
DO ADOLESCENTS REALLY RECOVER FROM ANOREXIA? OR THE LACK OF STANDARDISED DEFINITION MAY MASK THEIR PROCESS?: A SYSTEMATIC REVIEW

Alara Kerimler¹, Hakan Öğütlü^{2}, Darren Cutinha³*

¹ Beykent University Institute of Graduate Studies Department of Clinical Psychology

² Department of Child and Adolescent Psychiatry, Cognitive Behavioral Psychotherapies Association, Ankara, Turkey

³ Department of Child and Adolescent Psychiatry, South London and Maudsley NHS Foundation Trust, London, UK

Abstract

Anorexia Nervosa (AN) is an eating disorder that causes physical, behavioural and psychological deterioration. Although diagnostic criteria are clearly defined, there has been no consensus on what recovery is. This study aimed to review prior studies indicating recovery criteria, as well as recovery rates of patients with adolescent-onset AN.

Related studies were searched through databases MEDLINE, PsycINFO and CINAHL. A total of 15 English studies with patients had adolescent-onset and DSM-5/ICD-11 diagnoses. A systematic review was conducted by following the PRISMA expanded checklist and qualities of eligible articles were evaluated via the Quality Criteria Checklist (QCC).

Of the 15 studies, two studies mentioned only physical dimension of recovery, whereas rest of the 13 studies covered behavioural and psychological dimensions. EDE-Q was found as the most frequently used scale. Follow-up years of the patients fluctuated from one to 30 years, and the recovery rate varied from 30.6% to 72%.

There are some difficulties faced in defining recovered patients. Since no consensus was achieved, every researcher set their recovery criteria. Until the policymakers of field standardize definition of recovery from AN, researchers should be aware of the fact that inconsistencies in definition can affect results of their research.

Keywords: eating disorders, anorexia nervosa, adolescents, recovery, outcome.

* Correspondence concerning this article should be addressed to Hakan Öğütlü, MD, Department of Child and Adolescent Psychiatry, Cognitive Behavioral Psychotherapies Association, Karum Is Merkezi Iran Caddesi No: 21 06680 Gaziosmanpasa Mah. 06680 Cankaya Ankara, Turkey. ORCID: <https://orcid.org/0000-0002-1325-446X>.
E-mail address: hogutlu@gmail.com

Anorexia nervosa (AN) is an eating disorder (ED) characterized by deterioration in eating and eating-related thoughts and behaviours, which leads to physical (i.e. cardiovascular, dermatologic, endocrine, gastrointestinal) and psychological (i.e. depression, insomnia, irritability, impaired concentration) complications, the extent and severity of which patients are mostly unaware of (Moore & Bokor, 2023). Follow-up studies define multiple outcomes with different prevalence for adolescents with AN, consisting of persistence of AN, meeting diagnostic criteria for any other EDs or psychiatric condition, (partial or full) remission, (partial or full) recovery, relapse or death (Andrés-Pepiñá et al., 2020; Dobrescu et al. 2020; Graell et al., 2018; Steinhausen, 2002). Defining remission is more straightforward when it is compared to recovery since the DSM-5 suggested a definition of it. When the patients restore their weight but bear the trace of other behavioural and psychological sufferings, they are accepted as in partial remission. In addition, when none of the diagnostic criteria is met, it is called a full remission (American Psychiatric Association & Association, 2013). If the absence of diagnostic criteria is a remission, what is recovery? There is a longstanding lack of consensus on the definition of recovery which leads to a lack of harmonization in clinical practice and research studies (Bardone-Cone et al., 2018; Khalsa et al., 2017; Lock et al., 2013). Although Lock et al. (2013) indicated that the definition of recovery might be related to the aim of the treatment and the studies, clinicians require a standardised definition to understand the end goals of the interventions, answer patients' and carers' questions and provide hope for the sufferers that think recovery is impossible (Bardone-Cone et al., 2010; Bardone-Cone et al., 2018; Lock et al., 2013; Morgan & Hayward, 1988)

