



Universiteit  
Leiden  
The Netherlands

## **An overview and investigation of relapse predictors in anorexia nervosa: a systematic review and meta-analysis**

Rijk, E.S.J. de; Almirabi, D.; Robinson, L.; Schmidt, U.; Furth, E.F. van; Landt, M.C.T.S.O.

### **Citation**

Rijk, E. S. J. de, Almirabi, D., Robinson, L., Schmidt, U., Furth, E. F. van, & Landt, M. C. T. S. O. (2023). An overview and investigation of relapse predictors in anorexia nervosa: a systematic review and meta-analysis. *International Journal Of Eating Disorders*, 57(1), 3-26. doi:10.1002/eat.24059

Version: Publisher's Version  
License: [Creative Commons CC BY 4.0 license](#)  
Downloaded from: <https://hdl.handle.net/1887/3714455>

**Note:** To cite this publication please use the final published version (if applicable).

## REVIEW

# An overview and investigation of relapse predictors in anorexia nervosa: A systematic review and meta-analysis

Eline S. J. de Rijk MSc<sup>1,2</sup>  | Durr Almirabi MSc<sup>3</sup> | Lauren Robinson PhD<sup>3</sup>  |  
Ulrike Schmidt PhD<sup>3</sup>  | Eric F. van Furth PhD<sup>1,2</sup>  |  
Margarita C. T. Slof-Op 't Landt PhD<sup>1,2</sup> 

<sup>1</sup>GGZ Rivierduinen Eating Disorders Ursula, Leiden, the Netherlands

<sup>2</sup>Department of Psychiatry, Leiden University Medical Center, Leiden, the Netherlands

<sup>3</sup>Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

## Correspondence

Eline S. J. de Rijk, GGZ Rivierduinen Eating Disorders Ursula, PO Box 405, 2300 AK Leiden, The Netherlands.

Email: [e.derijk@rivierduinen.nl](mailto:e.derijk@rivierduinen.nl)

Action Editor: Kelly L. Klump

## Abstract

**Objective:** An extensive number of predictors has been examined across the literature to improve knowledge of relapse in anorexia nervosa (AN). These studies provide various recovery and relapse definitions, follow-up durations and relapse rates. The current study summarizes these values and predictors of relapse in AN in a review and meta-analysis.

**Method:** The study was executed according to PRISMA guidelines. Different databases were searched and studies in which participants did not receive an official clinical diagnosis were excluded. A quality analysis was performed using the National Institute of Health's Study Quality Assessment Tool. Random-effects meta-analyses were conducted to summarize data.

**Results:** Definitions of relapse and recovery were diverse. During an average follow-up period of 31 months an average relapse rate of 37% was found. Predictive variables from 28 studies were grouped in six categories: age and sex, symptoms and behaviors, AN subtype and duration, weight or weight change, comorbidity, and personality. The studies were characterized by non-significant and contradictory results. Meta-analyses were performed for the predictors age, AN duration, pre-treatment BMI, post-treatment BMI and depression. These yielded significant effects for post-treatment BMI and depression: higher pre-treatment depression (SMD = .40 CI [.21-.59] and lower post-treatment BMI (SMD = -.35 CI [-.63 to -.07]) increased relapse chances in AN.

**Discussion:** Our results emphasized a lack of sufficiently powered studies, consistent results, and robust findings. Solely post-treatment BMI and pre-treatment depression predicted relapse. Future research should use uniform definitions, larger samples and better designs, to improve our understanding of relapse in AN.

**Public significance:** Knowledge about predictors is important to understand high relapse rates. Our study performed a review and meta-analysis of relapse predictors in AN. Related to the heterogeneity in studies examining predictors, an overview of relapse and recovery definitions, follow-up durations and relapse rates for AN was provided. Significant effects were found for post-treatment BMI and pre-treatment

depression. More studies with uniform definitions are needed to improve clinical implications.

### Resumen

**Objetivo:** En la literatura se ha examinado un amplio número de predictores para mejorar el conocimiento de la recaída en la anorexia nerviosa (AN). Estos estudios proporcionan diversas definiciones de recuperación y recaída, duraciones del seguimiento y tasas de recaída. El presente estudio resume estos valores y predictores de recaída en AN en una revisión y metaanálisis.

**Método:** El estudio se realizó siguiendo las directrices PRISMA. Se realizaron búsquedas en diferentes bases de datos y se excluyeron los estudios en los que los participantes no recibieron un diagnóstico clínico oficial. Se realizó un análisis de calidad mediante la herramienta de evaluación de la calidad de los estudios del Instituto Nacional de Salud. Se realizaron metaanálisis de efectos aleatorios para resumir los datos.

**Resultados:** Las definiciones de recaída y recuperación fueron diversas. Durante un período de seguimiento promedio de 31 meses se encontró una tasa media de recaída del 37%. Las variables predictivas de 28 estudios se agruparon en seis categorías: edad y sexo, síntomas y conductas, subtipo y duración de la AN, peso o cambio de peso, comorbilidad y personalidad. Los estudios se caracterizaron por resultados no significativos y contradictorios. Se realizaron metaanálisis para los predictores edad, duración de la AN, IMC pretratamiento, IMC postratamiento y depresión. Éstos arrojaron efectos significativos para el IMC postratamiento y la depresión: una mayor depresión pretratamiento (DME =  $-.40$ ; IC: [21 a, 59]) y un menor IMC postratamiento (DME =  $-.35$ ; IC: [ $-.63$  a  $-.07$ ]) aumentaron las probabilidades de recaída en la AN.

**Discusión:** Nuestros resultados enfatizaron la falta de estudios con suficiente potencia, resultados consistentes y hallazgos robustos. Sólo el IMC postratamiento y la depresión pretratamiento predijeron la recaída. Las investigaciones futuras deberían utilizar definiciones uniformes, muestras más grandes y mejores diseños, para mejorar nuestra comprensión de la recaída en la AN.

### KEYWORDS

anorexia nervosa, BMI, depression, predictors, recovery, relapse

## 1 | INTRODUCTION

Anorexia Nervosa (AN) is associated with serious problems including the presence of comorbidity, high levels of psychological distress, a reduced quality of life, and increased mortality (Ágh et al., 2016; Jagielska & Kacperska, 2017; van Eeden et al., 2021). Although an extensive amount of research focusses on treatment for AN, outcome is often not satisfactory (Khalsa et al., 2017) and a substantial proportion of patients relapse. Estimates of AN relapse rates range between 30% and 50% (Berends et al., 2016; Khalsa et al., 2017; Steinglass et al., 2011), making relapse a significant concern.

Multiple studies have focused on predictors of relapse in AN, including physical, cognitive, and socio-emotional factors (Berends et al., 2018; Frostad et al., 2022). Reviewing these predictors could provide insight into what factors contribute to relapse in AN and might assist in the formation of more effective treatments and (relapse) prevention programs (Berends et al., 2018; Bodell & Mayer, 2011; Carter et al., 2012). However, high variability exists across studies examining predictors, including differences in designs (Berends et al., 2018; Glashouwer et al., 2018; Khalsa et al., 2017), as well as the use of different relapse and recovery definitions (Khalsa et al., 2017; Sala et al., 2023). Some studies defined relapse according to levels of BMI, while others provided a more extensive definition

including the presence of compensatory, binge and/or purging behaviors (Khalsa et al., 2017). This high variability in study designs and definitions, has resulted in different interpretations of predictive effects, complicating overall knowledge of factors preceding relapse in AN.

To the best of our knowledge, only two studies attempted to summarize predictors of relapse in AN. First, a systematic review conducted by Berends et al. (2018) showed that the factors: weight and shape concerns, comorbidity, compulsions, longer illness duration and intensity of treatment were predictive of relapse. In addition, their meta-analysis revealed a relapse rate of 31% among individuals with AN. A second meta-analysis examined predictors of relapse across all eating disorder (ED) types: AN, Bulimia Nervosa, Binge Eating Disorder and Other Specified Feeding or Eating Disorder (Sala et al., 2023). This meta-analysis quantified and compared predictors of relapse and examined moderators of these relationships. Significant predictors of relapse were found: psychiatric comorbidity, ED psychopathology, leptin levels, motivation, BMI, response to treatment, and ED subtype. Given the severe nature of AN (Jagielska & Kacperska, 2017; van Eeden et al., 2021), the conceivably higher relapse rate in comparison to the other EDs (Sala et al., 2023), and the predominant focus on AN in previous ED relapse studies, conducting an updated combined systematic review and meta-analysis that specifically focuses on predictors of relapse in AN emerges as a crucial next step to further understand the relapse risk in this disorder.

Therefore, the main goal of the current study was to perform a systematic review and meta-analysis examining predictors of relapse in AN. Predictors of relapse assessed in at least two independent studies were summarized in the systematic review. To improve the interpretation of effects and in comparison to the study of Sala et al. (2023), meta-analyses were solely performed for predictors that were evaluated in a minimum of five independent studies (Jackson & Turner, 2017). This approach improved methodological consistency and strengthened the validity of the synthesized findings of the meta-analysis. Furthermore, to acknowledge the large diversity in relapse studies performed so far, an overview of the different recovery and relapse definitions, relapse rates, follow-up durations and quality of included studies was provided.

## 2 | METHODS

### 2.1 | Search strategy

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) checklist (Moher et al., 2009). The protocol was registered in PROSPERO (January 17, 2021) and may be accessed under the registration number CRD42021230170 or on the PROSPERO website: [https://www.crd.york.ac.uk/prospero/display\\_record.php?ID=CRD42021230170](https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021230170). Using the Problem, Intervention, Control, Outcome (PICO; Huang et al., 2006) format, the research question was defined as: "What are predictors of relapse in people who have been previously diagnosed with an ED?". Although the primary goal of the study was to

examine this for all EDs, the article changed its scope to AN after the article of Sala et al. (2023) was published. While not mentioned as a separate goal in PROSPERO, meta-analyses were conducted if predictors were examined in a significant number of studies. Relevant studies focusing on keywords of the research question were identified by searching the following electronic databases: PubMed, MEDLINE, Embase, PsycINFO, PsycArticles, Psychology and Behavioral Sciences Collection, Web of Science, COCHRANE Library, Emcare, and Academic Search Premier. The following key terms were included in the search (for exact literature search used see Appendix S1): (eating disorder\* OR bulimi\* OR anorexi\* OR binge\* OR other specified eating disorder\*) AND (recur\* OR relaps\* OR rehospitalization OR setback) AND (predictor\* OR risk factor\* OR prognostic factor\*). No limitations were included for age, follow-up duration, treatment, design and definitions of relapse or recovery. Studies that only examined ED symptoms or in which participants did not receive an official clinical diagnosis were excluded from the analysis. The titles and abstracts of identified studies were screened by a random combination of two researchers (LR, MSOL, DA, ER) in the program Covidence (Veritas Health Innovation, 2021). A moderate level of inter rater agreement (Cohen's kappa = .43; McHugh, 2012) was found, and disagreements were addressed through discussions involving all four researchers. Subsequently, the full texts of all included abstracts were also screened by two researchers to determine eligibility for the review. A high inter rater agreement was found for included studies (Cohen's kappa = .82) and disagreements were resolved by means of discussion.

### 2.2 | Data extraction and synthesis

Studies exploring predictors of relapse in AN that dated from the earliest article until the search date (July 3, 2023) were included in the current study. One of three authors (LR, MSOL, and ER) extracted the following data from included studies: author, year, number of participants, in or outpatient treatment, treatment, age, sex (assigned at birth) or gender (MacDonald et al., 2022), ethnicity, test type, definition of recovery and relapse, results of predictors, relapse rate and time of follow up. To control extraction, one of the remaining authors (LR, MSOL, and ER) checked definitions, key numbers and statistics, any disagreements were discussed.

