, J. 1986. Anorexia nervosa. Treatment efficacy of cyproheptadine and amitriptyline. Archives of general psychiatry [Online], 43. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/543/CN-00041543/frame.html.

HAUSER, S. T., JACOBSON, A. M., LAVORI, P., WOLFSDORF, J. I., HERSKOWITZ, R. D., MILLEY, J. E., BLISS, R., WERTLIEB, D. & STEIN, J. 1990. Adherence among children and adolescents with insulin-dependent diabetes mellitus over a four-year longitudinal follow-up: II. Immediate and long-term linkages with the family milieu. J Pediatr Psychol, 15, 527-42.

HEDGES, D. W., REIMHERR, F. W., HOOPES, S. P., ROSENTHAL, N. R., KAMIN, M., KARIM, R. & CAPECE, J. A. 2003. Treatment of bulimia nervosa with topiramate in a randomized, double-blind, placebo-controlled trial, part 2: improvement in psychiatric measures. Journal of clinical psychiatry [Online], 64. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/087/CN-00460087/frame.html.

HENDERSON, M. & FREEMAN, C. P. L. 1987. A self-rating scale for bulimia: The 'BITE'. British Journal of Psychiatry, 150, 18-24.

HERPERTZ-DAHLMANN, B. 2013. Randomized controlled non-inferiority trial of day patient treatment in comparison to inpatient treatment among adolescent patients with anorexia nervosa. European Child and Adolescent Psychiatry, 1), S96.

HERPERTZ-DAHLMANN, B., SCHWARTE, R., KREI, M., EGBERTS, K., WARNKE, A., WEWETZER, C., PFEIFFER, E., FLEISCHHAKER, C., SCHERAG, A., HOLTKAMP, K., HAGENAH, U., BUHREN, K., KONRAD, K., SCHMIDT, U., SCHADE-BRITTINGER, C., TIMMESFELD, N. & DEMPFLE, A. 2014. Day-patient treatment after short inpatient care versus continued inpatient treatment in adolescents with anorexia nervosa (ANDI): A multicentre, randomised, open-label, non-inferiority trial. The Lancet, 383, 1222-1229.

HERSCOVICI, C. R., KOVALSKYS, I. & ORELLANA, L. 2015. An exploratory evaluation of the family meal intervention for adolescent anorexia nervosa. Family Process, No Pagination Specified.

HIBBS, R., MAGILL, N., GODDARD, E., RHIND, C., RAENKER, S., MACDONALD, P., TODD, G., ARCELUS, J., MORGAN, J., BEECHAM, J., SCHMIDT, U., LANDAU, S. & TREASURE, J. 2015. Clinical effectiveness of a skills training intervention for caregivers in improving patient and caregiver health following in-patient treatment for severe anorexia nervosa: pragmatic randomised controlled trial. British Journal of Psychiatry Open, 1, 56-66.

HIGGINS, J. P. T. & GREEN, S. 2011. Cochrane Handbook for Systematic Reviews of Interventions, The Cochrane Collaboration.

HILBERT, A., HILDEBRANDT, T., AGRAS, W., WILFLEY, D. E. & WILSON, G. 2015. Rapid Response in Psychological Treatments for Binge Eating Disorder. Journal of Consulting and Clinical Psychology Apr, No Pagination Specified.

HILBERT, A. & TUSCHEN-CAFFIER, B. 2004. Body image interventions in cognitive-behavioural therapy of binge-eating disorder: a component analysis. Behaviour Research & Therapy, 42, 1325-39.

HILDEBRANDT, T., LOEB, K., TROUPE, S. & DELINSKY, S. 2012. Adjunctive mirror exposure for eating disorders: A randomized controlled pilot study. Behaviour Research and Therapy, 50, 797-804.

HILL, D. M., CRAIGHEAD, L. W. & SAFER, D. L. 2011. Appetite-focused dialectical behavior therapy for the treatment of binge eating with purging: A preliminary trial. International Journal of Eating Disorders, 44, 249-261.

HODSOLL, J., RHIND, C., MICALI, N., HIBBS, R., GODDARD, E., NAZAR, B., SCHMIDT, U., GOWERS, S., MACDONALD, P., TODD, G., LANDAU, S. & TREASURE, J. 2016. A pilot multicentre randomised trial to explore the effect of two forms of carer skills training on carer wellbeing and clinical outcome in adolescent anorexia nervosa.

HOGDAHL, L. 2013. Evaluation of internet-based, guided, self-help, cognitive behavioural therapy for bulimia nervosa and similar eating disorders in a specialist outpatient setting: a randomized controlled trial. Http://www.controlled-trials.com/ISRCTN44999017 [Online]. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/565/CN-00974565/frame.html.

HOOPES, S. P., REIMHERR, F. W., HEDGES, D. W., ROSENTHAL, N. R., KAMIN, M., KARIM, R., CAPECE, J. A. & KARVOIS, D. 2003. Treatment of bulimia nervosa with topiramate in a randomized, double-blind, placebo-controlled trial, part 1: improvement in binge and purge measures. Journal of clinical psychiatry [Online], 64. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/355/CN-00459355/frame.html.

HOTZEL, K., VON BRACHEL, R., SCHMIDT, U., RIEGER, E., KOSFELDER, J., HECHLER, T., SCHULTE, D. & VOCKS, S. 2014. An internet-based program to enhance motivation to change in females with symptoms of an eating disorder: A randomized controlled trial. Psychological Medicine, 44, 1947-1963.

HOUSE, J., EISLER, I., SIMIC, M. & MICALI, N. 2008. Diagnosing eating disorders in adolescents: A comparison of the eating disorder examination and the development and well-being assessment. International Journal of Eating Disorders, 41, 535-541.

HOUSE, J., SCHMIDT, U., CRAIG, M., LANDAU, S., SIMIC, M., NICHOLLS, D., HUGO, P., BERELOWITZ, M. & EISLER, I. 2012. Comparison of specialist and nonspecialist care pathways for adolescents with anorexia nervosa and related eating disorders. International Journal of Eating Disorders, 45, 949-956.

HOYLE, D., SLATER, J., WILLIAMS, C., SCHMIDT, U. & WADE, T. D. 2013. Evaluation of a web-based skills intervention for carers of people with anorexia nervosa: A randomized controlled trial. International Journal of Eating Disorders, 46, 634-638.

HSU, L. K., RAND, W., SULLIVAN, S., LIU, D. W., MULLIKEN, B., MCDONAGH, B. & KAYE, W. H. 2001. Cognitive therapy, nutritional therapy and their combination in the treatment of bulimia nervosa. Psychological medicine [Online], 31. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/549/CN-00354549/frame.html.

HUDSON, J., MCELROY, S., FERREIRA-CORNWELL, C., RADEWONUK, J. & GASIOR, M. 2015. A double-blind, placebo-controlled, randomized-withdrawal study of lisdexamfetamine dimesylate in adults with moderate to severe binge eating disorder. Neuropsychopharmacology, 40, S135-S136.

HUDSON, J. I., MCELROY, S. L., RAYMOND, N. C., CROW, S., KECK, P. E., CARTER, W. P., MITCHELL, J. E., STRAKOWSKI, S. M., POPE, H. G., COLEMAN, B. S. & JONAS, J. M. 1998. Fluvoxamine in the treatment of binge-eating disorder: a multicenter placebo-controlled, double-blind trial. American journal of psychiatry [Online], 155. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/594/CN-00157594/frame.html.

HUGHES, E. K., LE GRANGE, D., COURT, A., YEO, M., CAMPBELL, S., WHITELAW, M., ATKINS, L. & SAWYER, S. M. 2014. Implementation of family-based treatment for adolescents with anorexia nervosa. Journal of pediatric health care : official publication of National Association of Pediatric Nurse Associates & Practitioners, 28, 322-330.

JACKSON, R., AMERATUNGA, S., BROAD, J., CONNOR, J., LETHABY, A., ROBB, G., WELLS, S., GLASZIOU, P. & HENEGHAN, C. 2006. The GATE frame: critical appraisal with pictures. Evid Based Med, 11, 35-8.

JACOBI, C., DAHME, B. & DITTMANN, R. 2002. Cognitive-behavioural, fluoxetine and combined treatment for bulimia nervosa: Short- and long-term results. European eating disorders review [Online], 10. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/296/CN-00443296/frame.html.

JADAD, A. R., MOORE, R. A., CARROLL, D., JENKINSON, C., REYNOLDS, D. J., GAVAGHAN, D. J. & MCQUAY, H. J. 1996. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials, 17, 1-12.

JAGER, B., LIEDTKE, R., KUNSEBECK, H., LEMPA, W., KERSTING, A. & SEIDE, L. 1996. Psychotherapy and bulimia nervosa: Evaluation and long-term follow-up of two conflict-orientated treatment conditions. Acta Psychiatrica Scandinavica, 93, 268-278.

JANAS-KOZIK, M., KRZYSTANEK, M., STACHOWICZ, M., KRUPKA-MATUSZCZYK, I., JANAS, A. & RYBAKOWSKI, J. K. 2011. Bright light treatment of depressive symptoms in patients with restrictive type of anorexia nervosa. Journal of Affective Disorders, 130, 462-465.

JONES, M., LUCE, K. H., OSBORNE, M. I., TAYLOR, K., CUNNING, D., DOYLE, A. C., WILFLEY, D. E. & TAYLOR, C. B. 2008. Randomized, controlled trial of an internet-facilitated intervention for reducing binge eating and overweight in adolescents. Pediatrics, 121, 453-62.

KAFANTARIS, V., LEIGH, E., HERTZ, S., BEREST, A., SCHEBENDACH, J., STERLING, W. M., SAITO, E., SUNDAY, S., HIGDON, C., GOLDEN, N. H. & MALHOTRA, A. K. 2011. A placebo-controlled pilot study of adjunctive olanzapine for adolescents with anorexia nervosa. Journal of Child and Adolescent Psychopharmacology, 21, 207-212.

KANERVA, R., RISSANEN, A. & SARNA, S. 1995. Fluoxetine in the treatment of anxiety, depressive symptoms, and eating-related symptoms in bulimia nervosa. Nordic Journal of Psychiatry, 49, 237-42.

KARACIC, M., WALES, J. A., ARCELUS, J., PALMER, R. L., COOPER, Z. & FAIRBURN, C. G. 2011. Changes in alcohol intake in response to transdiagnostic cognitive behaviour therapy for eating disorders. Behaviour Research and Therapy, 49, 573-577.

KAWAI, K., YAMASHITA, S., TAKAKURA, S., KOMAKI, G., KUBO, C. & SUDO, N. 2015. The outcome of treatment for anorexia nervosa inpatients who required urgent hospitalization. Psychotherapy and Psychosomatics, 84, 38.

KAYE, W. H., NAGATA, T., WELTZIN, T. E., HSU, L. K., SOKOL, M. S., MCCONAHA, C., PLOTNICOV, K. H., WEISE, J. & DEEP, D. 2001. Double-blind placebo-controlled administration of fluoxetine in restricting- and restricting-purging-type anorexia nervosa. Biological psychiatry [Online], 49. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/933/CN-00342933/frame.html.

KEEL, P. K., MITCHELL, J. E., DAVIS, T. L. & CROW, S. J. 2002. Long-term impact of treatment in women diagnosed with bulimia nervosa. International journal of eating disorders [Online], 31. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/141/CN-00379141/frame.html.

KELLS, M., DAVIDSON, K., HITCHKO, L., SCHUBERT-BOB, P. & MCCABE, M. 2012. Examining supervised meals in patients with restrictive eating disorders. Journal of Adolescent Health, 1), S42-S43.

KENARDY, J., MENSCH, M., BOWEN, K., GREEN, B. & WALTON, J. 2002. Group therapy for binge eating in Type 2 diabetes: A randomized trial. Diabetic Medicine, 19, 234-239.

KLIBANSKI, A., BILLER, B. M., SCHOENFELD, D. A., HERZOG, D. B. & SAXE, V. C. 1995. The effects of estrogen administration on trabecular bone loss in young women with anorexia nervosa. Journal of clinical endocrinology and metabolism [Online], 80. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/672/CN-00111672/frame.html.

KOHN, M. A., CARPENTER, C. R. & NEWMAN, T. B. 2013. Understanding the direction of bias in studies of diagnostic test accuracy. Acad Emerg Med, 20, 1194-206.

KONG, S. 2005. Day treatment programme for patients with eating disorders: Randomized controlled trial. Journal of Advanced Nursing, 51, 5-14.

KOSMERLY, S., WALLER, G. & LAFRANCE ROBINSON, A. 2015. Clinician adherence to guidelines in the delivery of family-based therapy for eating disorders. Int J Eat Disord, 48, 223-9.

KRAUTH, C. & BUSER, K. 2002. How high are the costs of eating disorders - Anorexia nervosa and bulimia nervosa - For german society? European Journal of Health Economics, 3, 244-250.

KRISTELLER, J., WOLEVER, R. Q. & SHEETS, V. 2014. Mindfulness-Based Eating Awareness Training (MB-EAT) for binge eating: A randomized clinical trial. Mindfulness, 5, 282-297.

LAESSLE, R. G., BEUMONT, P. J., BUTOW, P., LENNERTS, W., O'CONNOR, M., PIRKE, K. M., TOUYZ, S. W. & WAADT, S. 1991. A comparison of nutritional management with stress management in the treatment of bulimia nervosa. British journal of psychiatry [Online], 159. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/096/CN-00081096/frame.html.

LAVENDER, A., STARTUP, H., NAUMANN, U., SAMARAWICKREMA, N., DEJONG, H., KENYON, M., VAN DEN EYNDE, F. & SCHMIDT, U. 2012. Emotional and Social Mind Training: A Randomised Controlled Trial of a New Group-Based Treatment for Bulimia Nervosa. PLoS ONE, 7.

LE GRANGE, D., CROSBY, R. D., RATHOUZ, P. J. & LEVENTHAL, B. L. 2007. A randomized controlled comparison of family-based treatment and supportive psychotherapy for adolescent bulimia nervosa. Archives of General Psychiatry, 64, 1049-56.

LE GRANGE, D., HUGHES, E. K., COURT, A., YEO, M., CROSBY, R. D. & SAWYER, S. M. 2016. Randomized Clinical Trial of Parent-Focused Treatment and Family-Based Treatment for Adolescent Anorexia Nervosa. Journal of the American Academy of Child & Adolescent Psychiatry.

LE GRANGE, D., LOCK, J., AGRAS, W. S., BRYSON, S. W. & JO, B. 2015. Randomized Clinical Trial of Family-Based Treatment and Cognitive-Behavioral Therapy for Adolescent Bulimia Nervosa. Journal of the American Academy of Child and Adolescent Psychiatry, 54, 886-894.

LEE, N. F. & RUSH, A. 1986. Cognitive-behavioral group therapy for bulimia. International Journal of Eating Disorders, 5, 599-615.

LEITENBERG, H., ROSEN, J. C., GROSS, J., NUDELMAN, S. & VARA, L. S. 1988. Exposure plus response-prevention treatment of bulimia nervosa. Journal of consulting and clinical psychology [Online], 56. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/350/CN-00214350/frame.html.

LEOMBRUNI, P., AMIANTO, F., DELSEDIME, N., GRAMAGLIA, C., ABBATE-DAGA, G. & FASSINO, S. 2006. Citalopram versus fluoxetine for the treatment of patients with bulimia nervosa: a single-blind randomized controlled trial. Advances in Therapy, 23, 481-94.

