From Predictors to Processes: Exploring the Early Response Effect in the Psychological Treatment of Eating Disorders

Ammara Imtiaz

A thesis submitted in the partial fulfilment of the requirements for the award of Doctor of Clinical Psychology at the University of Sheffield

Clinical and Applied Psychology Unit

School of Psychology

The University of Sheffield

December 2025

Declaration

I, the author, declare his thesis has been submitted for the award of Doctorate in Clinical Psychology at the University of Sheffield. This thesis is my own original work, and all sources have been referenced accordingly. This work has not been submitted for any other degree or to any other institution. No funding has been received for this thesis. No conflicts of interests stated.

Structure and Word Count

Section One: Literature Review

Excluding references and tables: 7226

Including references and tables: 13440

Section Two: Empirical Study

Excluding references and tables: 5680

Including references and tables: 8154

Total Word Count

Excluding references and tables: 12906

Including references and tables: 21594

Publications

Empirical study: Imtiaz, A., Bell, C., Waller, G., & Delgadillo, J. (2026). An Investigation of Key Symptoms That Account for the Early Response Effect During Psychological Therapy for Eating Disorders. *European eating disorders review: the journal of the Eating Disorders Association*. Advance online publication. https://doi.org/10.1002/erv.70094

Lay Summary

Treatment outcomes for people with eating disorders remain poor, with only around half of those who receive recommended treatment achieving full recovery. Researchers have therefore sought to identify which factors best predict who will benefit most from psychotherapy. One of the strongest predictors is the *early response effect*: people who show meaningful symptomatic improvement in the first few sessions of therapy are more likely to recover by the end of treatment.

Despite this, it is often assumed that people with *severe and enduring eating disorders (SE-ED)* respond poorly to treatment and require alternative approaches. The first chapter of this thesis reviewed the evidence behind this assumption. Fourteen studies used the term SE-ED, yet none agreed on a consistent definition. Across 35 studies examining whether duration or severity predict treatment outcomes, neither factor consistently related to recovery. Three studies found that early response was not influenced by severity or chronicity of the eating disorder. Taken together, these findings suggest that early response may be a more reliable predictor of recovery than either chronicity or severity.

Building on this insight, the second chapter presents an empirical study examining which specific symptoms change early in treatment. Data from 232 people receiving outpatient psychotherapy for eating disorders were analysed using a supervised machine learning method called Bayesian network analysis. This identified six symptoms most strongly linked to recovery: avoiding others seeing one's body, feeling "fat", preoccupation with food or calories, fears of overeating, feeling unhappy with one's body shape, and following strict dietary rules. Early improvements in these areas during the first four sessions of therapy were strong predictors of recovery. These findings indicate that focusing on these six areas during the initial sessions of treatment may help clinicians better support recovery for people with eating disorders.

Acknowledgements

First and foremost, I would like to thank my research supervisor, Professor Jaime Delgadillo, for his guidance throughout my empirical study and for sharing his passion for machine learning. My gratitude also goes to Professor Matt Field for supervising my systematic review, and to Professor Glenn Waller for his contributions. My gratitude extends to my clinical collaborators, Caroline Bell and Joanne Ormiston-Booth, at Freed Beeches Eating Disorders Service, where all interventions and data collection took place, and to the patients who made this project possible.

I'd like to extend my thanks to my ASD study mentor, Dr Kate Morris, for helping me work with my brain and to my clinical and academic tutors, Dr Kath Boon and Dr Toria Tew, for their pastoral support throughout my training.

To my fellow trainees, Aaminah, Amareen and Sophia: It's been a privilege to have shared this journey with you, an honour to have met such wonderful psychologists, and a blessing to have found my chosen family along the way.

To my friends, especially my soul sister, Lucy—thank you for believing in me from the moment I first told you I wanted to become a clinical psychologist over a decade ago. I am here because of friends like you and promise to do this role justice.

I could not have done this without my incredible family, who have been with me from the very beginning. A special thank you to my dad, who has always been so proud of me, my mum for her love and prayers, and to my sister, Hinnah, for everything that cannot be captured on paper—and for being the second reviewer for my systematic review. And finally, to my husband and soulmate, Kamran: thank you for encouraging me to pursue this dream and for always finding a way to ground me.

And lastly, I want to dedicate this to my mum, who will be very happy that I'm now free to go shopping with her. Alhamdulillah.

Table of Contents

Section One: Literature Review

Predictors of psychological treatment outcomes in severe and enduring eating disorders: A systematic review of severity, duration, and early response

Abstract

Aims: *Severe and enduring eating disorders (SE-ED)* are commonly assumed to be associated with poorer treatment outcomes due to greater illness chronicity and severity. However, empirical support for this assumption remains unclear. This systematic review aimed to examine how SE-ED has been defined in the literature and evaluate whether illness duration and severity predict psychological treatment outcomes. As a secondary exploratory aim, the review examined whether early response, a robust predictor of outcome in the broader eating disorder literature, also has a prognostic value within SE-ED populations.

Methods: Following protocol registration with PROSPERO, a systematic search of PsycINFO, PubMed, and Scopus was conducted for studies published between 2015 and 2025, in line with PRISMA guidelines. Randomised controlled trials and cohort studies examining SE-ED definitions, or illness duration and/or severity as predictors of psychological treatment outcomes, were included. Risk of bias was assessed using the Critical Appraisal Skills Programme (CASP) checklists, and findings were synthesised using narrative methods.

Results: 44 studies met inclusion criteria. 14 studies provided explicit definitions of SE-ED, demonstrating substantial inconsistency in how the construct was operationalised. Across diagnostic subtypes and psychological treatment modalities, illness duration and illness severity was not consistently associated with treatment outcomes. No studies directly examined early response as a predictor of outcome in SE-ED; however, three studies included in the review reported that illness duration and severity did not influence early response.

Conclusion: SE-ED lacks a consistent definition and shows limited validity as a predictor of psychological treatment outcomes. Given that illness duration and severity did not predict outcomes, early response warrants investigation as a potentially more informative prognostic marker in this population.

Practitioner Points:

- The label severe and enduring eating disorder (SE-ED) is inconsistently defined in the literature.
- Illness duration and baseline severity do not reliably predict psychological treatment outcomes in eating disorders, which suggest SE-ED lacks predictive validity.
- Clinicians should use the SE-ED label with caution and should not assume individuals with long-standing or severe eating disorders are unlikely to benefit from evidence-based treatments.
- Early treatment response is a robust predictor of outcomes in the broader eating disorder population; however, its prognostic value in SE-ED has not been examined and warrants investigation.

Introduction

Eating disorders encompass several diagnostic subtypes, including anorexia nervosa, characterised by severe restriction of energy intake resulting in significantly low body weight; bulimia nervosa, marked by recurrent binge-eating episodes followed by compensatory behaviours such as self-induced vomiting or fasting; binge-eating disorder, involving recurrent binge-eating episodes without compensatory behaviours; and other specified feeding or eating disorder, which captures clinically significant eating pathology that does not meet full diagnostic criteria for the other disorders (American Psychiatric Association, 2013). Despite these differences in behavioural presentation, all eating disorder subtypes share a core underlying psychopathology: the overvaluation of body shape and weight, and the perceived ability to control them (Fairburn et al., 2003).

Current clinical guidelines indicate that several evidence-based psychological treatments are effective for eating disorders in adults. These include cognitive behavioural therapy (CBT), the Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA), and specialist supportive clinical management (SSCM) (National Institute for Health and Care Excellence [NICE], 2017). However, despite the availability of these interventions, meta-analytic data indicate that fewer than half of adults who receive recommended treatment achieve recovery or remission – defined as no longer meeting diagnostic criteria and experiencing improvements in psychosocial functioning and quality of life (Bardone-Cone et al., 2010; Monteleone et al., 2022). Moreover, up to 20% of individuals go on to develop a long-term, chronic form of the disorder (Keel & Brown, 2010). Given these limited outcomes, there is a growing need to understand the factors that influence differential recovery outcomes. As a result, researchers have increasingly focused on identifying predictors of treatment response and recovery.

The early response effect

Among the prognostic factors examined, the early response effect has emerged as one of the strongest and most consistent predictors of treatment outcomes, as documented in systematic and meta-analytic reviews of the literature (Chang et al., 2021; Linardon et al., 2016; Vall & Wade, 2015). Early response refers to statistically reliable improvement in symptoms during the early phase of psychological treatment (MacDonald et al., 2015). Although definitions of the early response timeframe differ across studies (e.g., Hilbert et al., 2019; MacDonald et al., 2015) and measurement approaches vary (categorical versus dimensional), approximately 65% of eating disorder studies define early response as occurring within the first four weeks of treatment and focus on the magnitude of symptom change rather than setting a fixed categorical threshold (Chang et al., 2021).

Despite this strong evidence base, clinical decision-making often fails to reflect the prognostic value of early symptom change. Instead, treatment decisions are frequently shaped by assumptions that greater illness severity and longer duration reduce engagement and responsiveness to psychological therapies (Hay & Touyz, 2015). Consequently, individuals with long-standing eating disorders are commonly labelled as having a ‘severe and enduring eating disorder’ (SE-ED) (Robinson, 2014) and are often offered alternative interventions aimed at improving quality of life rather than supporting full recovery (Broomfield et al., 2017).

Severe and enduring eating disorder

The conceptual basis of the SE-ED construct closely aligns with the clinical staging model of eating disorders, which describes illness progression from ‘high-risk’ to ‘severe and enduring illness’ (Treasure et al., 2014). Within this framework, maintaining mechanisms are proposed to become increasingly entrenched over time through neurobiological adaptation, habit learning, and cumulative social and cognitive impairment, leading to greater treatment

resistance at later stages.

Since the term SE-ED was introduced in the literature in 2008 (Arkell & Robinson, 2008), several studies examining treatment approaches and outcomes in this population have been published (e.g., Hay & Touyz, 2015; Kotilahti et al., 2020; Piñar-Gutiérrez et al., 2021; Treasure et al., 2015). However, despite growing research attention, there is still no clear consensus on how SE-ED should be defined. Proposed definitions of the 'enduring' aspect vary widely, ranging from as little as three years (Hay et al., 2012) to more than twenty years (Robinson et al., 2015). Similarly, the 'severe' dimension has been operationalised inconsistently, being variously defined by low body mass index (BMI) or persistent cognitive and behavioural impairment (Broomfield et al., 2017). A more recent proposal conceptualised SE-ED as persistent symptoms and functional impairment lasting longer than three years despite engagement in at least two evidence-based treatments (Hay & Touyz, 2018). Notably, most empirical work on SE-ED has focused almost exclusively on anorexia nervosa, with limited consideration of other eating disorder subtypes.

The SE-ED construct is further complicated by its underlying assumption that longer illness duration corresponds to greater illness severity (Robinson et al., 2015). However, existing evidence does not support this assumption (Austin et al., 2021; Dapelo et al., 2020), with some research suggesting that shorter illness duration may instead be associated with greater impairment (Davidsen et al., 2017). Furthermore, there is limited evidence that SE-ED status – regardless of how it is defined – predicts treatment outcomes. For example, Radunz et al.'s (2020) systematic review and meta-analysis found no association between illness duration and treatment outcome, and Vall and Wade (2015) similarly reported little evidence that illness severity reliably predicts treatment response. In the absence of definitional clarity and robust empirical support, it is therefore concerning that a small body of literature has even advocated for palliative approaches and assisted dying in individuals

labelled with SE-ED (Gaudiani et al., 2022; Yager, 2020), despite evidence that recovery remains possible and that these patients respond similarly to evidence-based treatments as those without this label (Raykos et al., 2018).

Current Systematic Review

In light of this ambiguity regarding SE-ED, a comprehensive and updated synthesis of the literature is needed. The most recent systematic review focusing on illness duration alone (Radunz et al., 2020) is now outdated and did not examine illness severity, despite its central role in many definitions of SE-ED. Several new studies have since been published, providing an opportunity to evaluate both constructs together and to consider the extent to which they meaningfully influence treatment outcomes.

Furthermore, revisiting this area enables an examination of whether the early response effect identified in the broader eating disorder population is also applicable to the SE-ED population. Establishing whether early symptom change is associated with recovery in this population is clinically important, as it may encourage clinicians to prioritise early change during psychological therapy rather than prematurely abandoning treatment approaches all together.

The aim of this systematic review was therefore threefold. First, it sought to clarify how SE-ED has been defined in empirical studies investigating psychological interventions for adults described as having SE-ED. Second, it aimed to synthesise findings from studies examining illness duration and illness severity as predictors of psychological treatment outcomes. In the present review, illness severity was operationalised primarily using the Eating Disorder Examination Questionnaire (EDE-Q; Fairburn & Beglin, 2008), the most widely used and psychometrically robust measure of eating disorder psychopathology. The EDE-Q assesses core cognitive and behavioural symptoms of eating disorders and aligns with transdiagnostic models that highlight the overvaluation of shape and weight as key

maintaining mechanisms (Fairburn et al., 2003). It demonstrates strong internal consistency, construct validity, and sensitivity to clinical change (Luce & Crowther, 1999; Ro et al., 2015), making it well suited for assessing severity across diagnostic groups in treatment-outcome research. BMI was additionally included as a severity indicator in anorexia nervosa, where low weight reflects both diagnostic status and illness severity (APA, 2013). Finally, a secondary exploratory aim was to examine whether any studies had investigated early response as a predictor of treatment outcomes within SE-ED samples.

Objectives:

The following research questions guided this review:

1. How are severe and enduring eating disorders characterised in the literature?
2. Does duration of eating disorder predict treatment outcomes?
3. Does severity of eating disorder predict treatment outcomes?
4. Does early response in therapy predict treatment outcomes for severe and enduring eating disorders (if such studies exist)?

Methods

Protocol registration

The protocol for this systematic literature review was registered in the PROSPERO database prior to conducting the literature search (https://www.crd.york.ac.uk/PROSPERO/view/CRD420251153052).

Search strategy and study selection

The present study was conducted and reported in line with the evidence-based guidelines for reporting systematic reviews (Moher et al., 2009). The inclusion and exclusion criteria that guided this review are listed in table 1. The criteria were developed following a PICOS framework, which has demonstrated greater sensitivity than the SPIDER framework and greater specificity than the PICO search tools (Methley et al., 2014).

Study design was restricted to randomised controlled trials (RCTs) and cohort studies, as these longitudinal designs permit examination of predictors of treatment outcome over time. Cross-sectional studies were excluded because they do not allow for evaluation of temporal associations between illness duration, symptom severity, and post-treatment outcomes. Restricting inclusion to longitudinal designs allowed this review to confidently draw conclusions regarding predictive relationships.

The review was restricted to adult samples (aged 18 years and above) in order to reduce developmental and treatment heterogeneity. Treatment approaches for non-adult samples predominately involve family-based interventions (Fisher et al., 2018), which differ substantially from individual psychological therapies typically delivered in adult services. Restricting inclusion to adults therefore enhanced comparability across studies and ensured applicability to adult treatment contexts.

Given the absence of a universally accepted definition of SE-ED, the search strategy intentionally incorporated a range of proxy terms including ‘long-standing,’ ‘chronic,’ ‘persistent,’ ‘treatment resistant,’ and ‘extreme.’ Studies were eligible if they explicitly defined SE-ED with duration or severity level thresholds or if they examined illness duration and/or symptom severity as predictors of psychological treatment outcome. This approach was adopted to reflect the heterogeneity in how severity and chronicity are operationalised within the literature. Studies including mixed samples (i.e., participants with varying levels of illness duration or severity) were included where duration or symptom severity was analysed as a predictor variable. This ensured that relevant evidence was captured even where samples were not exclusively characterised as SE-ED.

Key search terms (related to eating disorders, severe and enduring, psychological therapies and outcome predictors) were developed based on the Radunz et al. (2020) review, with potential terms tested and refined through a series of scoping searches which aimed to

elucidate the optimal balance between specificity and comprehension. The final terms were combined using Boolean operators and involved three separate searches.

The three separate searches were performed across three databases (SCOPUS, PSYCinfo and PubMed) on October 6, 2025, and the results were pooled together. Restrictions were applied relating to the date of publication (2015 – 2025). The decision to focus on more recent literature was made to build upon existing systematic reviews which had synthesised earlier research examining SE-ED, primarily in relation to illness duration. The present review therefore aimed to update and extend this body of work by incorporating more recent empirical findings and by broadening the scope beyond duration alone to also consider symptom severity, definitional approaches, and predictors of psychological treatment outcomes.

For the first search an additional restriction (title and abstract only) was applied to improve specificity in exploring SE-ED definitions; this restriction was not applied to the subsequent two searches (see table 2).

The first author screened titles and abstracts against the inclusion and exclusion criteria. A second reviewer screened 10% of these titles and abstract at random to check inter-rater agreement and any disagreements were resolved through discussion until consensus was reached. Agreement between reviewers was 98%, indicating high consistency. Full texts of the remaining articles were sourced and assessed for eligibility by the primary reviewer.

Table 1. *Study inclusion and exclusion criteria presented in a PICOS framework*

	Inclusion criteria	**Exclusion criteria**
Population	• Participants who have an eating disorder diagnosis and have received psychological treatment for their eating disorder. • Participants aged 18 or older.	• Participants who have not received psychological treatment for their eating disorder. • Participants under the age of 18.
Intervention	• Any psychological intervention used to treat eating disorders, in any modality (individual, group, online, etc.) and setting (outpatient, inpatient, day programme).	• Interventions that do not involve a psychological component.
Comparator	• Studies that provide a clearly defined criteria for severe and enduring eating disorders (could also be referred to as "long-standing", "chronic", "treatment resistant" or "extreme" eating disorders). • Studies that examine symptom severity (with the EDE-Q or BMI for anorexia nervosa) or illness duration as a predictor of treatment outcome in eating disorder treatment.	• Studies that do not provide a clearly defined criteria for severe and enduring eating disorders. • Studies that do not examine symptom severity or illness duration as a predictor of treatment outcome. • Studies that do not examine symptom severity with the EDE-Q or BMI.
Outcome	• The defining criteria used to characterise severe and enduring eating disorders. • The association between symptom severity and/or	• Studies where post-treatment outcomes of eating disorder treatment are not measured.

	illness duration to eating disorder treatment outcomes.	
Study design	• Randomised controlled trials and cohort studies • Studies published in peer-reviewed scientific journals.	• Grey literature (e.g. dissertations), studies not published in peer-reviewed scientific journals. • Studies not published in English. • Studies that are cross-sectional in design.

Table 2. *Search strategy*

Search	Search aim	Search strategy	Restrictions applied
1	To explore how severe and enduring eating disorders are characterised in the literature investigating psychological treatments for this population	(anorexia OR Eating Disorder* OR disordered eat* OR binge eat OR bulimia) AND (severe OR enduring OR chronic OR persistent OR longstanding OR treatment resistant OR treatment failure)	• Date of publication (2015 – 2025) • Field (title and abstract only).
2	To explore whether duration and/or severity predict psychological treatment outcomes	(anorexia OR Eating Disorder* OR disordered eat* OR binge eat OR bulimia) AND (treatment OR therapy OR psychotherapy OR intervention) AND (response OR outcome) AND (predictor OR predict)	• Date of publication (2015 – 2025)
3	To explore if any papers have	(anorexia OR Eating Disorder* OR disordered eat* OR binge	• Date of publication (2015 – 2025)

	examined early response as predictor of treatment outcomes in those characterised as having a severe and enduring eating disorder	eat OR bulimia) AND (severe OR enduring OR chronic OR persistent OR longstanding OR treatment resistant OR treatment failure) AND (treatment OR therapy OR psychotherapy OR intervention) AND (response OR outcome) AND (predictor OR predict).

Quality Assessment

The critical appraisal skills programme (CASP) checklist for RCT and longitudinal cohort studies (CASP, 2024; Appendix A and B) was used to assess the risk of bias. The CASP checklists assess validity of study design, sound methodology, clarity of results and the impact of results locally. As there is no guidance to categorise risk of bias from the CASP checklists, this study used the number of "yes" responses in the CASP checklist to categorise studies. For both checklists, a "yes" response would indicate something of good quality and thus the more "yes" responses, the lower risk of bias. For RCTs: 0-3 high risk; 4-6 medium to high risk; 7- 8 medium risk; 9-11 low to medium risk and 12-13 low risk. For cohort studies: 0-3 high risk; 3-6 medium to high risk; 7-10 medium risk; 11-12 low to medium risk and 13-14 low risk. All included articles were assessed by the first author, with 25% of articles selected at random to be independently assessed by a separate researcher. The first author and separate researcher subsequently compared their assessments and resolved any discrepancies. Interrater reliability before discrepancies were resolved was calculated using Cohen's kappa coefficient (Cohen, 1960) and this was categorised as "substantial" agreement, $k = 0.686$.

Data Extraction and Synthesis

Data were extracted and tabulated by the first reviewer. The primary outcome of interest was how SE-ED was defined and whether illness duration and severity predicted outcomes after psychological treatment. A secondary outcome of interest explored whether early response in therapy was a predictor for treatment outcomes among patients labelled with SE-ED. The following data was extracted from the papers providing SE-ED definition – study design, country, diagnostic subtype, number of participants, sample demographic, intervention, outcome measures and SE-ED definition. The following data was extracted from papers exploring eating disorder severity and/or duration as a predictor of treatment outcomes – study design, country, diagnostic subtype, number of participants, sample demographic, intervention, predictors explored, outcome measures, predictor analysis and findings.