Given the expected levels of variability, the measurement of a predictor needed to be present in a minimum of five articles to achieve enough power to conduct a random-effects meta-analysis (Jackson & Turner, 2017). Data could only be included if a variable was assessed by the same instrument at the same time point (pre or post-treatment), and means and standard deviations were present for a group of relapsed and non-relapsed people. These values were extracted to calculate the standardized mean difference (SMD), resembling the difference between means of the relapsed and the non-relapsed group divided by the pooled standard deviation (Brown et al., 2017). Furthermore, Cohen's *d* was used to interpret effects, a value of >.2 was perceived as a small effect, >.5 as moderate and .8 as big (Cohen (1988)). In total, 24 authors were contacted because data were not available: 5 of these responded that they were unable to retrieve data, 12 did

not respond and 7 authors sent the missing data. The meta-analyses were performed in STATA using the commands 'meta set' and 'meta summarize'. Heterogeneity was evaluated using the Higgins  $I^2$  statistic, which describes the percentage of variation caused by heterogeneity instead of chance (Higgins et al., 2003). A Higgins  $I^2$  value  $> 25\%$  was perceived as a low level of heterogeneity,  $> 50\%$  as moderate, and  $> 75\%$  as high (Huedo-Medina et al., 2006). Due to the expected heterogeneity in the data, including diversity in definitions of relapse and recovery as well as follow-up durations, random-effects models were used (Hedges & Vevea, 1998). Potential publication bias was examined by Egger's test (Egger et al., 1997), which corrects for small sample sizes. The output of the meta-analyses included: SMD, 95% confidence intervals, and heterogeneity statistics.

### 2.3 | Quality assessment

To assess risk of bias of included studies, the National Institute of Health's Study Quality Assessment Tool (National Heart, Lung, and Blood Institute (NHLBI), 2014) was used. This tool provides the possibility to examine diverse study types, and has shown relevancy for assessing predictive studies in the ED field (Glashouwer et al., 2019; Khalsa et al., 2017). The NHLBI tool does not generate a total score but guides researchers in the assessment of topics that contribute to study relevancy. It was decided to exclude irrelevant questions from the tool

(Glashouwer et al., 2019). Two reviewers (MSOL and ER) rated each study independently by answering the tools questions and disagreements were discussed. Besides completing the tool, we assessed the number of examined predictors in every study, whether people showed recovery according to the study's definition before experiencing relapse, and how recovery and relapse were defined in every study. A combination of the NHLBI tool and the additional items was used to determine overall study quality. Quality was deemed good if the definitions were clear, the study assessed whether people showed recovery before relapse, and a combination of the tool and the number of predictors did not result in more than two negative values. Studies were rated as fair if the definitions or the experience of recovery before relapse were positive and the combination of the tool and predictors did not reach above three negative values. Quality was deemed poor when the definitions and recovery before relapse were negative and the combination had four or more negative values.

## 3 | RESULTS

### 3.1 | Study characteristics

The PRISMA flowchart summarizes the stages of the current review in Figure 1. The first database search identified 20,206 studies, after duplicate removal 6,145 studies were left of which 125 were excluded

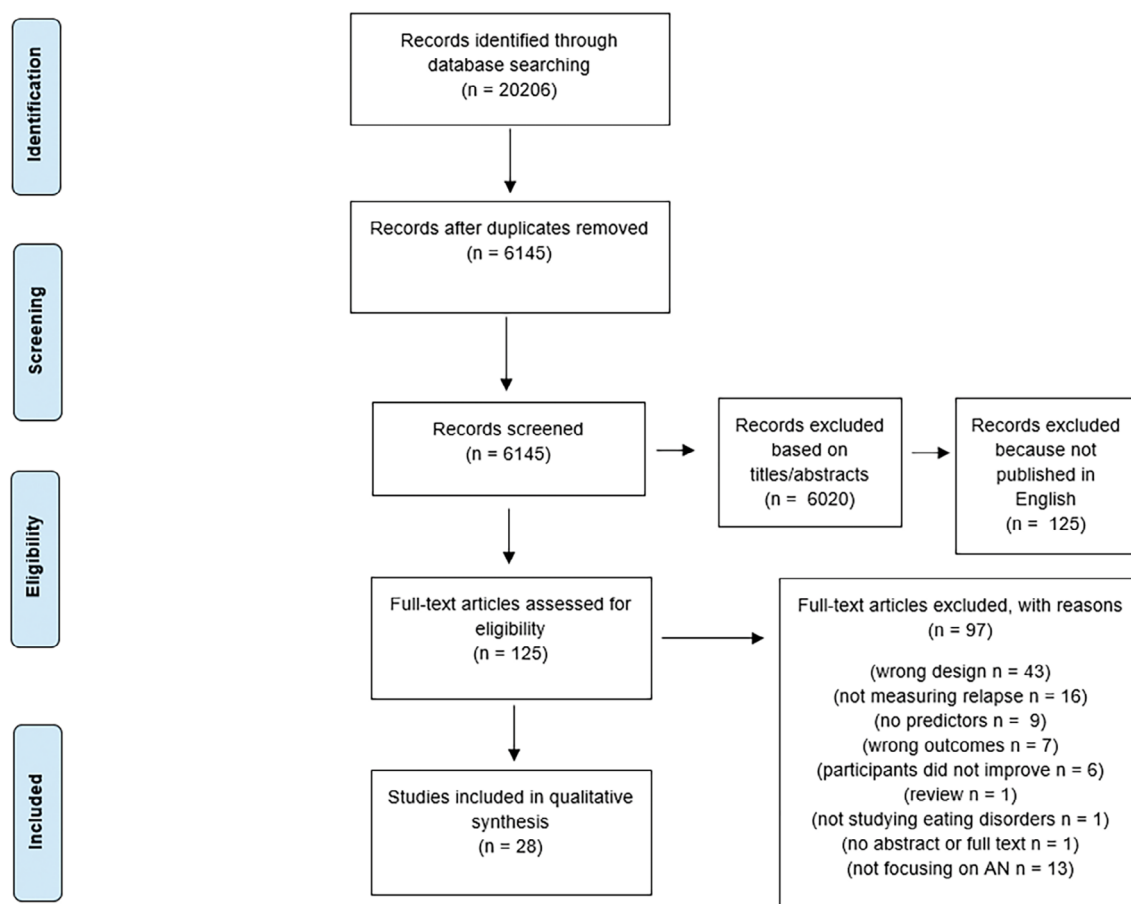


FIGURE 1 PRISMA flow diagram.

because they were not English (see Appendix S2). Titles and abstracts were screened, resulting in 125 studies eligible for full text review. Another 84 studies were excluded: for 43 studies the study design disabled the identification of predictors, 16 did not measure relapse, nine studies did not focus on predictors, seven did not focus on relapse as outcome, six studies did not show an improvement in AN symptoms before relapse, one was a review, one article did not study EDs, and for one study we were not able to find an abstract or full text. Additionally, 13 studies that did not focus on AN were excluded, resulting in 26 unique studies and two meeting abstracts that met inclusion criteria (we refer to this as 28 studies).

Table 1 summarizes the contents of the 28 included studies that were conducted between the years 1988 and 2023. The mean sample size was  $N = 91$  (range between  $N = 20$  and  $N = 234$ ) and participants' age ranged between 9 and 45 years old. In total, 21 studies focused completely on AN, four studies examined AN in combination with Bulimia Nervosa and three studies examined AN in combination with Bulimia Nervosa, Eating Disorder Not Otherwise Specified, and/or Binge Eating Disorder. The overall study quality was determined by the NHLBI tool and additional questions, showing an inter rater agreement of Cohen's kappa = .84. The outcomes for the quality assessment were poor for nine studies (32%), fair for 16 studies (57%), and good for three studies (11%). The absence of a clear recovery and/or relapse definition, the inclusion of too many predictors in relation to the sample size, and the absence of a power analysis most often contributed to the low quality of included studies.

### 3.2 | Definitions of relapse and recovery

As can be seen in Table 1, studies used different relapse and recovery definitions for AN. Relapse was frequently defined according to the Morgan–Russell (Morgan & Russell, 1975) category “poor” (Bodell & Mayer, 2011; Deter & Hertog, 1994; El Ghoch et al., 2016; Fichter et al., 2006), as rehospitalization/readmission (Avnon et al., 2018; Castro et al., 2004; Garber et al., 2015; Gualandi et al., 2017; Hetman et al., 2017; Kahn et al., 2020; Marzola et al., 2021; McCabe, 2008; McCormick et al., 2009; Pruccoli et al., 2023; Sebastian et al., 2019; Shachar-Lavie et al., 2022; Vall & Wade, 2017), according to DSM criteria (Berends et al., 2016; Castellini et al., 2011; MacDonald et al., 2022; Richard et al., 2005) or by psychiatric status rating (Herzog et al., 1999; Keel et al., 2005). Although most definitions were weight based, some also focused on the occurrence of bingeing and/or purging behaviors (Carter et al., 2004, 2012; Edwin et al., 1988).

To define recovery in AN, the majority of studies focused on BMI (Avnon et al., 2018; Berends et al., 2016; Bodell & Mayer, 2011; Carter et al., 2004, 2012; Castro et al., 2004; Edwin et al., 1988; El Ghoch et al., 2016; Hetman et al., 2017; Kahn et al., 2020; Kim et al., 2021; McCormick et al., 2009; Shachar-Lavie et al., 2022). Three studies determined recovery as the absence of symptoms according to DSM criteria (Castellini et al., 2011; MacDonald et al., 2022; Richard et al., 2005), two used the Morgan-Russell criteria “good” (Deter & Hertog, 1994; Fichter et al., 2006) and two used scores on

psychiatric status rating (Herzog et al., 1999; Keel et al., 2005). The current study acknowledged the high variability in relapse and recovery definitions. Besides this diversity, no value was added to the definitions and all studies were taken into account.

### 3.3 | Follow-up period and rate of relapse

The total average follow-up duration was 31 months (range 3–144 months). On average, rate of relapse was 37% for AN, relapse chances were similar for short (38%,  $\leq 1$  year) and long term follow-up periods (38%,  $> 5$  years follow-up), and slightly lower for intermediate (31%, 1–5 years) follow-up.

### 3.4 | Predictors of relapse

The 28 included studies investigated more than 150 unique predictors of relapse in AN, predictors examined in two or more studies are mentioned in Table 2. Related to this large number of predictors, only those predictors which were significantly replicated at least once, and/or were investigated in a minimum of five different studies (Jackson & Turner, 2017) will be discussed in the review below. The predictors were grouped into five different categories: age and sex (also including age at onset), symptoms and behaviors (severity, exercise), AN subtype and duration (AN duration, treatment duration, prior hospitalization, AN subtype), weight or weight change (pre-treatment BMI, post-treatment BMI, percentage of target weight, rate of weight gain), comorbidity (general comorbidity, obsessive compulsive disorders, anxiety, depression), and personality (self-esteem). For most investigated predictors no significant effects were found and if findings were significant they were often contradictory.