LEOMBRUNI, P., PIERO, A., LAVAGNINO, L., BRUSTOLIN, A., CAMPISI, S. & FASSINO, S. 2008. A randomized, double-blind trial comparing sertraline and fluoxetine 6-month treatment in obese patients with Binge Eating Disorder. Progress in Neuro-Psychopharmacology & Biological Psychiatry, 32, 1599-605.

LEXCHIN, J., BERO, L. A., DJULBEGOVIC, B. & CLARK, O. 2003. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ, 326, 1167-70.

LIU, C. Y., TSENG, M. C. M., CHEN, K. Y., CHANG, C. H., LIAO, S. C. & CHEN, H. C. 2015. Sex difference in using the SCOFF questionnaire to identify eating disorder patients at a psychiatric outpatient clinic. Comprehensive Psychiatry, 57, 160-166.

LJOTSSON, B., LUNDIN, C., MITSELL, K., CARLBRING, P., RAMKLINT, M. & GHADERI, A. 2007. Remote treatment of bulimia nervosa and binge eating disorder: a randomized trial of Internet-assisted cognitive behavioural therapy. Behaviour Research & Therapy, 45, 649-61.

LOCK, J. 2003. A health service perspective Anorexia Nervosa. Eating Disorders, 11, 197-207.

LOCK, J. 2015. An update on evidence-based psychosocial treatments for eating disorders in children and adolescents. Journal of Clinical Child and Adolescent Psychology, 44, 707-721.

LOCK, J., AGRAS, W. S., BRYSON, S. & KRAEMER, H. C. 2005. A comparison of short- and long-term family therapy for adolescent anorexia nervosa. Journal of the American Academy of Child & Adolescent Psychiatry, 44, 632-9.

LOCK, J., COUTURIER, J. & AGRAS, W. 2008. Costs of remission and recovery using family therapy for adolescent anorexia nervosa: a descriptive report. . Eating Disorders. , 16, 322-30.

LOCK, J., COUTURIER, J. & AGRAS, W. S. 2006. Comparison of long-term outcomes in adolescents with anorexia nervosa treated with family therapy. Journal of the American Academy of Child & Adolescent Psychiatry, 45, 666-72.

LOCK, J., LE GRANGE, D., AGRAS, W. S., FITZPATRICK, K. K., JO, B., ACCURSO, E., FORSBERG, S., ANDERSON, K., ARNOW, K. & STAINER, M. 2015. Can adaptive treatment improve outcomes in family-based therapy for adolescents with anorexia nervosa? Feasibility and treatment effects of a multi-site treatment study. Behav Res Ther, 73, 90-5.

LOCK, J., LE GRANGE, D., AGRAS, W. S., MOYE, A., BRYSON, S. W. & JO, B. 2010. Randomized clinical trial comparing family-based treatment with adolescent-focused individual therapy for adolescents with anorexia nervosa. Archives of General Psychiatry, 67, 1025-32.

LOEB, K. L., WILSON, G. T., GILBERT, J. S. & LABOUVIE, E. 2000. Guided and unguided self-help for binge eating. Behaviour research and therapy [Online], 38. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/569/CN-00270569/frame.html.

LUCK, A. J., MORGAN, J. F., REID, F., O'BRIEN, A., BRUNTON, J., PRICE, C., PERRY, L. & LACEY, J. 2002. The SCOFF questionnaire and clinical interview for eating disorders in general practice: Comparative study. BMJ: British Medical Journal, 325, 755-756.

LYNCH, F. L., STRIEGEL-MOORE, R. H., DICKERSON, J. F., PERRIN, N., DEBAR, L., WILSON, G. T. & KRAEMER, H. C. 2010. Cost-effectiveness of guided self-help treatment for recurrent binge eating. Journal of Consulting & Clinical Psychology, 78, 322-33.

MADDEN, S., MISKOVIC-WHEATLEY, J., WALLIS, A., KOHN, M., LOCK, J., LE GRANGE, D., JO, B., CLARKE, S., RHODES, P., HAY, P. & TOUYZ, S. 2014. A randomized controlled trial of in-patient treatment for anorexia nervosa in medically unstable adolescents. Psychological Medicine, 45, 415-427.

MAGILL, N., RHIND, C., HIBBS, R., GODDARD, E., MACDONALD, P., ARCELUS, J., MORGAN, J., BEECHAM, J., SCHMIDT, U., LANDAU, S. & TREASURE, J. 2016. Two-year Follow-up of a Pragmatic Randomised Controlled Trial Examining the Effect of Adding a Carer's Skill Training Intervention in Inpatients with Anorexia Nervosa. European Eating Disorders Review, 24, 122-130.

MASHEB, R. M. & GRILO, C. M. 2007. Rapid response predicts treatment outcomes in binge eating disorder: implications for stepped care. Journal of Consulting & Clinical Psychology, 75, 639-44.

MASHEB, R. M., GRILO, C. M. & ROLLS, B. J. 2011. An RCT for obesity and binge eating disorder: Low-energy-density nutrition counseling and cognitive behavioral therapy. Obesity, 19, S113-S114.

MASSON, P. C., VON RANSON, K. M., WALLACE, L. M. & SAFER, D. L. 2013. A randomized wait-list controlled pilot study of dialectical behaviour therapy guided self-help for binge eating disorder. Behaviour Research and Therapy, 51, 723-728.

MCCANN, U. D. & AGRAS, W. S. 1990. Successful treatment of nonpurging bulimia nervosa with desipramine: a double-blind, placebo-controlled study. American journal of psychiatry [Online], 147. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/673/CN-00070673/frame.html.

MCCLELLAND, J., KEKIC, M., BOZHILOVA, N., NESTLER, S., DEW, T., EYNDE, F., DAVID, A. S., RUBIA, K., CAMPBELL, I. C. & SCHMIDT, U. 2016. A randomised controlled trial of neuronavigated repetitive transcranial magnetic stimulation (rTMS) in anorexia nervosa. PloS one, 11.

MCELROY, S. L., ARNOLD, L. M., SHAPIRA, N. A., KECK, P. E., ROSENTHAL, N. R., KARIM, M. R., KAMIN, M. & HUDSON, J. I. 2003a. Topiramate in the treatment of binge eating disorder associated with obesity: a randomized, placebo-controlled trial. American journal of psychiatry [Online], 160. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/019/CN-00413019/frame.html.

MCELROY, S. L., CASUTO, L. S., NELSON, E. B., LAKE, K. A., SOUTULLO, C. A., KECK, P. E. & HUDSON, J. I. 2000. Placebo-controlled trial of sertraline in the treatment of binge eating disorder. American journal of psychiatry [Online], 157. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/089/CN-00297089/frame.html.

MCELROY, S. L., GUERDJIKOVA, A., KOTWAL, R., WELGE, J. A., NELSON, E. B., LAKE, K. A., KECK, P. E., JR. & HUDSON, J. I. 2007a. Atomoxetine in the treatment of binge-eating disorder: a randomized placebo-controlled trial. Journal of Clinical Psychiatry, 68, 390-8.

MCELROY, S. L., GUERDJIKOVA, A. I., BLOM, T. J., CROW, S. J., MEMISOGLU, A., SILVERMAN, B. L. & EHRICH, E. W. 2013. A placebo-controlled pilot study of the novel opioid receptor antagonist ALKS-33 in binge eating disorder. International Journal of Eating Disorders, 46, 239-245.

MCELROY, S. L., GUERDJIKOVA, A. I., MORI, N., BLOM, T. J., WILLIAMS, S., CASUTO, L. S. & KECK, P. E. 2015a. Armodafinil in binge eating disorder: A randomized, placebo-controlled trial. International Clinical Psychopharmacology, 30, 209-215.

MCELROY, S. L., GUERDJIKOVA, A. I., WINSTANLEY, E. L., O'MELIA, A. M., MORI, N., KECK JR, P. E. & HUDSON, J. I. 2011. Sodium oxybate in the treatment of binge eating disorder: An open-label, prospective study. International Journal of Eating Disorders, 44, 262-268.

MCELROY, S. L., HUDSON, J., FERREIRA-CORNWELL, M. C., RADEWONUK, J., WHITAKER, T. & GASIOR, M. 2016. Lisdexamfetamine Dimesylate for Adults with Moderate to Severe Binge Eating Disorder: Results of Two Pivotal Phase 3 Randomized Controlled Trials. Neuropsychopharmacology, 41, 1251-1260.

MCELROY, S. L., HUDSON, J. I., CAPECE, J. A., BEYERS, K., FISHER, A. C., ROSENTHAL, N. R. & TOPIRAMATE BINGE EATING DISORDER RESEARCH, G. 2007b. Topiramate for the treatment of binge eating disorder associated with obesity: a placebo-controlled study. Biological Psychiatry, 61, 1039-48.

MCELROY, S. L., HUDSON, J. I., MALHOTRA, S., WELGE, J. A., NELSON, E. B. & KECK, P. E., JR. 2003b. Citalopram in the treatment of binge-eating disorder: a placebo-controlled trial. J Clin Psychiatry, 64, 807-13.

MCELROY, S. L., HUDSON, J. I., MITCHELL, J. E., WILFLEY, D., FERREIRA-CORNWELL, M. C., GAO, J., WANG, J., WHITAKER, T., JONAS, J. & GASIOR, M. 2015b. Efficacy and safety of lisdexamfetamine for treatment of adults with moderate to severe binge-eating disorder: a randomized clinical trial. JAMA psychiatry, 72, 235-46.

MCELROY, S. L., KOTWAL, R., GUERDJIKOVA, A. I., WELGE, J. A., NELSON, E. B., LAKE, K. A., D'ALESSIO, D. A., KECK, P. E. & HUDSON, J. I. 2006. Zonisamide in the treatment of binge eating disorder with obesity: a randomized controlled trial. J Clin Psychiatry, 67, 1897-906.

MCINTOSH, V. V., JORDAN, J., CARTER, F. A., LUTY, S. E., MCKENZIE, J. M., BULIK, C. M., FRAMPTON, C. M. & JOYCE, P. R. 2005. Three psychotherapies for anorexia nervosa: a randomized, controlled trial. American Journal of Psychiatry, 162, 741-7.

MCINTOSH, V. V., JORDAN, J., CARTER, J. D., FRAMPTON, C. M., MCKENZIE, J. M., LATNER, J. D. & JOYCE, P. R. 2016. Psychotherapy for transdiagnostic binge eating: A randomized controlled trial of cognitive-behavioural therapy, appetite-focused cognitive-behavioural therapy, and schema therapy. Psychiatry Res, 240, 412-20.

MCIVER, S. 2010. Yoga may help manage binge eating disorder. Focus on Alternative and Complementary Therapies [Online], 15. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/648/CN-00758648/frame.html.

MEGUERDITCHIAN, C., SAMUELIAN-MASSAT, C., VALERO, R., BEGU-LE CORROLLER, A., FROMONT, I., MANCINI, J., SPARROW, J. D., POINSO, F. & VIALETTES, B. 2010. Inpatient treatment and anorexia nervosa outcomes. e-SPEN, 5, e40-e44.

MICALI, N., HAGBERG, K. W., PETERSEN, I. & TREASURE, J. L. 2013. The incidence of eating disorders in the UK in 2000-2009: findings from the General Practice Research Database. BMJ Open, 3.

MILANO, W., PETRELLA, C., SABATINO, C. & CAPASSO, A. 2004. Treatment of bulimia nervosa with sertraline: a randomized controlled trial. Advances in Therapy, 21, 232-7.

MILOS, G., SPINDLER, A., BUDDEBERG, C. & RUGGIERO, G. 2004. Eating Disorders Symptomatology and Inpatient Treatment Experience in Eating-Disordered Subjects. International Journal of Eating Disorders, 35, 161-168.

MILOS, G., SPINDLER, A., SCHNYDER, U. & FAIRBURN, C. G. 2005. Instability of eating disorder diagnoses: prospective study. Br J Psychiatry, 187, 573-8.

MISRA, M., KATZMAN, D., MILLER, K. K., MENDES, N., SNELGROVE, D., RUSSELL, M., GOLDSTEIN, M. A., EBRAHIMI, S., CLAUSS, L., WEIGEL, T., MICKLEY, D., SCHOENFELD, D. A., HERZOG, D. B. & KLIBANSKI, A. 2011. Physiologic estrogen replacement increases bone density in adolescent girls with anorexia nervosa. Journal of Bone and Mineral Research, 26, 2430-2438.

MISRA, M., KATZMAN, D. K., ESTELLA, N. M., MILLER, K. K., EDDY, K. T. & KLIBANSKI, A. 2013. Impact of estrogen replacement therapy on anxiety symptoms, body shape perception and eating attitudes in adolescent girls with anorexia nervosa. Endocrine Reviews, 1).

MITCHELL, J. E., AGRAS, S., CROW, S., HALMI, K., FAIRBURN, C. G., BRYSON, S. & KRAEMER, H. 2011. Stepped care and cognitive-behavioural therapy for bulimia nervosa: Randomised trial. British Journal of Psychiatry, 198, 391-397.

MITCHELL, J. E., CROSBY, R. D., WONDERLICH, S. A., CROW, S., LANCASTER, K., SIMONICH, H., SWAN-KREMEIER, L., LYSNE, C. & COOK MYERS, T. 2008. A randomized trial comparing the efficacy of cognitive-behavioral therapy for bulimia nervosa delivered via telemedicine versus face-to-face. Behaviour Research and Therapy, 46, 581-592.

MITCHELL, J. E., FLETCHER, L., HANSON, K., MUSSELL, M. P., SEIM, H., CROSBY, R. & AL-BANNA, M. 2001. The relative efficacy of fluoxetine and manual-based self-help in the treatment of outpatients with bulimia nervosa. Journal of clinical psychopharmacology [Online], 21. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/579/CN-00354579/frame.html.

MITCHELL, J. E., MYERS, T., CROSBY, R., O'NEILL, G., CARLISLE, J. & GERLACH, S. 2009. Health care utilization in patients with eating disorders. International Journal of Eating Disorders, 42, 571-574.

MITCHELL, J. E., PYLE, R., ECKERT, E. D. & HATSUKAMI, D. 1990a. The influence of prior alcohol and drug abuse problems on bulimia nervosa treatment outcome. Addictive Behaviors, 15, 169-73.

MITCHELL, J. E., PYLE, R. L., ECKERT, E. D., HATSUKAMI, D., POMEROY, C. & ZIMMERMAN, R. 1990b. A comparison study of antidepressants and structured intensive group psychotherapy in the treatment of bulimia nervosa. Archives of general psychiatry [Online], 47. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/558/CN-00065558/frame.html.

MITCHELL, J. E., PYLE, R. L., POMEROY, C., ZOLLMAN, M., CROSBY, R., SEIM, H., ECKERT, E. D. & ZIMMERMAN, R. 1993. Cognitive-behavioral group psychotherapy of bulimia nervosa: importance of logistical variables. International journal of eating disorders [Online], 14. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/021/CN-00098021/frame.html.

MOLINARI, E., BARUFFI, M., CROCI, M., MARCHI, S. & PETRONI, M. L. 2005. Binge eating disorder in obesity: comparison of different therapeutic strategies. Eating & Weight Disorders: EWD, 10, 154-61.

MORGAN, J. F., REID, F. & LACEY, J. H. 1999. The SCOFF questionnaire: Assessment of a new screening tool for eating disorders. British Medical Journal, 319, 1467-1468.

MOYA, T., FLEITLICH-BILYK, B., GOODMAN, R., NOGUEIRA, F. C., FOCCHI, P. S., NICOLETTI, M., PINZON, V., CORDAS, T. A. & NETO, F. L. 2005. The eating disorders

section of the development and well-being assessment (DAWBA): Development and validation. Revista Brasileira de Psiquiatria, 27, 25-31.