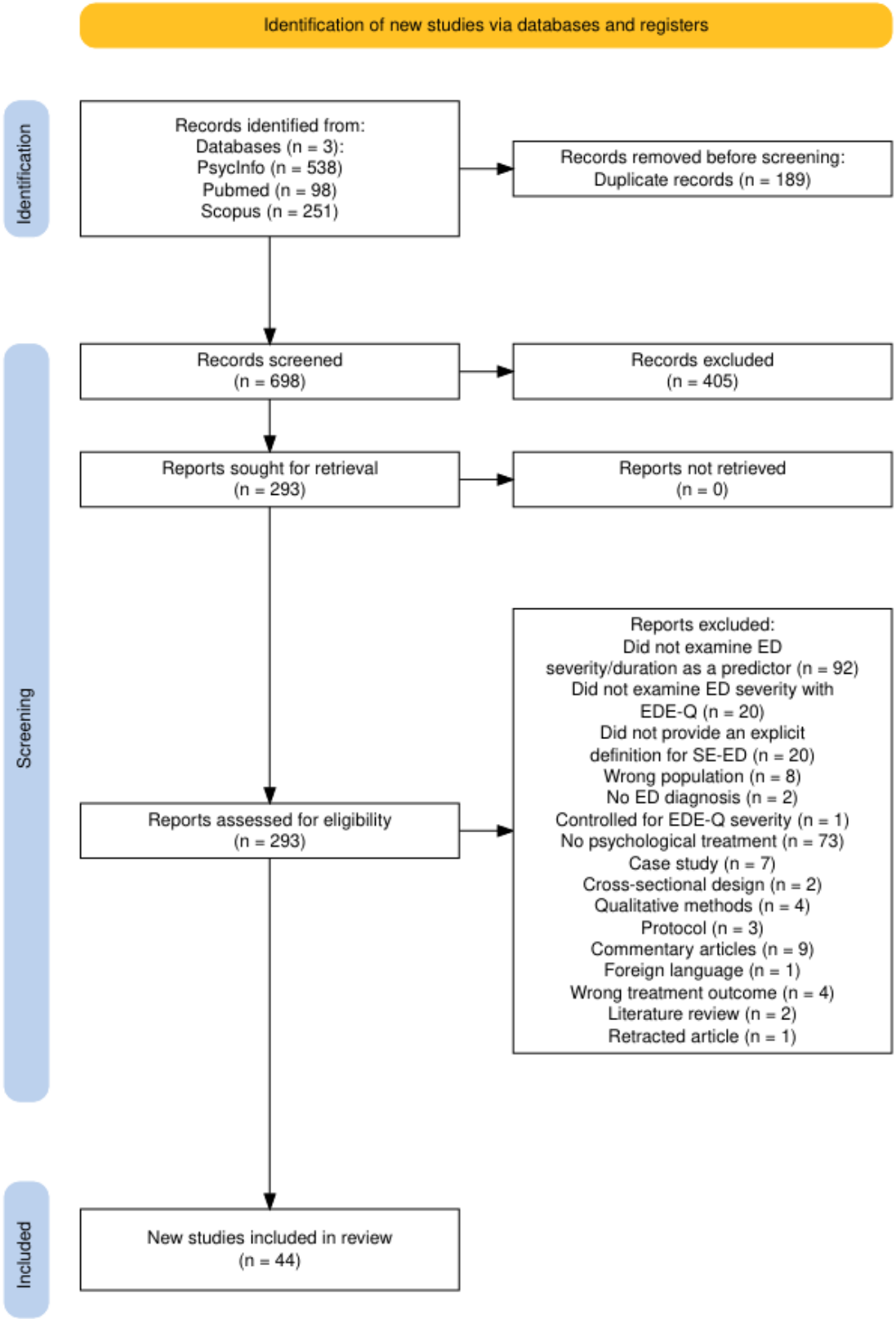
A narrative synthesis of the results were conducted for all the reviewed studies. Within this synthesis, illness duration and severity were treated as separate predictors of outcome, and studies were grouped under each predictor according to the behavioural presentation of the eating disorder (restrictive eating disorders, binge-eating/purging eating disorders or mixed eating disorder samples) to facilitate a structured comparison across diagnostic subtypes. A meta-analytic method for predictor variables was not conducted due to time constraints, under-reporting of statistical information, and heterogeneity of the studies included in the reviews.

Results

The search strategy identified 698 unique records. Screening of titles and abstracts led to the exclusion of studies involving adolescent or child samples, those without a clear psychological treatment component, and studies examining non–eating disorder populations. Following this stage, 293 full-text articles were assessed for eligibility. Common reasons for

exclusion at full-text assessment included a failure to examine illness duration or severity as predictors of treatment outcomes, the use of eating disorder severity measures other than the EDE-Q, the absence of an explicit definition of SE-ED, or the absence of an identifiable psychological intervention. Ultimately, 44 studies met eligibility criteria, fourteen of which provided an explicit definition of SE-ED and 35 which examined illness duration and/or severity as a predictor of psychological treatment outcomes. Figure 1 presents the PRISMA flow diagram for the screening and selection process.

Figure 1. *PRISMA diagram*



Quality Assessment

Of the RCTs one study was judged to have a low to medium risk of bias, nine studies were judged to have a medium risk of bias, and two studies were judged to have a medium to high risk of bias. A consistent methodological issue contributing to risk of bias was the lack of allocation concealment and blinding, which was largely due to the nature of the research. In addition, no power analyses were conducted specifically for the predictor analyses. Other sources of potential bias included missing statistical information (e.g., effect sizes), limited generalisability of samples (e.g., predominantly female and white participants), and missing information regarding the methods and interventions used.

For the longitudinal cohort studies, one study was considered to have a low risk of bias, eleven studies had a low to medium risk of bias, and twenty studies had a medium risk of bias. All studies used validated and reliable outcome measures; however, many relied on self-report data, which raises concerns about validity due to self-report bias. The most common methodological issue was the lack of follow-up or insufficient information regarding follow-up procedures. Another concern was the precision of the results, as many studies did not report confidence intervals or effect sizes. Finally, all studies demonstrated limited generalisability because their samples were predominantly female and of white ethnic background or were carried out in highly specialised settings which may not always reflect routine clinical practice (e.g. inpatient settings). Further details about all risk of bias assessments can be found in Appendix C.

SE-ED definitions across empirical studies on psychological therapies

Fourteen studies investigating psychological treatments provided an explicit definition of SE-ED when describing their sample (see Table 3). These studies were most commonly longitudinal cohort studies (12/14 - 85.71%), with a smaller number of RCTs (2/14 - 14.29%), and were conducted across Europe, North America, and Australia. The majority

focused on individuals with anorexia nervosa or underweight individuals with other specified feeding and eating disorder (11/14 – 78.57%), whereas a smaller number of studies examined SE-ED in bulimia nervosa (1/14 – 7.14%), binge eating disorder (1/14 – 7.14%) or mixed eating disorder presentations (1/14 – 7.14%). Samples were predominantly female and white, and interventions were largely delivered in inpatient or day-patient settings (8/14 – 57.14%), most often using CBT (10/14 – 71.43%). Other treatment modalities included family therapy, SSCM and dialectical behaviour therapy (DBT).

SE-ED definitions varied considerably across studies. The most frequently applied definition was an illness duration of at least seven years (9/14 – 64.29%), sometimes combined with markers of clinical severity such as BMI below 16 (1/14 – 7.14%), high psychological distress (1/14 – 7.14%), or failure to benefit from previous treatments (1/14 – 7.14%). Other studies operationalised SE-ED using shorter illness duration thresholds (1/14 – 7.14%), BMI-based severity criteria only (2/14 – 14.29%), or frequency and caloric amount of binge-eating episodes for individuals with bulimia nervosa and binge eating disorder (2/14 - 14.29%).

Table 3. *SE-ED Definitions across empirical studies on psychological therapies*

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Outcome	SE-ED Definition
Ambwani et al. (2020)	Cohort	UK	AN, OSFED (BMI > 18.5)	187	M age (SD) = 27.81 (9.30), 96.8% Female, 97.5% White	Evidence based psychological treatment following NICE guidelines	BMI, EDE-Q, WSAS, DASS-21. Recovery (BMI > 18.5 plus EDE-Q global < 2.77)	Illness duration ≥7 and severe psychological distress
Bamford et al. (2015)	RCT	UK & Australia	AN	63	M age = 33.4	CBT vs SSCM	BMI, EDE-Q, EDQOL, WSAS, BDI	Illness duration ≥7 years
Calugi et al. (2017)	Cohort	Italy	AN	66	M age (SD) = 26.1 (5.9), 96.97% Female	IP CBT-E followed by DP CBT-E	BMI, EDEQ, BSI	Illness duration ≥7 years
Calugi et al. (2021)	Cohort	Italy	AN	30	M age (SD) = 22.4 (7.9), 96.67% Female, 100% White	CBT-E	BMI, EDEQ, BSI, CIA. Good BMI outcome (BMI > 18.5); Full response (BMI > 18.5 + EDE-Q global < 2.77)	BMI < 16
Frostad et al. (2021)	Cohort	Norway	AN	21	M age (SD) = 25.5 (7.9)	CBT-E	BMI	BMI < 16
Garte et al. (2015)	Cohort	Norway	AN, BN, OSFED	62	M age (SD) = 27.61(7.23), 93.5% Female	DP CBT-E	Recovery (BMI > 18.5, EDEQ < 2.46)	Illness duration ≥7

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Outcome	SE-ED Definition
Lydecker et al. (2020)	RCT	USA	BED	521	M age (SD) = 46.35 (9.17), 73.3% Female, 77% White	CBT vs BWL vs multimodal behavioural + pharmacotherapy	EDEQ, SCID-I, BDI, Binge eating episodes (frequency), remission (no binge eating episodes over 28 days), % of weight loss, attainment of 5% weight loss	Severe binge eating disorder defined as 8-13 binge eating episodes per week
Marzola et al. (2019)	Cohort	Italy	AN	137	M age (SD) = 24.9 (9.7), 100% Female, 100% White	IP PT + psychoeducational group	EDE-Q, BDI, STAI, EQ-5D, WAI	Illness duration ≥7 years
Marzola et al. (2021)	Cohort	Italy	AN	169	M age (SD) = 23.9 (8.7), 87.57% Female	IP integrated psychological work informed by motivational interviewing	EDI-2, EDE-Q, BSQ, STAI, BDI, HAM-A, HAM-D, EQ-5D	Illness duration ≥7 years
Meule et al. (2023)	Cohort	Germany	AN	902	M age (SD) = 23.4 (9.35), 100% Female	IP CBT 'multimodal AN' model	BMI, GAF, change in CGI, EDI-2	Illness duration ≥ 3 years
Quadflieg & Fichter (2019)	Cohort	Germany	BN	1351	M age (SD) = 25.94	IP CBT	Remission (no ED for 3 months); Partial remission (ED syndrome); Poor outcome (full ED diagnosis for 3). Assessed using SIAB-S	Severity defined as binge-eating episodes > 3000 calories 2x a week

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Outcome	SE-ED Definition
Raykos et al. (2018)	Cohort	Australia	AN	134	M age (SD) = 22 (18.25), 97% Female, 76.9% Austrian born	CBT-E	Treatment completion, change in EDE-Q, change in QLESQ, change in BMI over first five treatment sessions (early response) and change in BMI at EoT	Illness duration ≥7 years and BMI < 16
Redgrave et al. (2021)	Cohort	USA	AN, OSFED	191	M age (SD) = 32.55 (12.29), 100% Female, 88.7% White	IP group CBT + group DBT + FT	BMI	Illness duration ≥7 years
Smith & Woodside (2021)	Cohort	Canada	AN	75	M age (SD) = 25.93 (8.35), 97.3%, Female, 39.4%	IP 'psychosocial interventions' + group therapy	BMI >18.5, abstinence of binge/purging behaviour	Illness duration > 7 years with two or more incomplete admission or no complete admissions

Note: AN = Anorexia nervosa; BDI = Beck Depression Inventory; BED = Binge eating disorder; BMI = Body Mass Index; BN = Bulimia nervosa; BSI = Brief Symptom Inventory; BSQ = Body Shape Questionnaire; BWL = Behavioural weight loss; CBT = Cognitive Behavioural Therapy; CBT-E = Enhanced Cognitive Behavioural Therapy; CGI = Clinical Global Impression-Improvement scale; CIA = Clinical Impairment Assessment; DAAS = Depression Anxiety Stress Scales; DBT = Dialectical Behaviour Therapy; DP = Day programme; Dx = Diagnosis; ED = Eating Disorder; EDE-Q = Eating Disorder Examination Questionnaire; EDI = Eating Disorder Inventory; EDQOL = Eating Disorders Quality of Life Questionnaire; EoT = End of Treatment; EQ-5D = EuroQoL Health Questionnaire; FT = Family Therapy; GAF = Global Assessment of Functioning scale; HAM-A = The Hamilton Anxiety Rating Scale; HAM-D = The Hamilton Depression Rating Scale; IP = Inpatient; M = mean; NICE = National Institute of Care and Excellence; OSFED = Other specified feeding and eating disorder; PT = Psychodynamic Therapy; QLESQ = Quality of Life Enjoyment and Satisfaction

Questionnaire; SCID-I = Structured Clinical Interview for DSM-IV Axis I Disorders; SD = standard deviation; SIAB = Structured Interview for Anorexic and Bulimic disorders; SSCM = Specialist Supportive Clinical Management; STAI = State-Trait Anxiety Inventory; UK = United Kingdom; USA = United States; WAI = The Working Alliance Inventory; WSAS = Weissman Social Adjustment Scale

Illness duration as a predictor of treatment outcomes

Restrictive Eating Disorders

Ten studies examined illness duration as a predictor of treatment response in individuals with restrictive eating disorders, primarily anorexia nervosa or other specified feeding and eating disorder closely resembling anorexia nervosa (see table 4). Treatments were typically CBT-based, often delivered in inpatient or day programme settings. Other treatments included DBT, MANTRA and family therapy. Across these studies, illness duration was most commonly measured in number of years since onset.

Measures of treatment outcome varied between studies and often combined weight-based, symptom-based, or diagnostic criteria. Several studies defined recovery using BMI thresholds, such as achieving a BMI above 18.5 or 19.5 at end of treatment or achieving BMI over 17.5 within the time expected to gain weight at or above 0.5 kg per week. Others defined recovery or positive outcomes through diagnostic thresholds (e.g., no longer meeting diagnostic criteria for anorexia nervosa), or through combined BMI and EDE-Q thresholds, such as BMI over 18.5 alongside an EDE-Q global score below clinical thresholds of 2.77. Some studies used reliable change indices, categorising participants into full recovery, partial improvement, or deterioration based on symptom change. Other studies did not employ reliable changes indices and measured outcomes dimensionally with the EDE-Q or BMI.

Across restrictive eating disorder studies, seven (70%) reported that illness duration did not predict treatment outcome. However, three studies (30%) found that longer illness duration was associated with poorer weight-related or functional outcomes, including smaller BMI increases during treatment, lower BMI at follow-up, or reduced likelihood of achieving predefined recovery criteria.

Binge-eating/purging eating disorders

Seven studies examined illness duration in individuals with bulimia nervosa, binge eating disorder, or other specified feeding and eating disorder characterised primarily by binge-eating/purging behaviours. All studies evaluated CBT-based interventions, delivered individually, in groups, online, or within inpatient settings. Illness duration was usually measured in years since symptom onset. Measures of treatment outcomes again varied substantially. Some studies defined recovery or positive treatment outcomes as the complete absence of binge-eating or purging for 28 days, while others used percentage reduction in binge-eating frequency or EDE-Q global scores. Several studies measured treatment outcomes dimensionally with EDE-Q global scores or binge-eating frequency while others categorised outcomes using pre-defined recovery criteria.

Across the binge-eating/purging eating disorder studies, five studies (71.43%) found no relationship between duration of illness and treatment outcome. Two studies (28.57%), however, reported that longer illness duration predicted poorer outcomes, including higher EDE-Q global scores at follow-up or a reduced likelihood of symptomatic remission.

Mixed eating disorder samples

Five studies examined illness duration in a mixed eating disorder sample including anorexia nervosa, bulimia nervosa, binge eating disorder, and other specified feeding and eating disorder. These studies consistently used CBT-based interventions delivered across inpatient, day programme, or outpatient settings. Measures of treatment outcomes varied but typically involved categorical remission (e.g., full remission defined as the absence of eating disorder symptoms for four weeks; partial remission defined as improvement with residual symptoms) or dimensional reductions in EDE-Q global scores.

Across all mixed eating disorder studies, illness duration did not predict treatment outcomes, regardless of intervention intensity or whether outcomes were measured categorically or dimensionally.

Illness severity as a predictor of treatment outcomes

Restrictive Eating Disorders

Eleven studies examined illness severity as a predictor of treatment outcome in restrictive eating disorder populations (see table 4). Severity was most commonly operationalised using BMI (10/11 – 90.91%), with a smaller number of studies using the EDE-Q either in addition to BMI or alone (6/11 – 54.55%). Interventions typically involved CBT delivered in inpatient settings. Other treatments included SSCM, DBT, family therapy and MANTRA.

As with studies that looked at illness duration, measures of treatment outcomes varied and frequently combined weight-based and symptom-based markers. Some studies defined recovery or positive treatment outcomes using pre-defined BMI criteria or combined pre-defined BMI criteria with diagnostic, behavioural and psychological markers, including EDE-Q clinical thresholds cut-offs or abstinence of eating disorder behaviours. Other studies relied on dimensional BMI change or EDE-Q improvement when measuring treatment outcomes.

Findings were mixed. Across studies examining the predictive value of BMI, six out of ten (60%) reported that lower baseline BMI was associated with poorer outcomes, including lower BMI after treatment or a reduced likelihood of meeting a pre-defined recovery criterion. Three studies (30%) found no association between baseline BMI and treatment outcomes, while one study (10%) found that lower baseline BMI predicted greater BMI improvement.

Studies examining the predictive value of the EDE-Q also produced mixed results. Half of the studies (3/6 – 50%) found no predictive association. Two studies (33.33%)

reported that lower baseline EDE-Q scores predicted more favourable treatment outcomes, whereas one study (16.67%) found that more severe baseline EDE-Q scores predicted favourable outcomes.

Binge-eating/purging eating disorders

Ten studies assessed illness severity as a predictor of treatment outcome in binge-eating/purging eating disorders, using the EDE-Q global score or specific subscales as indicators of severity. Interventions were commonly CBT based, delivered through various formats including therapist-led, group or guided self-help. Other treatments included integrated cognitive affective therapy and behavioural weight loss programmes.

Measures of treatment outcomes included reduction or cessation of binge-eating episodes, binge-eating frequency, no longer meeting diagnostic criteria, clinically significant symptom reduction, and end of treatment EDE-Q global scores. Other studies combined behavioural abstinence with EDE-Q clinical thresholds and weight loss/BMI. Some studies distinguished full remission, partial remission, and poor outcome depending on the persistence of behavioural symptoms and diagnostic features.

Across the binge-eating/purging eating disorder studies, findings examining the predictive value of the EDE-Q were heterogeneous. Two studies out of the ten (20%) found no association between EDE-Q severity and treatment outcomes. Four studies (40%) reported that higher baseline EDE-Q subscale scores, particularly the weight and shape concern or restraint subscales, predicted greater reductions in binge-eating behaviour. Conversely, four studies (40%) reported that higher EDE-Q severity predicted poorer outcomes, including higher symptom levels at follow-up or reduced likelihood of remission.

Mixed eating disorder samples

Six studies assessed illness severity in mixed eating disorders samples using EDE-Q scores, with three also incorporating BMI for underweight participants. All studies included

CBT, delivered in inpatient, day programme and outpatient settings. Other treatment modalities included DBT and family therapy.

Measures of treatment outcomes varied. Some studies measured treatment outcomes dimensionally with the EDE-Q and BMI. Other studies defined recovery or positive outcomes as reliable and clinically significant improvement on the EDE-Q, achieving an EDE-Q score within clinical cut-offs, absence of binge-purge behaviours for specified periods, or a combination of weight and symptom criteria. One study specifically looked at the need for higher level of care as an outcome.

Five out of the six studies (83.33%) found that baseline EDE-Q severity did not predict treatment outcomes. One study (16.67%) observed that higher baseline EDE-Q scores predicted higher symptom levels at end of treatment. Two out of the three studies (66.67%) incorporating BMI reported that severity defined by BMI did not predict outcomes, although one study (33.3%) reported that lower BMI predicted the need for a higher level of care.

Table 4. *Illness duration and severity as predictors of treatment outcomes*

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Baseline predictors explored	Outcome	Analysis	Findings
Restrictive ED										
Bamford et al. (2015)	RCT; Pre–post; 15 weeks; EoT; 6- and 12-month FU	UK & Australia	AN	63	M age = 33.4	CBT vs SSCM	BMI, EDE-Q	BMI, EDE-Q, EDQOL, WSAS, BDI	Mixed model analyses	Higher BMI and lower EDE-Q predicted better QoL.
Calugi et al. (2025)	Cohort; Pre-post, EoT, 20-week FU	Italy	AN	421	M age = 25.4, 96.2% Female	CBT-E in an IP setting followed by DP	Illness duration, BMI, EDE-Q	'Good BMI outcome' (BMI > 18.5); Full response (BMI > 18.5 & EDEQ global over 2.77); 'Full recovery' (no longer meeting diagnostic criteria, no ED behaviours, BMI > 18.5 & EDE-Q scores within 1SD of age matched community norms)	Multivariable logistic regression	Duration not predictive; higher BMI and lower EDE-Q predicted recovery.
Frank et al. (2024)	Cohort; Pre-post, EoT, 6 month FU	USA	AN	159	M age (SD) = 28.1 (10.0), 95.2% Female, 83.7% White	DP FT + DBT	Illness duration, EDE-Q	BMI	Machine learning models	Illness duration and EDE-Q not predictive.
Ibrahim et al. (2022)	Cohort; Pre-post, 1-year FU	UK	AN	212	M age (SD) = 28.9 (9.6), 97% Female	TAU vs Integrated CBT-E vis crisis IP vs IP CBT-E	BMI	'Good outcome' (BMI > 19.5 & no abnormal eating or compensatory behaviours) vs poor outcome (BMI < 19.5 &/or ongoing ED behaviours, readmission or death)	Stepwise linear regression	BMI not predictive
Kaufmann et al. (2021)	Cohort; Pre-post; EoT;, 1-6 year FU	Switzerland	AN	107	M age (SD) = 24.86 (8.44), 91.59% Female	Multimodal IP programme	Illness duration, BMI	BMI	Multiple linear regression	Duration not predictive; higher BMI predicted better BMI.
Li et al. (2020)	Cohort; Pre-post; EoT	UK	AN	476	Aged between 18 – 69, 97.90% Female, 87.61% White	IP CBT+CREST; DP CBT/CAT/MET/ MANTRA/MBT	Illness duration, EDE-Q, BMI	EDE-Q; Recovered (reliable improvement & normative range); Improved (reliable improvement only); Unchanged (normative range without reliable change); Deteriorated (reliable worsening).	Binary logistic regression	Duration not predictive; higher EDE-Q predicted improvement; BMI not predictive.