### 3.5 | Age and sex

The majority of the 16 studies that examined the predictive effect of age on relapse in AN found no effect (see Table 2). Furthermore, the reported predictive effects of age were contradictory; three studies found older age to be predictive (Berends et al., 2016; Keel et al., 2005; Pruccoli et al., 2023), whereas one found that younger age (Castro et al., 2004) was predictive. All four studies were comparable in sample size and provided treatment. However, Castro et al. (2004) included young people, while the samples of the other studies had a higher mean age, which might explain the differences in effects. As age was often included as predictor, a meta-analysis including nine studies could be performed (see Figure 2, Figure S1 and Table 3). Age, at start of treatment, was compared between a non-relapsed ( $N = 338$ ) and a relapsed group ( $N = 512$ ). A moderate level of heterogeneity (Higgins  $I^2 = 68.1\%$ ) was found and no correction for small study effects was needed (Egger's test  $p = .803$ ). This analysis showed that age was not a significant predictor for relapse (SMD =  $-.13$ , 95% CI  $[-.40$  to  $.13$ ],  $p = .323$ ). None of the five studies that examined



TABLE 1 Characteristics and findings of reviewed studies.

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex <sup>a</sup>	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
1. Avnon et al. (2018)	44	AN	Inpatient	Treatment has two phases: an intensive inpatient phase followed by a short period of a half-day outpatient program to promote patients' reintegration into the community.	M = 14.8 years	87% female	Israel	NR	Logistic regression	Discharge from treatment (target weight and attained 2 weeks outpatient treatment)	Rehospitalization	34%	12	Fair
2. Berends et al. (2016)	83	AN	In and out patient	Treatment focusses on three areas: (1) eating habits, body weight, and body image; (2) psychological aspects of functioning, such as self-esteem, perfectionism, and traumas (3) social functioning within the family system and in society.	M = 17.9 years	All female	The Netherlands	NR	Multi-variate	Successfully treatment completion, weight restored	BMI < 18.5 for adults and SD BMI < -1 for adolescents, full recurrence of symptoms of AN according to DSM-IV criteria	11%	18	Poor
3. Boddell & Mayer (2011)	22	AN	Inpatient	Structured behavioral program aimed at normalizing weight and eating behavior.	18–45 years	All female	America	NR	Binary logistic regression	90% weight restored as defined by the Metropolitan Life	Morgan-Russell criteria: good, fair, and poor. No DSM-IV criteria for AN for a minimum of 8 weeks	50%	12	Fair
4. Carter et al. (2004)	51	AN	Inpatient	Intensive group therapy program that is primarily directed at the normalization of eating and the restoration of body weight.	M = 26.9 years	All female	Canada	85% = Caucasian 10% = Asian 5% = Native Canadian Hispanic Ethnicity/SES: NR	Mantel-Cox log-rank test	Weight-restored (BMI > 20) following inpatient treatment	BMI below 17.5 for a period of at least 3 consecutive months and binge eating and/or purging at least weekly for a 3-month period	35%	18	Good
5. Carter et al. (2012)	100	AN	Inpatient	Specialized hospital-based program operated by a multi-disciplinary team that focuses on nutritional rehabilitation through provision of supervised meals, weight restoration to a BMI of 20, eradication of binge eating and purging	M = 25.4 years	95% female	Canada	94% = Caucasian 4% = Asian 2% = African-Canadian or East Indian Ethnicity/SES: NR	Multi-variate	Weight restored BMI of 20 for 2 weeks, max one binge purge episode over the previous 28 days	BMI less than or equal to 17.5 for 3 consecutive months or at least one binge purge episode a week for 3 consecutive months	4% at 3 months 23% at 6 months 36% at 9 months 41% at 12 months	12	Good

**TABLE 1** (Continued)

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex <sup>a</sup>	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
6. Castellini et al. (2011)	AN = 165	AN, BN, BED, EDNOS	Outpatient	Patients were provided an individual cognitive behavioral therapy.	M = 27 years	91% female	Italy	NR	Stepwise logistic regression	Not fulfilling the DSM-IV or DSM-V criteria for any ED	Return to a full syndromal criteria (DSM-IV and DSM-V) after a period of remission	26%	72	Fair
7. Castro et al. (2004)	101	AN	Inpatient and outpatient	Treatment during admission is based on a multidisciplinary approach that combines biologic management, nutritional rehabilitation, a behavioral program to improve eating patterns and weight, individual and group cognitive treatment, and individual and group parent counseling.	11–19 years	90% female	Spain	NR	Stepwise logistic regression	Weight-recovered (BMI-values within normal range)	Readmission based on inability to maintain weight, onset of bingeing or vomiting, conflictive family situation	25%	12	Poor
8. Deter & Hertog (1994)	84	AN	Inpatient	Treatment consisted of psychodynamic individual therapy and, if possible, systemic family therapy in the hospital setting over a mean period of 3 months.	NR	All female	Germany	SES: Morgan-Russell scores 3–7.6, no interpretation of scale provided Ethnicity/race: NR	Stepwise multiple regression	Morgan-Russell category: good	Morgan-Russell category: poor (average body weight 115% and amenorrhea)	22%	144	Fair
9. Edwin et al. (1988)	67	AN	Inpatient	Patients were referred for psychological testing as a routine part of admission evaluation.	M = 21 years	All female	America	NR	NR	Normal body weight at discharge inpatient treatment	Patients classified as not successful (definition maintained greater than 80% of matched	61%	60	Fair

(Continues)



TABLE 1 (Continued)

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex <sup>a</sup>	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
10. El Ghoch et al. (2016)	54	AN	Inpatient	Enhanced form of cognitive behavioral therapy for eating disorders, in which patients received dietician-assisted eating until they reached a BMI > 18.5 kg/m <sup>2</sup> .	Relapse M = 24.9 years Non-relapse M = 25.6 years	All female	Italy	NR	Binary logistic regression	BMI > 18.5 kg/m <sup>2</sup> for 4 or more weeks before discharge from the inpatient program	Morgan-Russell category poor  mean weight and bingeing less than once per week, and working or attending school at least part-time + socializing)	52%	12	Good
11. Fichter et al. (2006)	103	AN	Inpatient	Patients received broad-spectrum behavioral treatment that consisted of (1) antecedent events and (eating) behavior; (2) behavioral program to induce weight gain through reinforcement (3) nutritional counseling and psychoeducation; (4) training of interoceptive and emotional perception using body-oriented and emotion-provoking approaches; (5) training of social skills and emotional expression (6) cognitive-behavior therapy; (7) activation of one's own responsibility and initiative (8) relapse prevention	M = 25 years	All female	Germany	SES: Morgan-Russell scores range between 7.1 and 7.3 Ethnicity/race: NR	Stepwise logistic regression	Morgan-Russell outcome was defined as good if body weight maintained in the normal range (a score of 10–12) and regular menstrual cycles (a score of 10–12)	Morgan-Russell outcome was defined as poor based on weight and menstruation: values below 6 points (weight) or 4 points (menstruation) = poor outcome	40%	144	Fair
12. Garber et al. (2015) Meeting abstract	54	AN	Inpatient	Refeeding with three meals and three snacks was prescribed in a range of calories from 800	M = 16.2 years	NR	America	NR	Survival analysis	NR	Rehospitalization	-	12	Fair

**TABLE 1** (Continued)

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex <sup>a</sup>	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
13. Gualandi et al. (2017)	61	AN	Inpatient	NR	9–18 years old	NR	Italy	NR	Mann Whitney test	NR	Rehospitalization	21%	12	Poor
14. Herzog et al. (1999)	AN = 136	AN, BN	Unclear	NR	M = 25 years	All female	America	NR	Cox regression	Absence of symptoms or the presence of only residual symptoms for at least 8 consecutive weeks (psychological status rating (PSR) 1 or 2)	Return of full symptoms (PSR 5 or 6) for at least 8 weeks following a state of full recovery (PSR 1 or 2)	40%	90	Fair
15. Hetman et al. (2017)	51	AN	Inpatient	The treatment program has two phases: an intensive inpatient phase and a shorter phase of a half-day outpatient program. This two-phase structure emphasizes reintegration into the community.	M = 15 years	90% female	Israel	SES: High (3.9%) Mid (70.6%) Low (2%) Unknown (23.3%) Ethnicity/race: NR	Hierarchical binary logistic regressions	The target weight is the weight goal for discharge, and defined as a range of 2 kg from a minimal recommended weight	Rehospitalization	37%	12	Fair
16. Kahn et al. (2020)	56	AN	Inpatient	The treatment program promotes a comprehensive multidisciplinary approach directed to normalization of eating habits, weight restoration and healthy functioning. Patients receive intensive individual and group therapy, family therapy, nutritional management, and participate in daily schooling at the unit.	M = 15.4 years	88% female	Israel	High = 8.9% Mid-high = 35.7% Mid = 35.7% Low = 8.9% Ethnicity/race: NR	Generalized estimation equations	Attained expected body weight and consistent maintenance of it for 2 consecutive weeks	Rehospitalization	38%	12	Fair
17. Keel et al. (2005)	AN = 136	AN, BN	Inpatient and out-patient	NR	AN = 18.7–22.5 years	All female	America	NR	Multivariate	Period of 8 consecutive weeks in which no or minimal symptoms were present (PSR score 1 or 2)	Return to full syndromal criteria after a period of remission (PSR score of 5 or 6)	36%	12	Fair

(Continues)

TABLE 1 (Continued)

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex*	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
18. Kim et al. (2021)	41	AN	Inpatient	Treatment aims to normalize weight and eating behaviors until the patient can be weight-restored to at least 90% ideal body weight	M = 25 years	All female	America	NR	Binary logistic regression	BMI > 18.5 kg/m <sup>2</sup> for 8 weeks leading up to the 1-year follow-up	No successful weight maintenance	51%	12	Fair
19. MacDonald et al. (2022)	AN = 33	AN, BN	Inpatient and out-patient	Inpatient treatment involved hospital admission, and day treatment spanned 5 days per week for 7–8 h per day. Both services involved group therapy-based, action-oriented multimodal treatments, which included cognitive behavioral therapy, supervised meals, a contained environment, and staff support. Patients were expected to improve eating and interrupt behavioral ED symptoms quickly, gain weight when necessary, and use their time to maintain/practice their behavior changes, continue with food exposures, and work on relevant issues in groups	M = 28.5 years	Gender: 97% female 2% men 1% trans-gender	Canada	Caucasian: 85% African-Canadian: 3% Asian: 5% Latino/a: 1% Mixed: 3% Other: 3% Ethnicity/SES: NR	Multiple linear regression	Maximum of three binge eating and/or purging episodes in the last month of intensive treatment and BMI greater than 18.5 at discharge	The first month of follow-up in which binge eating/ purging was above threshold $\geq 4$ episodes per month or BMI < 18.5 according to DSM-5	45%	12	Fair
20. Marzola et al. (2021)	170	AN	Inpatient	Re-establish patients' clinical life-threatening conditions, work with the patients to foster their motivation for the subsequent therapeutic steps, deliver structured daily sessions on symptom management focusing on diet and body image	M = 24.8 years	NR	Italy	NR	Cox regression	NR	Rehospitalization/ readmission	39%	36	Fair

**TABLE 1** (Continued)

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex <sup>a</sup>	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
21. McCabe (2008)	147	AN	Inpatient	concerns, work psychologically to understand the possible causes of those factors that led to an emergency admission, and provide families with psychoeducation. The treatment program is highly structured, with an emphasis on shaping and reinforcement of appropriate behaviors. Weight gain is a primary treatment objective; there are five scheduled meals per day. Psychological interventions are delivered primarily in a group therapy format. Groups are intended to provide psychoeducation to patients regarding the various medical, psychological, and interpersonal aspects of eating disorders and to teach the skills designed to enhance the self-management of their disorder	M = 21.1 years	95% female	America	Caucasian 98% Other: NR Ethnicity/SES: NR	Logistic regression and survival analysis	NR	Readmission	27%	36	Fair
22. McCormick et al. (2009)	20	AN	Inpatient	NR	M = 27.6 years	All female	America	NR	Repeated measures analysis	Weight recovery at discharge	1-year post-hospitalization according to two criteria: (a) readmission to a partial or inpatient unit for ED related treatment; and/or (b) a drop in BMI below 17.5 kg/m <sup>2</sup>	35%	12	Fair