MUNSCH, S., BIEDERT, E., MEYER, A., MICHAEL, T., SCHLUP, B., TUCH, A. & MARGRAF, J. 2007. A randomized comparison of cognitive behavioral therapy and behavioral weight loss treatment for overweight individuals with binge eating disorder. International Journal of Eating Disorders, 40, 102-13.

MUSTAFA, R., SANTESSO, N., BROZEK, J., AKL, E., WALTER, S., NORMAN, G. J., KULASEGARAM, M., CHRISTENSEN, R., GUYATT, G. H., FALCK-YTTER, Y. & CHANG, S. 2013. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. Journal of Clinical Epidemiology, 66, 736-742.

NAKAHARA, T., NAGAI, N., TANAKA, M., MURANAGA, T., KOJIMA, S., NOZOE, S. & NARUO, T. 2006. The effects of bone therapy on tibial bone loss in young women with anorexia nervosa. International Journal of Eating Disorders, 39, 20-6.

NAUTA, H., HOSPERS, H. & JANSEN, A. 2001. One-year follow-up effects of two obesity treatments on psychological well-being and weight. British Journal of Health Psychology, 6, 271-284.

NAUTA, H., HOSPERS, H., KOK, G. & JANSEN, A. 2000. A comparison between a cognitive and a behavioral treatment for obese binge eaters and obese non-binge eaters. Behavior therapy [Online], 31. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/984/CN-00342984/frame.html.

NCCMH. 2015. Access and Waiting Time Standard for Children and Young People with an Eating Disorder. In: HEALTH, N. C. C. F. M. (ed.). Royal Collge of Psychiatry: Commisioned by NHS England.

NEVONEN, L. & BROBERG, A. G. 2005. A comparison of sequenced individual and group psychotherapy for eating disorder not otherwise specified. European eating disorders review [Online], 13. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/663/CN-00575663/frame.html.

NICE 2008b. Social Value Judgements. Principles for the Development of NICE Guidance. 2nd edition. London, UK: National Institute of Health and Care Excellence.

NICE 2012a. The Guidelines Manual. London, UK: National Institute of Health and Care Excellence.

NICE 2012b. Methods for the development of NICE public health guidance. Third Edition ed. London, UK: National Institute of Health and Care Excellence.

NICE 2014a. The Guideline Manual. London, UK: National Institute of Health and Care Excellence.

NICE 2014b. Process and methods guides. Developing NICE guidelines: the manual. . London, UK: National Institute of Health and Care Excellence.

NICKEL, C., TRITT, K., MUEHLBACHER, M., PEDROSA GIL, F., MITTERLEHNER, F. O., KAPLAN, P., LAHMANN, C., LEIBERICH, P. K., KRAWCZYK, J., KETTLER, C., ROTHER, W. K., LOEW, T. H. & NICKEL, M. K. 2005. Topiramate treatment in bulimia nervosa patients: a randomized, double-blind, placebo-controlled trial. International Journal of Eating Disorders, 38, 295-300.

O'BRIEN, J. A. & WARD, A. 2003. Cost of providing acute hospitalization for patients with anorexia nervosa and bulimia. Drug Benefit Trends, 15, 37-42.

O'CONNOR, G., NICHOLLS, D., HUDSON, L. & SINGHAL, A. 2016a. Refeeding Low Weight Hospitalized Adolescents With Anorexia Nervosa: A Multicenter Randomized Controlled Trial. Nutr Clin Pract.

O'CONNOR, K. L., SCISCO, J. L., SMITH, T. J., YOUNG, A. J., MONTAIN, S. J., PRICE, L. L., LIEBERMAN, H. R. & KARL, J. P. 2016b. Altered Appetite-Mediating Hormone Concentrations Precede Compensatory Overeating After Severe, Short-Term Energy Deprivation in Healthy Adults. The Journal of nutrition, 146, 209-217.

OLMSTED, M. P., DANEMAN, D., RYDALL, A. C., LAWSON, M. L. & RODIN, G. 2002. The effects of psychoeducation on disturbed eating attitudes and behavior in young women with type 1 diabetes mellitus. International Journal of Eating Disorders, 32, 230-239.

OLMSTED, M. P., DAVIS, R., GARNER, D. M., EAGLE, M., ROCKERT, W. & IRVINE, M. J. 1991. Efficacy of a brief group psychoeducational intervention for bulimia nervosa. Behav Res Ther, 29, 71-83.

PAGOTO, S., BODENLOS, J. S., KANTOR, L., GITKIND, M., CURTIN, C. & MA, Y. 2007. Association of major depression and binge eating disorder with weight loss in a clinical setting. Obesity, 15, 2557-2559.

PALMER, R. L., BIRCHALL, H., MCGRAIN, L. & SULLIVAN, V. 2002. Self-help for bulimic disorders: A randomised controlled trial comparing minimal guidance with face-to-face or telephone guidance. British Journal of Psychiatry, 181, 230-235.

PENDLETON, V. R., GOODRICK, G., POSTON, W. S., REEVES, R. S. & FOREYT, J. P. 2002. Exercise augments the effects of cognitive-behavioral therapy in the treatment of binge eating. International Journal of Eating Disorders, 31, 172-184.

PETERSON, C. B., MITCHELL, J. E., CROW, S. J., CROSBY, R. D. & WONDERLICH, S. A. 2009. The efficacy of self-help group treatment and therapist-led group treatment for binge eating disorder. American Journal of Psychiatry, 166, 1347-54.

PETERSON, C. B., MITCHELL, J. E., ENGBLOOM, S., NUGENT, S., MUSSELL, M. P., CROW, S. J. & THURAS, P. 2001. Self-help versus therapist-led group cognitive-behavioral treatment of binge eating disorder at follow-up. International journal of eating disorders [Online], 30. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/013/CN-00369013/frame.html.

PIKE, K. M., WALSH, B. T., VITOUSEK, K., WILSON, G. T. & BAUER, J. 2003. Cognitive behavior therapy in the posthospitalization treatment of anorexia nervosa. American journal of psychiatry [Online], 160. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/603/CN-00458603/frame.html.

POHJOLAINEN, V., RASANEN, P., ROINE, R. P., SINTONEN, H., WAHLBECK, K. & KARLSSON, H. 2010. Cost-utility of treatment of bulimia nervosa. International Journal of Eating Disorders, 43, 596-602.

POPE, H. G., KECK, P. E., MCELROY, S. L. & HUDSON, J. I. 1989. A placebo-controlled study of trazodone in bulimia nervosa. Journal of clinical psychopharmacology [Online], 9. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/001/CN-00062001/frame.html.

PSYCHIATRY, R. C. O. 2012. CR168. Junior MARSIPAN: Management of Really Sick Patients under 18 with Anorexia Nervosa.

PSYCHIATRY, R. C. O. 2014. CR189. MARSIPAN: Management of Really Sick Patients with Anorexia Nervosa [Online]. Available: http://www.rcpsych.ac.uk/usefulresources/publications/collegereports/cr/cr189.aspx [Accessed].

RAMSAY, R., WARD, A., TREASURE, J. & RUSSELL, G. F. M. 1999. Compulsory treatment in anorexia nervosa: Short-term benefits and long- term mortality. British Journal of Psychiatry, 175, 147-153.

REDGRAVE, G. W., COUGHLIN, J. W., SCHREYER, C. C., MARTIN, L. M., LEONPACHER, A. K., SEIDE, M., VERDI, A. M., PLETCH, A. & GUARDA, A. S. 2015. Refeeding and weight restoration outcomes in anorexia nervosa: Challenging current guidelines. International Journal of Eating Disorders, 48, 866-873.

REEVES, R. S., MCPHERSON, R. S., NICHAMAN, M. Z., HARRIST, R. B., FOREYT, J. P. & GOODRICK, G. K. 2001. Nutrient intake of obese female binge eaters. Journal of the American Dietetic Association [Online], 101. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/432/CN-00328432/frame.html.

RICCA, V., CASTELLINI, G., MANNUCCI, E., LO SAURO, C., RAVALDI, C., ROTELLA, C. M. & FARAVELLI, C. 2010. Comparison of individual and group cognitive behavioral therapy for binge eating disorder. A randomized, three-year follow-up study. Appetite, 55, 656-65.

RICCA, V., MANNUCCI, E., MEZZANI, B., MORETTI, S., BERNARDO, M., BERTELLI, M., ROTELLA, C. M. & FARAVELLI, C. 2001a. Fluoxetine and fluvoxamine combined with individual cognitive-behaviour therapy in binge eating disorder: a one-year follow-up study. Psychotherapy and psychosomatics [Online], 70. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/026/CN-00369026/frame.html.

RICCA, V., MANNUCCI, E., MEZZANI, B., MORETTI, S., DI BERNARDO, M., BERTELLI, M., ROTELLA, C. M. & FARAVELLI, C. 2001b. Fluoxetine and fluvoxamine combined with individual cognitive-behaviour therapy in binge eating disorder: A one-year follow-up study. Psychotherapy and Psychosomatics, 70, 298-306.

RIGAUD, D., BOULIER, A., TALLONNEAU, I., BRINDISI, M. C. & ROZEN, R. 2010. Body fluid retention and body weight change in anorexia nervosa patients during refeeding. Clinical Nutrition, 29, 749-755.

RIGAUD, D., BRAYER, V., BITON-JELIC, V., PAIS, V., PENNACCHIO, H. & BRUN, J. M. 2007a. [Nasogastric tube feeding in bulimia. Controlled study with follow-up at 3 months]. Presse Medicale, 36, 1354-63.

RIGAUD, D., BRONDEL, L., POUPARD, A. T., TALONNEAU, I. & BRUN, J. M. 2007b. A randomized trial on the efficacy of a 2-month tube feeding regimen in anorexia nervosa: A 1-year follow-up study. Clin Nutr, 26, 421-9.

RIGAUD, D., PENNACCHIO, H., ROBLOT, A., JACQUET, M., TALLONNEAU, I. & VERGES, B. 2009. Tube feeding at home in anorexia nervosa patients. [French]

Efficacite de la nutrition enterale a domicile chez 60 malades ayant une anorexie mentale. Presse medicale (Paris, France : 1983), 38, 1739-1745.

RIGAUD, D. J., BRAYER, V., ROBLOT, A., BRINDISI, M. C. & VERGES, B. 2011. Efficacy of tube feeding in binge-eating/vomiting patients: A 2-month randomized trial with 1-year follow-up. Journal of Parenteral and Enteral Nutrition, 35, 356-364.

RIVAS, T., BERSABE, R., JIMENEZ, M. & BERROCAL, C. 2010. The eating attitudes test (EAT-26): Reliability and validity in spanish female samples. Spanish Journal of Psychology, 13, 1044-1056.

RO, O., REAS, D. L. & STEDAL, K. 2015. Eating Disorder Examination Questionnaire (EDE-Q) in Norwegian Adults: Discrimination between Female Controls and Eating Disorder Patients. European Eating Disorders Review, 23, 408-412.

ROBB, A. S., SILBER, T. J., ORRELL-VALENTE, J. K., VALADEZ-MELTZER, A., ELLIS, N., DADSON, M. J. & CHATOOR, I. 2002. Supplemental nocturnal nasogastric refeeding for better short-term outcome in hospitalized adolescent girls with anorexia nervosa. American Journal of Psychiatry, 159, 1347-1353.

ROBERGEAU, K., JOSEPH, J. & SILBER, T. J. 2006. Hospitalization of Children and Adolescents for Eating Disorders in the State of New York. Journal of Adolescent Health, 39, 806-810.

ROBIN, A. L., SIEGEL, P. T., MOYE, A. W., GILROY, M., DENNIS, A. B. & SIKAND, A. 1999. A controlled comparison of family versus individual therapy for adolescents with anorexia nervosa. Journal of the American Academy of Child and Adolescent Psychiatry [Online], 38. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/981/CN-00264981/frame.html.

ROMANO, S. J., HALMI, K. A., SARKAR, N. P., KOKE, S. C. & LEE, J. S. 2002. A placebo-controlled study of fluoxetine in continued treatment of bulimia nervosa after successful acute fluoxetine treatment. American journal of psychiatry [Online], 159. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/700/CN-00376700/frame.html.

RUGGIERO, G. M., LAINI, V., MAURI, M. C., FERRARI, V. M. S., CLEMENTE, A., LUGO, F., MANTERO, M., REDAELLI, G., ZAPPULLI, D. & CAVAGNINI, F. 2001. A single blind comparison of amisulpride, fluoxetine and clomipramine in the treatment of restricting anorectics. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 25, 1049-1059.

RUSSELL, G. F., SZMUKLER, G. I., DARE, C. & EISLER, I. 1987. An evaluation of family therapy in anorexia nervosa and bulimia nervosa. Archives of general psychiatry [Online], 44. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/277/CN-00051277/frame.html.

RUWAARD, J., LANGE, A., BROEKSTEEG, J., RENTERIA-AGIRRE, A., SCHRIEKEN, B., DOLAN, C. V. & EMMELKAMP, P. 2013. Online cognitive-behavioural treatment of bulimic symptoms: a randomized controlled trial. Clinical psychology & psychotherapy, 20, 308-318.

SAFER, D. L., ROBINSON, A. H. & JO, B. 2010. Outcome from a randomized controlled trial of group therapy for binge eating disorder: comparing dialectical behavior therapy adapted for binge eating to an active comparison group therapy.[Erratum appears in Behav Ther. 2010 Sep;41(3):432 Note: Robinson, Athena Hagler [added]]. Behavior Therapy, 41, 106-20.

SALERNO, L., RHIND, C., HIBBS, R., MICALI, N., SCHMIDT, U., GOWERS, S., MACDONALD, P., GODDARD, E., TODD, G., LO COCO, G. & TREASURE, J. 2016. An examination of the impact of care giving styles (accommodation and skilful communication and support) on the one year outcome of adolescent anorexia nervosa: Testing the assumptions of the cognitive interpersonal model in anorexia nervosa. Journal of Affective Disorders, 191, 230-236.

SAMNALIEV, M., NOH, H., SONNEVILLE, K. & AUSTIN, S. 2015. The economic burden of eating disorders and related mental health comorbidities: An exploratory analysis using the US Medical Expenditures Panel Survey. . Preventive medicine reports, 2, 32-34.

SANCHEZ-ORTIZ, V. C., MUNRO, C., STARTUP, H., TREASURE, J. & SCHMIDT, U. 2011. The role of email guidance in internet-based cognitive-behavioural self-care treatment for bulimia nervosa. European Eating Disorders Review, 19, 342-348.

SCHMIDT, U. 2006. A pilot randomised controlled trial of the CD-ROM based intervention for carers of people with anorexia nervosa (AN) [ISRCTN33652725]. controlled-trials.com [Online]. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/240/CN-00596240/frame.html.

SCHMIDT, U., COOPER, P. J., ESSERS, H., FREEMAN, C. P., HOLLAND, R. L., PALMER, R. L., SHUR, E., RUSSELL, G. F., BOWLER, C., COKER, S., GEDDES, J. R., MACKENZIE, F., MUNRO, J., NEWTON, R., TILLER, J., TATTERSALL, M. L., VIZE, C. & WEBSTER, J. 2004. Fluvoxamine and graded psychotherapy in the treatment of bulimia nervosa: A randomized, double-blind, placebo-controlled, multicenter study of short-term and long-term pharmacotherapy combined with a stepped care approach to psychotherapy. Journal of Clinical Psychopharmacology, 24, 549-552.