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Baseline predictors explored	Outcome	Analysis	Findings
Meule et al. (2023)	Cohort; Pre–post; EoT	Germany	AN	902	M age (SD) = 23.4 (9.35); 100% female	IP CBT multimodal	Illness duration	BMI, GAF, CGI, EDI-2	Linear regression	Longer duration predicted poorer outcomes.
Raykos et al. (2018)	Cohort; Pre-post; EoT	Australia	AN	134	M age (SD) = 22 (18.25), 97% Female, 76.9% Austrian born	CBT-E	Illness duration, BMI, EDE-Q	Treatment completion, change in EDE-Q, change QLESQ, change in BMI	Logistic & longitudinal regression	Duration and EDE-Q not predictive; lower BMI predicted larger BMI improvement
Redgrave et al. (2021)	Cohort; pre-post; EoT; 6-month FU	US	AN, OSFED	191	M age (SD) = 32.55 (12.29), 100% Female, 88.7% White	IP CBT + DBT + FT	Illness duration, BMI	BMI	Binary logistic regression	Illness duration and BMI not predictive
Steinglass et al. (2024)	Cohort; Pre-post; EoT, 3-year FU	USA	AN	88	M age (SD) = 26.6 (8.3), 94.3% Female, 86.4% White,	IP behavioural treatment	Illness duration, BMI, EDE-Q	BMI, EDE-Q, Maintenance of BMI > 18.5, EDE-Q < 2	Linear regression analysis	Longer duration predicted lower BMI; lower BMI predicted lower BMI; EDE-Q not predictive.
Wales et al. (2016)	Cohort; Pre-post; EoT, 1-year FU	UK	AN	87	M age (SD) = 26.61 (9.06), 94.25% Female, 95.40% White	IP Group CBT +m mindfulness + psychoeducation	Illness duration, BMI	Positive outcome (achieving BMI > 17.5 within the time required to gain weight at ≥0.5 kg per week)	Linear and logistic regression	Illness duration not predictive; higher BMI predicted positive outcome
Wild et al. (2016)	RCT; Pre-post; EoT; 3-month FU; 1-year FU	Germany	AN	169	M age (SD) = 27.4 (7.8)	FPT vs CBT-E vs TAU	Illness duration, BMI	Recovered (PSR = 3 and BMI > 17.5), Full Syndrome (PSR = 1–2 and BMI < 17.5), Partial Syndrome (PSR = 2).	Linear and logistic regression	Duration not predictive; Higher BMI predicted recovery.
Binge-eating/purging ED										
Accurso et al. (2016)	RCT; Pre-post; EoT; 4-month FU	USA	BN	80	M age (SD) = 27.3 (9.6), 90% Female, 87.5% White	I-CAT vs CBT-E	EDE-Q	EDE-Q, Bulimic behaviours (frequency of binge eating & purging in past 28 days)	Generalised linear model	Higher shape concern and restraint predicted reductions in bulimic behaviours

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Baseline predictors explored	Outcome	Analysis	Findings
Anderson et al. (2020)	RCT; Pre-post; EoT, 6-month FU	USA	BED	112	M age (SD) = 35.1 (8.7), 82.1% women, 91.1% White	I-CAT vs GSH-CBT	EDE-Q	Binge eating frequency & abstinence in past 28 days	Generalised linear model	Lower restraint predicted greater binge reduction.
Camacho-Barcia et al. (2024)	Cohort; Pre-post; EoT	Spain	BN , OSFED	478	M age (SD) = 29.5 (9.6), 92.1% Female	Group CBT	Illness duration	Remission (no symptoms for 4 weeks); Partial remission (significant improvement with residual symptoms); non-remission (persistent significant symptoms); dropout	Stepwise logistic regression	Duration not predictive
Cooper et al. (2016)	RCT; Pre-post; EoT, 60-week FU	UK	BN , BED, OSFED	130	Mean age = 25.9 (7.7), 97.7% Female, 95.4% White	CBT-E vs IPT	Illness duration, EDE-Q	EDE-Q Global	Mixed effect logistic regression	Longer duration predicted worse outcome; higher shape concern predicted poorer outcomes
Dingemans et al. (2020)	Cohort; Pre-post, EoT	The Netherlands	BED	91	M age (SD) = 33.8 (9.5), 93% Female	Group CBT	Illness duration, EDE-Q	Recovery (≥50% reduction in EDE-Q &/or EDE-Q ≤2.17); binge-eating abstinence	Multivariate Cox regression	Duration not predictive; higher baseline EDE-Q predicted binge-eating abstinence
Forrest et al. (2021)	RCT, Pre-post, EoT	Treatment trial in USA	BED	191	M age (SD) = 48.4 (9.5), 71.2% Female, 84.29% White	BWL vs GSH-CBT vs medication	EDE-Q	EDE-Q, % binge-eating reduction, binge-eating abstinence and weight loss (% and >5%).	Logistic/linear regression, elastic net regression and random forests	Higher baseline weight/shape dissatisfaction and binge eating frequency predicted greater binge-eating reduction
Grilo et al. (2021)	RCT; pre-post, EoT	USA	BED	475	M age (SD) = 48.5 (11.2), 82.3% Female, 87.8% White	GSH-CBT vs therapist-led CBT	EDE-Q	Remission (zero binge-eating episodes in past month)	Logistic regression	EDE-Q not predictive

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Baseline predictors explored	Outcome	Analysis	Findings
Hilbert et al. (2020)	RCT, Pre-post; EoT; 6-month FU	Germany & Switzerland	BED	178	M age (SD) = 43.2(12.3), 87.6% Female	Face to face GSH-CBT vs internet GSH-CBT	EDE-Q	Full remission (zero objective binge eating episodes in past 28 days)	Network analysis	EDE-Q not predictive
Linardon & Fuller-Tyszkiewicz (2024)	RCT; Pre-post, 4-week & 8-week FU	Australia	BED	398	M age (SD) = 34.1 (10.2), 88% Women, 86% White	Internet GSH--CBT	EDE-Q	Remission (50% reduction in binge-eating episodes)	Multivariate analyses	Higher binge eating frequency predicted remission
Mannan et al. (2025)	RCT; Pre-post; EoT; 6 month & 12-month FU	Brazil	BN , BED	98	Mean age = 40.55 (1.18), 74.5% White	HAPIFED vs CBT-E	Illness duration, EDE-Q	EDE-Q, BMI, binge-eating frequency	Generalised linear mixed models	Illness duration not predictive; lower EDE-Q predicted better outcomes
Matherne et al. (2022)	RCT; Pre-post, EoT; 1 year FU	USA	BN	149	M age = 27.20, 97% Female	Face to face group CBT vs internet CBT	Illness duration	Abstinence from binge eating/purging over 28 days, binge-eating/ purging frequency	Generalised estimating equation models	Illness duration not predictive
Melisse et al. (2022)	Cohort; Pre-post, EoT, 20-week FU	The Netherlands	BN , BED, OSFED	625	M age = 28.28 (9.35), 95.1% Female	CBT-E	Illness duration, EDE-Q	EDE-Q Global	Bivariate and multiple regression	Illness duration not predictive; higher EDE-Q predicted higher EDE-Q at FU
Quadflieg & Fichter (2019)	Cohort; Pre-post, 11-years FU	Germany	BN	1351	M age (SD) = 25.94	IP CBT	Illness duration	Remission (no ED diagnosis for 3 months), Partial remission (ED syndrome for 3 months) Poor outcome (full ED diagnosis for 3 months). Assessed using SIAB-S.	Bivariate and logistic regression	Age at treatment predicted poor outcomes; duration dependent on age
Mixed ED Sample										
Arai et al. (2024)	Cohort; Pre-post. Mid-treatment and EoT	Australia	BN , OSFED, AN	126	M age (SD) = 23, 80.1% White	CBT-T	EDE-Q	EDE-Q, CIA	Linear regression	EDE-Q not predictive

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Baseline predictors explored	Outcome	Analysis	Findings
Austin et al. (2025)	Cohort; Pre-post, 12-month FU	UK	AN, BN , BED, OSFED	236	M age (SD) = 20.46 (2.40), 92.80% Female, 66.9% White	CBT or MANTRA	Illness duration	EDE-Q, BMI, Frequency of ED behaviours (binge eating, vomiting, laxative use or excessive exercise)	Regression	Illness duration not predictive
Day et al. (2025)	Cohort Pre-post, EoT	Australia	AN, BN , BED, OSFED	110	M age (SD) = 25.97 (7.04), 97.2% Women, 87.5% born in Australia	DP CBT + DBT vs IP CBT + DBT	EDE-Q, BMI (for pt with BMI > 18.5 only)	EDEQ, BMI (for patients with BMI > 18.5)	Stepwise hierarchal linear models	Higher EDE-Q predicted higher EDE-Q scores at EoT; BMI did predictive
Dingemans et al. (2016)	Cohort; Pre-post, EoT, 120 weeks FU	The Netherlands	AN, BN , BED, OSFED	1153	96% Female, 33%, 95% Dutch ethnic background	OP or DP or IP CBT, FT, SST, EMDR	Illness duration, EDE-Q, BMI (for pt with BMI > 18.5 only)	Reliable change in global EDE-Q, 50% symptom-reduction, clinical significance cut-off & combined RCSI criteria (50% reduction + clinical cut-off).	Cox regression models	Illness duration, EDE-Q and BMI not predictive
Fernandez-Aranda et al. (2021)	Cohort; Pre-post; EoT	Spain	AN, BN , BED, OSFED	1199	M age (SD) = 27.6 (9.0), 90.1% Female	CBT (DP for AN, OP for others)	Illness duration	Full remission (total absence of symptoms > 4 weeks), partial remission (substantial symptomatic improvement but with residual symptoms), non-remission	Stepwise logistic regressions	Illness duration not predictive
Keegan & Wade (2024)	Cohort; Pre-post; EoT; 1 month & 6 month FU	Australia	BN , OSFED, AN	37	M age (SD) = 27.10 (6.47), 94.50% Female, 86.50% White	CBT-T	EDE-Q	Good outcome (EDE-Q global score < 2.77), abstinence (no bulimic behaviours), remission (no bulimic behaviours plus EDEQ < 2.77)	Logistic regression	EDE-Q not predictive
Riesco et al. (2018)	Cohort; Pre-post; EoT	Spain	OSFED	176	M age (SD) = 26.41 (8.2), 100% Female, 94.31% Spanish	Group CBT	Illness duration	Full remission (total absence of symptoms meeting diagnostic criteria for at least 4 weeks), partial remission (substantial improvement but with residual symptoms), non-remission and drop-out	Stepwise logistic regression	Illness duration not predictive
Simpson et al. (2021)	Cohort; Pre-post; EoT	USA	AN, BN , BED, OSFED	788	M age (SD) = 20.87 (8.35), 91.6% Female, 73.3% White	DP DBT	Illness duration, EDE-Q, BMI (for pt with BMI > 18.5)	Higher level of care	Logistic regression models	Lower BMI predicted need for higher care; EDE-Q and illness duration not predictive

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Baseline predictors explored	Outcome	Analysis	Findings
Svendsen et al. (2023)	Cohort; Pre-post, mid-treatment; EoT	The Netherlands	AN, BN , BED, OSFED	593	M age (SD) = 27 (8.9), 98% Female,	Insight therapy + CBT	EDE-Q	Non-response (reliable change score of <-1.41 on the EDE-Q); Reliably improved (change score of >-1.41)	Machine learning	EDE-Q not predictive
Walker et al. (2020)	Cohort; Pre-post; EoT	USA	AN, BN , BED, OSFED	210	M age (SD) = 25.10 (10.43), 89.4% Female, 94.2% White,	DP CBT + DBT	Illness duration	EAT-26, ORS; Remission (EAT-26 > 18.34)	Regression analyses	Illness duration not predictive

Note. AN = Anorexia nervosa; BDI = Beck Depression Inventory; BED = Binge-eating disorder; BMI = Body mass index; BN = Bulimia nervosa; BWL = Behavioural weight loss; CAT = Cognitive analytic therapy; CBT = Cognitive behavioural therapy; CBT-E = Enhanced cognitive behavioural therapy; CBT-T = Ten-session cognitive behavioural therapy; CGI = Clinical global impression; CIA = Clinical impairment assessment; CREST = Cognitive remediation and emotional skills training; DBT = Dialectical behaviour therapy; DP = Day programme; EAT-26 = Eating attitude test; ED = Eating disorder; EDE-Q = Eating Disorder Examination Questionnaire; EDI-2 = Eating Disorder Inventory-2; EDQOL = Eating Disorder Quality of Life Scale; EMDR = Eye movement desensitisation and reprocessing; EoT = End of treatment; FPT = Focal psychodynamic therapy; FT = Family therapy; FU = Follow-up; GAF = Global assessment of functioning; GSH = Guided self-help; HAPIFED = Healthy Approach to weight management and Food in Eating Disorders; I-CAT = Integrated cognitive-affective therapy; IP = Inpatient; IPT = Interpersonal psychotherapy; M = mean; MANTRA = Maudsley model of anorexia nervosa treatment for adults; MBT = Mentalisation-based therapy; MET = Motivational enhancement therapy; OP = Outpatient; ORS = Outcome Rating Scale; OSFED = Other specified feeding or eating disorder; PSR = Psychiatric status rating; Pt = Patients; QLESQ = Quality of Life Enjoyment and Satisfaction Questionnaire; QoL = Quality of Life; RCT = Randomised controlled trial; SD = Standard Deviation; SIAB-S = Structured Interview for Anorexic and Bulimic Disorders–Self-report; SSCM = Specialist supportive clinical management; SST = Social skills training; TAU = Treatment as usual; UK = United Kingdom; USA = United States; WSAS = Work and Social Adjustment Scale

Early Response in severe and enduring eating disorders

No studies examining early response as a predictor of treatment outcomes in SE-ED was identified. However, as an exploratory step outside the original review protocol, the result sections of the included studies were examined to determine whether any had reported associations between illness duration or severity and early response.

Three included studies reported analyses examining whether illness duration or severity predicted early response (see table 5). These studies used varying definitions of early response, including categorical or dimensional measures of BMI or symptom change within varying timeframes (5 – 7 weeks).

Across all three studies, illness duration did not predict the likelihood of achieving early response. Illness severity, when measured with the EDE-Q, also did not reduce the likelihood of achieving early response. When illness severity was measured with BMI, findings were mixed. One study reported no difference, while one reported a lower BMI led to greater early change and one reported a lower BMI decreased the likelihood of achieving early response.

Table 5. Illness duration and severity as predictors of early response

Study	Study design	Country	Dx	n	Sample demographic	Intervention	Early response	Analysis	Findings
Raylos et al. (2018)	Cohort; Pre-post; EoT	Australia	AN	134	M age (SD) = 22 (18.25), 97% Female, 76.9% Austrian born,	CBT-E	Change in BMI over first 5 sessions	Logistic & longitudinal regression	Illness duration did not predict early BMI change or early symptom trajectory. Baseline EDE-Q did not predict early BMI change. Baseline BMI predicted greater early BMI change. Early change occurred equally across duration/severity levels - SE-ED (illness duration ≥7 years and BMI < 16) vs non SE-ED
Wales et al. (2016)	Cohort; Pre-post; EoT, 1-year FU	UK	AN	87	M age (SD) = 26.61(9.06), 94.25% Female, 95.40% White	IP Group CBT + mindfulness + psychoeducation	Weight gain of 0.5-1kg/week during first 6 weeks	Logistic regression and group comparisons between early responders vs non-responders	Illness duration did not predict early response. Symptom severity (EDE-Q) did not differentiate early responders from non-responders Higher baseline BMI increased likelihood of meeting early weight-gain criteria
Walker et al. (2020)	Cohort; Pre-post; EoT	USA	AN, BN, BED, OSFED	210	M age (SD) = 25.10 (10.43), 89.4% Female, 94.2% White	DP CBT + DBT	≥16.52% EAT-26 reduction by week 7	Logistic regression and ROC analyses testing	Illness duration did not predict early response. Baseline BMI did not differ between early responders and non-responders.

Note. anorexia nervosa = Anorexia nervosa; BED = Binge-eating disorder; BMI = Body mass index; BN = Bulimia nervosa; CBT = Cognitive behavioural therapy; CBT-E = Enhanced cognitive behavioural therapy; DBT = Dialectal behaviour therapy; DP = Day patient; EAT-26 = Eating attitude test; EDE-Q = Eating Disorder Examination Questionnaire; EoT = End of treatment; FU = Follow-up; IP = Inpatient; M = Mean; OSFED = Other specified feeding or eating disorder; ROC = Receiver operating characteristics; SD = Standard deviation; SE-ED = Severe and enduring eating disorder; UK = United Kingdom; USA = United States

Discussion

Summary of findings

This systematic review had two primary aims. First, it examined how the SE-ED construct has been characterised in the literature on psychological treatments for this population. Second, it examined whether duration or severity of an eating disorder influences psychological treatment outcomes. A secondary exploratory aim was to investigate whether the literature has examined the early response effect in SE-ED. The review synthesised findings from 44 studies, fourteen of which provided an explicit definition of SE-ED and 35 of which examined illness duration and/or severity as predictors of treatment outcomes.

Across the fourteen studies characterising SE-ED, most (78.57%) focused solely on anorexia nervosa or underweight other specified feeding and eating disorder. The most commonly used definition of SE-ED was an illness duration of at least seven years (64.29%); however, this definition was not applied consistently. Some studies combined this duration threshold with different markers of clinical severity, whereas others used entirely different criteria related to either duration or severity. The lack of conceptual and operational consensus regarding SE-ED is consistent with a previous review that reported similar findings (Broomfield et al., 2017). However, this earlier review focused exclusively on anorexia nervosa studies with varied methodologies and treatment approaches. To our knowledge, the present review is the first to synthesise how SE-ED is operationalised across diagnostic subtypes within empirical studies investigating psychological therapies.

A central assumption underlying many SE-ED frameworks is that longer illness duration and greater baseline severity lead to poorer treatment outcomes because the disorder and its maintaining mechanisms becomes deeply entrenched over time (Treasure et al., 2015). However, the findings of the present review do not support this assumption. Across studies examining illness duration as a predictor of treatment outcomes, chronicity was not

consistently associated with outcome. This remained true across all diagnostic subtypes and regardless of treatment modality or intensity. Similarly, illness severity was not consistently associated with treatment outcomes, irrespective of eating disorder diagnosis, treatment modality, treatment intensity, or whether severity was operationalised using EDE-Q or BMI. These findings are consistent with the meta-analysis and systematic review by Radunz et al. (2021). However, unlike the previous review, the present study examined severity and duration together and included additional, more recent studies, thereby providing further evidence that SE-ED lacks predictive value in relation to treatment outcomes.

With respect to the exploratory aim, the review identified no studies that directly investigated whether early response predicts treatment outcomes in patients explicitly labelled with SE-ED. As an additional, post hoc step outside the original protocol, the results of included studies were examined to determine whether any had explored associations between illness duration or severity and early response. Only three studies reported analyses relevant to this question. Across these studies, no significant associations were identified between illness duration or severity and early response, regardless of how early response was defined, treatment modality, or diagnostic subtype. Overall, the present review provides only limited and indirect evidence regarding the prognostic role of early response in SE-ED, and conclusions must therefore remain tentative.

Heterogeneity and quality across the studies

The lack of consensus identified across the studies must be interpreted in the context of substantial methodological heterogeneity. Considerable variation existed in study design, treatment modality, and clinical setting, limiting the direct comparability of findings. For example, while some studies evaluated outcomes within structured, protocol-driven psychological therapies, others employed more integrative and flexible interventions delivered in inpatient settings or day programmes. In many cases, insufficient detail regarding

the interventions was provided, making it difficult to determine whether variability in outcomes reflected differences in treatment content, intensity, or contextual factors rather than true differences in prognostic indicators.

Inconsistencies in the operationalisation of outcome measures further complicate interpretation. Several studies relied on dimensional indicators such as BMI or self-reported eating disorder psychopathology. However, such approaches are inherently sensitive to baseline values. Individuals with higher baseline BMI scores may appear to demonstrate more favourable outcomes at end of treatment simply because less change is required to reach a comparatively higher endpoint. Conversely, individuals starting at a lower BMI may achieve substantial and clinically meaningful weight gain yet still appear to have poorer outcomes when evaluated solely on absolute end-point values. This methodological limitation may have contributed to inconsistent conclusions regarding treatment response and predictor effects across studies.

One approach that addresses these limitations is the use of reliable and clinically significant improvement (RCSI – Jacobson & Traux, 1991), which requires observed change to exceed measurement error and reflect a shift from the clinical to non-clinical range. RCSI provides a more meaningful and standardised method of evaluating treatment response and permits more valid cross-study comparisons (Blampied, 2022). Despite its recognised advantages, only two studies in the current review applied the full RCSI criteria when evaluating outcomes (Dingemans et al., 2016; Li et al., 2020), highlighting a broader limitation in outcome measurement within this literature.

Risk of bias further limits the strengths of conclusions that can be drawn. The majority of included studies were rated as having a moderate risk of bias, with two studies classified as moderate-to-high risk. Common methodological concerns included the absence of power analyses for predictor models, incomplete reporting of statistical methods and

results, insufficient details regarding interventions, lack of follow-up assessments, and limited generalisability. Together, these issues limit confidence in the robustness and external validity of the findings and highlights the need for more rigorous and transparent research in this area.

Strengths and limitations

The present review addresses a notable gap in the existing literature by examining eating disorder severity alongside illness duration within the context of SE-ED. Previous reviews have primarily focused on illness duration in isolation (e.g., Radunz et al., 2020) or examined predictors of treatment outcomes more broadly (e.g., Gregersten et al., 2019; Vall & Wade, 2015). To our knowledge, this is the first systematic review to jointly evaluate severity and duration as potential predictors of treatment outcomes across all eating disorder diagnostic subtypes while also tentatively exploring the role of early response in SE-ED.

Several methodological strengths enhance the robustness of this review. The study protocol was pre-registered and published prior to search and data extraction, in line with best-practice guidelines, thereby reducing the risk of selective reporting and enhances transparency. The review was conducted in accordance with PRISMA guidelines (Moher et al., 2009), strengthening methodological quality and facilitating replicability. Additionally, a second reviewer independently screened titles and abstracts and contributed to the quality appraisal process, improving reliability and reducing the likelihood of selection and assessment bias.