(Continues)

TABLE 1 (Continued)

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex*	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
23. McFarlane et al. (2008)	AN = 16	AN, BN, EDNOS	Out-patient	Cognitive-behavioral group therapy program that is aimed primarily at normalizing eating, interrupting eating disorder symptoms, and restoring body weight when needed.	M = 29.8 years	All female	Canada	NR	Cox regression analysis	No more than two binge eating and/or purging episodes per month for 2 months, achieving and maintaining a BMI of 19.5 or higher for 2 months, and adhering to a prescribed normalized meal plan for 2 months	More than 8 binge eating and/or purging episodes per month for three consecutive months, or BMI below 18.5 for 3 consecutive months, or eating less than 1000 cal per day for 3 consecutive months	6 months 38% 1 year 41% 18 months 48%	18	Poor
24. Prucelli et al. (2023)	234	AN	Inpatient	A multidisciplinary psychological, psycho-pharmacological, and nutritional program consisting of multiple consultations by specialized neuropsychiatrists/psychologists and a dietitian. During their stay in the day hospital program, the patient consumed assisted meals, together with specialized operators.	M = 15.6 years	83% female	Italy	NR	Kaplan-Meier analysis and Cox regression model	NR	Rehospitalization	37%	12	Fair
25. Richard et al. (2005)	AN = 233	AN, BN	Inpatient	NR	M = 24.8 years	96% female	Germany	NR	Survival analysis	Recovery/remission = symptom free status for 12 months	Returning to full symptomatic status (DSMIII) after being in partial or full remission	33%	30	Poor
26. Sebastian et al. (2019)	AN = 54	AN, BN, EDNOS	Inpatient	NR	M = 14.8 years	83% female	America	NR	Multivariate	NR	Readmission	21%	16	Poor
27. Shachar-Lave et al. (2022)	202	AN, Atypical AN	Inpatient	A multidisciplinary treatment including individual and group therapies. The nutritional program is based on an individual daily diet designed by a clinical nutritionist and the patient. Group-supervised meals are a crucial	AN: M = 15.1 years, atypical AN; M = 16.1 years	83% female	Israel	NR	Three-way log-linear analysis	Achievement of target weight	Rehospitalization	39%	12	Fair

**TABLE 1** (Continued)

Author (year)	N	EDs	In/out patient treatment	Treatment	Age	Sex <sup>a</sup>	Study was conducted in country	Race, ethnicity, socio-economic status (SES)	Test type	Definition recovery	Definition relapse	Relapse rate	Max. follow-up in months	Quality analysis
28. Vall & Wade (2017)	40	AN	Inpatient	Beginning with management of medical instability. The program then graduates through the withdrawal of enteral feeds to a full oral feeding, establishing normal eating with families by gate passes as part of the transition from home to hospital. Patients and their families attend a family meeting with the treating team once a week.	M = 15.4 years	95% females	Australia	92.5% Caucasian Other: NR Ethnicity/SES: NR	Binary/logistic regression	Attaining medical stability, meeting a safe minimum weight, and showing evidence of behavioral change	Readmission	30%	3	Poor

Abbreviation: NR, not reported.  
<sup>a</sup>Sex is standardly mentioned as sex assigned at birth, if gender was reported this was specifically mentioned.

**TABLE 2** Complete overview of relapse predictors in AN, only includes predictors that were examined in two or more studies.

Predictor	Amount of studies in which predictor was significant	Amount of studies in which predictor was non-significant	Significant	Non-significant
<b>Age and sex</b>				
Age	4	12	Berends et al. (2016): $\beta = 2.34$ , $p = .04$ Castro et al. (2004): $\beta = -.36$ , $p = .007$ Keel et al. (2005): $\beta = .06$ , $p < .11$ Pruccoli et al. (2023): $\beta = .52$ , $p = .020$	Carter et al. (2004), Edwin et al. (1988), Gualandi et al. (2017), Herzog et al. (1999), Hetman et al. (2017), Kim et al. (2021), Kahn et al. (2020), Marzola et al. (2021), McCabe (2008), Richard et al. (2005), Sebastian et al. (2019), Vall and Wade (2017)
Age at onset	-	5	-	Berends et al. (2016), Carter et al. (2004), Castellini et al. (2011), Edwin et al. (1988), Fichter et al. (2006)
Sex Not discussed <sup>a</sup>	1	3	Kahn et al. (2020): $\chi^2 = 3.91$ , $p = .04$ (females higher chance R)	Castellini et al. (2011), Hetman et al. (2017), Sebastian et al. (2019)
<b>Symptoms and behaviors</b>				
Binging Not discussed	1	3	Castellini et al. (2011): HR = 1.16, 99% CI [1.06–1.27] (higher for R)	Carter et al. (2004), Herzog et al. (1999), McFarlane et al. (2008)
ED severity	4	6	Kahn et al. (2020), Eating Disorder Examination Questionnaire (EDEQ): $\beta = -.86$ , $p = .008$ (lower for N-R) Richard et al. (2005): Eating Disorder Inventory (EDI): $\beta = -1.02$ , $p < .05$ (lower for R) Castro et al. (2004): Eating Attitudes Test (EAT) $\beta = .026$ , $p = .020$ (higher for R) McCabe (2008)	Berends et al. (2016) EDEQ: Castellini et al. (2011), Vall and Wade (2017) EDE: Hetman et al. (2017) EDI: Fichter et al. (2006), Gualandi et al. (2017)
Exercise	2	1	Carter et al. (2004): EDE interview, $\chi^2 = 4.87$ , $p = .027$ (higher for R) Fichter et al. (2006): $t = 2.51$ , $p < .05$ (higher for R)	Vall and Wade (2017)
Vomiting/purging Not discussed	1	4	Deter and Hertog (1994): $p = .03$ (lower for R)	Carter et al. (2004), Fichter et al. (2006), Herzog et al. (1999), Vall and Wade (2017)
<b>ED expression</b>				
Duration of ED	4	10	Fichter et al. (2006): $t = 3.51$ , $p < .01$ (longer for R) Deter and Hertog (1994): $p = .005$ (longer for R) Marzola et al. (2021): $p = .042$ (shorter for R) Richard et al. (2005): $\beta = .52$ , $p < .05$	Berends et al. (2016), Carter et al. (2004), Castro et al. (2004), Edwin et al. (1988), Hetman et al. (2017), Kim et al. (2021), McCabe (2008), McCormick et al. (2009), Sebastian et al. (2019), Vall and Wade (2017)



**TABLE 2** (Continued)

ED subtype	1	7	Carter et al. (2012): HR = 2.87, $p = .003$ (BP subtype of AN at admission predicted R)	Berends et al. (2016), Carter et al. (2004), Fichter et al. (2006), Herzog et al. (1999), Hetman et al. (2017), Kahn et al. (2020), McCabe (2008)
Prior hospitalization <i>Not discussed</i>	1	3	Hetman et al. (2017): $\chi^2 = 5.78$ , $p < .05$ (longer for R)	Herzog et al. (1999), Keel et al. (2005), McCabe (2008)
Treatment duration	2		Berends et al. (2016): $p = .007$ (longer duration of inpatient predicted R) Fichter et al. (2006): $t = 2.8$ , $p < .01$ (longer duration for R)	-
Body dissatisfaction <i>Not discussed</i>	1	1	McCabe (2008): body dissatisfaction, $\beta = .056$ , $p < .05$	Marzola et al. (2021)
<b>Weight change</b>				
BMI admission	0	10	-	Avnon et al. (2018), Berends et al. (2018), Carter et al. (2004), Carter et al. (2012), Castellini et al. (2011), Fichter et al. (2006), Hetman et al. (2017), Kim et al. (2021), McCormick et al. (2009), Vall and Wade (2017)
BMI discharge	1	7	El Ghoch et al. (2016): $t = 3.86$ , $p < .001$ (higher for R)	Avnon et al. (2018), Berends et al. (2016), Carter et al. (2004), Carter et al. (2012), Gualandi et al. (2017), Hetman et al. (2017), McCabe (2008)
Percentage of target weight	3	2	Avnon et al. (2018): $t = 3.25$ , $p < .01$ (lower for R) Hetman et al. (2017): $t = 2.47$ , $p < .05$ (lower for R) Shachar-Lavie et al. (2022): $\chi^2 = 11.94$ , $p = .001$ (lower for R)	McCabe (2008), Richard et al. (2005)
Rate of weight gain	2	1	Castro et al. (2004): $B = -.008$ , $p = .009$ Sebastian et al. (2019): 1.81 odds, $p = .087$ (increase in weight gain for R)	Carter et al. (2012)
<b>Comorbidity</b>				
Anxiety disorders	2	2	Fichter et al. (2006): $t = 2.93$ , $p < .01$ (higher R) Gualandi et al. (2017)	Berends et al. (2016), Marzola et al. (2021)
Depression	3	10	Fichter et al. (2006): $t = 2.99$ , $p < .05$ (higher for R) Kahn et al. (2020): $\beta = -6.64$ , $p = .03$ (lower for N-R) Gualandi et al. (2017)	Berends et al. (2016), Carter et al. (2004, 2012), Herzog et al. (1999), Hetman et al. (2017), Keel et al. (2005), Marzola et al. (2021), McCabe (2008), McCormick et al. (2009), McFarlane et al. (2008)

(Continues)

TABLE 2 (Continued)

General comorbidity	2	3	Deter and Hertog (1994): $p < .08$ (higher for R) Carter et al. (2012): CI [.95–2.12], $p < .09$ (higher scores at discharge for R)	Berends et al. (2016), Castellini et al. (2011), Sebastian et al. (2019)
OCD	3	2	Carter et al. (2004): Higher washing Compulsions CI [.6–1.0], $p = .038$ ; and Rumination CI [1.0–1.5], $p = .030$ for R Carter et al. (2012): CI [1.02–1.13], $p = .012$ (higher levels for R) Fichter et al. (2006): $t = 3.08$ ; $p = .01$ (higher for R)	Berends et al. (2016), Castellini et al. (2011)
Personality disorders <i>Not discussed</i>	-	2	-	Berends et al. (2016), Castellini et al. (2011)
Suicide <i>Not discussed</i>	1	2	Carter et al. (2004): $\chi^2 = 4.85$ , $p = .028$ (a history of suicide for R)	Herzog et al. (1999), Hetman et al. (2017)
Personality Motivation <i>Not discussed</i>	1	2	Carter et al. (2012): HR = .91, $p = .010$ (reduced motivation for R)	Richard et al. (2005), Vall and Wade (2017)
Self-esteem	3	1	Carter et al. (2012): CI [.88–1.00], $p = .044$ (lower for R) Fichter et al. (2006), Structured Interview for Anorexia and Bulimia (SIAB): $t = 3.81$ , $p < .01$ (lower for R) McFarlane et al. (2008): $\beta = .051$ , $p < .05$	Carter et al. (2004)

Abbreviation: CI, confidence interval; HR, Hazard ratio; N-R, predicted non-relapse; R, predicted relapse.

<sup>a</sup>Not discussed: When a factor is followed by the phrase: 'not discussed', this means that a predictor is not discussed in the article.