SCHMIDT, U., LEE, S., BEECHAM, J., PERKINS, S., TREASURE, J., YI, I., WINN, S., ROBINSON, P., MURPHY, R., KEVILLE, S., JOHNSON-SABINE, E., JENKINS, M., FROST, S., DODGE, L., BERELOWITZ, M. & EISLER, I. 2007a. A randomized controlled trial of family therapy and cognitive behavior therapy guided self-care for adolescents with bulimia nervosa and related disorders. American Journal of Psychiatry, 164, 591-598.

SCHMIDT, U., LEE, S., BEECHAM, J., PERKINS, S., TREASURE, J., YI, I., WINN, S., ROBINSON, P., MURPHY, R., KEVILLE, S., JOHNSON-SABINE, E., JENKINS, M., FROST, S., DODGE, L., BERELOWITZ, M. & EISLER, I. 2007b. A randomized controlled trial of family therapy and cognitive behavior therapy guided self-care for adolescents with bulimia nervosa and related disorders. American Journal of Psychiatry, 164, 591-8.

SCHMIDT, U., MAGILL, N., RENWICK, B., KEYES, A., KENYON, M., DEJONG, H., LOSE, A., BROADBENT, H., LOOMES, R., YASIN, H., WATSON, C., GHELANI, S., BONIN, E. M., SERPELL, L., RICHARDS, L., JOHNSON-SABINE, E., BOUGHTON, N., WHITEHEAD, L., BEECHAM, J., TREASURE, J. & LANDAU, S. 2015. The Maudsley Outpatient Study of Treatments for Anorexia Nervosa and Related Conditions (MOSAIC): Comparison of the Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA) with specialist supportive clinical management (SSCM) in outpatients with broadly defined anorexia nervosa: A randomized controlled trial. Journal of Consulting and Clinical Psychology, 83, 796-807.

SCHMIDT, U., OLDERSHAW, A., JICHI, F., STERNHEIM, L., STARTUP, H., MCINTOSH, V., JORDAN, J., TCHANTURIA, K., WOLFF, G., ROONEY, M., LANDAU, S. & TREASURE, J. 2012. Out-patient psychological therapies for adults with anorexia nervosa: Randomised controlled trial. British Journal of Psychiatry, 201, 392-399.

SCHREYER, C. C., COUGHLIN, J. W., MAKHZOUMI, S. H., REDGRAVE, G. W., HANSEN, J. L. & GUARDA, A. S. 2015. Perceived coercion in inpatients with anorexia nervosa: Associations with illness severity and hospital course. International Journal of Eating Disorders, No Pagination Specified.

SCHÜNEMANN, H., BROŻEK, J., GUYATT, O. & OXMAN, A. D. 2013. GRADE Handbook for grading the quality of evidence and strength of recommendations using the GRADE approach: The GRADE Working Group.

SCHUNEMANN, H. J., OXMAN, A. D., BROZEK, J., GLASZIOU, P., JAESCHKE, R., VIST, G. E., WILLIAMS, J. W., JR., KUNZ, R., CRAIG, J., MONTORI, V. M., BOSSUYT, P., GUYATT, G. H. & GROUP, G. W. 2008. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ, 336, 1106-10.

SHAPIRO, J. R., REBA-HARRELSON, L., DYMEK-VALENTINE, M., WOOLSON, S. L., HAMER, R. M. & BULIK, C. M. 2007. Feasibility and acceptability of CD-ROM-based cognitive-behavioural treatment for binge-eating disorder. European Eating Disorders Review, 15, 175-84.

SIERVO, M., BOSCHI, V., PAPA, A., BELLINI, O. & FALCONI, C. 2005. Application of the SCOFF, Eating Attitude Test 26 (EAT 26) and Eating Inventory (TFEQ) Questionnaires in young women seeking diet-therapy. Eating and weight disorders : EWD, 10, 76-82.

SMINK, F. R., VAN HOEKEN, D. & HOEK, H. W. 2012. Epidemiology of eating disorders: incidence, prevalence and mortality rates. Curr Psychiatry Rep, 14, 406-14.

SMITH, C., FOGARTY, S., TOUYZ, S., MADDEN, S., BUCKETT, G. & HAY, P. 2014. Acupuncture and acupressure and massage health outcomes for patients with anorexia nervosa: Findings from a pilot randomized controlled trial and patient interviews. Journal of Alternative and Complementary Medicine, 20, 103-112.

SPETTIGUE, W., MARAS, D., OBEID, N., HENDERSON, K. A., BUCHHOLZ, A., GOMEZ, R. & NORRIS, M. L. 2015. A Psycho-Education Intervention for Parents of Adolescents With Eating Disorders: A Randomized Controlled Trial. Eating Disorders, 23, 60-75.

STEELE, A. L. & WADE, T. D. 2008. A randomised trial investigating guided self-help to reduce perfectionism and its impact on bulimia nervosa: a pilot study. Behaviour Research & Therapy, 46, 1316-23.

STEFFEN, K., ROERIG, J., MITCHELL, J. & CROSBY, R. 2006. A survey of herbal and alternative medication use among participants with eating disorder symptoms. International Journal of Eating Disorders, 39, 741-6.

STEINHAUSEN, H. C. & JENSEN, C. M. 2015. Time trends in lifetime incidence rates of first-time diagnosed anorexia nervosa and bulimia nervosa across 16 years in a Danish nationwide psychiatric registry study. Int J Eat Disord, 48, 845-50.

STRIEGEL-MOORE, R. H., DOHM, F. A., WILFLEY, D. E., PIKE, K. M., BRAY, N. L., KRAEMER, H. C. & FAIRBURN, C. G. 2004. Toward an understanding of health services use in women with binge eating disorder. Obesity research, 12, 799-806.

STRIEGEL-MOORE, R. H., WILSON, G. T., DEBAR, L., PERRIN, N., LYNCH, F., ROSSELLI, F. & KRAEMER, H. C. 2010. Cognitive behavioral guided self-help for the treatment of recurrent binge eating. Journal of Consulting & Clinical Psychology, 78, 312-21.

STROKOSCH, G. R., FRIEDMAN, A. J., WU, S. C. & KAMIN, M. 2006. Effects of an oral contraceptive (norgestimate/ethinyl estradiol) on bone mineral density in adolescent females with anorexia nervosa: a double-blind, placebo-controlled study. Journal of Adolescent Health, 39, 819-27.

STUHLDREHER, N., WILD, B., KONIG, H.-H., KONNOPKA, A., ZIPFEL, S. & HERZOG, W. 2014. Determinants of direct and indirect costs in anorexia nervosa. International Journal of Eating Disorders [Online]. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/842/CN-01038842/frame.html.

SU, J. & BIRMINGHAM, C. 2003. Anorexia nervosa: The cost of long-term disability. . Eating Weight Disorders 76-79.

SUNDGOT-BORGEN, J., ROSENVINGE, J. H., BAHR, R. & SCHNEIDER, L. S. 2002. The effect of exercise, cognitive therapy, and nutritional counseling in treating bulimia nervosa. Medicine and Science in Sports and Exercise, 34, 190-195.

SYSKO, R., GLASOFER, D. R., HILDEBRANDT, T., KLIMEK, P., MITCHELL, J. E., BERG, K. C., PETERSON, C. B., WONDERLICH, S. A. & WALSH, B. T. 2015. The eating disorder assessment for DSM-5 (EDA-5): Development and validation of a structured interview for feeding and eating disorders. International Journal of Eating Disorders, 48, 452-463.

TAKII, M., UCHIGATA, Y., KOMAKI, G., NOZAKI, T., KAWAI, H., IWAMOTO, Y. & KUBO, C. 2003. An integrated inpatient therapy for type 1 diabetic females with bulimia nervosa: A 3-year follow-up study. Journal of Psychosomatic Research, 55, 349-356.

TANTILLO, M., MACDOWELL, S., ANSON, E., TAILLIE, E. & COLE, R. 2009. Combining supported housing and partial hospitalization to improve eating disorder symptoms,

perceived health status, and health related quality of life for women with eating disorders. Eating Disorders, 17, 385-399.

TELCH, C. F., AGRAS, W. S., ROSSITER, E. M., WILFLEY, D. & KENARDY, J. 1990. Group cognitive-behavioral treatment for the nonpurging bulimic: an initial evaluation. Journal of consulting and clinical psychology [Online], 58. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/789/CN-00071789/frame.html.

THACKWRAY, D. E., SMITH, M. C., BODFISH, J. W. & MEYERS, A. W. 1993. A comparison of behavioral and cognitive-behavioral interventions for bulimia nervosa. Journal of Consulting and Clinical Psychology, 61, 639-645.

THIELS, C., SCHMIDT, U., TREASURE, J., GARTHE, R. & TROOP, N. 1998. Guided self-change for bulimia nervosa incorporating use of a self-care manual. American Journal of Psychiatry, 155, 947-953.

THURFJELL, B., EDLUND, B., ARINELL, H., HAGGLOF, B. & ENGSTROM, I. 2003. Psychometric properties of Eating Disorder Inventory for Children (EDI-C) in Swedish girls with and without a known eating disorder. Eating & Weight Disorders: EWD, 8, 296-303.

TOBIN, D. L., BANKER, J. D., WEISBERG, L. & BOWERS, W. 2007. I know what you did last summer (and it was not CBT): a factor analytic model of international psychotherapeutic practice in the eating disorders. Int J Eat Disord, 40, 754-7.

TOUYZ, S., LE GRANGE, D., LACEY, H., HAY, P., SMITH, R., MAGUIRE, S., BAMFORD, B., PIKE, K. M. & CROSBY, R. D. 2013. Treating severe and enduring anorexia nervosa: a randomized controlled trial. Psychological medicine, 43, 2501-2511.

TOUYZ, S. W., WILLIAMS, H., MARDEN, K., KOPEC-SCHRADER, E. & BEUMONT, P. J. 1994. Videotape feedback of eating behaviour in patients with anorexia nervosa: does it normalise eating behaviour? Australian Journal of Nutrition and Dietetics [Online], 51. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/142/CN-00285142/frame.html.

TRAVISS, G. D., HEYWOOD-EVERETT, S. & HILL, A. J. 2011. Guided self-help for disordered eating: A randomised control trial. Behaviour Research and Therapy, 49, 25-31.

TREASURE, J., MURPHY, T., SZMUKLER, G., TODD, G., GAVAN, K. & JOYCE, J. 2001. The experience of caregiving for severe mental illness: a comparison between anorexia nervosa and psychosis. Soc Psychiatry Psychiatr Epidemiol, 36, 343-7.

TREASURE, J. & RUSSELL, G. 2011. The case for early intervention in anorexia nervosa: theoretical exploration of maintaining factors. Br J Psychiatry, 199, 5-7.

TREASURE, J., SCHMIDT, U., TROOP, N., TILLER, J., TODD, G., KEILEN, M. & DODGE, E. 1994. First step in managing bulimia nervosa: controlled trial of therapeutic manual. BMJ (Clinical research ed.) [Online], 308. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/116/CN-00100116/frame.html.

TREASURE, J., SCHMIDT, U., TROOP, N., TILLER, J., TODD, G. & TURNBULL, S. 1996. Sequential treatment for bulimia nervosa incorporating a self-care manual. British Journal of Psychiatry, 168, 94-98.

TREASURE, J., TODD, G., BROLLY, M., TILLER, J., NEHMED, A. & DENMAN, F. 1995. A pilot study of a randomised trial of cognitive analytical therapy vs educational behavioral therapy for adult anorexia nervosa. Behaviour research and therapy [Online], 33. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/290/CN-00114290/frame.html.

TULLOCH, S. J., LELLIOT, P., BANNISTER, D., O'HERLIHY, A. & AYTON, A. 2008. The Costs, Outcomes and Satisfaction for Inpatient Child and Adolescent Psychiatric Services

(COSI-CAPS) study. National Co-ordinating Centre for NHS Service Delivery and Organisation R&D (NCCSDO).

TURNER, H., MARSHALL, E., STOPA, L. & WALLER, G. 2015. Cognitive-behavioural therapy for outpatients with eating disorders: Effectiveness for a transdiagnostic group in a routine clinical setting. Behaviour Research and Therapy, 68, 70-75.

TURNER, H., MARSHALL, E., WOOD, F., STOPA, L. & WALLER, G. 2016. CBT for eating disorders: The impact of early changes in eating pathology on later changes in personality pathology, anxiety and depression. Behav Res Ther, 77, 1-6.

TURNER, H., TATHAM, M., LANT, M., MOUNTFORD, V. A. & WALLER, G. 2014. Clinicians' concerns about delivering cognitive-behavioural therapy for eating disorders. Behaviour research and therapy, 57, 38-42.

VAN DEN EYNDE, F., CLAUDINO, A. M., MOGG, A., HORRELL, L., STAHL, D., RIBEIRO, W., UHER, R., CAMPBELL, I. & SCHMIDT, U. 2010. Repetitive transcranial magnetic stimulation reduces cue-induced food craving in bulimic disorders. Biological Psychiatry, 67, 793-5.

VAN HOEKEN, D., SEIDELL, J. & HOEK, H. W. 2003. Epidemiology. , John Wiley & Sons.

VANDER WAL, J. S., STEIN, R. I. & BLASHILL, A. J. 2011. The EDE-Q, BULIT-R, and BEDT as self-report measures of binge eating disorder. Eating Behaviors, 12, 267-271.

VANDEREYCKEN, W. & VANSTEENKISTE, M. 2009. Let eating disorder patients decide: Providing choice may reduce early drop-out from inpatient treatment. European Eating Disorders Review, 17, 177-183.

VENTURA, M. & BAUER, B. 1999. Empowerment of women with purging-type bulimia nervosa through nutritional rehabilitation. Eating and weight disorders [Online], 4. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/817/CN-00327817/frame.html.

WAGNER, G., PENELO, E., NOBIS, G., MAYERHOFER, A., SCHAU, J., SPITZER, M., IMGART, H. & KARWAUTZ, A. 2013. Is technology assisted guided self-help successful in treating female adolescents with bulimia nervosa? Neuropsychiatrie, 27, 66-73.

WALLER, G. 1992. Bulimic attitudes in different eating disorders: Clinical utility of the BITE. International Journal of Eating Disorders, 11, 73-78.

WALLER, G. 2009. Evidence-based treatment and therapist drift. Behav Res Ther, 47, 119-27.

WALLER, G. 2016. Treatment Protocols for Eating Disorders: Clinicians' Attitudes, Concerns, Adherence and Difficulties Delivering Evidence-Based Psychological Interventions. Current Psychiatry Reports, 18, 1-8.

WALLER, G., D'SOUZA WALSH, K. & WRIGHT, C. 2016. Impact of education on clinicians' attitudes to exposure therapy for eating disorders. Behaviour Research and Therapy, 76, 76-80.

WALLER, G., GRAY, E., HINRICHSEN, H., MOUNTFORD, V., LAWSON, R. & PATIENT, E. 2014a. Cognitive-behavioral therapy for bulimia nervosa and atypical bulimic nervosa: Effectiveness in clinical settings. International Journal of Eating Disorders, 47, 13-17.

WALLER, G., MICALI, N. & JAMES, A. 2014b. General Practitioners are poor at identifying the eating disorders. Adv Eat Disord, 2, 146-157.

WALLER, G., STRINGER, H. & MEYER, C. 2012. What cognitive behavioral techniques do therapists report using when delivering cognitive behavioral therapy for the eating disorders? J Consult Clin Psychol, 80, 171-5.

WALSH, B. T., FAIRBURN, C. G., MICKLEY, D., SYSKO, R. & PARIDES, M. K. 2004. Treatment of Bulimia Nervosa in a Primary Care Setting. American Journal of Psychiatry, 161, 556-561.

WALSH, B. T., GLADIS, M., ROOSE, S. P., STEWART, J. W. & GLASSMAN, A. H. 1987. A controlled trial of phenelzine in bulimia. Psychopharmacology bulletin [Online], 23. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/855/CN-00048855/frame.html.