There are two types of studies conducted as follows 1) quantitative studies and 2) qualitative to define recovery, and it is suggested to perceive them as a bridge for each other instead of separate approaches (Bardone-Cone et al., 2018). Quantitative studies follow two empirical methodologies to validate the definition of recovery (Bardone-Cone et al., 2018) whereas qualitative studies investigate patients' and their relatives' perspectives (Bachner-Melman et al., 2018; Bardone-Cone et al., 2018; de Vos et al., 2017). As a first methodology for quantitative studies, the researcher compares disordered eating psychopathologies (i.e. body image, food, eating or shape/weight obsessions, thin idealization) among the groups, consisting of the ones that fit the recovery criteria proposed by that study, the ones with current EDs and the ones without EDs. At the end of the comparisons, if the recovered patients present significantly better scores than patients with ED and do not differ from participants without ED, these results are accepted as evidence to validate the definition of recovery (Bachner-Melman et al., 2006; Bardone-Cone et al., 2010; Bardone-Cone et al., 2018). The second used way to test validity is to compare relapse rates of AN by using longitudinal studies. In this methodology, the recovery criteria that give low relapse rates are accepted as a useful and meaningful

definition for recovery (Bardone-Cone et al., 2010; Kordy et al., 2002). The earliest studies in this field associated recovery with a lack of physical (i.e. low BMI and disrupted menses) (Morgan & Russell, 1975) and behavioural (dietary restriction, compensatory behaviours, binge eating) symptomology; however, recent studies agreed on the necessity of psychological and cognitive evaluations to define recovery (Bardone-Cone et al., 2010; Bardone-Cone et al., 2018; Couturier & Lock, 2006; Khalsa et al., 2017). Although weight gain is a critical factor in improving ED pathology (Couturier & Lock, 2006; Lock et al., 2013) and earlier weight restoration decreases eating-related concerns and restrictive dietary patterns significantly, gaining weight alone remained incapable of improving weight and shape-related concerns (Accurso et al., 2014).

Recently, Khalsa et al. (2017) proposed a definition for both full and partial recovery from EDs; however, the importance of including support from higher authorities (i.e. Academy for Eating Disorders, Eating Disorder Research Society, National Eating Disorders Association etc.) was emphasized in their study. In the annual meeting of the International Eating Disorder Research Society (EDRS) (2018), the need for a standardised definition of recovery was discussed and participants agreed on the importance of a specific definition for recovering from EDs that combines the absence of diagnostic criteria and functional recovery across multiple domains such as psychological, emotional and social ones for over 12 months. In addition, they suggested further studies to investigate commonly used definitions and gain a better understanding of the differences in psychometrically and self-assessed acceptance of recovery (Wade & Lock, 2020).

Since there is no consensus on the definition of recovery from EDs (therefore AN), this study aims to examine which criteria have been used by researchers to accept adolescents with AN as recovered and present how recovery rates have fluctuated. In addition to previous reviews (Bardone-Cone et al., 2018; Khalsa et al., 2017), this study aims to decrease heterogeneity/differences in participant characteristics by focusing on adolescence age-onset AN, based on DSM-5/ ICD-11 diagnosis.

Methodology

The PRISMA 2020 expanded checklist was followed while conducting this systematic review (Page et al., 2021).

Search Strategy

To establish relevant studies, two different platforms and three databases were searched: MEDLINE (Ovid), PsycINFO (Ovid), and CINAHL (EBSCOhost)

Articles Section

from the beginning of November to the 28th of June (2022). The terms: “recover*”, “anorexia*” and “adolescen*” were utilized for both free-text and subheading searching, and the asterisk (*) was used to cover multiple suffixes of the terms. The detailed table for the search strategy is given below.