Despite these strengths, several limitations of this study must be acknowledged. One limitation relates to how symptom severity was operationalised. Although the EDE-Q offers a robust and widely used measure of eating disorder psychopathology, it relies on self-report and is therefore vulnerable to reporting biases and measurement error. BMI was also used to operationalise severity in anorexia nervosa due to its frequent use in diagnostic classification

systems, the centrality of weight gain as a treatment goal, and its role in determining thresholds for medical risk and intensive care (APA, 2013; NICE, 2017). However, several studies have identified weak or no association between BMI and psychopathology, suggesting BMI-based severity classifications may be a poor proxy for true clinical severity in anorexia nervosa (Dang et al., 2022).

A substantial proportion of included studies (n = 32, 72.73%) were rated as having moderate or higher risk of bias, which limits confidence in the reliability of the overall findings. Grey literature was excluded due to the absence of formal peer-review processes (Paez et al., 2017), increasing the possibility of publication bias. Furthermore, a secondary search strategy involving hand-searching reference lists was not undertaken due to time constraints and the large volume of eligible studies, increasing the risk of selection bias.

Risk of bias was assessed using the CASP checklists, selected for their suitability for the study designs included (RCTs and cohort studies). However, not all CASP items were fully applicable to the studies examining predictors, and the tool does not permit detailed appraisal of statistical modelling decisions or the appropriateness of predictor analyses. Finally, the exclusion of non-English-language publications may have introduced additional bias, and the predominance of Western samples limits the generalisability of the findings across cultures and countries.

Clinical and research implications

The findings of this review reveals that the construct of SE-ED lacks a consistent operational definition and demonstrates limited validity as a predictor of psychological treatment outcomes. The review also tentatively suggests that early response warrant further investigation within SE-ED populations. Taken together, these findings raise questions regarding the clinical utility of SE-ED as a meaningful treatment-stratification tool and suggests the need for critical revaluation of its conceptual foundations.

From a clinical perspective, these findings caution against the routine use of perceived chronicity or severity as a basis for limiting access to evidence-based psychological treatments. Individuals labelled as having SE-ED are often viewed as unlikely to benefit from standard interventions and may instead be offered approaches with limited empirical support (Wonderlich et al., 2012). However, the absence of consistent prognostic effects of duration and severity suggests that such assumptions may be unfounded. Clinicians should therefore avoid prematurely abandoning active, evidence-based treatments solely on the basis of illness chronicity or severity and instead prioritise prognostic factors that demonstrate stronger empirical support in the wider eating disorder literature, such as early response.

From a research standpoint, the lack of predictive validity associated with SE-ED highlights the need to reconsider the construct itself. The empirical basis for the commonly used criterion underlying SE-ED also appears uncertain. For example, the frequently cited seven-year illness duration criterion does not appear to originate from a clearly established empirical cut-off. This remains true for the 10-year criterion used by Arkell & Robinson (2008) when introducing SE-ED to the literature. Rather, it may have been loosely informed by patterns observed in long-term outcome studies, in which rates decline rapidly during the early years of illness before gradually plateauing, producing an asymptotic recovery curve approximately six to twelve years after onset (Robinson, 2014). Similarly, proposed severity thresholds appear to be derived largely from clinical risk indicators rather than prognostic evidence. Robinson (2014), for instance, suggested a body mass index (BMI) below 13 as representing the highest level of clinical risk based primarily on clinical experience rather than predictive validity.

Although early conceptualisations of SE-ED may have emerged from clinicians' attempts to better understand and support individuals experiencing prolonged illness, high medical risk, or repeated treatment non-response, more recent discourse suggests that the

construct is being shaped by broader sociopolitical pressures. The SE-ED label may increasingly function to justify the withdrawal of evidence-based treatment due to perceived financial burden on healthcare systems and wider society (Ahmed et al., 2024). In parallel, some authors have proposed that establishing a formal definition of SE-ED could facilitate access to medical assistance in dying (Gaudiani et al., 2022). Such proposals are particularly concerning in the context of limited empirical support and raise substantial ethical concerns, particularly given ongoing debate regarding decision-making capacity in the context of severe malnutrition and its associated cognitive impairments in eating disorders (Hirst et al., 2017).

Finally, future research should prioritise the investigation of early response as a predictor of outcome specifically within SE-ED populations and examine whether illness duration or severity predicts early response. If early response remains robust irrespective of illness duration and severity, researchers can then focus on identifying the mechanisms and symptom domains that account for early change to inform the refinement of therapeutic targets and enhance treatment outcomes in eating disorders.

Conclusions

In summary, this review highlights considerable inconsistency in how SE-ED is defined and finds little evidence that illness severity or duration reliably predicts psychological treatment outcomes. Evidence regarding early response was sparse and indirect; however, the limited available data suggest it may warrant further investigation as a potential predictor of treatment outcome, regardless of illness duration or baseline severity.

References

Accurso, E. C., Wonderlich, S. A., Crosby, R. D., Smith, T. L., Klein, M. H., Mitchell, J. E., Crow, S. J., Berg, K. C., & Peterson, C. B. (2016). Predictors and moderators of treatment outcome in a randomized clinical trial for adults with symptoms of bulimia nervosa. *Journal of Consulting and Clinical Psychology*, *84*(2), 178-184. https://doi.org/10.1037/ccp0000073*

Ahmed, M., Islam, M. D., Aouad, P., Miskovic-Wheatley, J., Touyz, S., Maguire, S., & Cunich, M. (2025). Global and Regional Economic Burden of Eating Disorders: A Systematic Review and Critique of Methods. *The International journal of eating disorders*, *58*(1), 91–116. https://doi.org/10.1002/eat.24302

Ambwani, S., Cardi, V., Albano, G., Cao, L., Crosby, R. D., Macdonald, P., Schmidt, U., & Treasure, J. (2020). A multicenter audit of outpatient care for adult anorexia nervosa: Symptom trajectory, service use, and evidence in support of "early stage" versus "severe and enduring" classification. *The International Journal of Eating Disorders*, *53*(8), 1337–1348. https://doi.org/10.1002/eat.23246*

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Anderson, L. M., Smith, K. M., Schaefer, L. M., Crosby, R. D., Cao, L., Engel, S. G., Crow, S. J., Wonderlich, S. A., & Peterson, C. B. (2020). Predictors and moderators of treatment outcome in a randomized clinical trial for binge-eating disorder. *Journal of Consulting and Clinical Psychology*, *88*(7), 631-642. https://doi.org/10.1037/ccp0000503*

Arai, M., Tonta, K. E., Erceg-Hurn, D. M., Raykos, B. C., Waller, G., & McEvoy, P. M. (2025). Eating-specific interpersonal difficulties: Changes and impacts on outcomes during ten-session cognitive behavioral therapy for eating disorders (CBT-T).

International Journal of Eating Disorders, *58*(2), 362-371. https://doi.org/10.1002/eat.24336*

Arkell, J., & Robinson, P., (2008). A pilot case series using qualitative and quantitative methods: Biological, psychological and social outcome in severe and enduring eating disorder (anorexia nervosa). *International Journal of Eating Disorders*, *41*(7), 650–656. https://doi.org/10.1002/eat.20546

Austin, A., Flynn, M., Richards, K., Allen, K., Mountford, V. A., Glennon, D., Brown, A., Grant, N., Franklin-Smith, M., & Schelhase, M. (2025). The relationship between duration of untreated symptoms and clinical outcomes in first episode eating disorders. *Psychiatry Research Communications*, *5*(2). https://doi.org/10.1016/j.psycom.2025.100217*

Austin, A., Flynn, M., Richards, K., Hodsoll, J., Duarte, T. A., Robinson, P., Kelly, J., & Schmidt, U. (2021). Duration of untreated eating disorder and relationship to outcomes: A systematic review of the literature. *European Eating Disorder Review, 29*(3), 329-345. https://doi.org/10.1002/erv.2745

Bamford, B., Barras, C., Sly, R., Stiles-Shields, C., Touyz, S., Le Grange, D., Hay, P., Crosby, R., & Lacey, H. (2015). Eating disorder symptoms and quality of life: where should clinicians place their focus in severe and enduring anorexia nervosa?. *The International journal of eating disorders*, *48*(1), 133–138. https://doi.org/10.1002/eat.22327*

Bardone-Cone, A. M., Harney, M. B., Maldonado, C. R., Lawson, M. A., Robinson, D. P., Smith, R., & Tosh, A. (2010). Defining recovery from an eating disorder: conceptualisation, validation, and examination of psychosocial functioning and psychiatric comorbidity. *Behaviour Research and Therapy, 48*(3), 194-202. https://doi.org/10.1016/j.brat.2009.11.001

Blampied, N. M. (2022). Reliable change and the reliable change index: still useful after all these years? *The Cognitive Behaviour Therapist, 25,* e50. https://doi.org/10.1017/S1754470X22000484

Broomfield, C., Stedal, K., Touyz, S., & Rhodes, P. (2017). Labelling and defining severe and enduring anorexia nervosa: A systematic review and critical analysis. *International Journal of Eating Disorders, 50*(6), 611-623. https://doi.org/10.1002.eat.22715

Calugi, S., Cattaneo, G., Chimini, M., Dalle Grave, A., Conti, M., & Dalle Grave, R. (2025). Predictors of Intensive Enhanced Cognitive Behavioral Therapy for Anorexia Nervosa. Prospective Cohort Study. *The International journal of eating disorders*, *58*(8), 1513–1522. https://doi.org/10.1002/eat.24455*

Calugi, S., El Ghoch, M., & Dalle Grave, R. (2017). Intensive enhanced cognitive behavioural therapy for severe and enduring anorexia nervosa: A longitudinal outcome study. *Behaviour Research and Therapy*, *89*, 41-48. https://doi.org/10.1016/j.brat.2016.11.006*

Calugi, S., Sartirana, M., Frostad, S., & Dalle Grave, R. (2021). Enhanced cognitive behavior therapy for severe and extreme anorexia nervosa: An outpatient case series. *International Journal of Eating Disorders*, *54*(3), 305-312. https://doi.org/10.1002/eat.23428*

Camacho-Barcia, L., Sánchez, I., Ibáñez-Caparrós, A., Ohsako, N., Granero, R., Artero, C., Crespo, J. M., Paslakis, G., Jiménez-Múrcia, S., & Fernandez-Aranda, F. (2024). The Impact of High Levels of Compensatory Exercise on Treatment Outcomes in Threshold and Subthreshold Bulimia Nervosa. *Nutrients*, *16*(14). https://doi.org/10.3390/nu16142337*

Chang, P. G. R. Y., Delgadillo, J. & Waller, G. (2021). Early response to psychological treatment for eating disorders: A systemic review and meta-analysis. *Clinical Psychology Review, 86*, 1-14. https://doi.org/10.1016/j.cpr.2021.102032

Cohen, K. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Cooper, Z., Allen, E., Bailey-Straebler, S., Basden, S., Murphy, R., O'Connor, M. E., & Fairburn, C. G. (2016). Predictors and moderators of response to enhanced cognitive behaviour therapy and interpersonal psychotherapy for the treatment of eating disorders. *Behaviour Research and Therapy*, *84*, 9-13. https://doi.org/10.1016/j.brat.2016.07.002*

Critical Appraisal Skills Programme. (2024). *CASP checklists*. CASP. https://casp-uk.net/casp-tools-checklists/

Dang, A. B., Giles, S., Fuller-Tyszkiewicz, M., Kiropoulos, L., & Krug, I. (2022). A systematic review and meta-analysis on the *DSM–5* severity ratings for eating disorders. *Clinical Psychology: Science and Practice, 29*(4), 325–344. https://doi.org/10.1037/cps0000078

Dapelo, M. M., Gill, A. A., Lacelle, L., & Vogel, M. (2020). Severity and endurance in eating disorders: An exploration of a clinical sample from Chile. *Frontiers in Psychiatry, 11*(869). https://doi.org/10.3389/fpsyt.2020.00869

Davidsen, A. H., Hoyt, W. T., Poulsen, S., Waaddegaard, M., & Lau, M. (2017). Eating disorder severity and functional impairment: moderating effects of illness duration in a clinical sample. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity, 22*(3), 499-507. https://doi.org/10.1007/s40519-016-0319-z

Day, S., Mitchison, D., Mannan, H., Tannous, W. K., Conti, J., Dearden, A., Doyle, A. K., Gill, K., Hannigan, A., Houlihan, C., Ramjan, L., Rankin, R., Valentine, N., & Hay, P. (2025). Residential versus day program treatment for eating disorders: A comparison of post-treatment outcomes and predictors. *Journal of affective disorders*, *371*, 177–186. https://doi.org/10.1016/j.jad.2024.11.054*

Dingemans, A. E., van Son, G. E., Aardoom, J. J., Bruidegom, K., Slof-Op't Landt, M. C. T., & Van Furth, E. F. (2016). Predictors of psychological outcome in patients with eating disorders: A routine outcome monitoring study. *International Journal of Eating Disorders*, *49*(9), 863-873. https://doi.org/10.1002/eat.22560*

Dingemans, A. E., van Son, G. E., Vanhaelen, C. B., & Furth, E. F. (2020). Depressive symptoms rather than executive functioning predict group cognitive behavioural therapy outcome in binge eating disorder. *European Eating Disorders Review*, *28*(6), 620-632. https://doi/org/10.1002/erv.2768*

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy, 41*(5), 509-528. https://doi.org/10.1016/s0005-7967(02)00088-8

Fairburn, C., & Beglin, S. J. (2008). Eating disorders examination questionnaire (EDE-Q 6.0) In: C. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders.* New York: Guilford Press.

Fernandez-Aranda, F., Treasure, J., Paslakis, G., Aguera, Z., Gimenez, M., Granero, R., Sanchez, I., Serrano-Troncoso, E., Gorwood, P., Herpertz-Dahlmann, B., Bonin, E.-M., Monteleone, P., & Jimenez-Murcia, S. (2021). The impact of duration of illness on treatment nonresponse and drop-out: Exploring the relevance of enduring eating disorder concept. *European Eating Disorders Review*, *29*(3), 499-513. https://doi.org/10.1002/erv.2822*

Fisher, C. A., Skocic, S., Rutherford, K. A., & Hetrick, S. E. (2018). Family therapy approaches for anorexia nervosa. *The Cochrane database of systematic reviews*, *10*(10), CD004780. https://doi-org/10.1002/14651858.CD004780.pub3

Forrest, L. N., Ivezaj, V., & Grilo, C. M. (2023). Machine learning v. traditional regression models predicting treatment outcomes for binge-eating disorder from a randomized controlled trial. *Psychological Medicine*, *53*(7), 2777-2788. https://doi.org/10.1017/S0033291721004748*

Frank, G. K., Stoddard, J. J., Brown, T., Gowin, J., & Kaye, W. H. (2024). Weight gained during treatment predicts 6-month body mass index in a large sample of patients with anorexia nervosa using ensemble machine learning. *International Journal of Eating Disorders*, *57*(8), 1653-1667. https://doi.org/10.1002/eat.24208*

Frostad, S., Calugi, S., Engen, C. B., & Dalle Grave, R. (2021). Enhanced cognitive behaviour therapy (CBT-E) for severe and extreme anorexia nervosa in an outpatient eating disorder unit at a public hospital: A quality-assessment study. *Journal of Eating Disorders, 9*(149). https://doi.org/10.1186/s40337-021-00499-1*

Garte, M., Hagen, B., Reas, D. L., Isdahl, P. J., Hinderaker, E., & Ro, O. (2015). Implementation of a day hospital treatment programme based on CBT-E for severe eating disorders in adults: An open trial. *Advances in Eating Disorders*, *3*(1), 48-62. https://doi.org/10.1080/21662630.2014.958510*

Gaudiani, J. L., Bogertz, A., & Yager, J. (2022). Terminal anorexia nervosa: three cases and proposed clinical characteristics. *Journal of Eating disorders, 10*(23).) https://doi.org/10.1186/s40337-022-00548-3

Gregertsen, E. C., Mandy, W., Kanakam, N., Armstrong, S., & Serpell, L. (2019). Pre-treatment patient characteristics as predictors of drop-out and treatment outcome in individual and family therapy for adolescents and adults with anorexia nervosa: A

systematic review and meta-analysis. *Psychiatry Research*, *271,* 484–501. https://doi.org/10.1016/j.psychres.2018.11.068

Grilo, C. M., Thompson-Brenner, H., Shingleton, R. M., Thompson, D. R., & Franko, D. L. (2021). Clinical moderators and predictors of cognitive-behavioral therapy by guided-self-help versus therapist-led for binge-eating disorder: Analysis of aggregated clinical trials. *International Journal of Eating Disorders*, *54*(10), 1875-1880. https://doi.org/10.1002/eat.23601*

Hay, P. J., Touyz, S., & Sud, R. (2012). Treatment for severe and enduring anorexia nervosa: A review. *Australian & New Zealand Journal of Psychiatry, 46*(12), 1136-1144. https://doi.org/10.1177/0004867412450469

Hay, P., & Touyz, S. (2015). Treatment of patients with severe and enduring eating disorders. *Current Opinion in Psychiatry, 28*(6), 473-477. https://doi.org/10.1097/YCO.0000000000000191

Hay, P., & Touyz, S. (2018). Classification challenges in the field of eating disorders: Can severe and enduring anorexia nervosa be better defined? *Journal of Eating Disorders,* 6(41). https://doi.org/10.1186/s40337-018-0229-8

Hilbert, A., Herpertz, S., Zipfel, S., Tuschen-Caffier, B., Friederich, H. C., Mayr, A., & de Zwaan, M. (2020). Psychopathological Networks in Cognitive-Behavioral Treatments for Binge-Eating Disorder. *Psychotherapy and Psychosomatics*, *89*(6), 379-385. https://doi.org/10.1159/000509458*

Hilbert, A., Herpertz, S., Zipfel, S., Tuschen-Caffier, B., Friederich, H. C., Mayr, A., Crosby, R. D., & de Zwaan, M. (2019). Early change trajectories in cognitive-behavioral therapy for binge-eating disorder. *Behavior Therapy, 50*, 115–125. https://doi.org/10.1016/j.beth.2018.03.013

Hirst, R. B., Beard, C. L., Colby, K. A., Quittner, Z., Mills, B. M., & Lavender, J. M. (2017). Anorexia nervosa and bulimia nervosa: A meta-analysis of executive functioning. *Neuroscience & Biobehavioural Reviews, 83,* 678-690. https://doi.org/10.1016/j.neubiorev.2017.08.011

Ibrahim, A., Ryan, S., Viljoen, D., Tutisani, E., Gardner, L., Collins, L., & Ayton, A. (2022). Integrated enhanced cognitive behavioural (I-CBTE) therapy significantly improves effectiveness of inpatient treatment of anorexia nervosa in real life settings. *Journal of Eating Disorders, 10*(98). https://doi.org/10.1186/s40337-022-00620-y*

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. https://doi.org/10.1037//0022-006x.59.1.12

Kaufmann, L.-K., Moergeli, H., & Milos, G. F. (2021). Lifetime weight characteristics of adult inpatients with severe anorexia nervosa: Maximal lifetime BMI predicts treatment outcome. *Frontiers in Psychiatry, 12.* https://doi.org/10.3389/fpsyt.2021.682952*

Keegan, E., & Wade, T. D. (2024). Early change predicts outcome in 10-session cognitive behavioural therapy for non-underweight patients with eating disorders: A secondary data analysis. *Clinical Psychologist*, *28*(1), 70-74. https://doi.org/10.1080/13284207.2024.2306934*

Keel, P. K., & Brown, T. A. (2010). Update on course and outcome in eating disorders. *International Journal of Eating Disorders, 43*(3), 195-204. https://doi.org/10.1002/eat.20810

Kotilahti, E., West, M., Isomaa, R., Karhunen, L., Rocks, T., & Ruusunen, A. (2020). Treatment interventions for Severe and Enduring Eating Disorders: Systematic review. *International Journal of Eating Disorders*, *53*(8), 1280–1302.

https://doi.org/10.1002/eat.23322

Li, Z., Dandil, Y., Toloza, C., Carr, A., Oyeleye, O., Kinnaird, E., & Tchanturia, K. (2020). Measuring clinical efficacy through the lens of audit data in different adult eating disorder treatment programmes. *Frontiers in Psychiatry, 11*. https://doi.org/10.3389/fpsyt.2020.599945*

Linardon, J., & Fuller-Tyszkiewicz, M. (2024). Exploration of the individual and combined effects of predictors of engagement, dropout, and change from digital interventions for recurrent binge eating. *International Journal of Eating Disorders*, *57*(5), 1202-1212. https://doi.org/10.1002/eat.24175*

Linardon, J., Brennan, L., & de la Piedad Garcia, X. (2016). Rapid response to eating disorder treatment: A systematic review and meta-analysis. *International Journal of Eating Disorders, 49*(10), 905-919. https://doi.org/10.1002/eat.22595

Luce, K. H., & Crowther, J. H. (1999). The reliability of the eating disorder examination – self report questionnaire version (EDE-Q). *International Journal of Eating Disorder, 23*(3), 349-351. https://doi.org/10.1002/(SICI)1098-108X(199904)25:3<349::AID-EAT15>3.0.CO;2-M

Lydecker, J. A., Ivezaj, V., & Grilo, C. M. (2020). Testing the validity and clinical utility of the severity specifiers for binge-eating disorder for predicting treatment outcomes. *Journal of Consulting and Clinical Psychology*, *88*(2), 172-178. https://doi.org/10.1037/ccp0000464*

MacDonald, D. E., Trottier, K., McFarlane, T., & Olmsted, M. P. (2015). Empirically defining rapid response to intensive treatment to maximize prognostic utility for bulimia nervosa and purging disorder. *Behaviour Research and Therapy, 68,* 48-53. https://doi.org/10.1016/j.brat.2015.03.007

Mannan, H., Palavras, M. A., Claudino, A., & Hay, P. J. (2025). Predictors and Moderators of

Outcomes in a Trial of Cognitive Behavioural Therapy Integrated with Behavioural Weight Loss for High Weight Individuals with Disorders of Recurrent Binge Eating. *Nutrients*, *17*(7). https://doi.org/10.3390/nu17071288*

Marzola, E., Albini, E., Delsedime, N., Fassino, S., & Abbate-Daga, G. (2019). Therapeutic alliance in inpatients with severe anorexia nervosa. *European Eating Disorders Review*, *27*(6), 671-681. https://doi.org/10.1002/erv.2687*

Marzola, E., Martini, M., Brustolin, A., & Abbate-Daga, G. (2021). Inpatients with severe-enduring anorexia nervosa: Understanding the "enduringness" specifier. *European Psychiatry, 64*(1). https://doi.org/10.1192/j.eurpsy.2021.2218*