WALSH, B. T., HADIGAN, C. M., DEVLIN, M. J., GLADIS, M. & ROOSE, S. P. 1991. Long-term outcome of antidepressant treatment for bulimia nervosa. American journal of psychiatry [Online], 148. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/825/CN-00077825/frame.html.

WALSH, B. T., KAPLAN, A. S., ATTIA, E., OLMSTED, M., PARIDES, M., CARTER, J. C., PIKE, K. M., DEVLIN, M. J., WOODSIDE, B., ROBERTO, C. A. & ROCKERT, W. 2006. Fluoxetine after weight restoration in anorexia nervosa: a randomized controlled trial. Jama [Online], 295. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/966/CN-00556966/frame.html.

WALSH, B. T., WILSON, G. T., LOEB, K. L., DEVLIN, M. J., PIKE, K. M., ROOSE, S. P., FLEISS, J. & WATERNAUX, C. 1997. Medication and psychotherapy in the treatment of bulimia nervosa. American journal of psychiatry [Online], 154. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/289/CN-00138289/frame.html.

WARD, A., RAMSAY, R., RUSSELL, G. & TREASURE, J. 2015. Follow-up mortality study of compulsorily treated patients with anorexia nervosa. International Journal of Eating Disorders, 48, 860-865.

WATSON, T. L., BOWERS, W. A. & ANDERSEN, A. E. 2000. Involuntary treatment of eating disorders. American Journal of Psychiatry, 157, 1806-1810.

WHITE, M. A. & GRILO, C. M. 2013. Bupropion for overweight women with binge-eating disorder: A randomized, double-blind, placebo-controlled trial. Journal of Clinical Psychiatry, 74, 400-406.

WHITING, P. F., RUTJES, A. W., WESTWOOD, M. E., MALLETT, S., DEEKS, J. J., REITSMA, J. B., LEEFLANG, M. M., STERNE, J. A., BOSSUYT, P. M. & GROUP, Q.-. 2011. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med, 155, 529-36.

WHITNEY, J., MURPHY, T., LANDAU, S., GAVAN, K., TODD, G., WHITAKER, W. & TREASURE, J. 2012. A practical comparison of two types of family intervention: An exploratory RCT of family day workshops and individual family work as a supplement to inpatient care for adults with anorexia nervosa. European Eating Disorders Review, 20, 142-150.

WILFLEY, D. E., AGRAS, W., TELCH, C. F., ROSSITER, E. M., SCHNEIDER, J. A., COLE, A. G., SIFFORD, L. & RAEBURN, S. D. 1993. Group cognitive-behavioral therapy and group interpersonal psychotherapy for the nonpurging bulimic individual: A controlled comparison. Journal of Consulting and Clinical Psychology, 61, 296-305.

WILFLEY, D. E., WELCH, R. R., STEIN, R. I., SPURRELL, E. B., COHEN, L. R., SAELENS, B. E., DOUNCHIS, J. Z., FRANK, M. A., WISEMAN, C. V. & MATT, G. E. 2002. A randomized comparison of group cognitive-behavioral therapy and group interpersonal psychotherapy for the treatment of overweight individuals with binge-eating disorder. Archives of general psychiatry [Online], 59. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/603/CN-00390603/frame.html.

WILLIAMSON, D. A., THAW, J. M. & VARNADO-SULLIVAN, P. J. 2001. Cost-effectiveness analysis of a hospital-based cognitive-behavioral treatment program for eating disorders. Behavior Therapy, 32, 459-477.

WILSON, G. T. 1996. Manual-based treatments: the clinical application of research findings. Behav Res Ther, 34, 295-314.

WILSON, G. T., WILFLEY, D. E., AGRAS, W. S. & BRYSON, S. W. 2010. Psychological treatments of binge eating disorder. Archives of General Psychiatry, 67, 94-101.

WOLF, E. M. & CROWTHER, J. H. 1992. An evaluation of behavioral and cognitive-behavioral group interventions for the treatment of bulimia nervosa in women. Int J Eating Disord [Online], 11. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/522/CN-00196522/frame.html.

ZEECK, A., HERZOG, T. & HARTMANN, A. 2004. Day clinic or inpatient care for severe bulimia nervosa? European Eating Disorders Review, 12, 79-86.

ZEECK, A., MAIER, J., HARTMANN, A., WETZLER-BURMEISTER, E., WIRSCHING, M. & HERZOG, T. 2008. Treatment components of inpatient and day clinic treatment of anorexia nervosa: The patient's perspective. PPmP: Psychotherapie Psychosomatik Medizinische Psychologie, 58, No Pagination Specified.

ZEECK, A., WEBER, S., SANDHOLZ, A., JOOS, A. & HARTMANN, A. 2011. Stability of long-term outcome in bulimia nervosa: A 3-year follow-up. Journal of Clinical Psychology, 67, 318-327.

ZEECK, A., WEBER, S., SANDHOLZ, A., WETZLER-BURMEISTER, E., WIRSCHING, M. & HARTMANN, A. 2009a. Inpatient versus day clinic treatment for Bulimia Nervosa: A randomized trial. Psychotherapy and Psychosomatics, 78, 152-160.

ZIPFEL, S., WILD, B., GROB, G., FRIEDERICH, H. C., TEUFEL, M., SCHELLBERG, D., GIEL, K. E., DE ZWAAN, M., DINKEL, A., HERPERTZ, S., BURGMER, M., LOWE, B., TAGAY, S., WIETERSHEIM, J. V., ZEECK, A., SCHADE-BRITTINGER, C., SCHAUENBURG, H. & HERZOG, W. 2014. Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): Randomised controlled trial. The Lancet, 383, 127-137.

12 Glossary

Term	Definition
Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.
Amenorrhea	An abnormal absence of menstruation.
Arm (of a clinical study)	Subsection of individuals within a study who receive one particular intervention, for example placebo arm.
Association	Statistical relationship between 2 or more events, characteristics or other variables. The relationship may or may not be causal.
Attrition bias	Systematic differences between comparison groups for withdrawal or exclusion of participants from a study.
AUC	Area under the curve
Available case analysis (ACA)	Analysis of data that is available for participants at the end of follow up.
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable) with which subsequent results are compared.
Before-and-after study	A study that investigates the effects of an intervention by measuring particular characteristics of a population both before and after taking the intervention, and assessing any change that occurs.
Bias	Influences on a study that can make the results look better or worse than they really are. Bias can occur by chance, deliberately or as a result of systematic errors in the design and execution of a study. It can also occur at different stages in the research process, for example during the collection, analysis, interpretation, publication or review of research data. For examples see Confounding factor, Performance bias, Publication bias Selection bias.
Carer (caregiver)	Someone who looks after family, partners or friends in need of help because they are ill, frail or have a disability.
Case-control study	A study to find out the cause(s) of a disease or condition. This is done by comparing a group of patients who have the disease or condition (cases) with a group of people who do not have it (controls) but who are otherwise as similar as possible (in characteristics thought to be unrelated to the causes of the disease or condition). This means the researcher can look for aspects of their lives that differ to see if they may cause the condition. Such studies are retrospective because they look back in time from the outcome to the possible causes of a disease or condition.
Case series	Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients.
Clinical audit	A systematic process for setting and monitoring standards of clinical care. Whereas ‘guidelines’ define what the best clinical practice should be, ‘audit’ investigates whether best practice is being carried out. Clinical audit can be described as a cycle or spiral. Within the cycle there are stages that follow a systematic process of establishing best practice, measuring care against specific criteria, taking action to improve care and monitoring to sustain improvement. The spiral suggests that as the process continues, each cycle aspires to a higher level of quality.
Clinical effectiveness	How well a specific test or treatment works when used in the 'real world' (for example when used by a doctor with a patient at home), rather than in a carefully controlled clinical trial. Trials that assess clinical effectiveness are sometimes called management trials. Clinical effectiveness is not the same as efficacy.
Clinical efficacy	The extent to which an intervention is active when studied under

Term	Definition
	controlled research conditions.
Clinician	A healthcare professional who provides patient care. For example a doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of RCTs prepared by the Cochrane Collaboration).
Cohort study	A study with 2 or more groups of people – cohorts – with similar characteristics. One group receives a treatment, is exposed to a risk factor or has a particular symptom and the other group does not. The study follows their progress over time and records what happens.
Comorbidity	A disease or condition that someone has in addition to the health problem being studied or treated.
Concealment of allocation	The process used to ensure that the person deciding to enter a participant into an RCT does not know the comparison group into which that individual will be allocated. This is distinct from blinding and is aimed at preventing selection bias. Some attempts at concealing allocation are more prone to manipulation than others and the method of allocation concealment is used as an assessment of the quality of a trial.
Confidence interval (CI)	There is always some uncertainty in research. This is because a small group of patients is studied to predict the effects of a treatment on the wider population. The confidence interval is a way of expressing how certain we are about the findings from a study, using statistics. It gives a range of results that is likely to include the 'true' value for the population. The CI is usually stated as '95% CI', which means that the range of values has a 95 in 100 chance of including the 'true' value. For example, a study may state that “based on our sample findings, we are 95% certain that the 'true' population blood pressure is not higher than 150 and not lower than 110”. In such a case the 95% CI would be 110 to 150. A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment – often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example if a large number of patients have been studied).
Confounding factor	Something that influences a study and can result in misleading findings if it is not understood or appropriately dealt with. For example, a study of heart disease may look at a group of people who exercise regularly and a group who do not exercise. If the ages of the people in the 2 groups are different, then any difference in heart disease rates between the 2 groups could be because of age rather than exercise. Therefore age is a confounding factor.
Continuous outcome	Data with a potentially infinite number of possible values within a given range. Height, weight and blood pressure are examples of continuous variables.
Control group	A group of people in a study who do not receive the treatment or test being studied. Instead, they may receive the standard treatment (sometimes called 'usual care') or a dummy treatment (placebo). The results for the control group are compared with those for a group receiving the treatment being tested. The aim is to check for any differences. Ideally, the people in the control group should be as similar as possible to those in the treatment group, to make it as easy as possible to detect any effects due to the treatment.
Cost–benefit analysis (CBA)	Cost-benefit analysis is one of the tools used to carry out an economic evaluation. The costs and benefits are measured using the same monetary units (for example UK pounds) to see whether the benefits exceed the costs.

Term	Definition
Cost–consequence analysis (CCA)	Cost-consequence analysis is one of the tools used to carry out an economic evaluation. This compares the costs (such as treatment and hospital care) with the consequences (such as health outcomes) of a test or treatment with a suitable alternative. Unlike cost–benefit analysis or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (such as the quality adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out.
Cost-effectiveness analysis (CEA)	Cost-effectiveness analysis is one of the tools used to carry out an economic evaluation. The benefits are expressed in non-monetary terms related to health, such as symptom-free days, heart attacks avoided, deaths avoided or life years gained (that is, the number of years by which life is extended as a result of the intervention).
Cost-effectiveness model	An explicit mathematical framework which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.
Cost–utility analysis (CUA)	Cost–utility analysis is one of the tools used to carry out an economic evaluation. The benefits are assessed in terms of both quality and duration of life, and expressed as quality adjusted life years (QALYs). See also Utility.
COX proportional hazard model	In survival analysis, a statistical model that asserts that the effect of the study factors (for example the intervention of interest) on the hazard rate (the risk of occurrence of an event) in the study population is multiplicative and does not change over time.
Credible interval (CrI)	The Bayesian equivalent of a confidence interval.
Decision analysis	An explicit quantitative approach to decision-making under uncertainty, based on evidence from research. This evidence is translated into probabilities, and then into diagrams or decision trees which direct the clinician through a succession of possible scenarios, actions and outcomes.
Dichotomous outcomes	Outcome that can take one of 2 possible values, such as dead/alive, smoker/non-smoker, present/not present (also called binary data).
Discounting	Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present.
Dominance	A health economics term. When comparing tests or treatments, an option that is both less effective and costs more is said to be 'dominated' by the alternative.
Drop-out	A participant who withdraws from a trial before the end.
Economic evaluation	An economic evaluation is used to assess the cost effectiveness of healthcare interventions (that is, to compare the costs and benefits of a healthcare intervention to assess whether it is worth doing). The aim of an economic evaluation is to maximise the level of benefits – health effects – relative to the resources available. It should be used to inform and support the decision-making process; it is not supposed to replace the judgement of healthcare professionals. There are several types of economic evaluation: cost–benefit analysis, cost–consequence analysis, cost-effectiveness analysis, cost-minimisation analysis and cost–utility analysis. They use similar methods to define and evaluate costs, but differ in the way they estimate the benefits of a particular drug, programme or intervention.
Effect (as in effect	A measure that shows the magnitude of the outcome in 1 group

Term	Definition
measure, treatment effect, estimate of effect, effect size)	compared with that in a control group. For example, if the absolute risk reduction is shown to be 5% and it is the outcome of interest, the effect size is 5%. The effect size is usually tested, using statistics, to find out how likely it is that the effect is a result of the treatment and has not just happened by chance.
Effectiveness	How beneficial a test or treatment is under usual or everyday conditions.
Efficacy	How beneficial a test, treatment or public health intervention is under ideal conditions (for example in a laboratory).
Ego-syntonic beliefs	Beliefs, values, and feelings consistent with one's sense of self
Epidemiological study	The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (for example infection, diet) and interventions.
EQ-5D (EuroQol 5 dimensions)	A standardised instrument used to measure health-related quality of life. It provides a single index value for health status.
Equivalence study	A trial designed to determine whether the response to 2 or more treatments differs by an amount that is clinically unimportant. This is usually demonstrated by showing that the true treatment difference is likely to lie between a lower and an upper equivalence level of clinically acceptable differences.
Evidence	Information on which a decision or guidance is based. Evidence is obtained from a range of sources including RCTs, observational studies, expert opinion (of clinical professionals or patients).
Exclusion criteria (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Extended dominance	If Option A is both more clinically effective than Option B and has a lower cost per unit of effect when both are compared with a do-nothing alternative, then Option A is said to have extended dominance over Option B. Option A is therefore more cost effective and should be preferred, other things remaining equal.
Extrapolation	An assumption that the results of studies of a specific population will also hold true for another population with similar characteristics.
False negative	A diagnostic test result that incorrectly indicates that an individual does not have the disease of interest, when they do actually have it.
False positive	A diagnostic test result that incorrectly indicates that an individual has the disease of interest, when they actually do not have it.
Fixed-effect model	In meta-analysis, a model that calculates a pooled effect estimate using the assumption that all observed variation between studies is caused by random sample variability. Studies are assumed to estimating the same overall effect.
Follow up	Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables.
Forest plot	A graphical representation of the individual results of each study included in a meta-analysis together with the combined meta-analysis result. The plot also allows readers to see the heterogeneity among the results of the studies. The results of individual studies are shown as squares centred on each study's point estimate. A horizontal line runs through each square to show each study's confidence interval. The overall estimate from the meta-analysis and its confidence interval are shown at the bottom, represented as a diamond. The centre of the diamond represents the pooled point estimate, and its horizontal tips represent the confidence interval.