Table 1. The detailed search strategy

Ovid MEDLINE(R) ALL	Ovid APA PsycInfo	EBSCOhost CINAHL
1 adolescen*.tw. 321982	1 adolescen*.tw. 283485	S1 (MH"Adolescence") 583,497
2 young people.tw 34171	2 young person*.tw. 2993	S2 TI (adolescen* or teen* or youth* or "young person*" or "young people*") OR AB (adolescen* or teen* or youth* or "young person*" or "young people*") 223,677
3 young person.tw 1244	3 young people.tw. 33169	S3 TI anorexia* OR AB anorexia* 8,865
4 teen*.tw. 33430	4 teen*.tw. 24909	S4 (MH "Anorexia Nervosa") OR (MH "Anorexia") 7,144
5 youth*.tw. 88771	5 youth*.tw. 116166	S5 TI recover* OR AB recover* 118,467
6 adolescent/ or child/ 305813 5	6 early adolescence/ 2585	S6 (MH "Recovery") 37,181
7 1 or 2 or 3 or 4 or 5 or 6 315660 9	7 1 or 2 or 3 or 4 or 5 or 6 376442	S7 S1 OR S2 645,069
8 anorexia*.tw. 32477	8 anorexia*.tw. 17103	S8 S3 OR S4 10,693
9 Anorexia Nervosa/ or Anorexia/ 19214	9 anorexia nervosa/ 11997	S9 S5 OR S6 135,370
10 recover*.tw. 754946	10 8 or 9 17657	S10 S7 AND S8 AND S9 286
11 7 and 9 and 10 683	11 recover*.tw. 92717	
	12 7 and 10 and 11 361	

2.2. Inclusion and Exclusion Criteria

After the studies were identified from databases, all studies were added to EndNote 20 (version 20.3.0.177.87.) software to delete duplicated ones and scan titles and abstracts to determine eligibility. In addition, the full texts of the articles that could not be eliminated by looking at the title and abstract went through further examination for conformity. Table 2 summarizes the inclusion and exclusion criteria followed during this review.

Table 2. Inclusion and Exclusion Criteria

INCLUSION CRITERIA	EXCLUSION CRITERIA
<p>- Participants Studies in which AN-onset is based on adolescence term* and AN was diagnosed according to DSM-V and/or ICD-11 Both genders (male/ female) Studies with mixed ED patients</p> <p>- Study Journal Articles that give criteria of recovery with quantitative methodologies**</p> <p>Date: Studies from 2013 to 2022****</p> <p>- Language: English studies</p>	<p>- Participants Studies in which the diagnosis of AN is not formal (i.e. clinician subjective thoughts) or based on previous DSM (DSM-IV, DSM-III) or ICD (ICD-10) forms Studies with a lack of information related to the age of AN-onset Animal studies</p> <p>- Study Journal Articles that define recovery by using only qualitative methodologies (i.e. patients' and/or parents' perspectives), reviews (both literature and systematic), books and dissertations The studies used the term of “weight-recovery/recovered” or “menstrual recovery” instead recovered/recovery***</p> <p>Date: Studies before 2013</p> <p>- Language: Non-English studies</p>

*Adolescence age was accepted as 10-19 years (World Health Organization, 2022a). **Since receiving or not receiving therapy did not affect the definition of recovery, whether the participant group received psychotherapy or not was not evaluated while compiling the articles giving the definition. ***As it has been previously recognised that physical recovery is insufficient to ameliorate body and shape concerns, articles that use the terms of “weight-recovery/recovered” or “menstrual recovery” were excluded (Accurso et al., 2014). ****To cover more recent studies, and decrease the differences in participant characteristics, only the studies that made the diagnosis by the updated version of the manuals (DSM-V and ICD-11) were included. Since the DSM-V and ICD-11 were published in 2013 and 2019, respectively, studies before 2013 were excluded.

Data Extraction

During the data extraction process, all the relevant information regarding the eligibility criteria (i.e. who are the patients, what is the AN age-onset, how patients were diagnosed etc.) and the concept of recovery (i.e. how was it defined, when patients were followed and what was the rate of recovery) were gathered primarily and presented in Table 4.