Matherne, C. E., Watson, H., Fassnacht, D. B., Ali, K., Zerwas, S., Peat, C., Runfola, C., Levine, M. D., Marcus, M. D., Zimmer, B., Moessner, M., Crosby, R., & Bulik, C. M. (2022). An exploratory investigation of predictors of outcome in face-to-face and online cognitive-behavioural therapy for bulimia nervosa. *European eating disorders review : the journal of the Eating Disorders Association*, *30*(4), 373–387. https://doi.org/10.1002/erv.2898*

Melisse, B., Dekker, J., van den Berg, E., de Jonge, M., Van Furth, E. F., Peen, J., & de Beurs, E. (2022). Comparing the effectiveness and predictors of cognitive behavioural therapy-enhanced between patients with various eating disorder diagnoses: A naturalistic study. *Cognitive Behaviour Therapist*, *15*. https://doi.org/10.1017/S1754470X22000174*

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*, *14*(1), 1–10. https://doi.org/10.1186/s12913-014-0579-0

Meule, A., Kolar, D. R., Rauh, E., & Voderholzer, U. (2023). Comparing illness duration and

age as predictors of treatment outcome in female inpatients with anorexia nervosa. *Eating Disorders*, *31*(3), 274-284. https://doi.org/10.1080/10640266.2022.2114586*

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Monteleone, A. M., Pellegrino, F., Croatto, G., Carfagno, M., Hilbert, A., Treasure, J., Wade, T., Bulik, C. M., Zipfel, S., Hay, P., Schmidt, U., Castellini, G., Favaro, A., Fernandez-Aranda, F., Il Shin, J., Voderholzer, U., Ricca, V., Moretti, D., Busatta, D., … Solmi, M. (2022). Treatment of eating disorders: A systematic meta-review of meta-analyses and network meta-analyses. *Neuroscience and Biobehavioral Reviews*, *142*, 104857. https://doi.org/10.1016/j.neubiorev.2022.104857

National Institute for Health and Care Excellence. (2017). *Eating disorders: recognition and treatment* (NICE Guideline NG69). https://www.nice.org.uk/guidance/ng69

Paez, A. (2017). Grey literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine, 10*(3), 233-240. https://doi.org/10.1111/jebm.12265

Piñar-Gutiérrez, A., Dios-Fuentes, E., Remón-Ruiz, P., Del Can-Sánchez, D., Vázquez-Morejón, A., López-Narbona, M., Dastis-Rodríguez De Guzmán, J., Venegas-Moreno, E., & Soto-Moreno, A.. (2021). Description of characteristics and outcomes of a cohort of patients with severe and enduring eating disorders (SE-ED). *Journal of Eating Disorders*, *9*(1). https://doi.org/10.1186/s40337-021-00492-8

Quadflieg, N., & Fichter, M. M. (2019). Long-term outcome of inpatients with bulimia nervosa—Results from the Christina Barz Study. *International Journal of Eating Disorders*, *52*(7), 834-845. https://doi.org/10.1002/eat.23084*

Radunz, M., Keegan, E., Osenk, I., & Wade, T. D. (2020). Relationship between eating disorder duration and treatment outcome: Systematic review and meta-analysis. *International Journal of Eating Disorders, 53*(11), 1761-1773. https://doi.org/10.1002/eat.23373

Raykos, B. C., Erceg-Hurn, D. M., McEvoy, P. M., Fursland, A., & Waller, G. (2018). Severe and enduring anorexia nervosa? Illness severity and duration are unrelated to outcomes from cognitive behaviour therapy. *Journal of Consulting and Clinical Psychology*, *86*(8), 702-709. https://doi.org/10.1037/ccp0000319*

Redgrave, G. W., Schreyer, C. C., Coughlin, J. W., Fischer, L. K., Pletch, A., & Guarda, A. S. (2021). Discharge body mass index, not illness chronicity, predicts 6-month weight outcome in patients hospitalized with anorexia nervosa. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.641861*

Riesco, N., Agüera, Z., Granero, R., Jiménez-Múrcia, S., Menchón, J. M., & Fernandez-Aranda, F. (2018). Other Specified Feeding or Eating Disorders (OSFED): Clinical heterogeneity and cognitive-behavioral therapy outcome. *European Psychiatry*, *54*, 109-116. https://doi.org/doi:10.1016/j.eurpsy.2018.08.001*

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*(5), 408-412. https://doi.org/10.1002/erv.2372

Robinson, P. (2014). Severe and enduring eating disorders: recognition and management. *Advances in psychiatric treatment, 20,* 392-401. https://doi.org/10.1192/apt.bp.113.011841

Robinson, P. H., Kukucska, R., Guidetti, G., & Leavey, G. (2015). Severe and enduring anorexia nervosa (SEED-AN a qualitative study of patients with 20+ years of

anorexia nervosa. *European Eating Disorders Review: The Journal of the Eating Disorders Association, 23*(4), 318-326, https://doi.org/10.1002/erv.2367

Simpson, C. C., Towne, T. L., Karam, A. M., Donahue, J. M., Hadjeasgari, C. F., Rockwell, R., & Kaye, W. H. (2021). Predictors of stepping up to higher level of care among eating disorder patients in a partial hospitalization program. *Frontiers in Psychology 12(*667868). https://doi.org/10.3389/fpsyg.2021.667868*

Smith, S., & Woodside, D. (2021). Characterizing treatment-resistant anorexia nervosa. *Frontiers in Psychiatry, 11*. https://doi.org/10.3389/fpsyt.2020.542206*

Steinglass, J. E., Fei, W., Foerde, K., Touzeau, C., Ruggiero, J., Lloyd, C., Attia, E., Wang, Y., & Walsh, B. (2024). Change in food choice during acute treatment and the effect on longer-term outcome in patients with anorexia nervosa. *Psychological Medicine*, *54*(6), 1133-1141. https://doi.org/10.1017/S0033291723002933*

Svendsen, V. G., Wijnen, B. F. M., de Vos, J. A., Veenstra, R., Evers, S. M. A. A., & Lokkerbol, J. (2023). A roadmap for applying machine learning when working with privacy-sensitive data: predicting non-response to treatment for eating disorders. *Expert Review of Pharmacoeconomics and Outcomes Research*, *23*(8), 933-949. https://doi.org/10.1080/14737167.2023.2230368*

Treasure, J., Stein, D., & Maguire, S. (2015). Has the time come for a staging model to map the course of eating disorders from high risk to severe enduring illness? An examination of the evidence. *Early Intervention in Psychiatry*, *9*(3), 173–184. https://doi.org/10.1111/eip.12170

Vall, E., & Wade, T. D. (2015). Predictors of treatment outcome in individuals with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders, 48*(7), 946-971. https://doi.org/10.1002/eat.22411

Vall, E., & Wade, T. D. (2015). Predictors of treatment outcome in individuals with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders, 48*(7), 946-971. https://doi.org/10.1002/eat.22411

Wales, J., Brewin, N., Cashmore, R., Haycraft, E., Baggott, J., Cooper, A., & Arcelus, J. (2016). Predictors of positive treatment outcome in people with anorexia nervosa treated in a specialized inpatient unit: The role of early response to treatment. *European Eating Disorders Review*, *24*(5), 417-424. https://doi.org/10.1002/erv.2443*

Walker, D. C., Donahue, J. M., Heiss, S., Gorrell, S., Anderson, L. M., Brooks, J. M., Ehrlich, E. P., Morison, J. N., & Anderson, D. A. (2020). Rapid response is predictive of treatment outcomes in a transdiagnostic intensive outpatient eating disorder sample: a replication of prior research in a real-world setting. *Eating and Weight Disorders*, *26*(5), 1345-1356. https://doi.org/10.1007/s40519-020-00939-y*

Wild, B., Friederich, H.-C., Zipfel, S., Resmark, G., Giel, K., Teufel, M., Schellberg, D., Lowe, B., de Zwaan, M., Zeeck, A., Herpertz, S., Burgmer, M., von Wietersheim, J., Tagay, S., Dinkel, A., & Herzog, W. (2016). Predictors of outcomes in outpatients with anorexia nervosa-Results from the ANTOP study. *Psychiatry Research*, *244*, 45-50. https://doi.org/10.1016/j.psychres.2016.07.002*

Wonderlich, S., Mitchell, J. E., Crosby, R. D., Myers, T. C., Kadlec, K., Lahaise, K., Swan-Kremeier, L., Dokken, J., Lange, M., Dinkel, J., Jorgensen, M., & Schander, L.. (2012). Minimizing and treating chronicity in the eating disorders: A clinical overview. *International Journal of Eating Disorders*, *45*(4), 467–475. https://doi.org/10.1002/eat.20978

Yager, J. (2020). Managing patients with severe and enduring anorexia nervosa: When is enough, enough? *The Journal of Nervous and Mental Disease, 208*(4). 277-282. https://doi.org/10.1097/NMD.0000000000001124

Appendices

Appendix A. CASP Checklist for Randomised Control Trials

Appendix B. CASP Checklist for Cohort Studies

Appendix C. Risk of Bias Assessment

Appendix A. CASP Checklist for Randomised Control Trials

CASP Checklist:
For Randomised Controlled Trials (RCTs)

Reviewer Name:	
Paper Title:	
Author:	
Web Link:	
Appraisal Date:	

During critical appraisal, never make assumptions about what the researchers have done. If it is not possible to tell, use the "Can't tell" response box. If you can't tell, at best it means the researchers have not been explicit or transparent, but at worst it could mean the researchers have not undertaken a particular task or process. Once you've finished the critical appraisal, if there are a large number of "Can't tell" responses, consider whether the findings of the study are trustworthy and interpret the results with caution.

Section A Is the basic study design valid for a randomised controlled trial?	
1. Did the study address a clearly formulated research question?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: *Was the study designed to assess the outcomes of an intervention?* *Is the research question 'formulated' in terms of:* • *Population studied* • *Intervention given* • *Comparator chosen* • *Outcomes measured?*	
2. Was the assignment of participants to interventions randomised?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *How was randomisation carried out? Was the method appropriate?* • *Was randomisation sufficient to eliminate systematic bias?* • *Was the allocation sequence concealed from investigators and participants?*	
3. Were all participants who entered the study accounted for at its conclusion?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Were losses to follow-up and exclusions after randomisation accounted for?* • *Were participants analysed in the study groups to which they were randomised (intention-to-treat analysis)?* • *Was the study stopped early? If so, what was the reason?*	
Section B Was the study methodologically sound?	
4. (a) Were the participants 'blind' to intervention they were given?	☐ Yes ☐ No ☐ Can't Tell

(b) Were the investigators 'blind' to the intervention they were giving to participants?	☐ Yes ☐ No ☐ Can't Tell
(c) Were the people assessing/analysing outcome/s 'blinded'?	☐ Yes ☐ No ☐ Can't Tell
5. Were the study groups similar at the start of the randomised controlled trial?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Were the baseline characteristics of each study group (e.g. age, sex, socio-economic group) clearly set out?* • *Were there any differences between the study groups that could affect the outcome/s?*	
6. Apart from the experimental intervention, did each study group receive the same level of care (that is, were they treated equally)?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Was there a clearly defined study protocol?* • *If any additional interventions were given (e.g. tests or treatments), were they similar between the study groups?* • *Were the follow-up intervals the same for each study group?*	
Section C: What are the results?	
7. Were the effects of intervention reported comprehensively?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Was a power calculation undertaken?* • *What outcomes were measured, and were they clearly specified?*	

• *How were the results expressed? For binary outcomes, were relative and absolute effects reported?* • *Were the results reported for each outcome in each study group at each follow-up interval?* • *Was there any missing or incomplete data?* • *Was there differential drop-out between the study groups that could affect the results?* • *Were potential sources of bias identified?* • *Which statistical tests were used?* • *Were p values reported?*	
8. Was the precision of the estimate of the intervention or treatment effect reported?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Were confidence intervals (CIs) reported?*	
9. Do the benefits of the experimental intervention outweigh the harms and costs?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *What was the size of the intervention or treatment effect?* • *Were harms or unintended effects reported for each study group?* • *Was a cost-effectiveness analysis undertaken? (Cost-effectiveness analysis allows a comparison to be made between different interventions used in the care of the same condition or problem.)*	
Section D: Will the results help locally?	
10. Can the results be applied to your local population/in your context?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Are the study participants similar to the people in your care?* • *Would any differences between your population and the study participants alter the outcomes reported in the study?* • *Are the outcomes important to your population?*	

• *Are there any outcomes you would have wanted information on that have not been studied or reported?* • *Are there any limitations of the study that would affect your decision?*

11. Would the experimental intervention provide greater value to the people in your care than any of the existing interventions?	☐ Yes ☐ No ☐ Can't Tell

CONSIDER: • *What resources are needed to introduce this intervention taking into account time, finances, and skills development or training needs?* • *Are you able to disinvest resources in one or more existing interventions in order to be able to re-invest in the new intervention?*

APPRAISAL SUMMARY: *List key points from your critical appraisal that need to be considered when assessing the validity of the results and their usefulness in decision-making.*		
Positive/Methodologically sound	**Negative/Relatively poor methodology**	**Unknowns**

Referencing recommendation:

CASP recommends using the Harvard style referencing, which is an author/date method. Sources are cited within the body of your assignment by giving the name of the author(s) followed by the date of publication. All other details about the publication are given in the list of references or bibliography at the end.

Example:

Critical Appraisal Skills Programme (2024). CASP (insert name of checklist i.e. randomised controlled trials (RCTs) Checklist.) [online] Available at: insert URL. Accessed: insert date accessed.

Creative Commons

Need further training on evidence-based decision making? Our online training courses are helpful for healthcare educational researchers and any other learners who:

- Need to critically appraise and stay abreast of the healthcare research literature as part of their clinical duties.
- Are considering carrying out research & developing their own research projects.
- Make decisions in their role, whether that be policy making or patient facing.

Benefits of CASP Training:

⇒ Affordable – courses start from as little as £6
⇒ Professional training – leading experts in critical appraisal training
⇒ Self-directed study – complete each course in your own time
⇒ 12 months access – revisit areas you aren't sure of and revise
⇒ CPD certification - after each completed module

Scan the QR code below or visit https://casp-uk.net/critical-appraisal-online-training-courses/ for more information and to start learning more.

Appendix B. CASP Checklist for Cohort Studies

CASP Checklist:
For Cohort Studies

Reviewer Name:	
Paper Title:	
Author:	
Web Link:	
Appraisal Date:	

During critical appraisal, never make assumptions about what the researchers have done. If it is not possible to tell, use the "Can't tell" response box. If you can't tell, at best it means the researchers have not been explicit or transparent, but at worst it could mean the researchers have not undertaken a particular task or process. Once you've finished the critical appraisal, if there are a large number of "Can't tell" responses, consider whether the findings of the study are trustworthy and interpret the results with caution.

Section A: Are the results valid?	
1. Did the study address a clearly focused issue?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: *A question can be 'focused' in terms of* • *the population studied* • *the risk factors studied* • *is it clear whether the study tried to detect a beneficial or harmful effect* • *the outcomes considered*	
2. Was the cohort recruited in an acceptable way?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Look for selection bias which might compromise the generalisability of the findings:* • *was the cohort representative of a defined population* • *was there something special about the cohort* • *was everybody included who should have been*	
3. Was the exposure accurately measured to minimise bias?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: *Look for measurement or classification bias:* • *did they use subjective or objective measurements* • *do the measurements truly reflect what you want them to (have they been validated)* • *were all the subjects classified into exposure groups using the same procedure*	
4. Was the outcome accurately measured to minimise bias?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: *Look for measurement or classification bias:* • *did they use subjective or objective measurements* • *do the measurements truly reflect what you want them to (have they been validated)* • *has a reliable system been established for detecting all the cases (for measuring disease occurrence)* • *were the measurement methods similar in the different groups* • *were the subjects and/or the outcome assessor blinded to exposure (does this matter)*	
5. (a) Have the authors identified all important confounding factors?	☐ Yes ☐ No ☐ Can't Tell

CONSIDER: • *list the ones you think might be important, and ones the author missed*	
b) Have they taken account of the confounding factors in the design and/or analysis?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *look for restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors*	
6. a) Was the follow up of subjects complete enough?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *the persons that are lost to follow-up may have different outcomes than those available for assessment* • *in an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the cohort*	
b) Was the follow up of subjects long enough?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *the good or bad effects should have had long enough to reveal themselves*	
Section B: What are the results?	
7. What are the results of this study?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *what are the bottom line results* • *have they reported the rate or the proportion between the exposed/unexposed, the ratio/rate difference* • *how strong is the association between exposure and outcome (RR)* • *what is the absolute risk reduction (ARR)*	
8. How precise are the results?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER:	

• *look for the range of the confidence intervals, if given*	
9. Do you believe the results?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *big effect is hard to ignore* • *can it be due to bias, chance or confounding* • *are the design and methods of this study sufficiently flawed to make the results unreliable* • *Bradford Hills criteria (e.g. time sequence, dose-response gradient, biological plausibility, consistency)*	
Section C: Will the results help locally?	
10. Can the results be applied to the local population?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *Is a cohort study the appropriate method to answer this question* • *If the subjects covered in this study could be sufficiently different from your population to cause concern* • *If your local setting is likely to differ much from that of the study* • *If you can quantify the local benefits and harms*	
11. Do the results of this study fit with other available evidence?	☐ Yes ☐ No ☐ Can't Tell
12. What are the implications of this study for practice?	☐ Yes ☐ No ☐ Can't Tell
CONSIDER: • *one observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making* • *for certain questions, observational studies provide the only evidence* • *recommendations from observational studies are always stronger when supported by other evidence*	

APPRAISAL SUMMARY: *List key points from your critical appraisal that need to be considered when assessing the validity of the results and their usefulness in decision-making.*		
Positive/Methodologically sound	**Negative/Relatively poor methodology**	**Unknowns**

Referencing recommendation:

CASP recommends using the Harvard style referencing, which is an author/date method. Sources are cited within the body of your assignment by giving the name of the author(s) followed by the date of publication. All other details about the publication are given in the list of references or bibliography at the end.

Example:

Critical Appraisal Skills Programme (2024). CASP (insert name of checklist i.e. qualitative studies Checklist.) [online] Available at: insert URL. Accessed: insert date accessed.

Creative Commons

Need further training on evidence-based decision making? Our online training courses are helpful for healthcare educational researchers and any other learners who:

- Need to critically appraise and stay abreast of the healthcare research literature as part of their clinical duties.
- Are considering carrying out research & developing their own research projects.
- Make decisions in their role, whether that be policy making or patient facing.

Benefits of CASP Training:

⇒ Affordable – courses start from as little as £6
⇒ Professional training – leading experts in critical appraisal training
⇒ Self-directed study – complete each course in your own time
⇒ 12 months access – revisit areas you aren't sure of and revise
⇒ CPD certification - after each completed module

Scan the QR code below or visit https://casp-uk.net/critical-appraisal-online-training-courses/ for more information and to start learning more.

Appendix C. Risk of Bias Assessments

	Did the research address a clearly focused issue?	Was the cohort recruited in an acceptable way?	Was the exposure accurately measured to minimise bias?	Was the outcome accurately measured to minimise bias?	Have the authors identified all important confounding factors?	Have the authors accounted for all confounding factors in the design/analysis?	Was follow-up complete enough?	Was follow-up long enough?	Are the results clearly reported?	How precise are the results?	Do you believe the results?	Can the results be applied locally?	Do the results fit with other available evidence?	Are the implications of practice clear?	Overall risk of bias
Ambwani et al. (2020)	Y	Y	CT	Y	N	N	N	Y	Y	N	Y	N	Y	Y	M
Arai et al. (2024)	Y	Y	Y	Y	N	N	N	N	Y	N	Y	N	Y	Y	M
Austin et al. (2025)	Y	Y	N	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	M
Calugi et al. (2021)	Y	Y	Y	Y	N	N	N	Y	Y	N	Y	N	Y	Y	M
Calugi et al. (2017)	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	N	Y	Y	M
Calugi et al. (2025)	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	L/M

Camacho-Barcia et al. (2024)	Y	Y	N	Y	N	N	N	Y	Y	N	Y	N	Y	Y	M
Day et al. (2025)	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	N	Y	Y	M
Dingemans et al. (2016)	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	L/M
Dingemans et al. (2020)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	L
Fernández-Aranda et al. (2021)	Y	Y	CT	Y	Y	CT	Y	N	Y	Y	Y	N	Y	Y	M
Frank et al. (2024)	Y	Y	Y	Y	Y	CT	N	Y	Y	Y	Y	N	Y	Y	L/M
Frostad et al. (2021)	Y	Y	Y	Y	N	N	Y	Y	Y	CT	Y	N	Y	Y	M
Garte et al. (2015)	Y	CT	Y	Y	CT	N	N	Y	Y	N	Y	N	Y	Y	M
Ibrahim et al. (2022)	Y	Y	CT	Y	N	CT	Y	Y	Y	Y	Y	N	Y	Y	M
Kaufmann et al. (2021)	Y	Y	Y	N	Y	N	N	Y	Y	N	Y	N	Y	N	M
Keegan & Wade (2024)	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Y	Y	L/M
Li et al. (2020)	Y	Y	Y	Y	Y	Y	N	N	Y	N	Y	N	Y	Y	M
Marzola et al. (2019)	Y	N	N	Y	Y	N	N	N	Y	CT	Y	N	Y	Y	M

Marzola et al. (2021)	Y	Y	N	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	M
Melisse et al. (2022)	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Y	Y	L/M
Meule et al. (2023)	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	L/M
Quadflieg & Fichter (2019)	Y	Y	Y	N	Y	N	N	Y	Y	Y	Y	N	Y	Y	M
Raykos et al. (2018)	Y	Y	Y	Y	N	N	Y	N	Y	Y	Y	N	Y	Y	M
Redgrave et al. (2021)	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	L/M
Riesco et al. (2018)	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	Y	M
Simpson et al. (2021)	Y	Y	Y	CT	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	L/M
Smith & Woodside (2021)	Y	Y	Y	N	Y	N	N	Y	Y	N	Y	N	Y	Y	M
Steinglass et al. (2024)	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	CT	L/M
Svenden et al. (2023)	Y	Y	Y	Y	Y	Y	N	Y	Y	CT	Y	N	Y	N	M
Wales et al. (2016)	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	L/M
Walker et al. (2021)	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	N	Y	Y	L/M

Note. Y = Yes, N = No, CT = Can't tell, L = Low, L/M = Low to Medium, M = Medium, M/H = Medium to High, H = High.