Term	Definition
Generalisability	The extent to which the results of a study hold true for groups that did not participate in the research.
Gold standard	A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease.
GRADE, GRADE profile	A system developed by the GRADE Working Group to address the short-comings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.
Guided self-help	This term is used interchangeably with self-help with support
Harms	Adverse effects of an intervention.
Hazard ratio	A hazard is the rate at which events happen, so that the probability of an event happening in a short time interval is the length of time multiplied by the hazard. Although the hazard may vary with time, the assumption in proportional hazard models for survival analysis is that the hazard in one group is a constant proportion of the hazard in the other group. This proportion is the hazard ratio.
Health economics	Study or analysis of the cost of using and distributing healthcare resources.
Health-related quality of life (HRQoL)	A measure of the effects of an illness to see how it affects someone's day-to-day life.
Heterogeneity	The term is used in meta-analyses and systematic reviews to describe when the results of a test or treatment (or estimates of its effect) differ
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect.
Incidence	The incidence of a disease is the rate at which new cases occur in a population during a specified period.
Inclusion criteria (clinical study)	Specific criteria that define who is eligible to participate in a clinical study.
Inclusion criteria (literature review)	Explicit criteria used to decide which studies should be considered as potential sources of evidence.
Incremental cost	The extra cost linked to using one test or treatment rather than another. Or the additional cost of doing a test or providing a treatment more frequently.
Incremental cost effectiveness ratio (ICER)	The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest for one treatment compared with another.
Incremental net benefit (INB)	The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000×QALYs gained) minus incremental cost.
Indirectness	The available evidence is different to the review question being addressed, in terms of population, intervention, comparison and outcome (PICO).
Intention-to-treat analysis (ITT)	An assessment of the people taking part in a clinical trial, based on the group they were initially (and randomly) allocated to. This is regardless of whether or not they dropped out, fully complied with the treatment or switched to an alternative treatment. Intention-to-treat analyses are often used to assess clinical effectiveness because they mirror actual practice: that is, not everyone complies with treatment and the treatment people receive may be changed according to how they respond to it.
Intervention	In medical terms this could be a drug treatment, surgical procedure,

Term	Definition
	diagnostic or psychological therapy. Examples of public health interventions could include action to help someone to be physically active or to eat a more healthy diet.
Kappa statistic	A statistical measure of inter-rater agreement that takes into account the agreement occurring by chance
Length of stay	The total number of days a patient stays in hospital.
Licence	See Product licence.
Life years gained	Mean average years of life gained per person as a result of the intervention compared with an alternative intervention.
Likelihood ratio	The likelihood ratio combines information about the sensitivity and specificity. It tells you how much a positive or negative result changes the likelihood that a patient would have the disease. The likelihood ratio of a positive test result (LR+) is sensitivity divided by (1 minus specificity).
Loss to follow up	Patients who have withdrawn from the clinical trial at the point of follow up.
Markov model	A method for estimating long-term costs and effects for recurrent or chronic conditions, based on health states and the probability of transition between them within a given time period (cycle).
Mean	An average value, calculated by adding all the observations and dividing by the number of observations.
Mean difference	In meta-analysis, a method used to combine measures on continuous scales (such as weight), where the mean, standard deviation and sample size in each group are known. The weight given to the difference in means from each study (for example how much influence each study has on the overall results of the meta-analysis) is determined by the precision of its estimate of effect.
Median	The value of the observation that comes half-way when the observations are ranked in order.
Meta-analysis	A method often used in systematic reviews. Results from several studies of the same test or treatment are combined to estimate the overall effect of the treatment.
Minimal important difference (MID)	Threshold for clinical importance which represents the minimal important difference for benefit or for harm; for example the threshold at which drug A is less effective than drug B by an amount that is clinically important to patients.
Monte Carlo	A technique used to approximate the probability of certain outcomes by running multiple simulations using random variables.
Multivariate model	A statistical model for analysis of the relationship between 2 or more predictors, (independent) variables and the outcome (dependent) variable.
Net monetary benefit (NMB)	The value (usually in monetary terms) of an intervention net of its cost. The NMB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the NMB is calculated as: (£20,000×QALYs gained) minus cost.
Network meta-analysis	Meta-analysis in which multiple treatments (that is, 3 or more) are being compared using both direct comparisons of interventions within RCTs and indirect comparisons across trials based on a common comparator.
Non-inferiority trial	A trial designed to determine whether the effect of a new treatment is not worse than a standard treatment by more than a pre-specified amount. A one-sided version of an equivalence trial.
Number needed to treat (NNT)	The average number of patients who need to be treated to get a positive outcome. For example, if the NNT is 4, then 4 patients would have to be treated to ensure 1 of them gets better. The closer the NNT is to 1, the

Term	Definition
	better the treatment. For example, if you give a stroke prevention drug to 20 people before 1 stroke is prevented, the number needed to treat is 20.
Observational study	Individuals or groups are observed or certain factors are measured. No attempt is made to affect the outcome. For example, an observational study of a disease or treatment would allow 'nature' or usual medical care to take its course. Changes or differences in one characteristic (for example whether or not people received a specific treatment or intervention) are studied without intervening. There is a greater risk of selection bias than in experimental studies.
Odds ratio (OR)	Odds are a way to represent how likely it is that something will happen (the probability). An odds ratio compares the probability of something in one group with the probability of the same thing in another. An odds ratio of 1 between 2 groups would show that the probability of the event (for example a person developing a disease, or a treatment working) is the same for both. An odds ratio greater than 1 means the event is more likely in the first group. An odds ratio less than 1 means that the event is less likely in the first group. Sometimes probability can be compared across more than 2 groups – in this case, one of the groups is chosen as the 'reference category' and the odds ratio is calculated for each group compared with the reference category. For example, to compare the risk of dying from lung cancer for non-smokers, occasional smokers and regular smokers, non-smokers could be used as the reference category. Odds ratios would be worked out for occasional smokers compared with non-smokers and for regular smokers compared with non-smokers. See also Confidence interval, Relative risk.
Opportunity cost	The loss of other healthcare programmes displaced by investment in or introduction of another intervention. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention.
Outcome	The impact that a test, treatment, policy, programme or other intervention has on a person, group or population. Outcomes from interventions to improve the public's health could include changes in knowledge and behaviour related to health, societal changes (for example a reduction in crime rates) and a change in people's health and wellbeing or health status. In clinical terms, outcomes could include the number of patients who fully recover from an illness or the number of hospital admissions, and an improvement or deterioration in someone's health, functional ability, symptoms or situation. Researchers should decide what outcomes to measure before a study begins.
p value	The p value is a statistical measure that indicates whether or not an effect is statistically significant. For example, if a study comparing 2 treatments found that one seems more effective than the other, the p value is the probability of obtaining these results by chance. By convention, if the p value is below 0.05 (that is, there is less than a 5% probability that the results occurred by chance) it is considered that there probably is a real difference between treatments. If the p value is 0.001 or less (less than a 1% probability that the results occurred by chance), the result is seen as highly significant. If the p value shows that there is likely to be a difference between treatments, the confidence interval describes how big the difference in effect might be.
Performance bias	Systematic differences between intervention groups in care provided apart from the intervention being evaluated. Blinding of study participants (both the recipients and providers of care) is used to protect against performance bias.
Placebo	A fake (or dummy) treatment given to participants in the control group of

Term	Definition
	a clinical trial. It is indistinguishable from the actual treatment (which is given to participants in the experimental group). The aim is to determine what effect the experimental treatment has had over and above any placebo effect caused because someone has received (or thinks they have received) care or attention.
Placebo effect	A beneficial (or adverse) effect produced by a placebo and not due to any property of the placebo itself.
Post-hoc analysis	Statistical analyses that are not specified in the trial protocol and are generally suggested by the data.
Power (statistical)	The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.
Prevalence	The prevalence of a disease is the proportion of a population that are cases at a point in time.
Primary care	Healthcare delivered outside hospitals. It includes a range of services provided by GPs, nurses, health visitors, midwives and other healthcare professionals and allied health professionals such as dentists, pharmacists and opticians.
Primary outcome	The outcome of greatest importance, usually the one in a study that the power calculation is based on.
Product licence	An authorisation from the Medicines and Healthcare Products Regulatory Agency (MHRA) to market a medicinal product.
Prognosis	A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.
Prospective study	A research study in which the health or other characteristic of participants is monitored (or 'followed up') for a period of time, with events recorded as they happen. This contrasts with retrospective studies.
Protocol (review)	A document written prior to commencing a review that details exactly how evidence to answer a review question will be obtained and synthesised. It defines in detail the population of interest, the interventions, the comparators/controls and the outcomes of interest (PICO).
Publication bias	Publication bias occurs when researchers publish the results of studies showing that a treatment works well and don't publish those showing it did not have any effect. If this happens, analysis of the published results will not give an accurate idea of how well the treatment works. This type of bias can be assessed by a funnel plot.
Quality of life	See Health-related quality of life.
Quality adjusted life year (QALY)	A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality-of-life. One QALY is equal to 1 year of life in perfect health. QALYS are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality-of-life score (on a scale of 0 to 1). It is often measured in terms of the person's ability to perform the activities of daily life, and freedom from pain and mental disturbance.
Random effect model	In meta-analysis, a model that calculates a pooled effect estimate using the assumption that each study is estimating a different true treatment effect due to real differences between studies. Observed variation in effects are therefore caused by a combination of random sample variability (within-study variation) and heterogeneity between studies (between-study variation). The overall effects is an average of the estimated true study effects.

Term	Definition
Randomisation	Assigning participants in a research study to different groups without taking any similarities or differences between them into account. For example, it could involve using a random numbers table or a computer-generated random sequence. It means that each individual (or each group in the case of cluster randomisation) has the same chance of receiving each intervention.
Randomised controlled trial (RCT)	A study in which a number of similar people are randomly assigned to 2 (or more) groups to test a specific drug or treatment. One group (the experimental group) receives the treatment being tested, the other (the comparison or control group) receives an alternative treatment, a dummy treatment (placebo) or no treatment at all. The groups are followed up to see how effective the experimental treatment was. Outcomes are measured at specific times and any difference in response between the groups is assessed statistically. This method is also used to reduce bias.
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Relative risk (RR)	The ratio of the risk of disease or death among those exposed to certain conditions compared with the risk for those who are not exposed to the same conditions (for example the risk of people who smoke getting lung cancer compared with the risk for people who do not smoke). If both groups face the same level of risk, the relative risk is 1. If the first group had a relative risk of 2, subjects in that group would be twice as likely to have the event happen. A relative risk of less than 1 means the outcome is less likely in the first group. Relative risk is sometimes referred to as risk ratio.
Reporting bias	See Publication bias.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.
Retrospective study	A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected.
Review question	The plan or set of steps to be followed in a study. A protocol for a systematic review describes the rationale for the review, the objectives and the methods that will be used to locate, select and critically appraise studies, and to collect and analyse data from the included studies.
Secondary care	Care provided in hospitals.
Secondary outcome	An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes.
Selection bias	Selection bias occurs if: • The characteristics of the people selected for a study differ from the wider population from which they have been drawn; or • There are differences between groups of participants in a study in terms of how likely they are to get better.
Self-help	This term is used interchangeably with self-help without support
Self-help with support	This term was used interchangeably with guided self-help
Sensitivity	How well a test detects the thing it is testing for. If a diagnostic test for a disease has high sensitivity, it is likely to pick up all cases of the disease in people who have it (that is, give a 'true positive' result). But if a test is too sensitive it will sometimes also give a positive result in people who don't have the disease (that is, give a 'false positive'). For example, if a test were developed to detect if a woman is 6 months pregnant, a very sensitive test would detect everyone who was 6 months pregnant but would probably also include those who are 5 and 7 months pregnant. If

Term	Definition
	the same test were more specific (sometimes referred to as having higher specificity), it would detect only those who are 6 months pregnant and someone who was 5 months pregnant would get a negative result (a 'true negative'). But it would probably also miss some people who were 6 months pregnant (that is, give a 'false negative'). Breast screening is a 'real-life' example. The number of women who are recalled for a second breast screening test is relatively high because the test is very sensitive. If it were made more specific, people who don't have the disease would be less likely to be called back for a second test but more women who have the disease would be missed.
Sensitivity analysis	A means of representing uncertainty in the results of an analysis. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. • One-way simple sensitivity analysis (univariate analysis) – each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. • Multi-way simple sensitivity analysis (scenario analysis) – 2 or more parameters are varied at the same time and the overall effect on the results is evaluated. • Threshold sensitivity analysis – the critical value of parameters above or below which the conclusions of the study will change are identified. • Probabilistic sensitivity analysis – probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (for example Monte Carlo simulation).
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p<0.05$).
Specificity	The proportion of true negatives that are correctly identified as such. For example, in diagnostic testing the specificity is the proportion of non-cases correctly diagnosed as non-cases. In terms of literature searching a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers. See also Sensitivity.
Stakeholder	An organisation with an interest in a topic on which NICE is developing a clinical guideline or piece of public health guidance. Organisations that register as stakeholders can comment on the draft scope and the draft guidance. Stakeholders may be: • manufacturers of drugs or equipment • national patient and carer organisations • NHS organisations • organisations representing healthcare professionals.
Standard deviation (SD)	A measure of the spread or dispersion of a set of observations, calculated as the average difference from the mean value in the sample.
Subgroup analysis	An analysis in which the intervention effect is evaluated in a defined subset of the participants in a trial, or in complementary subsets.
Systematic review	A review in which evidence from scientific studies has been identified, appraised and synthesised in a methodical way according to predetermined criteria. It may include a meta-analysis.
Time horizon	The time span over which costs and health outcomes are considered in a decision analysis or economic evaluation.
Treatment allocation	Assigning a participant to a particular arm of a trial.
True negative	A diagnostic test result that correctly indicates that an individual does not have the disease of interest when they actually do not have it.

Term	Definition
True positive	A diagnostic test result that correctly indicates that an individual has the disease of interest when they do actually have it.
Univariate	Analysis which separately explores each variable in a data set.
Utility	In health economics, a utility is the measure of the preference or value that an individual or society places upon a particular health state. It is generally a number between 0 (representing death) and 1 (perfect health). The most widely used measure of benefit in cost-utility analysis is the quality-adjusted life year, but other measures include disability-adjusted life years (DALYs) and healthy year equivalents (HYEs).