2.4. Quality Assessment

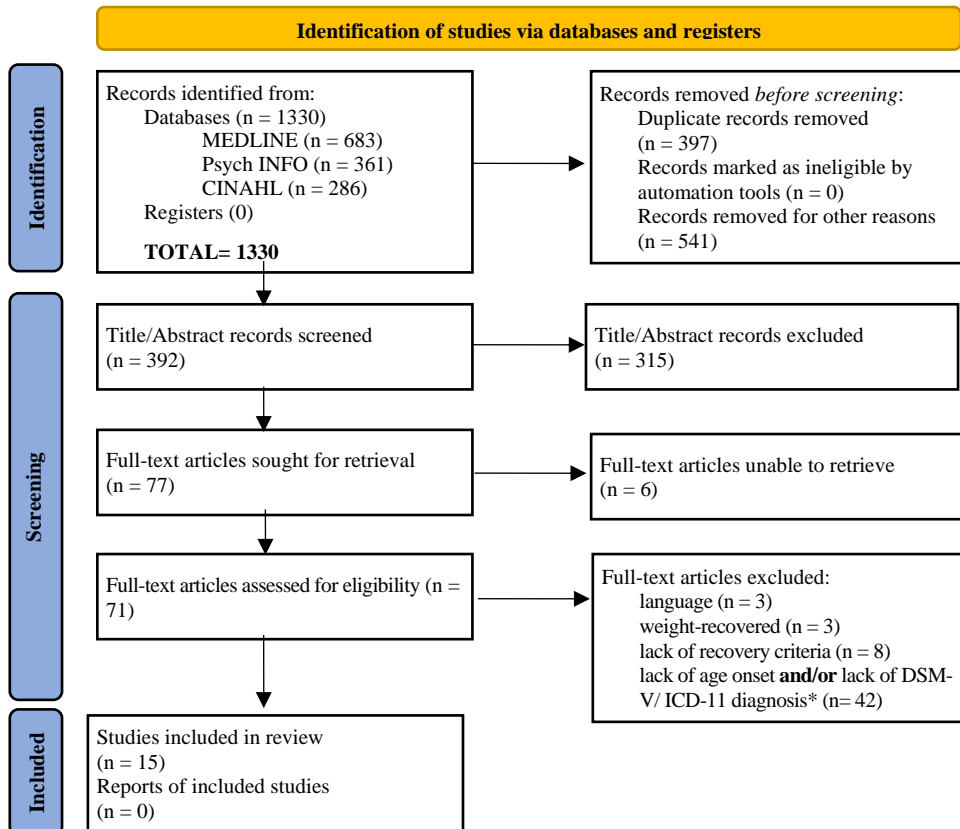
The quality assessment of each eligible article was performed by using the Quality Criteria Checklist (QCC) (Academy of Nutrition and Dietetics, 2022). In more detail, all questions within the scale have 4 options: Yes, No, Unclear and N/A. After

the first 4 relevancy questions were answered, 10 validity questions were followed for each eligible study. The articles with six or more “No” answers to validity questions were designated with a minus symbol (-). When the answers to the 2nd, 3rd, 6th and 7th questions did not indicate that the study was exceptionally strong, they were indicated as neutral (∅). Articles containing at least one “Yes” answer, in addition to the 2nd, 3rd, 6th and 7th questions, were marked with a plus symbol (+).

Results

Study Selection

15 studies were found eligible for this and the details of the screening process were presented in Figure 1.



*The number of articles caused by lack of AN age-onset and DSM-V diagnosis were not reported separately due to the existence of articles that did not meet both criteria.

Figure 1. PRISMA flow diagram of screening and inclusion of eligible studies.

Table 3. Characteristics and Key Findings of The Eligible Studies

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
Breithaupt, L., et al. (2022) Design: Cohort Study Country: New England, USA Published: United Kingdom	Total (n): 82 AN-R (n): 29 Age: 19.5 ± 2.3 Illness duration: 5.6 ± 5.2 AN-BP (n): 11 Age: $21.6 (1.3)$ Illness duration: 9.1 ± 0.6 Atypical AN-R (n): 19 Age: 19.9 ± 3.0 Illness duration: 5.4 ± 6.5 Atypical AN-BP (n): 7 Age: 20.1 ± 1.3 Illness duration: 7.6 ± 9.1 ARFID (n): 16 Age: 15.3 ± 4.9 Illness duration: 1.3 ± 0.7 Participants were recruited through advertisements, flyers, health care providers, outpatient practices, and higher level of care programs.	Not Specified as years, but enough knowledge was given to affirm “adolescent-onset”.	EDE 17.0 LIFE-EAT-3	Physical: Not specified. Behavioural and/ or Psychological: Absent or minimal/residual symptoms and no functional impairment due to ED cognitions and/or behaviours. Scale: LIFE-EAT-3 PSR score of 1 or 2. Duration: Not specified.	2 follows up were conducted: at 9 th and 10 th months after baseline.	The likelihood of stable recovery was stated as varying between 0.00 and 0.36. Persistence in recovery (from 9 months to 10 months) was unlikely across all individuals: AN-R, 0.08; AN-BP, 0.00; Atypical AN-R, 0.24; Atypical AN-BP, 0.33; ARFID, 0.36.
Wentz, E., et al. (2021) Design: Cross-sectional study Country:	Total (n): 72 AN (n): 34 Age: 44.2 Comparison Group (n): 38	Not Specified as years, but indicated as attending to eight-grade	MINI 6.0, SCID-I, DSM-V checklist GAF Morgan-Russell Scale	Physical: No weight deviation. Behavioural and/ or Psychological: Being free of all criteria of AN, BN or BED. Not having compensatory behaviours and absence of weight phobia.	30 years (30.13 ± 1.62)	20 out of 34 AN patients recovered (58.8%).