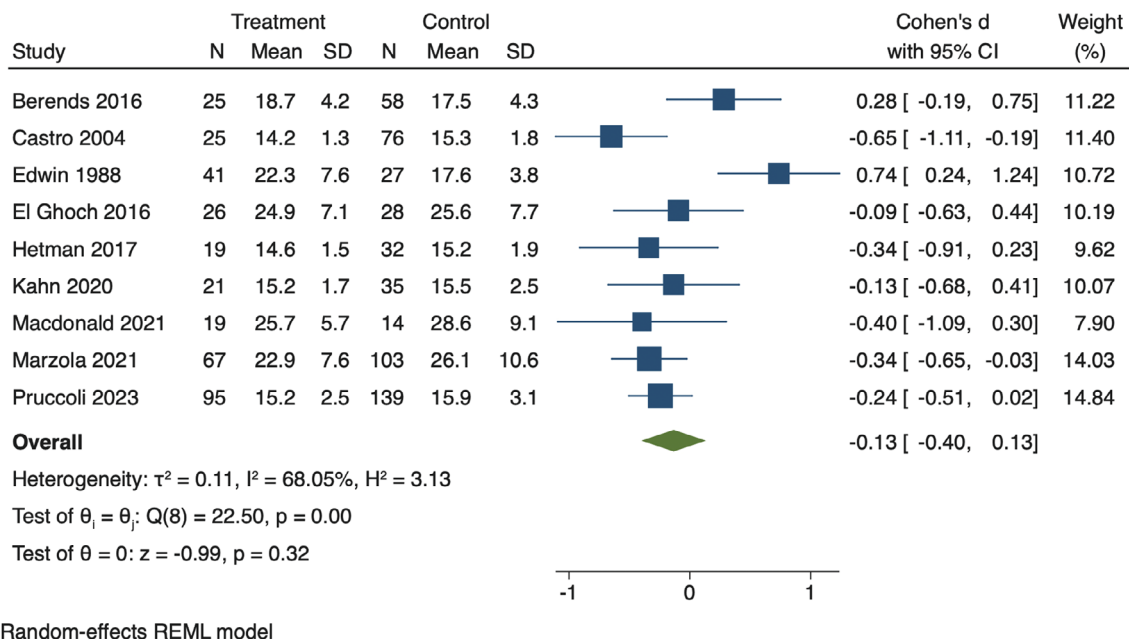
age at onset found it to be a significant predictor (Berends et al., 2016; Carter et al., 2004; Castellini et al., 2011; Edwin et al., 1988; Fichter et al., 2006). Overall this indicated that age is not a good predictor of relapse in AN.

### 3.6 | Symptoms and behaviors

For AN severity, only four studies found severity to be predictive of relapse, while most studies did not (see Table 2). Findings showed discrepant results: both higher (Castro et al., 2004; Kahn et al., 2020; McCabe, 2008) and lower (Richard et al., 2005) levels of severity predicted relapse. The contradictory results might be explained by the more stringent recovery definition of Richard et al. (2005) including a full symptom free status for 12 months, than the definitions used in the studies of Castro et al. (2004) and Kahn et al. (2020) which determined recovery by weight status. Another explanation could be

derived from the different severity measurements across these studies (Eating Disorder Inventory, Eating Disorder Examination Questionnaire, Eating Disorder Examination). As suggested by Richard et al. (2005) the EDI might not provide enough accurate information about the severity of illness in terms of outcome, therefore lower symptoms might have incorrectly predicted relapse.

Two studies found higher levels of exercise predictive of relapse in AN (Carter et al., 2004; Fichter et al., 2006), whereas one study did not find an effect (Vall & Wade, 2017). The Vall and Wade (2017) study had a lower mean age compared to the other two studies (15 years old vs. 25 years old). Furthermore, the follow-up duration of this study was short (3 months), in comparison to the Fichter et al. (2006) (144 months) and Carter et al. (2004) (18 months) studies. It might be that exercise only predicts relapse in the longer term, or is more important among adults. In relation to severity and exercise we cannot provide solid conclusions about the direction of effects.



**FIGURE 2** Pre-treatment age. Forest plot of the effect size of the difference in pre-treatment age between the relapsed and the non-relapsed group. The condition 'treatment' describes the relapsed group and 'control' the non-relapsed group.

**TABLE 3** Meta-analysis results for the predictors age, duration, BMI at admission, BMI at discharge and depression.

	N (relapse, non-relapse)	SMD	95% CI	Z	p
Age	850 (338, 512)	-.13	-.40; .13	-.99	.323
ED duration	527 (203, 324)	-.14	-.63; .35	-.57	.571
BMI admission	557 (217, 340)	-.02	-.20; .16	-.23	.822
BMI discharge	393 (161, 232)	-.35	-.63; -.07	-2.47	.013
Depression	490 (172, 318)	.55	.21; .89	4.17	<.001

### 3.7 | Expression

The majority of studies found no effect for duration of AN (see Table 2). Contradicting results were found as both longer (Deter & Hertog, 1994; Fichter et al., 2006; Richard et al., 2005) and shorter duration (Marzola et al., 2021) predicted relapse. Marzola et al. (2021) defined relapse as rehospitalization, while in the other three studies relapse was based on DSM, PSR or Morgan–Russell criteria, which might have resulted in a different effect. Furthermore, it must be mentioned that Marzola et al. (2021) stated that although shorter duration predicted relapse, results were not maintained when including confounding variables. We performed a meta-analysis for duration with six studies (see Figure 3, Figure S2, and Table 3), with  $N = 203$  participants in the relapsed group and  $N = 324$  in the non-relapsed group. Higgins  $I^2$  showed high variability across these studies (85.1%). No correction for small effects had to be conducted (Egger's test  $p = .067$ ). Based on the meta-analysis, AN duration had no effect on relapse (SMD =  $-.14$ , 95% CI [ $-.63$  to  $.35$ ],  $p = .571$ ).

Besides duration of AN, treatment duration was also examined, with two studies showing that longer duration predicted relapse (Berends et al., 2016; Fichter et al., 2006) although with different relapse rates

(40% and 11%, respectively) and mean age (17.9 instead of 25 years). For the predictor AN subtype, one study showed that people with a diagnosis of AN binge-purge subtype at admission had an increased chance of relapse (Carter et al., 2012) while others (Berends et al., 2016; Carter et al., 2004; Fichter et al., 2006; Herzog et al., 1999; Hetman et al., 2017; Kahn et al., 2020; McCabe, 2008) did not find significant results for AN subtype. The occurrence of effects in Carter et al. (2012) might be explained by the specific focus on the eradication of binge and purging symptoms into treatment, while the non-significant studies did not mention this. For this category we can conclude that insufficient evidence is available for AN subtype and duration to predict relapse, while suggestive effects were found for longer treatment duration to predict AN relapse.

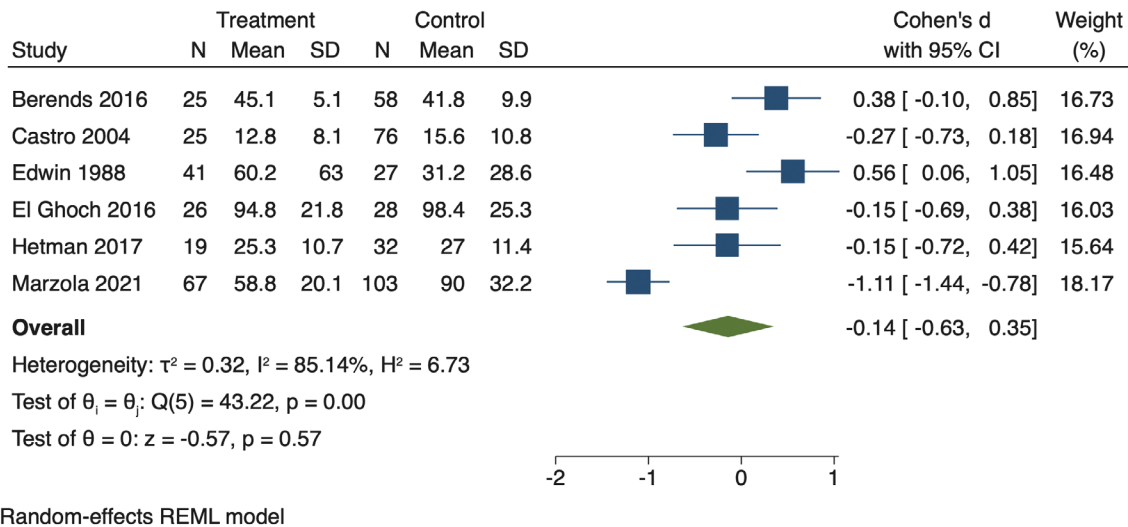
### 3.8 | Weight or weight change

Both pre-treatment BMI and post-treatment BMI have been examined. No study found that pre-treatment BMI had a predictive effect on relapse (see Table 2). Higher post-treatment BMI predicted AN relapse in one study (El Ghoch et al., 2016) while most studies did not find an effect (see Table 2). Interestingly, El Ghoch et al. (2016)

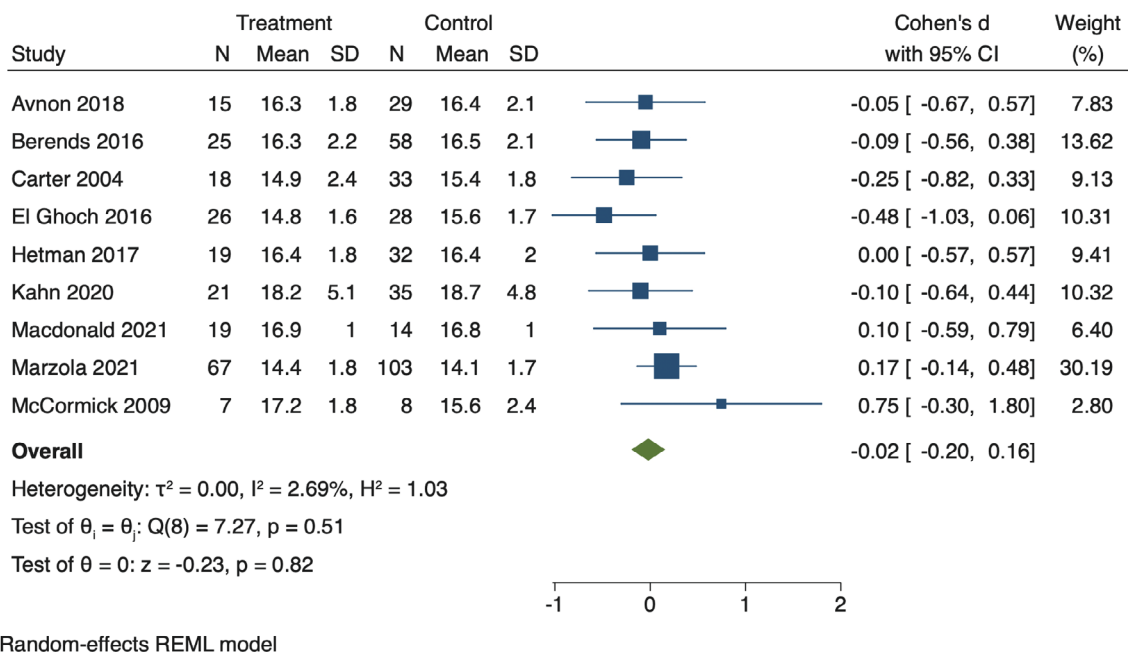
found a relatively high percentage of relapse (52%) in comparison to the other studies, which might have increased the chance of finding an effect. Both pre- and post BMI measurements were examined in a meta-analysis. Nine studies were entered in a meta-analysis for pre-treatment BMI, with  $N = 217$  participants in the relapsed group and  $N = 340$  in the non-relapsed group. Higgins  $I^2$  was 2.7% for pre-treatment BMI and no correction for small studies was needed as Egger's test was not significant ( $p = .923$ ). Non-significant effects for pre-treatment BMI (SMD =  $-.02$ , 95% CI [ $-.20$  to  $.16$ ],  $p = .822$ ), Figure 4, Figure S3, and Table 3) were found. For post-treatment

BMI again nine studies were entered in a meta-analysis, with  $N = 161$  participants in the relapsed group and  $N = 232$  in the non-relapsed group. The level of heterogeneity was relatively low for post-treatment BMI (Higgins  $I^2 = 41%$ ) and Egger's test was not significant ( $p = .371$ ). Post-treatment BMI did predict relapse (SMD =  $-.35$ , 95% CI [ $-.63$  to  $-.07$ ],  $p = .013$ , Figure 5, Figure S4 and Table 3).