	Did the study address a clearly formulated research question?	Was the assignment of participants to interventions randomised?	Were all participants who entered the study accounted for at its conclusion?	Were the participants 'blind' to the intervention they were given?	Were the investigators blind to the intervention they were giving?	Were the people assessing/analysing outcomes 'blinded'?	Were the study groups similar at the start of the RCT?	Were the groups treated equally aside from the intervention?	Were the effects of the intervention reported comprehensively?	Was the precision of the estimate of the intervention or treatment effect reported?	Do the benefits of the experimental intervention outweigh the harms and costs?	Can the results be applied to your local population/context?	Would the experimental intervention provide greater value than any existing intervention?	Overall risk of bias
Accurso et al. (2016)	Y	Y	N	N	N	CT	Y	Y	Y	N	Y	N	Y	M
Anderson et al. (2020)	Y	Y	Y	N	N	CT	Y	Y	Y	Y	Y	N	CT	M
Bamford et al. (2015)	Y	Y	CT	N	N	CT	CT	Y	N	N	CT	N	Y	L/M
Cooper et al. (2016)	Y	Y	Y	N	N	CT	N	Y	Y	Y	Y	N	Y	M
Forrest et al. (2021)	Y	Y	Y	N	N	CT	Y	Y	Y	N	Y	N	CT	M
Grillo et al. (2021)	Y	Y	N	N	N	CT	Y	Y	Y	N	Y	N	CT	M

Hilbert et al. (2020)	Y	Y	CT	N	N	CT	CT	CT	CT	N	CT	Y	Y	M/H
Linardon & Fuller-Tyszkiewicz (2024)	Y	Y	Y	N	N	CT	CT	Y	Y	Y	Y	N	N	M
Lydecker et al. (2020)	Y	Y	N	N	N	CT	N	CT	Y	Y	Y	Y	CT	M/H
Mannan et al. (2025)	Y	Y	Y	Y	N	CT	Y	Y	Y	Y	CT	N	CT	M
Matherne et al. (2022)	Y	Y	Y	N	N	CT	Y	Y	Y	Y	Y	N	CT	M
Wild et al. (2016)	Y	Y	Y	N	N	CT	Y	Y	Y	Y	Y	CT	CT	M

Note. Y = Yes, N = No, CT = Can't tell, L = Low, L/M = Low to Medium, M = Medium, M/H = Medium to High, H = High.

Section Two: Empirical Study

Key symptoms that account for the early response effect in eating disorder treatment:

A Bayesian network analysis of longitudinal clinical data

Abstract

Aims: The *early response effect,* defined as a reliable symptomatic improvement during the initial phase of treatment, is the most robust predictor of recovery following eating disorder treatment. Identifying the symptoms that account for the early response effect could inform the development of tailored treatment strategies and enhance treatment efficacy. This study aimed to investigate which symptom domains mostly influence the early response effect.

Methods: Following ethical approval, data from N = 232 patients treated in an outpatient eating disorder psychotherapy service was randomly partitioned into training (N = 161) and test (N = 71) samples. A Bayesian network model was developed in the training sample, modelling early changes (sessions 1 – 4) and interactions among symptoms measured by the Eating Disorder Examination Questionnaire (EDE-Q). A variable selection approach was applied to only include the most important variables in the model (i.e. reliable predictors of recovery). The trained model was externally validated by applying it to predict post-treatment recovery status (reliable and clinically significant improvement; RCSI) in the test sample. Prediction accuracy was evaluated using the AUC statistic.

Results: The model identified a network of six interrelated eating disorder symptoms which were the most important predictors of recovery. The model was reliable in predicting recovery status and showed adequate generalisability to a test sample (AUC = 0.81 vs. 0.77 respectively).

Conclusion: Early changes in six area (ranked by importance) reliably predict recovery after psychotherapy: [1] avoidance of body exposure; [2] feelings of ‘fatness’; [3] preoccupation with food, eating or calories; [4] fear of losing control over eating [5] dissatisfaction with body shape; [6] dietary rules.

Practitioner Points:

- Six specific symptom areas account for the early response effect and are therefore key targets for intervention during the initial phase of treatment for eating disorders:
    1. Avoidance of body exposure
    2. Feelings of ‘fatness’
    3. Preoccupation with food, eating, or calories
    4. Fear of losing control over eating
    5. Dissatisfaction with body shape
    6. Rigid dietary rules
- Addressing these symptoms early in treatment may enhance recovery outcomes, as these domains were the most influential predictors in a validated predictive model.
- Early treatment techniques should involve exposure-based work, cognitive restructuring around feeling ‘fat’, and behavioural experiences targeting eating-related fears and rules.

Introduction

Eating disorders are characterised by a preoccupation with weight or body shape and by abnormal eating behaviours (Jansen, 2016). They are typically categorised into several diagnostic subtypes, including anorexia nervosa, bulimia nervosa, binge eating disorder and other specified feeding and eating disorder (American Psychiatric Association, 2013). Eating disorders have a serious impact on quality of life (Le et al., 2019), are associated with high mortality rates (Miskovic-Wheatley et al., 2023) and can become disabling if not treated effectively (Klump et al., 2009).

Clinical guidelines recommend cognitive behavioural therapy (CBT), Maudsley model of anorexia nervosa treatment for adults (MANTRA) and specialist supportive clinical management (SSCM) as first line treatments for eating disorders in adults (National Institute for Health and Care Excellence (NICE), 2017). Other psychological treatments, such as interpersonal psychotherapy (Zhang et al., 2024) and dialectical behavioural therapy (Rahmani et al., 2018) also show some preliminary evidence. However, the efficacy of current treatment for adults varies. For anorexia nervosa, no single psychological treatment outperforms treatment as usual (Solmi et al., 2021), while CBT has a large effect on recovery and outperforms other active treatments for bulimia nervosa and binge eating disorder (Monteleone et al., 2022). Recovery and remission definitions differ across studies, but centre on patients no longer meeting diagnostic criteria alongside improved levels of psychosocial functioning and quality of life (Bardone-Cone et al., 2010).

The early response effect

The most robust predictor of positive outcomes at both end of treatment and follow-up is the early response effect, as documented in systematic and meta-analytic reviews of available literature (Chang et al., 2021; Linardon et al., 2016; Vall & Wade, 2015). This aligns with wider evidence that early response predicts positive psychological treatment outcomes

in other mental health disorders such as anxiety and depression (Beard & Delgadillo, 2019). Early response is defined as a statistically reliable improvement in symptoms during the early phase of psychological treatment (MacDonald et al., 2015). Proposed definitions of the timeframe for early response vary among authors, ranging from the first week of treatment to ten weeks (e.g., Hilbert et al., 2019; Macdonald et al., 2015), as well as the approach to measuring early response (a categorical or dimensional approach). However, 65% of eating disorder focused studies conceptualise the early response timeframe as being within the first four weeks of treatment, and a similar proportion use a dimensional approach, focusing on the magnitude of change rather than setting a fixed, categorical threshold (Chang et al., 2021).

Only half of patients with eating disorders demonstrate an early response to psychological treatment (Chang et al., 2021). Understanding the process that supports early response is of clear clinical importance. It is plausible that making changes in some areas of eating disorder pathology may be particularly important to effect change in other areas. This assumption is built into treatment protocols which recommend targeting some symptom domains before others. For example, the transdiagnostic CBT protocol (Fairburn et al., 2003) is conceptualised as a modular treatment that starts by stabilising disordered eating patterns, and which later targets areas such as over-evaluation of shape/weight, rigid dietary rules, maladaptive body checking/avoidance, etc. An empirical examination of the relative importance of some symptom domains during the early phase of treatment could potentially help to inform the sequencing of interventions and treatment targets. A network model has the potential to assess such processes.

Testing the network model of eating disorders

Network models conceptualise mental health disorders as systems of interacting symptoms, rather than manifestations of a single, underlying latent cause (McNally, 2016).

The latent variable model assumes that an underlying ‘disease’ (e.g. anorexia nervosa) causes a set of symptoms (e.g. body dissatisfaction, restrictive eating, fear of weight gain and low body weight) (Borsboom & Cramer, 2013) (Figure 1). In contrast, the network model views these symptoms as dynamically influencing each other. For example, body dissatisfaction might increase dietary restriction, leading to low body weight and cognitive impairment from starvation, which in turn heightens body dissatisfaction and preoccupation with weight, forming feedback loops that sustain the disorder (see Figure 2 for a theoretical illustration). This model suggests that targeting central symptoms (i.e. those most centrally connected to others) has the greatest potential to disrupt the whole network and effect recovery. That information would allow those ‘nodal’ symptoms to be addressed, with the maximum effect on the eating disorder.

Figure 1. A latent variable model of anorexia nervosa

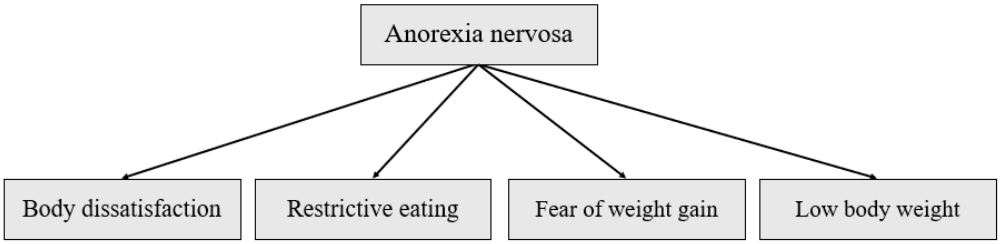
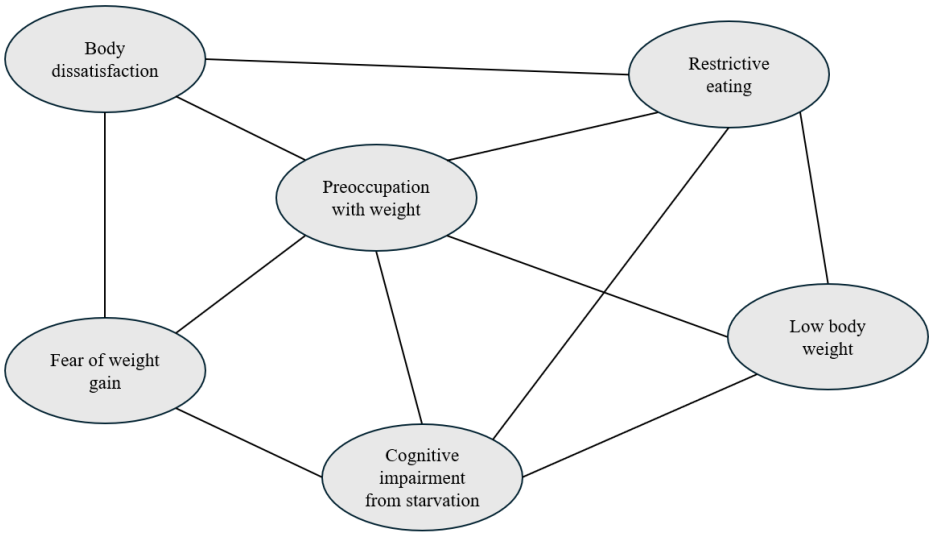


Figure 2. A theoretical illustration of the network model of anorexia nervosa



Network analysis uses advanced quantitative methods to test network theory based on data. Bayesian network analysis, which uses machine learning methods (a computational approach capable of drawing upon many variables to build predictive models in a data-driven manner - Breiman, 2001), is one of several approaches. Bayesian network analysis uses Bayesian statistics to estimate probabilistic relationships among symptoms by visually showing potential causal pathways (edges) between symptoms (nodes) (Briganti et al., 2023). This then result in the development of predictive algorithms, which can be cross validated in an independent sample.

By analysing the structure and strengths of connections between symptoms, key nodes (i.e. central symptoms) that drive or maintain the disorder can be identified as potential targets for intervention, thus disrupting the network (Borsboom, 2017; Levinson et al., 2023).

Aims of current study

This study was designed to answer three questions: [1] Which symptoms of eating disorder pathology account for the early response effect? [2] How do these symptoms interact during the early phase of therapy? [3] Are early changes in some symptoms more important than others?

These questions were approached empirically, using a data-driven analysis without pre-specified hypotheses, using outcome data from a sample of patients who accessed psychological interventions for eating disorders.

Methods

Design

Setting and Interventions

The study was based on a secondary analysis of a dataset collected between 2016 and 2022. Data collection took place as part of routine care within the service. The dataset included fully anonymised clinical and demographic data from patients who accessed

psychological treatment at a specialist outpatient eating disorder service in Nottinghamshire, England. The present study was approved by the University of Sheffield Research Ethics Committee (26/02/2024, Ref: 058515, Appendix A).

The service provides various psychological interventions, and the type and duration of treatment is decided on an individual case basis. Most patients receive integrative psychotherapy (41.5%) or CBT (18.3%). Depending on specific needs, others receive person-centred counselling (12.2%), cognitive analytic therapy (12.2%), gestalt therapy (11.6%), or eye-movement desensitization and reprocessing for comorbid post-traumatic stress disorder (4.3%) (Bell et al., 2017).

Participants

All patients had an initial assessment by the clinical lead or senior psychotherapist using a semi-structured diagnostic interview guided by standard diagnostic criteria (American Psychiatric Association, 2013), supplemented with a self-reported measure of eating disorder pathology (Eating Disorder Examination Questionnaire; EDE-Q; Fairburn & Beglin, 2008; Appendix B) to confirm suitability for psychological treatment. Body mass index (BMI) was calculated at baseline and post-treatment assessments, based on height/weight measurements taken at the clinic. The present study sample included data for patients who: were over 18 years old; completed a baseline EDE-Q measure; completed the EDE-Q at session 4 to measure early response; attended at least five therapy sessions (allowing early response to be measured separately to the final treatment outcome); and completed the EDE-Q again during or after the last attended session to measure their final treatment outcome. The eligible study sample included $N = 238$ cases.

Sample Size Calculation

Following Riley et al.'s (2018) guideline for the minimum required sample size when developing multivariable prediction models to enhance development and reduce the potential

for overfitting, $N = 367$ would be required, assuming the number of predictor parameters of 22 items of the EDE-Q. This minimum sample size is essential to attain a global shrinkage factor of ≥ 0.9, maintain a small absolute difference of ≤ 0.05 between the apparent and adjusted R^2, and ensure precise estimation with a margin of error ≤ 10% of the true value of the model's residual standard deviation and precise estimation of the mean predicted outcome (model intercept). Following Archer et al.'s (2020) rationale for the minimum sample size essential for external validation of the prediction model containing the number of predictor parameters of 22 items of the EDE-Q, $N = 336$ would be required to achieve the target precise estimates of $R^2 = 0.90$, achieve strong calibration properties and obtain the variance of observed outcome values = 0.09.

The dataset available to this study was less than the minimum sample size required to develop the prediction model and externally validate the model. Therefore, the model was simplified by reducing the number of predictor parameters from 22 items of the EDE-Q to 17 items, reducing the risk of overfitting, as recommended by Riley et al. (2018). The five predictor parameters removed were the items in the weight concern subscale, as these items strongly correlate with the items in the shape concern subscale in both clinical and non-clinical samples, showing the factors are not entirely distinctive (White et al., 2014). Despite these recommended steps that were taken, the sample size was still less than the minimum sample size required to develop the prediction model ($N = 283$) and to externally validate the model ($N = 235$). This was acknowledged as a major limitation and the model was retained in the present study to allow for preliminary exploration of potential predictors. Classification accuracy was assessed, as described below, to test generalisability and reliability before any conclusions were drawn.

Measures and data sources

The EDE-Q version 6.0 (Fairburn & Beglin, 2008) is a self-report assessment tool consisting of 28 items derived from the Eating Disorder Examination interview (Fairburn & Cooper, 1993). It assesses core cognitive and behavioural symptoms of eating disorders and aligns with the transdiagnostic model of eating disorders, which highlight the overvaluation of shape and weight as key maintaining mechanisms (Fairburn et al., 2003). The EDE-Q produces two types of data: severity and frequency. Twenty-two items assessing severity are rated on a 7-point forced choice scale, with higher scores indicating greater severity. These severity items are used to generate a global score and four subscale scores: restraint (five items), eating concern (five items), shape concern (eight items), and weight concern (five items). Six additional items assess the frequency of key behavioural components of eating disorder pathology, such as binge eating, self-induced vomiting, laxative misuse, diuretic misuse, and excessive exercise. The behavioural frequency items do not contribute to the subscale scores but provide clinically useful information to inform diagnosis and treatment.

The EDE-Q demonstrates satisfactory internal consistency, as evidenced by Cronbach's alpha coefficients ranging from 0.70-0.83 in clinical samples and from 0.78-0.93 in general population samples for the global scale and each subscale (Aardoom et al., 2012; Luce & Crowther, 1999; Rose et al., 2013). Likewise, its test-retest reliability is robust, with a Spearman's rho of 0.86 or greater (Rose et al., 2013). Furthermore, a receiver operating characteristics analysis (ROC) demonstrated an AUC of 0.93 (95% CI = 0.91–0.94), which strongly supports the discriminant validity of the EDE-Q (Ro et al., 2015).

Data analysis

All data were analysed using IBM SPSS Modeler version 18.3 using a supervised machine learning approach called Bayesian network analysis. A series of three distinct stages of data analysis were followed – the machine learning pipeline.

1. *Pre-processing.* The outcomes of interests were defined, and all available information was organised into a **training sample** to train the predictive algorithm and a testing sample to cross-validate it.
2. *Training and variable selection:* Specific eating disorder symptoms networks undergoing early change and relevant to the prediction of recovery were identified. This resulted in the development of an algorithm predicting recovery.
3. *External cross-validation:* The algorithm was externally validated in a **testing sample**.

Pre-processing. The initial stage involved randomly partitioning the original data set into two datasets using a 70:30 ratio for an adequate training sample (n = 162) and a statistically independent test sample (n = 76). Next, predictor and outcome variables were derived from available EDE-Q measures. Five out of 22 EDE-Q items from the 'weight concern' subscale were removed to avoid redundancy in item-level analyses, as these items strongly correlate with the items in the 'shape concern' subscale. This resulted in 17 items as potential predictors of treatment outcomes. Early response was modelled using a change score for each of the EDE-Q items (score at session 1 – session 4, where a positive value denotes improvement and a negative value denotes deterioration). A binary variable was employed to categorise cases achieving symptomatic recovery after the end of treatment, based on the criteria for reliable and clinical significant improvement (RCSI) proposed by Jacobson and Truax (1991). This outcome definition requires the following two conditions to be met: [1] Pre-treatment EDE-Q global scores should be above the clinical threshold (2.77 according to Fairburn & Beglin, 1994), and post-treatment scores need to be below that threshold in order to show *clinically significant change.* The EDE-Q global score was calculated by first computing the mean score for each of the four subscales and then calculating the overall mean of these subscale scores. [2] To ensure the observed change between session 1 and the final therapy session denotes statistically *reliable improvement* and

not simply measurement error, the change needs to be greater than the reliable change index of 1.13, as reported by Fairburn and Beglin (1994).

Training and variable selection. The model-training stage applied a Tree-Augmented Naïve Bayes (TAN) algorithm adjusting for small cell counts (Friedman et al., 1997). The TAN algorithm is a supervised machine learning approach that models a network of relationships between predictors and their joint influence over a target outcome. Unlike conventional multivariable logistic regression models that consider main effects only, or that require pre-defined assumptions about interactions, the TAN method offers a data-driven method to model a network of relationships between predictors (e.g., EDE-Q item-level change scores between sessions 1-4) and their joint influence over a target outcome (post-treatment recovery). TAN generates a Bayesian Network Model where each predictor is allowed to depend on one additional predictor, thus modelling multiple two-way interactions among variables. Variable selection was performed using the Markov blanket feature selection method (Fu & Desmarais, 2010). This method aims to select a subset of features (variables) that are sufficient to predict the target variable, minimising redundancy and irrelevance. This process is achieved by iteratively testing features for conditional independence (e.g., via chi-square, mutual information tests) to determine which features belong to the Markov blanket for the target variable (outcome). This resulted in the development of a Bayesian network model capturing observed relationships between early symptom changes and their association with recovery after therapy, visually represented by a Directed Acyclic Graph (DAG). Furthermore, the TAN algorithm generated a binary predicted classification for each patient (1 = predicted to recover' 0 = not predicted to recover) along with a predicted probability of recovery (0 – 100%).

External cross-validation. The Bayesian network model developed in the training sample was then used to make predictions using data from the test sample (N=76). The area

under the curve (AUC) was used to assess classification accuracy in both the training and test samples, to assess generalisability and prediction shrinkage. If the accuracy were to drop substantially across the two stages (below the pre-specified 10% margin), then the model would be judged to have poor generalisability.

Results

Sample characteristics are reported in Tables 1. As the dataset was collected during routine clinical practice, there was a degree of missing data. The extent of missing demographic and clinical information is also reported. Most patients were White British (97.9%) females (90.8%) with a mean age of 29.97 and various eating disorder diagnoses, of which the most frequent one was OSFED (39.5%). Most patients received either integrative psychotherapy (52.1%) or CBT (38.2%). While many patients had a reduction of eating disorder symptom severity after therapy (55.5% finished treatment with below-threshold EDE-Q scores), only 33.6% of patients fully recovered using stringent RCSI criteria.

Table 1. *Sample characteristics, n = 238*

Demographics	Mean (SD) / %	% of missing data
Age (Years)	29.97 (10.67)	.8%
Female	90.8%	.8%
White	97.9%	.8%
Unemployed	59.2%	10.1%
Clinical Characteristics	Mean (SD) / %	% of missing data
Diagnosis		0%
Anorexia nervosa	15.1%	
Bulimia nervosa	21.0%	
Binge eating disorder	24.4%	
Other specified feeding and eating disorder	39.5%	
Antidepressant usage		1.7%
Using antidepressants	37.4%	
Psychotherapy		5.9%

Integrative psychotherapy	52.1%	
Cognitive behavioural therapy	38.2%	
Psychology	7.6%	
Person centred counselling	1.7%	
Body Mass Index (BMI)		
Baseline BMI	27.88 (10.57)	5.9%
Baseline BMI (anorexia nervosa) (N=36)	18.78 (2.21)	0%
Posttreatment BMI	25.25 (22.82)	21.0%
Posttreatment BMI (anorexia nervosa) (N=36)	19.98 (5.10)	0%
Eating Disorder Examination Questionnaire (EDE-Q)		
Baseline EDE-Q GLOBAL score	4.28 (1.05)	0%
Session 4 EDE-Q GLOBAL score	3.77 (1.29)	0%
Session 4 EDE-Q GLOBAL change score	0.44 (0.80)	0%
Posttreatment EDE-Q GLOBAL change score	1.23(1.32)	0%
Reliable and clinically significant improvement (RCSI)		
Baseline EDE-Q Caseness	89.5%	0%
Posttreatment EDE-Q Caseness	55.5%	0%
Recovered based on RCSI criteria	33.6%	0%

Note. SD = Standard Deviation; 'Caseness' = Cases with EDE-Q GLOBAL scores above the clinical threshold of 2.77; RCSI criteria = Baseline EDE-Q Global scores are above clinical threshold of 2.77, Posttreatment EDE-Q Global scores are below clinical threshold of 2.77 and the observed change is greater than the reliable change index of 1.13.