13 Abbreviations

ABW	Average body weight
AN	Anorexia nervosa
ANIS	Anorexia nervosa inventory for self-rating
AS	Anteroposterior
BED	Binge eating disorder
BITE	Bulimic investigatory test Edinburgh
BMD	Bone mineral density
BMI	Body mass index
BN	Bulimia nervosa
BT	Behavioural therapy
BT	Behavioural therapy
CAMHS	Child and adolescent mental health services
CAT	Cognitive analytic therapy
CBT	Cognitive behavioural therapy
CDI	Clinical diagnostic interview
CI	Confidence interval
COC	Combined oral contraceptive pill
CPA	Care programme approach
CRT	Cognitive remediation therapy
DAWBA	Development and well-being assessment
DBT	Dialectical behaviour therapy
DHEA	Dehydroepiandrosterone
DM,	Diabetes mellitus
DSM	Diagnostic and Statistical Manual of Mental Disorders
EAT	Eating attitudes test
ECHO	Experienced carers helping others
ED	Eating disorder
EDA-5	The eating disorder assessment for DSM-5
EDE	Eating disorder examination
EDI	Eating disorder interview
EDI-C	Eating disorder inventory for children
EDNOS	Eating disorder not otherwise stated
ESM	Emotional and social mind training
ESP	Eating disorders screen for primary care
FBT	Family based therapy
FN	Femoral neck
FPP	Focal psychodynamic psychotherapy
FT	Family therapy
FU	Follow up
GAF	Global assessment of functioning
GP	General practice
GRADE	Grading of recommendations, assessment, development and evaluation

GSH or gSH	Guided self-help
HRT	Hormone replacement therapy
ICAT	Integrative cognitive-affective therapy
ICAT	integrative cognitive affective therapy
IFT –	intensive family coaching
IFW	Individual family work
IGF	Insulin-like growth factor
IP	Inpatient
IPT	Interpersonal therapy
IPT	Interpersonal psychotherapy
ITT	Intention to treat
LS	Lumbar spine
MANTRA	Maudsely anorexia treatment for adults
MARISPAN	Management of really sick patients with anorexia nervosa
MD	Mean difference
MHA	Mental health act.
MINI	Mini-international neuropsychiatric interview
N	Number
NA	Not available
NDRI	Norepinephrine-dopamine reuptake Inhibitor
Non-Sp	Non-specialist
NR	Not reported
NRI	Norepinephrine reuptake inhibitor
OAO	Overcoming anorexia online
OSFED	Other specified feeding or eating disorder
PE	Psychoeducation
QUADAS	Quality assessment of diagnostic accuracy studies
RCT	Randomised control trial
ROC	Receiver operating characteristics
RR	Relative risk
rTMS	Repetitive transcranial magnetic stimulation
SATA	Substance abuse treatment agent
SCAN	Schedules for clinical assessment in neuropsychiatry
SCID-I	Structured clinical interview for axis I disorders
SCOFF	Sick control one fat food
Se	Sensitivity
SEED	Severe and enduring anorexia nervosa
SFT-AN	Systematic family therapy for anorexia nervosa
SH	Self help
SIAB-EX	Structured Interview for anorexic and bulimic Syndromes for expert rating
SIAB-S	Structured Interview for anorexic and bulimic syndromes- self-rated questionnaire
SMD	Standard mean difference
SNRI	Serotonin norepinephrine reuptake inhibitor.
SOS	Speed of sound

Sp	Specialist
Sp	Specificity
SPT-BN	Supportive psychotherapy for young people bulimia nervosa
SSCM	Specialist supportive clinical management
SSRI	Selective serotonin reuptake inhibitor
TAU	Treatment as usual
TCA	Tricyclic antidepressants
Vs.	Versus
WLC	Wait list control

14 Recommendations

1. Do not use screening tools (for example SCOFF) as the sole method to determine whether or not people have an eating disorder.
2. Be aware that eating disorders present in a range of settings, including:
    - primary and secondary health care
    - social care
    - education
3. Think about the possibility of an eating disorder in children and young people with poor growth (for example a low weight or height for their age).
4. Think about the possibility of an eating disorder in people with one or more of the following:
    - an unusually low or high BMI or body weight for their age
    - dieting or restrictive eating practices (such as dieting when they are underweight) that are worrying them, their family members or carers, or professionals
    - family members or carers report a change in eating behaviour
    - other mental health problems
    - a disproportionate concern about their weight (for example, concerns about weight gain as a side effect of contraceptive medication)
    - problems managing a chronic illness that affects diet, such as diabetes
    - menstrual or other endocrine disturbances, or unexplained gastrointestinal symptoms
    - physical signs of:
        - starvation, such as poor circulation, dizziness, palpitations, fainting or pallor
        - compensatory behaviours, such as laxative misuse, vomiting or excessive exercise
        - dental erosion
    - taking part in activities associated with a high risk of eating disorders (for example, professional sport, fashion, dance, or modelling).
5. When assessing for an eating disorder, think about all of the points in recommendation 3 regardless of the person's gender, ethnicity or socio-economic background.
6. Think about the possibility of an eating disorder in children and young people with poor growth (for example a low weight or height for their age).
7. Be aware that the risk of eating disorders is highest in young women (13 to 17 years), and that young men are also at greater risk between 13 and 17 years than at other ages.
8. Do not use single measures such as BMI or duration of illness to determine whether to offer treatment for an eating disorder.

9. Professionals in primary and secondary mental health settings should assess the following in people with a suspected eating disorder:
    - their physical health, including checking for any physical effects of starvation or of compensatory behaviours such as vomiting
    - the presence of mental health problems commonly associated with eating disorders, including depression, anxiety, self-harm and obsessive compulsive disorder
    - the possibility of alcohol or substance misuse.
    - the need for emergency care in people whose physical health is compromised or who have a suicide risk.
10. Be aware that people with an eating disorder may:
    - avoid contact with and find it difficult or distressing to interact with healthcare professionals, staff and other service users
    - be vulnerable to stigma and shame.
11. Ensure that people with an eating disorder and their parents or carers (as appropriate) get equal access to treatments for eating disorders, regardless of:
    - gender or gender identity (including people who are transgender)
    - sexual orientation
    - religion, belief, culture or family origin
    - where they live and who they live with
    - any mental or physical health problems or disabilities.
12. Take particular care to ensure services are well coordinated when:
    - a young person moves from children's to adult services (see the NICE guideline on transition from children's to adults' services)
    - more than one service is involved (such as inpatient and outpatient services, or when a comorbidity is being treated by a separate service)
    - people need care in different places at different times of the year (for example, university students).
13. If an eating disorder is still suspected after the initial assessment, refer without delay to:
    - a community based, age-appropriate eating disorders service for an assessment and treatment (if possible) **or**
    - day patient or inpatient services for people with clinical signs in the concern or alert ranges (see recommendations 36 and 48).
14. When assessing a person with a suspected eating disorder, find out what they and their family members or carers (as appropriate) know about eating disorders and address any misconceptions.
15. Ensure that people with an eating disorder and their parents or carers (as appropriate) understand the purpose of any meetings and the reasons for sharing information about their care with others.
16. Offer people with eating disorders and their family members or carers (as appropriate) education and information on:

- the nature and risks of their eating disorder and how it is likely to affect them
- the treatments available and their likely benefits and limitations.

17. When communicating with people with an eating disorder and their family members or carers (as appropriate):
    - check that they understand what is being said
    - be sensitive when discussing a person's weight and appearance
    - be aware that family members or carers may feel guilty and responsible for the eating disorder
    - show empathy, compassion and respect
18. For children and young people, assess the impact of their home, education, work and wider social environment on their eating disorder. Ensure that their emotional, education and social needs are met throughout treatment.
19. If appropriate, encourage family members, carers, teachers, and peers of children and young people to support them during their treatment.
20. For children with an eating disorder, consider using the treatments recommended for young people with the same eating disorder.
21. When working with people with an eating disorder and their family members or carers (as appropriate):
    - hold discussions in places where confidentiality, privacy and dignity can be respected
    - explain the limits of confidentiality (that is, which health and social care professionals have access to information about their care, and when this may be shared with others).
22. When seeking consent for assessments or treatments for children or young people under 16, respect Gillick competence if they do not want their family members or carers involved.
23. Health, social care and education professionals working with children and young people with an eating disorder should be trained and skilled in:
    - negotiating and working with parents and carers
    - managing issues around information sharing and confidentiality
    - safeguarding
    - working with multidisciplinary teams
24. Professionals who assess and treat eating disorders should be competent to do this for the age groups they care for.
25. Base the content, structure and duration of psychological treatments on relevant manuals that focus on eating disorders.
26. Professionals who provide interventions for treating eating disorders should:
    - receive appropriate supervision
    - use standardised outcome measures, for example the Eating Disorder Examination Questionnaire (EDE-Q), bulimic behaviours or weight
    - monitor their competence (for example, by using recordings and external audit and scrutiny)

- monitor treatment adherence in people who use their service.

27. Healthcare professionals assessing children and young people with eating disorders should be alert throughout assessment and treatment to signs of bullying, teasing, abuse (emotional, physical and sexual) and neglect. For guidance on when to suspect child maltreatment, see the NICE guideline on child maltreatment.
28. Provide advice and education to women with an eating disorder who plan to conceive, to increase the likelihood of conception and to reduce the risk of miscarriage. This may include information on the importance of:
    - maintaining good mental health and wellbeing
    - ensuring adequate nutrient intake and a healthy body weight
    - stopping behaviours such as bingeing, vomiting, laxatives and excessive exercise.
29. Nominate a dedicated professional (such as a GP or midwife) to monitor and support pregnant women with an eating disorder during pregnancy and in the post-natal period, because of:
    - concerns they may have specifically about gaining weight
    - possible health risks to the mother and child.
    - the high risk of mental health problems in the perinatal period
30. For guidance on providing advice to pregnant women about healthy eating and feeding their baby, see the NICE guideline on maternal and child nutrition.
31. Consider more intensive prenatal care for pregnant women with current or remitted anorexia nervosa, to ensure adequate prenatal nutrition and foetal development.
32. When prescribing medication for people with an eating disorder and comorbid mental or physical health conditions, take into account the impact malnutrition and compensatory behaviours can have on the effectiveness and the risk of side effects.
33. When prescribing for people with an eating disorder and a comorbidity, assess how the eating disorder will affect medication adherence (for example, for medication that can affect body weight).
34. When prescribing for people with an eating disorder, take account of the risks of medication that can compromise physical health because of prexisting medical complications.
35. Offer ECG monitoring for people with an eating disorder who are taking medication that can compromise cardiac functioning (for example, bradycardia below 50 beats per minute or a prolonged QT interval).
36. GPs should assess fluid and electrolyte balance in people with an eating disorder who are using compensatory behaviours, such as vomiting, taking laxatives or diuretics, or water or salt loading.
37. GPs, paediatricians or psychiatrists should think about the need for acute medical care (including emergency admission) for people with severe electrolyte imbalance, dehydration or signs of incipient organ failure.
38. For people with continued unexplained electrolyte imbalance, GPs, eating disorder specialists, paediatricians or dieticians should assess whether it could be caused by another condition.

39. For people who need supplements to restore electrolyte balance, GPs, eating disorder specialists or dieticians should offer these orally unless the person has problems with gastrointestinal absorption.
40. GPs, eating disorder specialists, paediatricians, psychiatrists or cardiologists should assess whether ECG monitoring is needed, based on the following risk factors:
    - rapid weight loss
    - excessive exercise
    - severe purging behaviours, such as laxative or diuretic use or vomiting
    - bradycardia
    - hypotension
    - excessive caffeine (including from energy drinks)
    - prescribed or non-prescribed medications
    - muscular weakening
    - electrolyte imbalance
    - previous abnormal heart rhythm.
41. GPs, eating disorder specialists or dieticians should encourage people who are vomiting to:
    - have regular dental and medical reviews
    - avoid brushing teeth immediately after vomiting
    - rinse with non-acid mouthwash after vomiting
    - avoid highly acidic foods and drinks.
42. GPs, eating disorder specialists or dieticians should advise people who are misusing laxatives:
    - that laxatives do not reduce calorie absorption and so do not help with weight loss.
    - to gradually reduce and stop laxative use.
43. For guidance on identifying, assessing and managing overweight and obesity, see the NICE guideline on obesity.
44. GPs should offer a physical and mental health review at least annually to people with anorexia nervosa who are not receiving ongoing treatment for their eating disorder. The review should include:
    - weight or BMI
    - blood pressure
    - relevant blood tests
    - mood
    - any problems with daily functioning
    - assessment of risk (related to both physical and mental health)
    - an ECG, for people with purging behaviours and/or significant weight changes
    - discussion of treatment options.

45. Monitor physical and mental health (including weight and indicators of increased risk) in people who are having psychological interventions for anorexia nervosa.
46. For people with an eating disorder and compromised physical health, consider inpatient treatment or appropriate day patient care for medical stabilisation and to initiate refeeding if these cannot be done in an outpatient setting.
47. Children and young people with an eating disorder who need inpatient treatment or day patient care should be admitted to age-appropriate facilities that are as near to their home as possible and that have the capacity to provide appropriate educational activities.
48. For people with acute mental health risk (such as suicide risk), consider psychiatric crisis care or inpatient treatment
49. When deciding whether to use day-patient or inpatient care, take the following into account:
    - the person's BMI or weight, and whether either of these are below the safe range and rapidly dropping (for example more than 1 kg per week; be aware that there is no absolute weight or BMI threshold for admission)
    - whether several medical risk parameters (such as blood tests, physical observations and ECG [for example bradycardia below 50 beats per minute or a prolonged QT interval]) have values and/or rates of change in the concern or alert ranges (refer to Box 1 in MARSIPAN or Guidance 1 and 2 in junior MARSIPAN).
    - the person's current physical health and whether this is declining
    - whether the parents or carers of children and young people can support them and keep them from significant harm.
50. If a person is admitted for physical health problems caused by an eating disorder, start or continue psychological treatments for the eating disorder if appropriate.
51. Do not use inpatient care solely to provide psychological treatment for eating disorders.
52. Inpatient services should collaborate with other teams (including the community team) and the person's family members or carers (as appropriate), to help with treatment and transition.
53. Make a care plan for each person with an eating disorder, to cover the care they need after discharge.
54. Within one month of admission, review with the referring team, the person with an eating disorder and their parents or carers (as appropriate) whether inpatient care should be continued, stepped down to a less intensive setting, or stopped.
55. As part of the review:
    - assess whether enough progress has been made towards the goals agreed at admission (such as medical progress)
    - take into account the risk that people with an eating disorder can become institutionalised, and that a lack of change in their condition could indicate that inpatient treatment is harmful
    - consider seeking an independent second opinion.

56. Reaching a healthy weight should not be used as the only reason for discharging people with an eating disorder.
57. Be aware that a key goal of treatment for anorexia nervosa is to help people reach a healthy body weight or BMI for their age.
58. When weighing people with anorexia, consider sharing the results with them and (if appropriate) their family members or carers.
59. Consider either individual eating-disorder-focused cognitive behavioural therapy (CBT-ED) or eating-disorder-focused focal psychodynamic therapy for adults with anorexia nervosa.
60. Individual CBT-ED programmes for adults with anorexia nervosa should:
    - use a CBT-ED manual.
    - consist of up to 40 sessions over 40 weeks
    - aim to reduce the risk to physical health and any other symptoms of the eating disorder.
    - encourage reaching a healthy body weight and healthy eating
    - cover nutrition, relapse prevention, cognitive restructuring, mood regulation, social skills, body image concern and self-esteem.
    - create a personalised treatment plan based on the processes that appear to be maintaining the eating problem.
    - explain the risks of starvation and being underweight.
    - enhance self-efficacy
    - include self-monitoring
    - include homework, to help the person practice what they have learned in their daily life**.**
61. Eating-disorder-focused focal psychodynamic therapy programmes for adults with anorexia nervosa should:
    - use a focal psychodynamic manual specific to eating disorders
    - consist of up to 40 sessions over 40 weeks
    - include psychoeducation about nutrition and the effects of starvation
    - make a patient-centred focal hypothesis that is specific to the individual and addresses:
        - what the symptoms mean to the person
        - how the symptoms affect the person
        - how the symptoms influence the person's relationships with others and with the therapist.
    - in the first phase, focus on developing the therapeutic alliance between the therapist and person with anorexia nervosa, addressing pro-anorexic behaviour and ego-syntonic beliefs (beliefs, values and feelings consistent with the person's sense of self) and building self-esteem
    - in the second phase, focus on relevant relationships with other people and how these affect eating behaviour

- in the final phase, focus on transferring the therapy experience to situations in everyday life and address any concern the person has about what will happen when treatment ends.

62. If individual CBT-ED or focal psychodynamic-ED is ineffective, not available or not acceptable for adults with anorexia nervosa, consider specialist supportive clinical management (SSCM) or the Maudsley Anorexia Treatment for Adults (MANTRA).

63. Consider anorexia-nervosa-focused family therapy for young people with anorexia nervosa, delivered as single- or multi-family therapy and with sessions provided either:

- separately for the young person and for their family members and carers or
- for the young person and their family together.