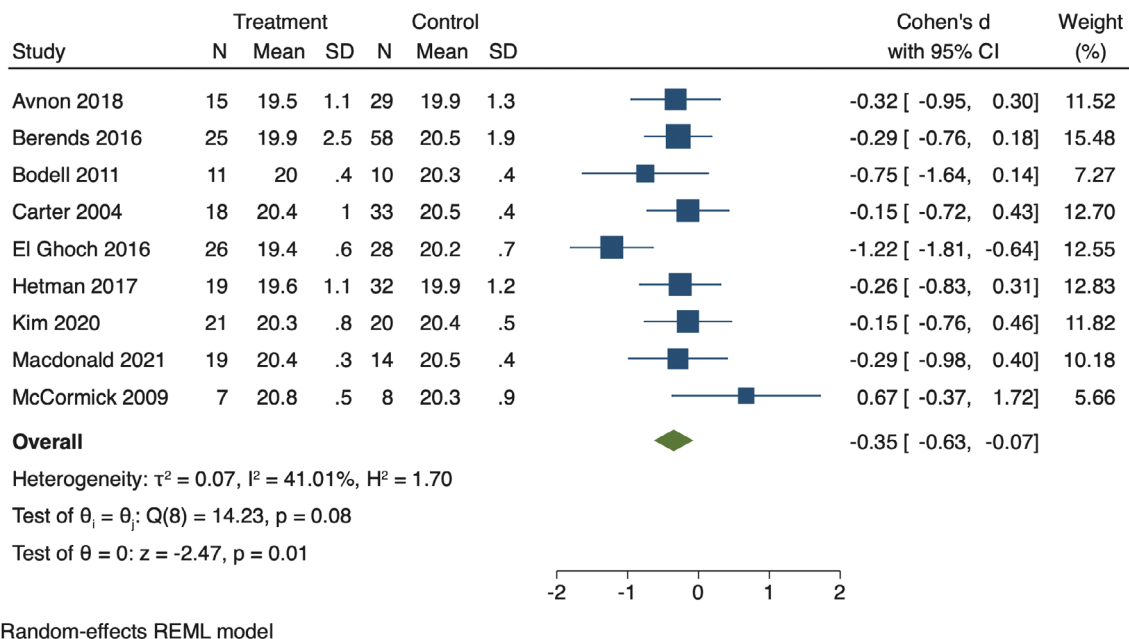
In addition to BMI, percentage of target weight has also been examined. Three studies found lower levels of target weight to predict relapse (Avnon et al., 2018; Hetman et al., 2017; Shachar-Lavie



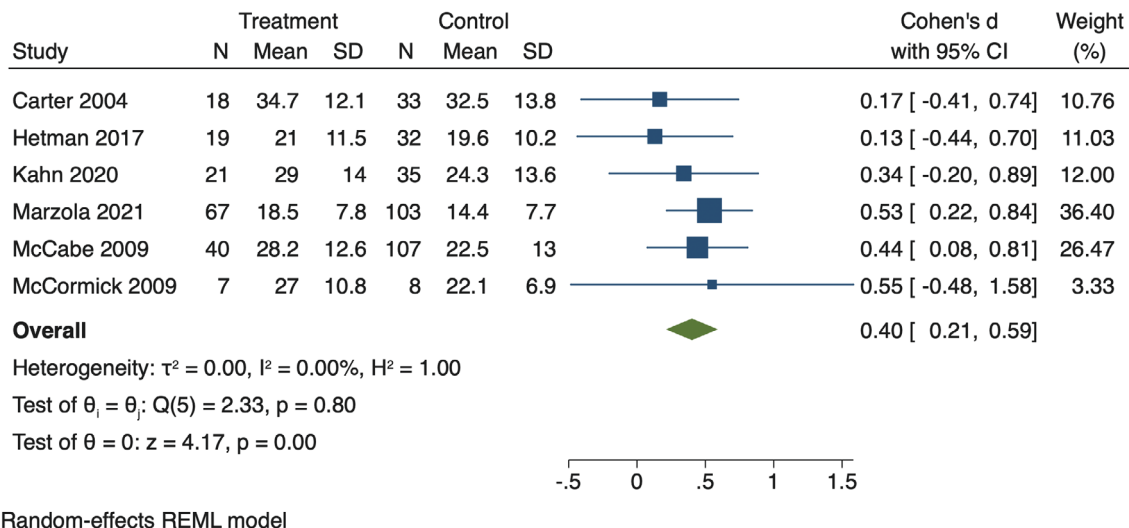
**FIGURE 3** AN duration. Forest plot of the effect size of the difference in AN duration (in months) between the relapsed and the non-relapsed group.



**FIGURE 4** Pre-treatment BMI. Forest plot of the effect size of the difference in pre-treatment BMI between the relapsed and the non-relapsed group.



**FIGURE 5** Post-treatment BMI. Forest plot of the effect size of the difference in post-treatment BMI between the relapsed and the non-relapsed group.



**FIGURE 6** Depression. Forest plot of the effect size of the difference in pre-treatment depression scores between the relapsed and the non-relapsed group.

et al., 2022) and two studies did not find an effect (McCabe, 2008; Richard et al., 2005). This discrepancy might be explained by a difference in age of the samples; studies that found a significant effect focused on adolescents whereas the other studies (McCabe, 2008; Richard et al., 2005) focused on adults. This indicated that percentage of target weight might especially be of importance in relapse among adolescents. For the related predictor rate of weight gain, one study showed that lower rate of weight gain was predictive of relapse (Castro et al., 2004), one that higher rate of weight gain was predictive

(Sebastian et al., 2019) and one found no effect (Carter et al., 2012). Despite the similarities in methodology among the three studies, Castro et al. (2004) and Sebastian et al. (2019) were deemed of low quality. This was primarily attributed to the inadequate definitions and lack of clarity concerning full recovery before relapse, potentially resulting in false positive outcomes. Based on the predictors which focused on weight, our review showed that a lower percentage of target weight might predict relapse. Stronger evidence was found for BMI showing that lower post-treatment BMI predicted relapse in AN.

### 3.9 | Comorbidity

Although three studies did not find a predictive effect for general levels of comorbidity, others found that higher general levels of comorbidity predicted relapse in AN (Carter et al., 2012; Deter & Hertog, 1994). These studies differed in the relapse definitions used, based on either Morgan-Russell criteria (Carter et al., 2012; Deter & Hertog, 1994) or DSM criteria (Berends et al., 2016; Castellini et al., 2011), which were more stringent. These findings suggest that variations in definitions could lead to different predictive effects.

Three studies showed that higher levels of obsessive compulsive behaviors (Carter et al., 2004, 2012; Fichter et al., 2006) predicted relapse in AN, while two studies did not find an effect (Berends et al., 2016; Castellini et al., 2011). In the latter studies the relapse rates were lower. Furthermore, the Berends et al. (2016) and Castellini et al. (2011) studies based their definitions on the DSM-5 while the others focused on Morgan-Russell criteria or weight and binge-purging behaviors, again indicating that predictive effects are conditional on the definitions used. Two studies found that an anxiety disorder diagnosis (Fichter et al., 2006; Gualandi et al., 2017) predicted relapse while two others could not replicate these findings (Berends et al., 2016; Marzola et al., 2021). Both Marzola et al. (2021) and Berends et al. (2016) included a family component into treatment, suggesting that this might protect against the negative effects of anxiety disorders on relapse.

Lastly, several studies did not find that levels of depression predicted relapse (see Table 2) while others showed that higher levels predicted relapse in AN (Fichter et al., 2006; Gualandi et al., 2017; Kahn et al., 2020). As the volume of studies examining depression was high, it became increasingly difficult to discern the reasons behind the significance of some studies while others remained inconclusive. We conducted a meta-analysis for pre-treatment levels of depression with six studies (see Figure 6, Figure S5 and Table 3). All these studies measured depression with the Beck Depression Inventory (Beck et al., 1988). The relapsed group contained a total of  $N = 172$  participants and the non-relapsed group  $N = 318$ . Higgins  $I^2$  showed zero variability and no correction for small effects had to be conducted (Egger's test  $p = .430$ ). Based on the meta-analysis, higher levels of pre-treatment depression predicted relapse in AN (SMD = .40, 95% CI [.21 to .59],  $p < .001$ ). Although there was not enough evidence for general comorbidity, the significant results overall suggested that higher levels predicted relapse. We could only provide a solid conclusion for pre-treatment levels of depression being predictive of relapse in AN.

### 3.10 | Personality

For the variable self-esteem, one study found no effect on relapse (Carter et al., 2004) and three found lower self-esteem predictive of relapse in AN (Carter et al., 2012; Fichter et al., 2006; McFarlane et al., 2008). While Carter et al. (2004) defined recovery as maintenance of a healthy weight, the other articles included factors like

absence of binge and purging or normal menses. Therefore people might have reached recovery faster in the Carter et al. (2004) study, which could have resulted in the absence of effects for self-esteem. In conclusion, we can state that self-esteem could play an important role when predicting relapse in AN.

## 4 | DISCUSSION

This study presented a systematic review and meta-analysis of the literature on predictors of relapse in AN. Our search identified 28 studies that investigated predictors, with an average follow-up duration of 31 months and a 37% relapse rate. A number of predictors were examined, showing contradictory, significant and non-significant results. The results were mixed and there were few replicated findings. Our systematic review discussed predictors across six different categories: age and sex, symptoms and behaviors, AN subtype and duration, weight or weight change, comorbidity, and personality. Despite the diversity in findings from our systematic review for age, severity, exercise, AN duration, AN subtype, and rate of weight gain, our review also showed that lower self-esteem, longer treatment duration, lower percentage of target weight, and higher levels of comorbidity might predict increased chances of relapse. Furthermore, random effects meta-analyses were conducted with the predictors age, pre-treatment BMI, post-treatment BMI, AN duration and depression; showing that higher levels of pre-treatment depression and lower levels of post-treatment BMI predicted increased chances of relapse in AN.