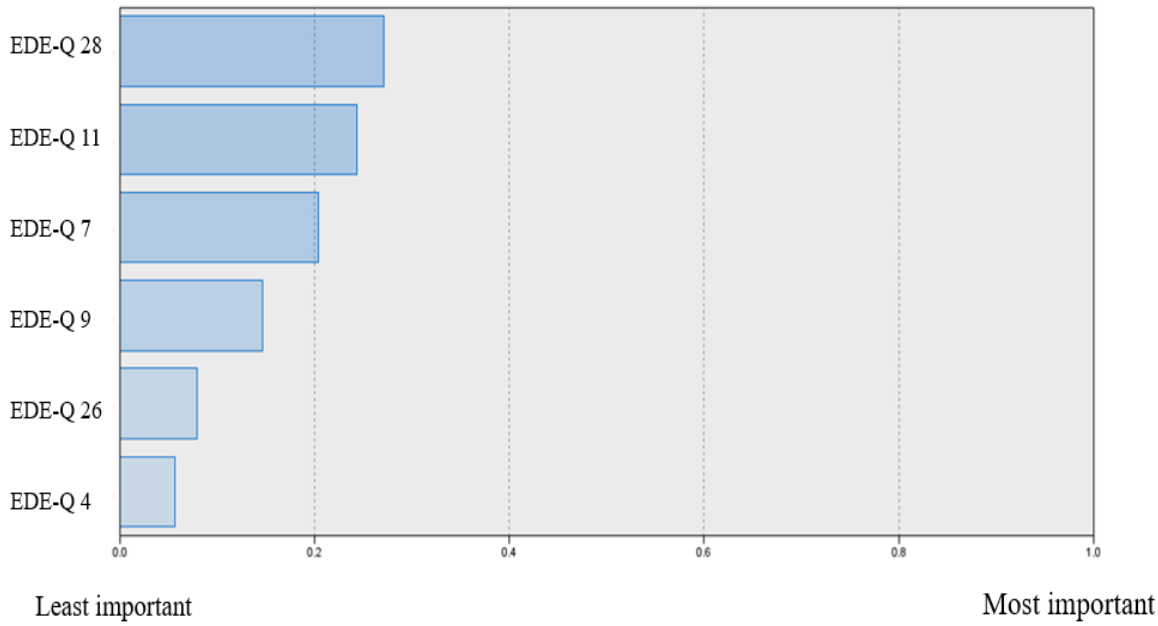
Bayesian Network Model

The variable selection converged on a Bayesian network model including only six predictors of recovery, representing change in specific EDE-Q items during the early phase of treatment. Ranked according to the variable importance plot illustrated in Figure 3, the selected items were:

1. *avoidance of body exposure* (EDE-Q item 28).
2. *feelings of 'fatness'* (EDE-Q item 11),

3. *preoccupation with food, eating or calories* (EDE-Q item 7),
4. *fear of losing control over eating* (EDE-Q item 9),
5. *dissatisfaction with body shape* (EDE-Q item 26) and
6. *rigidity of dietary rules* (EDE-Q item 4),

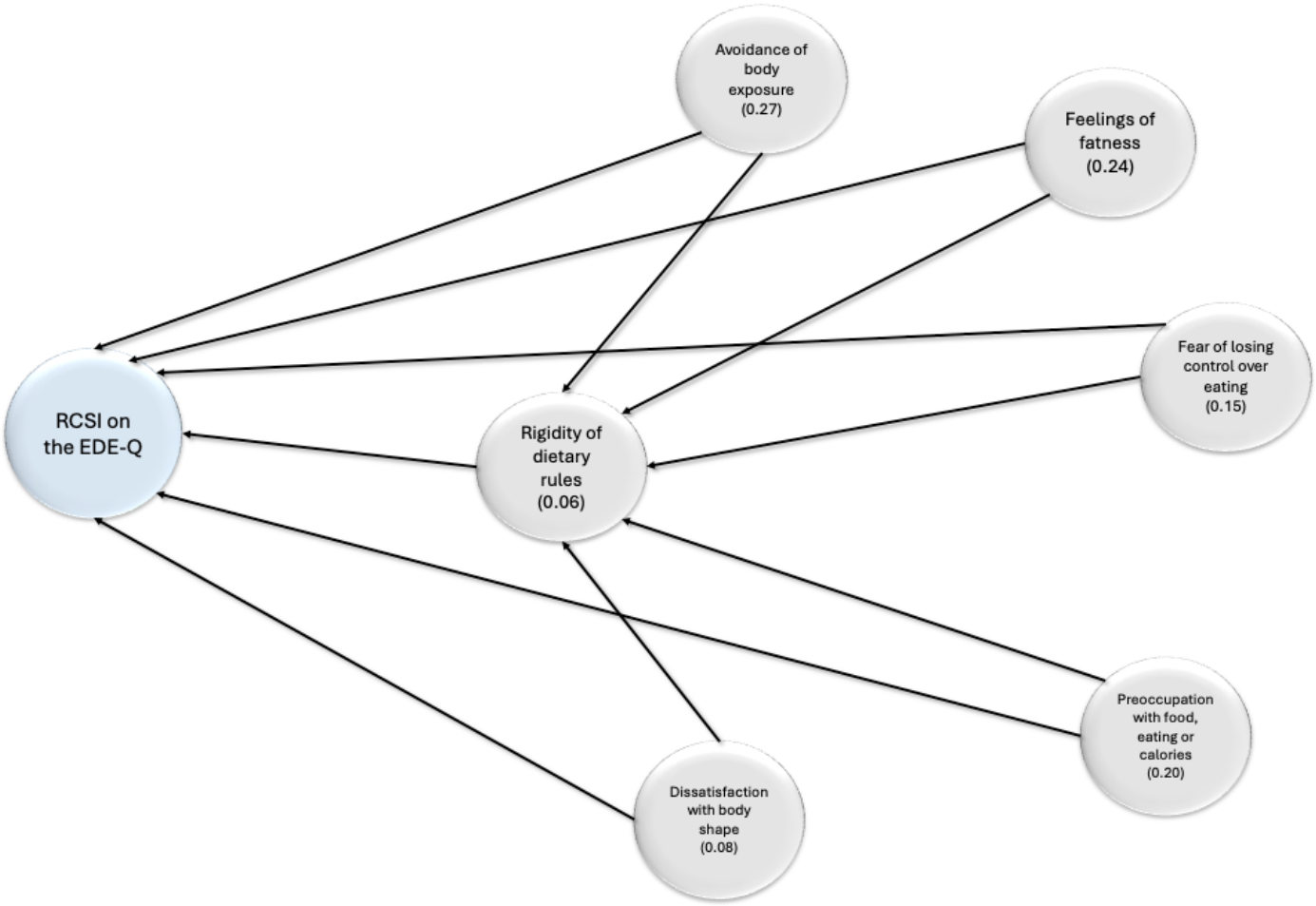
Figure 3. *Variable importance plot*



The Bayesian network model in Figure 4 illustrates the interactions between early change in these symptoms (each represented with a grey node) and their directed (causal) relationship with the outcome of interest (RSCI). The edges (arrows) connecting each node to the outcome indicate a main effect of early changes in that symptom domain as a reliable predictor of recovery. Furthermore, the edges between predictors represent conditional probabilities based on interactions observed in the data. In this model, the effect of node 4 (dietary rules) 'is conditional on' changes in all other nodes, and hence it is less important than other domains. The Bayesian network model differs from the theoretical model presented in the introduction in that it includes only the variables most important for recovery, rather than illustrating the multiple connections between symptoms that maintain a disorder.

Overall, the model correctly classified 68.52% of cases in the training sample, with prediction accuracy (AUC = 0.81).

Figure 4. *Bayesian Network Model of early change in key symptoms as predictors of recovery after therapy*



Note. Each grey circle represents an item-level predictor (node) in the network. The blue circle represents the outcome of interest (RCSI on the EDE-Q). Nodes are connected by arrows (edges). In this way, interactions between nodes exerts an influence on the outcome of interest (joint modelling of conditional probabilities).

External Cross-Validation of Bayesian Network Model Predicting RCSI

Twenty-seven of the 76 patients (38%) in the test sample fully recovered after therapy. The model generalised well to the test sample, with adequate prediction accuracy (AUC = 0.77), correctly classifying 65.7% of cases. There was a small margin of prediction shrinkage (2.73%) relative to the model's performance in the training sample.

Discussion

Summary of findings

Prior research in the field of eating disorders has examined early change measures as a composite score pooling all items on the EDE-Q and treating them as if they have equal weight, as reflected in the global or subscale scores. While informative at a broad level, this methodology does not reveal specific targets for clinicians to address in the early phase of therapy. The Bayesian network method applied in this study enabled the identification of the most important targets for therapeutic change in the earliest phase of treatment.

According to the results, early changes in six areas (ranked by importance) reliably predict recovery after therapy: [1] avoidance of body exposure; [2] feelings of 'fatness'; [3] preoccupation with food, eating or calories; [4] fear of losing control over eating; [5] dissatisfaction with body shape; [6] dietary rules. All of these areas are key targets for change, although it seems that the influence of dietary rules depends on changes in all other domains. Moreover, modifying avoidance of body exposure was found to be the most important area to focus on during the early sessions of therapy.

The performance of the machine learning model was high in the training sample (AUC = 0.81) and generalised well to the external sample (AUC = 0.77). Therefore, general inferences about eating disorder patients can be made based on this model, as it predicts outcomes for new cases with comparably good accuracy, and this -to our knowledge- is the first demonstration of out-of-sample validation of a model of early response targeting specific

symptoms.

In terms of overall treatment outcomes, more than half (55.5%) of patients in the sample had below-threshold scores (EDE-Q < 2.77) after treatment, which is comparable to samples from other eating disorder studies (e.g., Carter et al., 2010; Poulsen et al., 2014; Grilo et al., 2011). However, few studies used the stringent RCSI criteria, based on which only a third (33.6%) of patients in this sample fully recovered. This moderate effect may be due to the inconsistent use of empirically-supported treatment protocols in a large part of the sample, who received integrative and other therapies with a limited evidence base. Such deviations from evidence-based treatment are not uncommon in routine clinical practice and may reflect broader patterns of therapy drift, whereby clinicians gradually depart from structured treatment protocols over time (Waller & Turner, 2016). Furthermore, research suggests that clinicians may hold ambivalent attitudes towards evidence-based practice, often favouring integrative or individually tailored approaches despite limited empirical support (Lilienfield et al., 2013). Finally, the effectiveness of therapy in routine care samples could also be limited by the inclusion of patients with comorbidities that are usually excluded from clinical trial samples, such as post-traumatic stress disorder or substance use disorders.

Relationship with psychological theory and previous research

The Bayesian network model developed and applied in this study suggests that specific cognitive and behavioural symptoms related to body image and food are the most important symptoms to target during the early stages of therapy. The central role of these symptoms in this model aligns with the cognitive-behavioural transdiagnostic theory of eating disorders (Fairburn et al., 2003; Fairburn, 2008). According to this theory, the core psychopathology underpinning all eating disorders is the overvaluation of body shape, weight, and the perceived ability to control them. Individuals with eating disorders often base their self-worth on their body image and control over eating which gives rise to maladaptive

behaviours such as dietary restraint, body checking (e.g. frequent weighing or mirror checking) and body avoidance (e.g. avoiding mirrors or being seen in certain contexts). Although these behaviours may temporarily reduce anxiety or provide a sense of achievement, they often contribute to physiological and psychological consequences, including binge eating, preoccupation with food and preoccupation with body image, which further perpetuates the cycle of disordered eating.

The cognitive-behavioural transdiagnostic theory has been validated by several studies using various approaches such as confirmatory latent-trait analysis (Hoiles et al., 2012; Lampard et al., 2012) and network analysis (Calugi et al., 2020; DuBois et al., 2017; Forrest et al., 2018; Mares et al., 2022; Wang et al., 2018). These studies consistently confirm that overvaluation of weight, shape and eating is central to all eating disorders, highlighting the importance of targeting both the behavioural and cognitive expressions of the underlying overvaluation that maintains the disorder.

In addition to providing empirical support for the cognitive-behavioural transdiagnostic theory for eating disorders, this study's Bayesian network model also suggests an interdependent relationship between specific cognitive and behavioural symptoms related to body image and eating with dietary rules (EDE-Q 4). This is further supported by previous studies, which have identified relationships between dietary restraint and symptoms such as feelings of 'fatness' (Linardon et al., 2018), body dissatisfaction (Andrés & Saldaña, 2014), food preoccupation (Timmerman et al., 2003) and fear of losing control ever eating (Ricca et al., 2012).

Clinical Implications

These findings have several clinical implications, including the potential use of the variable importance plot as an empirically derived guide to plan the sequence of therapeutic techniques targeting each of the identified symptoms. For example, the first symptom area

relating to avoidance of body exposure may be responsive to behavioural experiments, such as body image surveys that test what other people think (Murray et al., 2022). Exposure-based elements (e.g., Becker et al., 2019; Butler at al., 2024) may also be beneficial for body avoidance, fear of losing control over eating and body dissatisfaction and may take the form of imaginal exposure (e.g., imagining losing control while eating), graded exposure (e.g., stepwise changes in clothing or facing situations feared to lead to a loss of control around eating) or flooding (mirror exposure).

The second area, feelings of 'fatness' has been linked to deficits in interoceptive awareness (Pink et al., 2021), which is the ability to identify and distinguish between bodily signals or between body sensations and emotional states (Khalsa et al., 2018). Feelings of 'fatness' is also associated with depressive symptoms (Morales et al., 2022). Therefore, early treatment approaches may focus on increasing interoceptive awareness, addressing depressive symptoms, and correcting cognitive biases such as the mislabelling of emotions and other somatic sensations as 'fatness' (Fairburn, 2008).

Finally, the third area relating to 'preoccupation with food, eating or calories', could benefit from value-driven activity scheduling to restore a sense of pleasure and achievement through other means than through dietary and weight control (Turgon et al., 2019). Research suggests that while initially a psychological consequence of dietary restraint (Polivy, 1996), food preoccupation can subsequently influence eating behaviour (Calitri et al., 2012), contributing to a self-perpetuating cycle of preoccupation and disordered eating. Therefore, targeting this area may facilitate early change in dietary rules, as suggested by the Bayesian network model developed in this study.

Strengths, limitations and further research

A strength of the study was the novel application of the TAN algorithm in the field of eating disorders. While this form of machine learning has been used to develop predictive

models of eating disorder onset (e.g. Rodgers, et al., 2019; Yuan et al., 2024), this study is the first to apply TAN to identify specific symptom networks that undergo early change and predict recovery following treatment.

Another major strength of the study was the development of a predictive model with high prediction accuracy that generalised well to an external test sample. Additionally, the use of real-life clinical data represents a further strength, as it captures the complexity and heterogeneity often excluded from clinical trials. Consequently, these findings are likely to be more generalisable to the broader population.

This study has several limitations and considerations. One consideration relates to the potential for shared measurement variance between the predictors and outcomes. Recovery in this study was determined using the RCSI criteria based on the global score of the EDE-Q. Because the predictive model also used individual EDE-Q items as input variables, it could be argued that early changes in some items may partly reflect autocorrelation with the outcome measure. From a strictly deterministic perspective, improvement in several items contributing to the global score could appear to predict recovery simply because they form part of the same composite measure. However, the Bayesian network model represents probabilistic rather than deterministic relationships between variables. In other words, early change in these six symptoms increases the estimated probability of recovery. Importantly, the EDE-Q contains a broader range of cognitive and behavioural items, yet only six were identified as important predictors within the network. The absence of the other symptoms from the predictive network suggests that not all symptoms contribute equally to recovery. The model therefore estimates the likelihood of recovery conditional on patterns of early symptom change, rather than if change in these items inevitably leads to recovery.

An additional consideration relates to the use of the EDE-Q as the primary outcome measure, which emphasises symptoms such as dietary restraint, eating concerns, weight

concern, and shape concern. Consequently, the EDE-Q primarily captures changes in these cognitive and behavioural symptom domains. While this allows for the assessment of symptom change consistent with the transdiagnostic model of eating disorders, it does not directly measure broader aspects of recovery such as overall psychological wellbeing, psychosocial functioning, or quality of life. As such, it may be limited as a measure of recovery, as it does not encompass the wider range of psychological and functional outcomes that may form part of a more holistic conceptualisation of recovery.

Further limitations related to the study not being pre-registered, and therefore, the findings should be interpreted with caution, particularly in relation to potential biases such as unintentional selective reporting. To support transparency, all analyses conducted have been reported.

An important limitation concerns the sample size, which was limited by the total number of eligible clinical records available in the service's database. Machine learning studies in the field of mental health have indicated that sample sizes of around 300 training cases are necessary to optimise prediction accuracy (Luedtke et al., 2019; Giesemann et al., 2023). Nevertheless, the above evidence of adequate generalisability and minimal shrinkage in a test sample provides reassurance that the results are not simply an illusion of overfitting in a small sample. Despite this, the study warrants replication with a larger sample size.

Another limitation concerns the analytical approach which assumes that each symptom (node) in the network can only have a maximum of two connections (edges) when developing Bayesian network models. In reality, some nodes are likely connected to more than two other nodes, meaning that the model may oversimplify the complexity of interactions between symptoms. Nonetheless, this approach was chosen for its ability to represent uncertainty, develop more parsimonious models (thereby enhancing their utility and reducing the risk of overfitting) and interpret potential causal relationships (Needham et al.,

2007).

An additional limitation is that the data used to develop the predictive model and assess patients' outcomes were derived exclusively from responses to a self-report questionnaire, the EDE-Q, rather than a more comprehensive semi-structured interview. As self-report measures are subject to measurement error and reporting biases, this may have affected the validity and reliability of the resulting symptom network and its subsequent predictions. Nonetheless, the EDE-Q is a well validated tool that is widely used in both clinical and research settings, and it has been shown to correlate strongly with the eating disorder examination interview (Berg et al., 2012).

A further limitation was the exclusion of the weight concern subscale items to simplify the model due to the small sample size. However, this exclusion was justified, as the weight concern subscale items strongly correlate with those in the shape concern subscale (White et al., 2014). Future research could replicate this study with a larger sample size that permits the inclusion of the weight concern subscale to determine whether the resulting model differs.

An added concern is that the research sample consisted predominantly white British female patients and included a very small proportion of patients diagnosed with anorexia nervosa. Due to the small sample size, it was not possible to examine whether different diagnostic subtypes resulted in distinct Bayesian network structures. Moreover, no information was available regarding co-occurring mental health or neurodevelopmental diagnoses, making it impossible to assess their potential impact on the model. Given these factors, the generalisability of the model to patients with different demographic and clinical characteristics may be limited. Future research should aim to include a more demographically diverse sample and examine whether the model differs across diagnostic subtypes and in the presence of co-existing mental health or neurodevelopmental conditions.

A final limitation of this study relates to the measurement of reliable and clinically significant improvement over time. As follow-up data were not available in the dataset used, patient outcomes could only be assessed at the last observed time point. Consequently, it was not possible to evaluate the predictive performance of the Bayesian network model beyond end-of-treatment. Future research should incorporate follow-up assessments to determine whether the model can predict longer-term outcomes and sustained recovery.

Conclusions

In conclusion, the present study involved the novel application of TAN to evaluate which individual EDE-Q items are useful indicators of overall early change. This approach identified that early change in a network of six specific symptom domains predicted recovery after psychological therapy for eating disorders. This observation could be informative for the sequencing of treatment during the earliest sessions of therapy.