Articles Section

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
Conducted: Gothenburg, Sweden Published: Germany	Age: 44.3 Participants were recruited from Gothenburg, Sweden cohort*.	when they diagnosed.	SF-36	Scale: Not specified. Duration: Minimum of 6 consecutive months. Other: Recovery was indicated as “full symptom recovery”.		
Kerr-Gaffney, J. et al. (2021) Design: Cross-sectional study Country: England, UK Conducted: England, UK Published: England, UK	Total (n): 218 AN (n): 64 Age: 21.53 ± 4.15 Recovered (n): 46 Age: 22.21 ± 3.47 ASD (n): 41 Age: 20.56 ± 8.42 TD (Controls) (n): 67 Age: 21.16 ± 3.60 Participants were recruited from three pre-existing data sets – no further information provided throughout the article.	AN Group: 16.80 ± 3.62 Recovered Group: 15.56 ± 2.52 ASD and TD Groups: N/A	ADOS-2 or AQ-10, SRS-2: AN, REC, ASD, TD WASI-II (or WAIS-R, WAID-III, WAIS-IV, WISC-III) or NART: AN, REC, TD SCID-5-RV: AN, REC, TD EDE-Q: AN, REC, TD HADS: AN, REC, TD BAI/BDI: AN, REC, TD BYI-II: ASD	Physical: BMI between 18.5 and 27 kg/m2 (%IBW > 85 for participants under 18 years). Behavioural and/ or Psychological: Not specified. Scale: Not specified. Duration: Minimum of a year.	N/A	N/A
Castro-Fornieles, J. Et al. (2021) Design: Cross-sectional study Country: Barcelona, Spain Conducted: Barcelona, Spain Published: Germany	Total (n): 54 ED Group (n): 11 Age: 37.3 ± 5.7 Age at first diagnosis: 14.7 ± 2.0 Recovery Group (n): 15 Age: 35.6 ± 2.5 Age at first diagnosis: 14.1 ± 1.7	ED Group: 13.5 ± 1.3 Recovered Group: 13.4 ± 1.8 Control Group: N/A	SCID-I, DSM-V Checklist EDE (Spanish version) OBI	Physical: Not specified. Behavioural and/ or Psychological: Not having current ED diagnosis. Scale: Not specified. Duration: Not specified.	20 (17-25) years	15 out of 26 AN patients recovered (%57.69).