The results of our systematic review suggesting that higher levels of comorbidity might predict relapse in AN were in accordance with the meta-analysis of Sala et al. (2023) and the systematic review of Berends et al. (2016). There was a 40.7% overlap in the articles included in our study and those in the work by Sala et al. (2023), and a 33.3% overlap with the study conducted by Berends et al. (2018). Specifically, the study of Sala et al. (2023) indicated that depression was the only comorbid disorder which predicted increased relapse chances across EDs. Our meta-analysis indeed revealed that higher levels of pre-treatment depression also predicted relapse for AN specifically. We were unable to conduct a meta-analysis on depression levels after treatment, preventing us from exploring effects on relapse following treatment completion. Furthermore, only studies which examined depression with the Beck Depression Inventory (Beck et al., 1988) were included in our meta-analysis, complicating comparison to research which included people with a formal diagnosis of depression. Although our results did show that higher depression levels might predict increased rates of relapse, data were limited and power was low (Cohen's  $d = .40$ ), therefore we should be careful when interpreting this primary conclusion.

Although the current systematic review rarely found significant results for BMI, the meta-analytic findings showed that lower levels of post-treatment BMI predicted higher relapse chances in AN. These findings were in accordance with a recent systematic review, which found comparable results of post-treatment BMI for AN (Frostd



et al., 2022). Our literature search resulted in the inclusion of six additional studies (Avnon et al., 2018; Berends et al., 2016; Hetman et al., 2017; MacDonald et al., 2022; Olmsted et al., 2015; Shachar-Lavie et al., 2022), and increased evidence of post-treatment BMI as predictor of relapse. Sala et al. (2023) also showed that lower BMI increased relapse chances for individuals with AN but not for other ED types, indicating that this predictor might indeed only predict relapse in AN. Although more articles measuring post-treatment BMI could be included in the meta-analysis than for depression, results should still be interpreted with caution as effects were small (Cohen's  $d = .35$ ).

Recognizing that our quality analysis revealed that only 11% of the included studies had good quality is crucial, as this impacted predictive findings. Multiple factors might have contributed to this low quality. First, the different operationalizations of relapse and recovery not only complicated the comparison of predictors, but also contributed to diverse relapse rates and non-significant results. For example, the use of more stringent relapse definitions, based on the DSM-5 (Berends et al., 2016; Castellini et al., 2011; Richard et al., 2005), frequently resulted in lower relapse rates than in studies with less stringent criteria. As the relapsed group was smaller and had more severe symptoms, certain predictors might not have reached significance. Second, for more than half of the included studies, the number of examined predictors in relation to the number of participants included, was substantially out of balance (VanVoorhis & Morgan, 2007). As a result, a significant number of studies had a poor study quality and lacked power to detect an effect. Third, although meta-analyses were only performed when sufficient data were available (Jackson & Turner, 2017), the statistical power for every analysis remained low. Related to this limitation, results should be interpreted with caution and the clinical implications given should be seen as a possible first step in the right direction. Fourth, despite considerable diversity in designs and definitions of the included studies, it is noteworthy that the majority of the study samples primarily consisted of young females and only a few studies assessed race, ethnicity and socio-economic status. This homogeneity in the demographics, combined with the limited assessment of the latter factors across most studies, restricted the generalizability of findings to the whole ED field. In addition to these limitations, our study possessed several noteworthy strengths. First, this study was unique in its focus on predictors of AN using both systematic review and meta-analytic methods; these dual perspectives mitigated the challenges posed by the significant heterogeneity in study designs and definitions. Second, meta-analyses were solely performed when predictors were assessed at the same point in time (admission or discharge) and were measured with the same instruments.

Research on predictors of relapse in AN would be enhanced with the recruitment and retention of larger, more diverse samples, considering the assessment of race, ethnicity and socio-economic status, and studied over the course of treatment and long-term follow up (Castellini et al., 2011; Fairburn et al., 2003; McFarlane et al., 2008). More immediately feasible improvements for predictive research include performing a-priori power analyses to ensure the predictor-participant ratio is sufficient for finding reliable predictive effects. In addition, it is important for future studies focusing on AN to adopt a uniform definition of recovery and relapse, as the varying definitions

across the literature likely contribute to divergent findings, including the current article (Khalsa et al., 2017).

To establish these standardized definitions, two crucial and intertwined aspects must be considered: the severity of symptoms and the duration of symptoms (De Young et al., 2020; Frank et al., 1991). A recent study by Khalsa et al. (2017) provided a framework including severity and duration criteria for AN which was based on Psychiatric Status Rating scores. In accordance with suggestions from the depression field (Frank et al., 1991) the severity criteria in the framework of Khalsa et al. (2017) were based on objective measures, including values like BMI, restricting, bingeing, and purging behaviors. Additionally, subjective measures were assessed (fear of gaining weight, a disturbed body image) as well as standardized ratings on the Eating Disorder Examination. Duration criteria were used to further distinguish these symptoms, resulting in different definitions (De Young et al., 2020). In the end, six definitions were suggested: partial recovery (6 months no symptoms), full recovery (12 months no symptoms), partial remission (1 month improved symptoms), full remission (3 months improved symptoms), partial relapse (1 month deterioration of symptoms) and full relapse (3 months deterioration of symptoms). The use of the proposed standardized definitions and follow-up periods will enhance comprehension of the timing of predictors, helps to distinguish the periods during which people are the most vulnerable to returning symptoms, will aid the comparability of predictive effects, and facilitates assessment by clinicians and researchers (Bardone-Cone et al., 2018; Khalsa et al., 2017).

As the studies focusing on relapse are highly heterogeneous in operationalizations and predictors of relapse, results from the current systematic review and meta-analysis should be interpreted with caution. Our findings indicated that depression assessment in AN should begin early, not only to address the risks associated with high levels of severity (e.g., suicide), but also to address the complications depression may present in sustaining improvement in AN over time. However, no causal link has been established between baseline depression and risk of AN relapse. As both the current study, Sala et al. (2023) and Frostad et al. (2022) have found that BMI was a predictor of relapse, it may be wise to work on a more extensive relapse prevention plan when BMI remains low after treatment completion.

Our review confirmed the high relapse rate for AN (37%, Sala et al. (2023)), emphasizing the need for a better understanding of relapse predictors. Results showed that higher pre-treatment depression levels and lower levels of post-treatment BMI predicted increased chances of relapse in AN. Based on the current state of the field, the most important recommendation for future studies is to use standardized definitions of recovery and relapse. This is essential to improve our understanding and interpretation of predictive effects and enables the discovery of important predictors of deterioration during remission or recovery. Ultimately, this can contribute to the development of prevention strategies and maintenance interventions, thereby increasing focus on the sustainability of recovery in AN.

#### AUTHOR CONTRIBUTIONS

**Eline S. J. de Rijk:** Conceptualization; formal analysis; writing – original draft; writing – review and editing. **Durr Almirabi:** Conceptualization; writing – review and editing. **Lauren Robinson:**



Conceptualization; formal analysis; writing – review and editing. **Ulrike Schmidt:** Conceptualization; writing – review and editing. **Eric F. van Furth:** Conceptualization; writing – review and editing. **Rita C Slof-Op 't Landt:** Conceptualization; formal analysis; writing – original draft; writing – review and editing.

## ACKNOWLEDGMENTS

Ulrike Schmidt is supported by the Medical Research Council/Arts and Humanities Research Council/Economic and Social Research Council Adolescence, Mental Health and the Developing Mind initiative as part of the EDIFY programme (grant number MR/W002418/1). Lauren Robinson and Ulrike Schmidt receive(d) salary support from the National Institute of Health Research (NIHR) Biomedical Research Centre (BRC) at the South London and Maudsley NHS Foundation Trust (SLaM) and King's College London (KCL). The views expressed are those of the authors and not necessarily those of the Health Foundation, NHS, the NIHR or the Department of Health.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## OPEN RESEARCH BADGES



This article has earned a Preregistered Research Designs badge for having a preregistered research design, available at <https://osf.io/8j4sc/>.

## DATA AVAILABILITY STATEMENT

Data will be available upon reasonable request.

## ORCID

Eline S. J. de Rijk <https://orcid.org/0000-0003-4312-142X>  
 Lauren Robinson <https://orcid.org/0000-0001-9550-2940>  
 Ulrike Schmidt <https://orcid.org/0000-0003-1335-1937>  
 Eric F. van Furth <https://orcid.org/0000-0002-1131-0886>  
 Margarita C. T. Slof-Op 't Landt <https://orcid.org/0000-0001-6135-2163>

## REFERENCES

- Ágh, T., Kovács, G., Supina, D., Pawaskar, M., Herman, B. K., Vokó, Z., & Sheehan, D. V. (2016). A systematic review of the health-related quality of life and economic burdens of anorexia nervosa, bulimia nervosa, and binge eating disorder. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 21, 353–364.
- Avnon, A., Orkaby, N., Hadas, A., Berger, U., Brunstein Klomek, A., & Fennig, S. (2018). Inpatient weight curve trajectory as a prognostic factor among adolescents with anorexia nervosa: A preliminary report. *Eating and Weight Disorders*, 23(5), 645–651. <https://doi.org/10.1007/s40519-017-0415-8>
- Bardone-Cone, A. M., Hunt, R. A., & Watson, H. J. (2018). An overview of conceptualizations of eating disorder recovery, recent findings, and future directions. *Current Psychiatry Reports*, 20(9), 79. <https://doi.org/10.1007/s11920-018-0932-9>
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck depression inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100. [https://doi.org/10.1016/0272-7358\(88\)90050-5](https://doi.org/10.1016/0272-7358(88)90050-5)
- Berends, T., Boonstra, N., & van Elburg, A. (2018). Relapse in anorexia nervosa: A systematic review and meta-analysis. *Current Opinion in Psychiatry*, 31(6), 445–455. <https://doi.org/10.1097/YCO.0000000000000453>
- Berends, T., van Meijel, B., Nugteren, W., Deen, M., Danner, U. N., Hoek, H. W., & van Elburg, A. A. (2016). Rate, timing and predictors of relapse in patients with anorexia nervosa following a relapse prevention program: A cohort study. *BMC Psychiatry*, 16(1), 316. <https://doi.org/10.1186/s12888-016-1019-y>
- Bodell, L. P., & Mayer, L. E. (2011). Percent body fat is a risk factor for relapse in anorexia nervosa: A replication study. *The International Journal of Eating Disorders*, 44(2), 118–123. <https://doi.org/10.1002/eat.20801>
- Brown, M., Robinson, L., Campione, G. C., Wuensch, K., Hildebrandt, T., & Micali, N. (2017). Intolerance of uncertainty in eating disorders: A systematic review and meta-analysis. *European Eating Disorders Review*, 25(5), 329–343. <https://doi.org/10.1002/erv.2523>
- Carter, J. C., Blackmore, E., Sutandar-Pinnock, K., & Woodside, D. B. (2004). Relapse in anorexia nervosa: A survival analysis. *Psychological Medicine*, 34(4), 671–679. <https://doi.org/10.1017/S0033291703001168>
- Carter, J. C., Mercer-Lynn, K. B., Norwood, S. J., Bewell-Weiss, C. V., Crosby, R. D., Woodside, D. B., & Olmsted, M. P. (2012). A prospective study of predictors of relapse in anorexia nervosa: Implications for relapse prevention. *Psychiatry Research*, 200(2–3), 518–523. <https://doi.org/10.1016/j.psychres.2012.04.037>
- Castellini, G., Lo Sauro, C., Mannucci, E., Ravaldi, C., Rotella, C. M., Faravelli, C., & Ricca, V. (2011). Diagnostic crossover and outcome predictors in eating disorders according to DSM-IV and DSM-V proposed criteria: A 6-year follow-up study. *Psychosomatic Medicine*, 73(3), 270–279. <https://doi.org/10.1097/PSY.0b013e31820a1838>
- Castro, J., Gila, A., Puig, J., Rodriguez, S., & Toro, J. (2004). Predictors of rehospitalization after total weight recovery in adolescents with anorexia nervosa. *The International Journal of Eating Disorders*, 36(1), 22–30. <https://doi.org/10.1002/eat.20009>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- De Young, K. P., Kambanis, P. E., Bottera, A. R., Mancuso, C., Thomas, J. J., Franko, D. L., Herzog, D. B., Walker, D. C., Anderson, D., & Eddy, K. T. (2020). Identifying duration criteria for eating-disorder remission and recovery through intensive modeling of longitudinal data. *The International Journal of Eating Disorders*, 53(8), 1224–1233. <https://doi.org/10.1002/eat.23249>
- Deter, H. C., & Hertog, W. (1994). Anorexia nervosa in a long-term perspective: results of the Heidelberg-Mannheim study. *Psychosomatic Medicine*, 56(1), 20–27. <https://doi.org/10.1097/00006842-199401000-00003>
- Edwin, D., Andersen, A. E., & Rosell, F. (1988). Outcome prediction by MMPI in subtypes of anorexia nervosa. *Psychosomatics*, 29(3), 273–282. [https://doi.org/10.1016/s0033-3182\(88\)72363-4](https://doi.org/10.1016/s0033-3182(88)72363-4)
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- El Ghoch, M., Calugi, S., Chignola, E., Bazzani, P. V., & Dalle Grave, R. (2016). Body mass index, body fat and risk factor of relapse in anorexia nervosa. *European Journal of Clinical Nutrition*, 70(2), 194–198. <https://doi.org/10.1038/ejcn.2015.164>
- Fairburn, C. G., Cooper, Z. C., & 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](https://doi.org/10.1016/s0005-7967(02)00088-8)