References

Aardoom, J. J., Dingemans, A. E. Slof Op't Landt, M. C. T., & Van Furth, E. F. (2012). Norms and discriminative validity of the Eating Disorder Examination Questionnaire (EDE-Q). *Eating Behaviours, 13,* 305-309. https://dx.doi.org/10.1016/j.eatbeh.2012.09.002

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Andrés, A., & Saldaña, B., (2014). Body dissatisfaction and dietary restraint influence binge eating behaviour. *Nutrition Research, 34*(11), 944-950. https://doi.org/10.1016/j.nutres.2014.09.003

Archer, L., Snell, K. I. E., Ensor, J., Hudda, M. T., Collins, G. S., & Riley, R. D. (2020). Minimum sample size for external validation of a clinical prediction model with a continuous outcome. *Statistics in Medicine, 40*(1), 133-146. https://doi.org/10.1002/sim.8766

Bardone-Cone, A. M., Harney, M. B., Maldonado, C. R., Lawson, M. A., Robinson, D. P., Smith, R., & Tosh, A. (2010). Defining recovery from an eating disorder: conceptualisation, validation, and examination of psychosocial functioning and psychiatric comorbidity. *Behaviour Research and Therapy, 48*(3), 194-202. https://doi.org/10.1016/j.brat.2009.11.001

Beard, J. I., & Delgadillo, J. (2019). Early response to psychological therapy as predictor of depression and anxiety treatment outcomes: A systemic review and meta-analysis. *Depression and Anxiety, 36*(9), 866-878. https://doi.org/10.1002/da.22931

Black Becker, C., Farrell, N. R., & Waller, G. (2019). *Exposure Therapy for Eating Disorders*. Oxford University Press. https://doi.org/10.1093/med-psych/9780190069742.001.0001

Bell, C., Waller, G., Shafran, R. & Delgadillo, J. (2017). Is there an optimal length of psychological treatment for eating disorder pathology? *International Journal of Eating Disorder, 50,* 687-692. https://doi.org/10.1002/eat.22660

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*(3), 428–438. https://doi.org/10.1002/eat.20931

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry, 16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology, 9,* 91-121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199-231. https://doi.org/10.1214/ss/1009213726

Briganti, G., Scutari, M., & McNally, R. J. (2023). A tutorial on bayesian networks for psychopathology researchers. *Psychological Methods*, *28(*4), 947–961. https://doi.org/10.1037/met0000479

Butler, R., M., Crumby, E. K., Christian, C., Brosof, L. C., Vanzhula, I. A., & Levinson, C. A. (2024). Facing eating disorder fears: An open trial adapting prolonged exposure to the treatment of eating disorders. *Behavior Therapy, 55*(2), 347-360. https://doi.org/10.1016/j.beth.2023.07.008

Calitri, R., Pothos, E. M., Tapper, K., Brunstorm, J. M., & Rodgers, P. J. (2012). Cognitive biases to healthy and unhealthy food words predict change in BMI. *Obesity, 18,* 2282-2287. https://doi.org/10.1038/oby.2010.78

Calugi, S., Sartirana, M., Misconel, A., Boglioli, C., & Dalle Grave, R. (2020). Eating disorder psychopathology in adults and adolescents with anorexia nervosa: A network approach. *International Journal of Eating Disorders, 53*(5), 690-701. https://doi.org/10.1002/eat.23270

Carter, F. A., Jordan, J., McIntosh, V. V. W., Luty, S. E., Mckenzie, J. M., Frampton, C. A. A., Bulik, C. M., & Joyce, P. R. (2010). The long-term efficacy of three psychotherapies for anorexia nervosa: A randomized, controlled trial. *International Journal of Eating Disorders, 44*(7), 647-654. https://doi.org/10.1002/eat.20879

Chang, P. G. R. Y., Delgadillo, J., & Waller G. (2021). Early response to psychological treatment for eating disorders: A systemic review and meta-analysis. *Clinical Psychology Review, 86,* 1-14. https://doi.org/10.1016/j.cpr.2021.102032

DuBois, R. H., Rodgers, R. F., Franko, D. L., Eddy, K. T., & Thomas, J. J. (2017). A network analysis investigation of the cognitive-behavioral theory of eating disorders. *Behaviour research and therapy, 97*, 213-221. https://doi.org/10.1016/j.brat.2017.08.004

Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders.* Guilford Press.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363-370. https://doi.org/10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy, 41*(5), 509-528. https://doi.org/10.1016/s0005-7967(02)00088-8

Fairburn, C., & Beglin, S. J. (2008). Eating disorders examination questionnaire (EDE-Q 6.0) In: C. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders.* New York:

Guilford Press.

Fairburn, C., & Cooper, Z. (1993). The eating disorder examination. In C. Fairburn & G. Wilson (Ed.), *Binge eating: Nature, assessment, and treatment.* New York. Guilford Press.

Forrest, L. N., Jones, P. J., Ortiz, S. N., & Smith, A. R. (2018). Core psychopathology in anorexia nervosa and bulimia nervosa: A network analysis. *International Journal of Eating Disorders*, *51*(7), 668–679. https://doi.org/10.1002/eat.22871

Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine Learning, 29,* 131-163. https://10.1023/A:1007465528199

Fu, S., & Desmarais, M. C. (2010). Markov blanket based feature selection: a review of past decade. *In Proceedings of the world congress on engineering* (Vol. 1, pp. 321-328). Newswood Ltd.

Giesemann, J., Delgadillo, J., Schwartz, B., Bennemann, B., & Lutz, W. (2023). Predicting dropout from psychological treatment using different machine learning algorithms, resampling methods, and sample sizes. *Psychotherapy Research, 33*(6), 683-695. https://doi.org/10.1080/10503307.2022.2161432

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*(5), 675–685. https://doi.org/10.1037/a0025049

Hilbert, A., Herpertz, S., Zipfel, S., Tuschen-Caffier, B., Friederich, H. C., Mayr, A., Crosby, R. D., & de Zwaan, M. (2019). Early change trajectories in cognitive-behavioral therapy for binge-eating disorder. *Behavior Therapy, 50*, 115–125. https://doi.org/10.1016/j.beth.2018.03.013

Hoiles, K. J., Egan, S. J., & Kane, R. T. (2012). The validity of the transdiagnostic cognitive behavioural model of eating disorders in predicting dietary restraint. *Eating Behaviours, 13*(2), 123-126. https://doi.org/10.1016/j.eatbeh.2011.11.007

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. https://doi.org/10.1037//0022-006x.59.1.12

Jansen, A. (2016). Eating disorders need more experimental psychopathology. *Behaviour Research and Therapy, 86,* 2-10. https://doi.org/10.1016/j.brat.2016.08.004

Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. D., Feusner, J. D., Garfinkel, S. N., Lane, R. D., Mehling, W. E., Meuret, A., E., Nemeroff, C. B., Oppenheimer, S., Petzschner, F. H., Pollatos, O., Rhudy, J. L., Schramm, L. P., Simmons, W. K., Stein, M. B., Stephan, K. E., & Zucker, N. (2018). Interoception and mental health: A roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(6), 501-513. https://doi.org/10.1016/j.bpsc.2017.12.004

Klump, K. L., Bulik, C. M., Kaye, W. H., Treasure, J., &Tyson, E. (2009). Academy for Eating Disorders position paper: Eating disorders are serious mental illnesses. *International Journal of Eating Disorders, 42,* 97-103. https://doi.org/10.1002/eat.20589

Lampard, A. M., Tasca, G. A., Balfour, L., Bissada, H. (2012). An evaluation of the transdiagnostic cognitive-behavioural model of eating disorders. *European Eating Disorders Review, 21*(2), 99-107. https://doi.org/10.1002/erv.2214

Le, L. K., Mihalopoulos, C., Engel, L., Touyz, S., Gonzalez-Chica, D. A., Stocks, N., & Hay, P. (2019). Burden and health state utility values of eating disorders: Results from a population-based survey. *Psychological Medicine, 31,* 1-9.

https://doi.org/10.1017/S0033291719003064

Levinson, C. A., Cusack, C., Brown, M. L., & Smith, A. R. (2023). A network approach can improve eating disorder conceptualization and treatment. *Nature Reviews Psychology, 1*(7), 419-430. https://doi.org/10.1038/s44159-022-00062-y

Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzman, R. D. (2013). Why many clinical psychologists are resistant to evidence-based practice: root causes and constructive remedies. *Clinical psychology review*, *33*(7), 883–900. https://doi.org/10.1016/j.cpr.2012.09.008

Linardon, J., Brennan, L., & de la Piedad Garcia, X. (2016). Rapid response to eating disorder treatment: A systematic review and meta-analysis. *International Journal of Eating Disorders, 49*(10), 905-919. https://doi.org/10.1002/eat.22595

Lindardon, J., Phillipou, A., Castle, D., Newton, R., Harrison, P., Cistullo, L. L., Griffiths, S., Hindle, A., & Brennan, L. (2018). Feeling fat in eating disorders: Testing the unique relationships between feeling fat and measures of disordered eating in anorexia nervosa and bulimia nervosa. *Body Image, 25,* 163-167. https://doi.org/10.1016/j.bodyim.2018.04.001

Luce, K. H., & Crowther, J. H. (1999). The reliability of the eating disorder examination – self report questionnaire version (EDE-Q). *International Journal of Eating Disorder, 23*(3), 349-351. https://doi.org/10.1002/(SICI)1098-108X(199904)25:3<349::AID-EAT15>3.0.CO;2-M

Luedtke, A., Sadikova, E., & Kessler, R. C. (2019). Sample size requirements for multivariate models to predict between-patient differences in best treatments of major depressive disorder. *Clinical Psychological Science, 7*(3), 445-461. https://doi.org/10.1177/2167702618815466

MacDonald, D. E., Trottier, K., McFarlane, T., & Olmsted, M. P. (2015). Empirically defining

rapid response to intensive treatment to maximize prognostic utility for bulimia nervosa and purging disorder. *Behaviour Research and Therapy, 68,* 48-53. https://doi.org/10.1016/j.brat.2015.03.007

Mares, S. H. W., Burger, J., Lemmens, L. K. J. M., van Elburg, A. A., & Vroling, M. S. (2022). Evaluation of the cognitive behavioural theory of eating disorders: A network analysis investigation. *Eating Behaviors, 44,* 101590. https://doi.org/10.1016/j.eatbeh.2021.101590

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy, 86,* 95-104. https://doi.org/10.1016/j.brat.2016.06.006

Miskovic-Wheatley, J., Bryant, E., Ong, S. H., Vatter, S., Le, A., Touyz, S., & Maguire, S. (2023). Eating disorder outcomes: findings from a rapid review of over a decade of research. *Journal of Eating Disorders, 11*(85), 1-25. https:doi.org/10.1186/s40337-023-00801-3

Monteleone, A. M., Pellegrino, F., Croatto, G., Carfagno, M., Hilbert, A., Treasure, J., Wade, T., Bulik, C. M., Zipfel, S., Hay, P., Schmidt, U., Castellini, G., Favaro, A., Fernandez-Aranda, F., Il Shin, J., Voderholzer, U., Ricca, V., Moretti, D., Busatta, D., … Solmi, M. (2022). Treatment of eating disorders: A systematic meta-review of meta-analyses and network meta-analyses. *Neuroscience and Biobehavioral Reviews*, *142*, 104857. https://doi.org/10.1016/j.neubiorev.2022.104857

Morales, C., Dolan, S. C., Anderson, D. A., Anderson, L. M., & Reilly, E. E. (2022). Exploring the contributions of affective constructs and interoceptive awareness to feeling fat. *Eating and Weight Disorders – Studies on Anorexia, Bulimia and Obesity, 27,* 3533-3541. https://doi.org/10.1007/s40519-022-01490-8

Murray, H., Kerr, A., Warnock-Parkes, E., Wild, J., Grey, N., Clark, D. M., & Ehlers, A. (2022). What do others think? The why, when and how of using surveys in CBT. *The*

Cognitive Behaviour Therapist, 15, e42. https://doi.org/10.1017/S1754470X22000393

National Institute for Health and Care Excellence. (2017). *Eating disorders: recognition and treatment* (NICE Guideline NG69). https://www.nice.org.uk/guidance/ng69

Needham, C. J., Bradford, J. R., Bulpitt, A. J., & Westhead, D. R. (2007). A primer on learning in Bayesian networks for computational biology. *PLoS Computational Biology, 31*(3), e129. https://doi.org/10.1371/journal.pcbi.0030129

Pink, A. E., Williams, C., Lee, M., Young, H. A., Harrison, S., Davies, A. E., & Price, M. (2021). Manipulating the sensation of feeling fat: The role of alexithymia, interoceptive sensitibility and perfectionism. *Physiology & Behavior, 239,* 113501. https://doi.org/10.1016/j.physbeh.2021.113501

Polivy, J. (1996). Psychological consequences of food restriction. *Journal of the American Dietetic Association, 96*(6), 589-592. https://doi.org/10.1016/S0002-8223(96)00161-7

Poulsen, S., Lunn, S., Daniel, S. I. F., Folke, S., Mathiesen, B. B., Katznelson, H., & Fairburn, C. G. (2014). A randomized controlled trial of psychoanalytic psychotherapy or cognitive-behavioural therapy for bulimia nervosa. *American Journal of Psychiatry, 171*(1). https://doi.org/10.1176/appi.ajp.2013.12121511

Rahmani, M., Omidi, A., Asemi, Z., & Akbari, H. (2018). The effect of dialectical behavioural therapy on binge eating, difficulties in emotion regulation and BMI in overweight patients with binge-eating disorder: A randomised controlled trial. *Mental Health & Prevention, 9,* 13-18. https://doi.org/10.1016/j.mhp.2017.11.002

Ricca, V., Castellini, G., Fioravanti, G., Sauro, C. L., Rotella, F., Ravaldi, C., Lazzeretti, L., & Faravelli, C. (2012). Emotional eating in anorexia nervosa and bulimia nervosa. *Comprehensive psychiatry, 53*(3), 245-251. https://doi.org/10.1016/j.comppsych.2011.04.062

Riley, R. D., Snell, K. I., Ensor, J., Burke, D. L., Harrell Jr, F. E., Moons, K. G., & Collins, G.

S.. (2019). Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Statistics in Medicine*, *38*(7), 1276–1296. https://doi.org/10.1002/sim.7992

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*(5), 408-412. https://doi.org/10.1002/erv.2372

Rodgers, R. F., DuBois, R., Thiebaut, S., Jaussent, I., Maimoun, L., Seneque, M., Lefebvre, P., Renard, E., Courtet, P., & Guillaume, S. (2019). Structural differences in eating disorder psychopathology after history of childhood abuse: Insights from a Bayesian network analysis. *Journal of Abnormal Psychology, 128*(8), 795–805. https://doi.org/10.1037/abn0000470

Rose, J. S., Vaewsorn, A., Rosselli-Navarra, F., Wilson, G. T., & Weissman, R. S. (2013). Test-retest reliability of the eating disorder examination-questionnaire (EDE-Q) in a college sample. *Journal of Eating Disorders, 1*(42), 1-10. https://doi.org/10.1186/2050-2974-1-42

Solmi, M., Wade, T. D., Byrne, S., Giovane, C. D., Fairburn, C. G., Ostinelli, E. G., Crescenzo, F. D., Johnson, C., Schmidt, U., Treasure, J., Favaro, A., Zipfel, S., & Cipriani, A. (2021). Comparative efficacy and acceptability of psychological interventions for the treatment of adult outpatients with anorexia nervosa: a systematic review and network meta-analysis. *The Lancet Psychiatry, 8*(3), 215-224. https://doi.org/10.1016/S2215-0366(20)30566-6

Timmerman, G. M., & Gregg, E. K. (2003). Dieting, perceived deprivation, and preoccupation with food. *Western Journal of Nursing Research*, *25*(4), 405-418. https://doi.org/10.1177/0193945903025004006

Turgon, R., Ruffault, A., Juneau, C., Blatier, C., & Shankland, R. (2019). Eating disorder treatment: a systematic review and meta-analysis of the efficacy of mindfulness-based programs. *Mindfulness, 10*(11), 2225-2244. https://doi.org/10.1007/s12671-019-01216-5

Vall, E., & Wade, T. D. (2015). Predictors of treatment outcome in individuals with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders, 48*(7), 946-971. https://doi.org/10.1002/eat.22411

Waller, G., & Turner, H. (2016). Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour research and therapy*, *77*, 129–137. https://doi.org/10.1016/j.brat.2015.12.005

Wang, S. B., Jones, P. J., Dreier, M., Elliott, H., & Grilo, C. M. (2018). Core psychopathology of treatment-seeking patients with binge-eating disorder: a network analysis investigation. *Psychological Medicine*, *49*(11), 1923–1928. https://doi.org/10.1017/S0033291718002702

White, H. J., Haycraft, E., Goodwin, H., & Meyer, C. (2014). Eating disorder examination questionnaire: factor structure for adolescent girls and boys. *International Journal of Eating Disorders, 47*(1), 99-105. https://doi.org/10.1002/eat.22199

Yuan, Y., Du, B., Lu, H., Zhang, Y., Jiang, K., Jeong, J., & Tu, Y. (2024). Masculinity level predicts the risk of eating disorders among adolescents in China: a study based on a Bayesian network. *Current Psychology, 43,* 21940-21951. https://doi.org/10.1007/s12144-024-05996-5

Zhang, K., Xie, Q., Fan, C., Hu, X., Lei., J., Kong, J., Liang, M., Luo, J., & Li, X. (2024). The effectiveness of interpersonal psychotherapy versus cognitive behavioural therapy for eating disorders: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy, 31*(1), e2953, https://doi.org/10.1002/cpp.2953

Appendices

Appendix A: Ethical approval letter

Appendix B: Eating Disorder examination questionnaire (EDE-Q 6.0)

Appendix A: Ethical approval letter

The University Of Sheffield.

Downloaded: 02/03/2024
Approved: 26/02/2024

Ammara Imtiaz
Registration number: 220237734
Psychology
Programme: Doctor of Clinical Psychology

Dear Ammara

PROJECT TITLE: A bayesian network analysis of the symptoms driving the early response effect in eating disorder treatment
APPLICATION: Reference Number 058515

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 26/02/2024 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 058515 (form submission date: 23/02/2024); (expected project end date: 31/05/2025).

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Your responsibilities in delivering this research project are set out at the end of this letter.

Yours sincerely

Department Of Psychology Research Ethics Committee
Ethics Administrator
Psychology

Please note the following responsibilities of the researcher in delivering the research project:

- The project must abide by the University's Research Ethics Policy: https://www.sheffield.ac.uk/research-services/ethics-integrity/policy
- The project must abide by the University's Good Research & Innovation Practices Policy: https://www.sheffield.ac.uk/polopoly_fs/1.671066!/file/GRIPPolicy.pdf
- The researcher must inform their supervisor (in the case of a student) or Ethics Administrator (in the case of a member of staff) of any significant changes to the project or the approved documentation.
- The researcher must comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data.
- The researcher is responsible for effectively managing the data collected both during and after the end of the project in line with best practice, and any relevant legislative, regulatory or contractual requirements.

Appendix B: Eating Disorder examination questionnaire (EDE-Q 6.0)

Eating Disorder examination questionnaire (EDE-Q 6.0)

Instructions: The following questions are concerned with the past four weeks (28 days) only.

Please read each question carefully. Please answer all the questions. Thank you.

Questions 1 to 12: Please circle the appropriate number on the right. Remember that the questions only refer to the past four weeks (28 days) only.

	ON HOW MANY OF THE PAST 28 DAYS ...	NO DAYS	1-5 DAYS	6-12 DAYS	13-15 DAYS	16-22 DAYS	23-27 DAYS	EVERY DAY
1	Have you been deliberately **trying** to limit the amount of food you eat to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
2	Have you gone for long periods of time (8 waking hours or more) without eating anything at all in order to influence your shape or weight?	0	1	2	3	4	5	6
3	Have you **tried** to exclude from your diet any foods that you like in order to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
4	Have you **tried** to follow definite rules regarding your eating (for example, a calorie limit) in order to influence your shape or weight (whether or not you have succeeded)?	0	1	2	3	4	5	6
5	Have you had a definite desire to have an **empty** stomach with the aim of influencing your shape or weight?	0	1	2	3	4	5	6
6	Have you had a definite desire to have a **totally fla**t stomach?	0	1	2	3	4	5	6
7	Has thinking about **food, eating or calories** made it very difficult to concentrate on things you are interested in (for example, working, following a conversation, or reading)?	0	1	2	3	4	5	6
8	Has thinking about **shape or weight** made it very difficult to concentrate on things you are interested in (for example, working, following a conversation, or reading)?	0	1	2	3	4	5	6
9	Have you had a definite fear of losing control over eating?	0	1	2	3	4	5	6
10	Have you had a definite fear that you might gain weight?	0	1	2	3	4	5	6
11	Have you felt fat?	0	1	2	3	4	5	6
12	Have you had a strong desire to lose weight?	0	1	2	3	4	5	6

PAGE 1/3 PLEASE GO TO THE NEXT PAGE

Eating Disorder examination questionnaire (EDE-Q 6.0)

Questions 13-18: Please fill in the appropriate number in the boxes on the right. Remember that the questions only refer to th past four weeks (28 days).

Over the past four weeks (28 days)….

13	Over the past 28 days, how many **times** have you eaten what other people would regards as an **unusually large amount of food** (given the circumstances)?	
14	… On how many of these times did you have a sense of having lost control over your eating (at the time you were eating)?	
15	Over the past 28 days, on how many **DAYS** have such episodes of overeating occurred (i.e. you have eaten an unusually large amount of food **and** have had a sense of loss of control at the time)?	
16	Over the past 28 days, how many **times** have you made yourself sick (vomit) as a means of controlling your shape or weight?	
17	Over the past 28 days, how many **times** have you taken laxatives as a means of controlling your shape or weight?	
18	Over the past 28 days, how many **times** have you exercised in a "driven" or "compulsive" way as a means of controlling your weight, shape or amount of fat, or to burn off calories?	

Questions 19 to 21: Please circle the appropriate number. <u>Please note that for these questions the term "binge eating" means</u> eating what others would regard as an unusually large amount of food for the circumstances, accompanied by a sense of having lost control over eating.

		NO DAYS	1-5 DAYS	6-12 DAYS	13-15 DAYS	16-22 DAYS	23-27 DAYS	EVERY DAY
19	Over the past 28 days, on how many days have you eaten in secret (ie, furtively)? … Do not count episodes of binge eating.	0	1	2	3	4	5	6
		None of the times	A few of the times	Less than half	Half of the times	More than half	Most of the time	Every time
20	On what proportion of the times that you have eaten have you felt guilty (felt that you've done wrong) because of its effect on your shape or weight? … Do not count episodes of binge eating.	0	1	2	3	4	5	6
			Not at all	Slightly	Moderately		Markedly	
21	Over the past 28 days, how concerned have you been about other people seeing you eat? … Do not count episodes of binge eating.	0	1	2	3	4	5	6

Eating Disorder examination questionnaire (EDE-Q 6.0)

Questions 22 to 28: Please circle the appropriate number on the right. Remember that the questions only refer to the past four weeks (28 days).

	ON HOW MANY OVER THE PAST 28 DAYS ...	NOT AT ALL	SLIGHTLY		MODERATELY		MARKEDLY	
22	Has your **weight** influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
23	Has your **shape** influenced how you think about (judge) yourself as a person?	0	1	2	3	4	5	6
24	How much would it have upset you if you had been asked to weigh yourself once a week (no more, or less, often) for the next four weeks?	0	1	2	3	4	5	6
25	How dissatisfied have you been with your **weight**?	0	1	2	3	4	5	6
26	How dissatisfied have you been with your **shape**?	0	1	2	3	4	5	6
27	How uncomfortable have you felt seeing your body (for example, seeing your shape in the mirror, in a shop window reflection, while undressing or taking a bath or shower)?	0	1	2	3	4	5	6
28	How uncomfortable have you felt about **others** seeing your shape or figure (for example, in communal changing rooms, when swimming, or wearing tight clothes)?	0	1	2	3	4	5	6

What is your weight at present? (Please give your best estimate.): .....

What is your height? (Please give your best estimate.): .....

If female: Over the past three to four months have you missed any menstrual periods?: YES ○ NO ○

If so, how many?: ○

Have you been taking the "pill"?: YES ○ NO ○

PAGE 3/3

THANK YOU