64. Anorexia-nervosa-focused family therapy for young people with anorexia nervosa should:

- use family-based treatment for eating disorders manual
- consist of 18 – 20 sessions over at most one year
- review the needs of the young person four weeks after treatment begins and then every three months, to establish how regular sessions should be and how long treatment should last
- emphasise the role of the family in helping the young person to recover
- not blame the young person or their family members or carers
- include psychoeducation about nutrition and the effects of starvation
- in the first phase, aim to establish a good therapeutic alliance with the young person, their parents or carers and other family members
- help the parents or carers take charge of the young person's eating and return control to the young person when they are ready
- in the final phase:
  - support the young person (with help from their parents or carers) to establish a level of independence appropriate for their level of development
  - focus on plans for when treatment ends (including any concerns the young person and their family have) and on relapse prevention.

65. Consider support for family members who are not involved in the family therapy, to help them to cope with distress caused by the condition.

66. Consider giving young people with anorexia nervosa additional appointments separate from their family members or carers.

67. If family therapy is unacceptable, contraindicated or ineffective for young people with anorexia nervosa, consider individual CBT-ED or adolescent focused eating disorder therapy.

68. Assess whether family members or carers (as appropriate) need support if the young person with anorexia nervosa is having therapy on their own.

69. Be aware that the family members or carers of a person with an eating disorder may experience severe distress. Offer them an assessment of their own needs, including:
    - what impact the eating disorder has on them
    - what support they need, including practical support and emergency plans for increasing medical or psychiatric risk.
70. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.
71. Only offer dietary counselling as part of a multidisciplinary approach.
72. Encourage people with anorexia nervosa to take an age-appropriate oral multi-vitamin and multi-mineral supplement until their diet includes enough to meet their dietary reference values.
73. Include family members or carers (as appropriate) in any dietary education or meal planning for children and young people with anorexia nervosa who are having therapy on their own.
74. Offer individualised supplementary dietary advice to children and young people with anorexia nervosa and their parents or carers (if appropriate) to help them meet their nutritional needs for growth and development (particularly during puberty).
75. Do not offer medication as the sole treatment for anorexia nervosa.
76. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitization, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
77. Explain to women with anorexia nervosa that the primary aim of prevention and treatment of a low bone mineral density is to achieve and maintain a healthy body weight or BMI for their age.
78. Do not routinely offer oral or transdermal oestrogen therapy to treat low bone mineral density in children or young people with anorexia nervosa.
79. Seek specialist paediatric or endocrinological advice before starting any hormonal treatment for a low bone mineral density. Coordinate any treatment with the eating disorders team.
80. Consider transdermal 17-β-estradiol (with cyclic progesterone) for young women (aged 13-17 years) with anorexia nervosa who have long-term low body weight and low bone mineral density with a bone age over 15.
81. Consider incremental physiological doses of oestrogen in young women (aged 13-17 years) with anorexia nervosa who have delayed puberty, long-term low body weight and low bone mineral density with a bone age under 15.
82. Consider bisphosphonates for women (18 years and over) with anorexia nervosa who have long-term low body weight and low bone mineral density. Discuss the benefits and risks (including risk of teratogenic effects) with women before starting treatment.
83. Advise people with anorexia nervosa and osteoporosis or related bone disorders to avoid high-impact physical activities and activities that significantly increase the chance of falls or fractures.
84. Offer a bone mineral density scan:
    - after 6 months of amenorrhea in young women (aged 13 to 17) and yearly after this even if the person gains weight

- after 12 months of amenorrhea in adult women (18 and above) and every 2 years after this even if the person gains weight.

Continue to offer scans until either menses has resumed or bone mineral density is within healthy limits.

85. Monitor growth and development in children and young people with anorexia nervosa who have not completed puberty (for example, not reached menarche or final height).

86. For guidance on osteoporosis risk assessment, see the NICE guideline on assessing the risk of fragility fractures in osteoporosis.

87. Seek specialist paediatric or endocrinology advice for delayed physical development or stunted growth in children and young people with an eating disorder.

88. Ensure that staff of inpatient services for people with eating disorders are trained to recognise the symptoms of refeeding syndrome and how to manage it.

89. Use a standard operating procedure for refeeding that emphasises the need to avoid under-nutrition and refeeding syndrome. Refer to existing national guidance, for example Management of Really Sick Patients with Anorexia Nervosa (MARSIPAN) and junior MARSIPAN.

90. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.

91. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Substance and medication misuse

92. For people with an eating disorder who are misusing substances, or over the counter or prescribed medication, provide treatment for the eating disorder unless the substance misuse is interfering with this treatment.

93. If substance misuse or medication is interfering with treatment, consider a multi-disciplinary approach with substance misuse services.

94. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.

95. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Diabetes

96. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.

97. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.

98. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes.

99. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.

100. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.

101. Offer people with an eating disorder who are misusing insulin the following treatment plan:

- a low carbohydrate diet, so that insulin can be started at a low level
- gradually increasing insulin doses to reduce blood glucose levels
- adjusted total glycaemic load and carbohydrate distribution to meet their individual needs and prevent rapid weight gain
- carbohydrate counting when adjusting their insulin dose (including via pumps)
- a diabetic educational intervention such as DAFNE
- education about the problems caused by misuse of diabetes medication.

102. For more guidance on managing diabetes, refer to the NICE guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults

103. Consider bulimia-nervosa-focused guided self-help for adults with bulimia nervosa.

104. Bulimia-nervosa-focused guided self-help programmes for adults with bulimia nervosa should:

- use a cognitive behavioural self-help book for eating disorders
- supplement the self-help programme with brief supportive sessions (for example four to nine sessions lasting 20 minutes each over 16 weeks running weekly at first)
- be delivered by a practitioner who is competent in delivering the treatment.

105. If bulimia-nervosa-focused guided self-help is ineffective after four weeks or is not acceptable, consider individual eating-disorder-focused cognitive behavioural therapy (CBT-ED).

106. Individual CBT-ED for adults with bulimia nervosa should:

- follow a CBT-ED manual
- consist of up to 20 sessions over 20 weeks, with sessions held twice-weekly in the first phase
- in the first phase focus on:
  - engagement and education
  - establishing a pattern of regular eating, and providing encouragement, advice and support while people do this
- follow by addressing the eating disorder psychopathology (that is, the extreme dietary restraint, the concerns about body shape and weight, and the tendency to binge in response to difficult thoughts and feelings)

- towards the end of treatment, spread appointments further apart and focus on maintaining positive changes and minimising the risk of relapse
- if appropriate, involve significant others to help with one-to-one treatment.

107. Explain to people with bulimia nervosa that psychological treatments have a limited effect on body weight.

108. Offer bulimia-nervosa-focused family therapy to young people with bulimia nervosa

109. Bulimia-nervosa-focused family therapy for young people with bulimia nervosa should:

- use a bulimia-nervosa-focused family therapy manual
- consist of 18-20 sessions over six months
- support and encourage the family to help the young person recover
- not blame the young person or their family members or carers
- include information about regulating body weight, dieting and the adverse effects of controlling weight with self-induced vomiting or laxatives
- establish a good therapeutic relationship with the young person and their family members or carers
- use a collaborative approach between the parents and the young person to establish regular eating patterns and minimize compensatory behaviours
- include regular meetings with the young person on their own throughout the treatment
- include self-monitoring of bulimic behaviours and discussions with family members or carers
- in later phases of treatment, support the young person and their family members or carers to establish a level of independence appropriate for their level of development
- in the final phase of treatment, focus on plans for when treatment ends (including any concerns the young person and their family have), and on relapse prevention.

110. If family therapy is ineffective, or is not acceptable, consider bulimia-nervosa-focused guided self-help for young people with bulimia nervosa.

111. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.

112. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitisation, weight training, yoga or warming therapy) as part of the treatment for eating disorders.

113. Do not offer medication as the sole treatment for bulimia nervosa.

114. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.

115. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Diabetes

116. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.

117. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.

118. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes.

119. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.

120. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.

121. Offer people with an eating disorder who are misusing insulin the following treatment plan:

- a low carbohydrate diet, so that insulin can be started at a low level
- gradually increasing insulin doses to reduce blood glucose levels
- adjusted total glycaemic load and carbohydrate distribution to meet their individual needs and prevent rapid weight gain
- carbohydrate counting when adjusting their insulin dose (including via pumps)
- a diabetic educational intervention such as DAFNE
- education about the problems caused by misuse of diabetes medication

122. For more guidance on managing diabetes, refer to the NICE guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults

123. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.

124. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Substance and medication misuse

125. For people with an eating disorder who are misusing substances, or over the counter or prescribed medication, provide treatment for the eating disorder unless the substance misuse is interfering with this treatment.

126. If substance misuse or medication is interfering with treatment, consider a multi-disciplinary approach with substance misuse services.

127. Offer a binge-eating-focused guided self-help programme to adults with binge eating disorder.

128. Binge-eating-focused guided self-help programmes for adults should:
    - use a cognitive behavioural self-help book
    - focus on adherence to the self-help programme
    - supplement the self-help programme with brief supportive sessions (for example four to nine sessions lasting 20 minutes each over 16 weeks that are first run weekly):
        - delivered by a practitioner who is competent in delivering the treatment
        - that focus exclusively on helping the person follow the programme.
129. If guided self-help is ineffective after four weeks or is not acceptable, offer group eating-disorder-focused cognitive behavioural therapy (CBT-ED)
130. Group CBT-ED programmes for adults with binge eating disorder should:
    - use a CBT-ED manual
    - consist of 16 weekly 90-minute group sessions over four months
    - focus on psychoeducation, self-monitoring of the eating behaviour and helping the person analyse their problems and goals
    - include making a daily food intake plan and identifying binge eating cues
    - include body exposure training and helping the person to identify and change negative beliefs about their body
    - help with avoiding relapses and coping with current and future risks and triggers.
131. Explain to people with binge eating disorder that psychological treatments aimed at treating binge eating have a limited effect on body weight and that weight loss is a post-therapy target. Refer to the NICE guideline on obesity identification, assessment and management for guidance on weight loss and bariatric surgery.
132. For young people with binge eating disorder, offer the same treatments recommended for adults with binge eating disorder.
133. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.
134. Do not offer medication as the sole treatment for binge eating disorder.
135. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitization, weight training, yoga or warming therapy) as part of the treatment for eating disorders.
136. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.
137. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Diabetes

138. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.

139. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.

140. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes

141. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.

142. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.

143. Offer people with an eating disorder who are misusing insulin the following treatment plan:

    - a low carbohydrate diet, so that insulin can be started at a low level
    - gradually increasing insulin doses to reduce blood glucose levels
    - adjusted total glycaemic load and carbohydrate distribution to meet their individual needs and prevent rapid weight gain
    - carbohydrate counting when adjusting their insulin dose (including via pumps)
    - a diabetic educational intervention such as DAFNE
    - education about the problems caused by misuse of diabetes medication.

144. For more guidance on managing diabetes, refer to the NICE guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults.

145. When deciding which order to treat an eating disorder and a comorbid mental health condition (in parallel, as part of the treatment or one after the other), take the following into account:

    - the severity and complexity of the eating disorder and comorbidity
    - the person's level of functioning
    - the patient's preference

146. Refer to the NICE guidelines on specific mental health problems for further guidance on treatment.

147. For people with OSFED, consider using the treatments for the eating disorder it most closely resembles.

148. Be aware that the family members or carers of a person with an eating disorder may experience severe distress. Offer them an assessment of their own needs, including

    - • what impact the eating disorder has on them
    - • what support they need, including practical support and emergency plans for increasing medical or psychiatric risk.

149. If appropriate, provide written information for family members or carers who cannot attend meetings with their child for assessment or treatment of an eating disorder.

150. Do not offer a physical therapy (such as transcranial magnetic stimulation, acupuncture, eye movement desensitization, weight training, yoga or warming therapy) as part of the treatment for eating disorders.

151. Eating disorder specialists and other care teams should collaborate when caring for people with physical or mental health comorbidities that may be affected by their eating disorder.

152. When collaborating, teams should use outcome measures for both the eating disorder and the physical and mental health comorbidities, to monitor the effectiveness of treatments for each condition and the potential impact they have on each other.

Diabetes

153. Eating disorder teams and diabetes teams should collaborate to explain the importance of physical health monitoring to people with an eating disorder and diabetes.

154. Consider involving family members and carers (as appropriate) in the treatment programme to help the person with blood glucose control.

155. Agree between the eating disorder and diabetes teams who has responsibility for monitoring the physical health of people with an eating disorder and diabetes.

156. Explain to the person and their diabetes team that they may need to monitor their blood glucose control more closely during the treatment for the eating disorder.

157. Address insulin misuse as part of any psychological treatments for eating disorders in people with diabetes.

158. Offer people with an eating disorder who are misusing insulin the following treatment plan:

- a low carbohydrate diet, so that insulin can be started at a low level
- gradually increasing insulin doses to reduce blood glucose levels
- adjusted total glycaemic load and carbohydrate distribution to meet their individual needs and prevent rapid weight gain
- carbohydrate counting when adjusting their insulin dose (including via pumps)
- a diabetic educational intervention such as DAFNE
- education about the problems caused by misuse of diabetes medication.

159. For more guidance on managing diabetes, refer to the NICE guidelines on type 1 and type 2 diabetes in children and young people, type 1 diabetes in adults, and type 2 diabetes in adults.

160. For people with an eating disorder who are misusing substances, or over the counter or prescribed medication, provide treatment for the eating disorder unless the substance misuse is interfering with this treatment.

161. If substance misuse or medication is interfering with treatment, consider a multi-disciplinary approach with substance misuse services.

162. If a person's physical health is at serious risk due to their eating disorder, they do not consent to treatment, and they can only be treated safely in an inpatient setting, use an appropriate legal framework for compulsory treatment (for example the Mental Health Act 1983)

163. If a child or young person lacks capacity, their physical health is at serious risk and they do not consent to treatment, ask their parents or carers to consent on their behalf and if necessary, use an appropriate legal framework for compulsory treatment (such as the Mental Health Act 1983 or the Children Act 1989).

164. Feeding people without their consent should only be done by multidisciplinary teams who are competent in doing so.

15 Research recommendations

- Research recommendation: “how effective are the current guideline recommendations in improving symptoms and remission rates for men (aged over 18 years) with an eating disorder?”
- Research recommendation: “evaluate the effectiveness of stepped care for psychological treatment of eating disorders for people of all-ages.
- Research recommendation: “what is the effectiveness of adapted treatments in those with anorexia nervosa who are not responding to treatment?”
- Research recommendation: “does exercise in addition to a recommended psychotherapy add any benefit to those with bulimia nervosa or binge eating disorder?”
- Research recommendation: “do treatments need to be modified for people of all ages with an eating disorder and a comorbidity?”
- Research recommendation: “compare the clinical and cost-effectiveness of individual eating-disorder focused cognitive behavioural therapy (CBT-ED) with guided self-help and group CBT-ED for adults with binge eating disorder, including complex binge eating disorder.”
- Research recommendation: “are shorter psychological treatment lengths equally effective compared with the treatment lengths recommended in this guideline for children, young people and adults with an eating disorder? “

16 Appendices

The appendices are contained in separate documents:

Appendix A: Scope

Appendix B: Declarations of interest

Appendix C: Special advisors

Appendix D: Stakeholders

Appendix E: Researchers contacted

Appendix F: Review questions

Appendix G: Research recommendations

Appendix H: Search strategies

Appendix I: HE search strategies

Appendix J: Excluded studies list

Appendix K: Flow charts

Appendix L: Grade evidence profiles

Appendix M: Forest plots

Appendix N: Bulimia nervosa TSU report

Appendix O: HE evidence check lists

Appendix P: HE evidence tables

Appendix Q: HE evidence profiles

Appendix R: NMA methodology and results

Appendix S: Clinical evidence of intervention for people with BN and BED