- Fichter, M. M., Quadflieg, N., & Hedlund, S. (2006). Twelve-year course and outcome predictors of anorexia nervosa. *The International Journal of Eating Disorders*, 39(2), 87–100. <https://doi.org/10.1002/eat.20215>
- Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., Rush, A. J., & Weissman, M. M. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: Remission, recovery, relapse, and recurrence. *Archives of General Psychiatry*, 48(9), 851–855.
- Frostad, S., Rozakou-Soumalia, N., Darvari, S., Foruzesh, B., Azkia, H., Larsen, M. P., Rowshandel, E., & Sjogren, J. M. (2022). BMI at discharge from treatment predicts relapse in anorexia nervosa: A systematic scoping review. *Journal of Personalized Medicine*, 12(5), 353–364. <https://doi.org/10.3390/jpm12050836>
- Garber, A. K., Hilton, J. F., Scarpulla, M. T., Buckelew, S. M., & Moscicki, A.-B. (2015). 20. The association of refeeding calorie level and recovery over one-year follow-up in adolescents with anorexia nervosa. *Journal of Adolescent Health*, 56(2), S11.
- Glashouwer, K. A., Bennik, E. C., de Jong, P. J., & Spruyt, A. (2018). Implicit measures of actual versus ideal body image: Relations with self-reported body dissatisfaction and dieting behaviors. *Cognitive Therapy and Research*, 42(5), 622–635. <https://doi.org/10.1007/s10608-018-9917-6>
- Glashouwer, K. A., van der Veer, R. M. L., Adipatria, F., de Jong, P. J., & Vocks, S. (2019). The role of body image disturbance in the onset, maintenance, and relapse of anorexia nervosa: A systematic review. *Clinical Psychology Review*, 74, 101771. <https://doi.org/10.1016/j.cpr.2019.101771>
- Gualandi, P., Lucia, M., Capece, S., Lattarulo, S., Quisisana, F., Rocchi, R., Lividini, A., Gentile, V., Malaspina, E., & Franzoni, E. (2017). Rate of weight gain as a predictor of readmission in adolescents with eating disorders. *Eating and Weight Disorders*, 22(3), 58.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, 3(4), 486–504. <https://doi.org/10.1037/1082-989X.3.4.486>
- Herzog, D. B., Dorer, D. J., Keel, P. K., Selwyn, S. E., Ekeblad, E. R., Flores, A. T., Greenwood, D. N., Burwell, R. A., & Keller, M. B. (1999). Recovery and relapse in anorexia and bulimia nervosa: A 7.5-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38(7), 829–837.
- Hetman, I., Klomek, A. B., Goldzweig, G., Hadas, A., Horwitz, M., & Fennig, S. (2017). Percentage from target weight (PFTW) predicts re-hospitalization in adolescent anorexia nervosa. *Israel Journal of Psychiatry*, 54(3), 28–34. <https://doi.org/10.1007/s12020-017-1309-9>
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Huang, X., Lin, J., & Demner-Fushman, D. (2006). Evaluation of PICO as a knowledge representation for clinical questions. Paper presented at the AMIA annual symposium proceedings.
- Huedo-Medina, T. B., Sánchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I<sup>2</sup> index? *Psychological Methods*, 11(2), 193–206. <https://doi.org/10.1037/1082-989X.11.2.193>
- VH Innovation. (2021). Covidence systematic review software. [www.covidence.org](http://www.covidence.org)
- Jackson, D., & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research Synthesis Methods*, 8(3), 290–302. <https://doi.org/10.1002/jrsm.1240>
- Jagielska, G., & Kacperska, I. (2017). Outcome, comorbidity and prognosis in anorexia nervosa. *Psychiatria Polska*, 51(2), 205–218.
- Kahn, M., Brunstein-Klomek, A., Hadas, A., Snir, A., & Fennig, S. (2020). Early changes in depression predict outcomes of inpatient adolescent anorexia nervosa. *Eating and Weight Disorders*, 25(3), 777–785. <https://doi.org/10.1007/s40519-019-00686-9>
- Keel, P. K., Dorer, D. J., Franko, D. L., Jackson, S. C., & Herzog, D. B. (2005). Postremission predictors of relapse in women with eating disorders. *American Journal of Psychiatry*, 162(12), 2263–2268. <https://doi.org/10.1176/appi.ajp.162.12.2263>
- Khalsa, S. S., Portnoff, L. C., McCurdy-McKinnon, D., & Feusner, J. D. (2017). What happens after treatment? A systematic review of relapse, remission, and recovery in anorexia nervosa. *Journal of Eating Disorders*, 5, 20. <https://doi.org/10.1186/s40337-017-0145-3>
- Kim, Y., Hersch, J., Bodell, L. P., Schebendach, J., Hildebrandt, T., Walsh, B. T., & Mayer, L. E. S. (2021). The association between leptin and weight maintenance outcome in anorexia nervosa. *The International Journal of Eating Disorders*, 54(4), 527–534. <https://doi.org/10.1002/eat.23407>
- MacDonald, D. E., McFarlane, T., Trottier, K., Mahan, M., & Olmsted, M. P. (2022). Maintenance treatment for eating disorders following inpatient or day treatment: Outcomes of intensive outpatient group and individual CBT treatments. *Eating Disorders*, 30(4), 453–469. <https://doi.org/10.1080/10640266.2021.1969787>
- Marzola, E., Longo, P., Sardella, F., Delsedime, N., & Abbate-Daga, G. (2021). Rehospitalization and “revolving door” in anorexia nervosa: Are there any predictors of time to readmission? *Frontiers in Psychiatry*, 12, 694223. <https://doi.org/10.3389/fpsy.2021.694223>
- McCabe, E. B. (2008). *Predictors of readmission in hospitalized anorexia nervosa patients*. University of Pittsburgh.
- McCormick, L. M., Keel, P. K., Brumm, M. C., Watson, D. B., Forman-Hoffman, V. L., & Bowers, W. A. (2009). A pilot study of personality pathology in patients with anorexia nervosa: Modifiable factors related to outcome after hospitalization. *Eating and Weight Disorders*, 14(2–3), e113–e120. <https://doi.org/10.1007/BF03327808>
- McFarlane, T., Olmsted, M. P., & Trottier, K. (2008). Timing and prediction of relapse in a transdiagnostic eating disorder sample. *The International Journal of Eating Disorders*, 41(7), 587–593. <https://doi.org/10.1002/eat.20550>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–269. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>
- Morgan, H. G., & Russell, G. (1975). Value of family background and clinical features as predictors of long-term outcome in anorexia nervosa: Four-year follow-up study of 41 patients. *Psychological Medicine*, 5(4), 355–371. <https://doi.org/10.1017/S0033291700056981>
- National Heart, Lung, and Blood Institute (NHLBI). (2014). Study quality assessment tool. <https://www.nhlbi.nih.gov/health-topics/study-quality-assessmenttools>
- Olmsted, M. P., MacDonald, D. E., McFarlane, T., Trottier, K., & Colton, P. (2015). Predictors of rapid relapse in bulimia nervosa. *The International Journal of Eating Disorders*, 48(3), 337–340. <https://doi.org/10.1002/eat.22380>
- Pruccoli, J., Pugliano, R., Pranzetti, B., & Parmeggiani, A. (2023). Premenstrual anorexia nervosa: Clinical features, psychopharmacological interventions, and rehospitalization analysis in a 1-year follow-up, controlled study. *European Journal of Pediatrics*, 182(6), 2855–2864.
- Richard, M., Bauer, S., & Kordy, H. (2005). Relapse in anorexia and bulimia nervosa—A 2.5-year follow-up study. *European Eating Disorders Review*, 13(3), 180–190. <https://doi.org/10.1002/erv.638>
- Sala, M., Keshishian, A., Song, S., Moskowitz, R., Bulik, C. M., Roos, C. R., & Levinson, C. A. (2023). Predictors of relapse in eating disorders: A meta-analysis. *Journal of Psychiatric Research*, 158, 281–299. <https://doi.org/10.1016/j.jpsychires.2023.01.002>
- Sebastian, M. R., Wiemann, C. M., & Hergenroeder, A. C. (2019). Rate of weight gain as a predictor of readmission in adolescents with eating disorders. *International Journal of Adolescent Medicine and Health*, 33(4), 1–5. <https://doi.org/10.1515/ijamh-2018-0228>
- Shachar-Lavie, I., Segal, H., Oryan, Z., Halifa-Kurtzman, I., Bar-Eyal, A., Hadas, A., Tamar, T., Benaroya-Milshstein, N., & Fennig, S. (2022). Atypical anorexia nervosa: Rethinking the association between target

- weight and rehospitalization. *Eating Behaviors*, 46. <https://doi.org/10.1016/j.eatbeh.2022.101649>
- Steinglass, J. E., Sysko, R., Glasofer, D., Albano, A. M., Simpson, H. B., & Walsh, B. T. (2011). Rationale for the application of exposure and response prevention to the treatment of anorexia nervosa. *International Journal of Eating Disorders*, 44(2), 134–141.
- Vall, E., & Wade, T. D. (2017). Predictors and moderators of outcomes and readmission for adolescent inpatients with anorexia nervosa: A pilot study. *Clinical Psychologist*, 21(2), 143–152. <https://doi.org/10.1111/cp.12091>
- van Eeden, A. E., van Hoeken, D., & Hoek, H. W. (2021). Incidence, prevalence and mortality of anorexia nervosa and bulimia nervosa. *Current Opinion in Psychiatry*, 34(6), 515–524.
- VanVoorhis, C. W., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology*, 3(2), 43–50. <https://doi.org/10.20982/tqmp.03.2.p043>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** de Rijk, E. S. J., Almirabi, D., Robinson, L., Schmidt, U., van Furth, E. F., & Slof-Op 't Landt, M. C. T. (2024). An overview and investigation of relapse predictors in anorexia nervosa: A systematic review and meta-analysis. *International Journal of Eating Disorders*, 57(1), 3–26. <https://doi.org/10.1002/eat.24059>