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
	Control Group (n): 28 Age: 36.5 ± 3.0					
	Patients were recruited from the Department of Child and Adolescent Psychiatry and Psychology of the Hospital Clinic of Barcelona's records (from 1987 to 1993).					
Silen et al. (2021) Design: Cohort study (population based) Country: Helsinki, Finland Conducted: Helsinki, Finland Published: New Jersey, US	Total (n): 145 AN (n): 46 Illness duration: 4.6 ± 3.4 BN (n): 18 Illness duration: 4.2 ± 2.9 BED (n): 6 Illness duration: 4.0 ± 2.4 OSFED (n): 33 Illness duration: 3.9 ± 2.5 UFED (n): 42 Illness duration: 3.2 ± 2.5 Any ED Illness duration: 4.0 ± 2.9 Participants were recruited from the FinnTwin12 birth cohort** (wave 4).	Any type of EDs: 16.5 ± 2.9 (Detailed information for subtypes was not provided and only 140 patients' knowledge about the age-onset of EDs was present.)	SCID-I, DSM-V Checklist	Physical: BMI ≥ 18.5 kg/m ² Behavioural and/ or Psychological: Participants themselves had to express that they thought they no longer suffered from an ED (both behaviourally and psychologically) to be accepted recovered. Scale: Not specified. Duration: Not specified. Other: Recovery was indicated as "full recovery".	4.0 ± 2.9 (0.5 - 13) years	5-year recovery rates were indicated as 40.7% from any ED, in more detail, 41.5% from AN, 23.1 % from BN, 40.0% from BED, 43.1% from OSFED, 42.6 from UFED.
Garriz, M. et al. (2021)	Total (n): 58 Current AN (n): 11 Age: 37.73 ± 5.75	Current ED Group: 13.18 ± 1.83	SCID-I, DSM-V Checklist	Physical: BMI > 18.5 kg/m ² Behavioural and/ or Psychological: No longer meeting the criteria for an ED.	22 (17-29) years	18 out of 29 patients recovered (62.06%).

Articles Section

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
Design: Case Control Study Country: Barcelona, Spain Conducted: Germany Published:	Recovered AN (n): 18 Age: 35.89 ± 2.76 Control Group (n): 29 Age: 36.55 ± 2.55 Patients were recruited from the Department of Child and Adolescent Psychiatry and Psychology of the Hospital Clinic of Barcelona's records (from 1987 to 1993).	Recovered Group: 13.61 ± 1.24	PDQ-4+ (Spanish version) OBSI-R EDI-2	Abstinence from bingeing, purging and fasting. Scale: EDE sub-scale scores all within 1 SD of healthy, age-matched population norms. Duration: Minimum of 3 months. Other: Recovery was indicated as "full recovery".		
Dobrescu, S. R. Et al. (2020) Design: Cohort Study Country: Gothenburg, Sweden Conducted: Gothenburg, Sweden Published: England, UK	Total (n): 98 AN (n): 47 Age: 44.42 ± 1.84 Control (n):51 Age: 44.22 ± 1.77 Participants were recruited Gothenburg, Sweden cohort* (study 5).	14.3 (10.0 - 17.2), without considering dropouts.	MINI 6.0., SCID-I, DSM-V Checklist GAF Morgan-Russell Scale SF-36	Physical: Sustained absence of weight deviation. Behavioural and/ or Psychological: Being free of all criterion symptoms of AN, BN and BED, absence of compensatory behaviours, and deviant attitudes regarding weight and shape, including weight phobia. Scale: Not specified. Duration: Minimum of 6 months. Other: Recovery was indicated as "full recovery".	30 years (30.13 ± 1.62)	30 out of 47 AN patients recovered (%64).
Dinkler, L. et al. (2019) Design: Cross-sectional study Country: Gothenburg, Sweden Conducted: Gothenburg, Sweden Published: New Jersey, USA	Total (n): 57 Recovered AN Group without ASD (n): 20 AGE: 44.2 ± 1.9 Recovered AN Group with ASD (n): 6 Age: 44.9 ± 2.3 Comparison Group (n): 31 Age: 44.2 ± 1.7	Recovered AN Group without ASD: 14.1 ± 1.9 Recovered AN Group with ASD Age of AN onset: 15.2 ± 1.2	MINI 6.0, SCID-I, DSM-V Checklist GAF Morgen-Russell Scale SF-36 WAIS-R Facial Emotional Recognition Task	Physical: Sustained absence of weight deviation (BMI > 17.5 kg/m2). Behavioural and/ or Psychological: Being free of all criterion symptoms of AN, BN and BED, absence of compensatory behaviours, and deviant attitudes regarding weight and shape, as well as weight phobia. Scale: Not specified. Duration: Minimum of 6 months	30 years (30.13 ± 1.62)	N/A

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
	Participants were recruited from Gothenburg, Sweden cohort* (study 5).	Recovered AN Group: 14.4 ± 1.8		Other: Recovery was indicated as “long term recovery”.		
Kurotori et al. (2019) Design: Cohort Study (retrospective chart review) Country: Japan Conducted: New Zealand	Total (n): 92 R-AN (n): 79 Age at admission: 12.7 ± 1.4 ARFID (n): 13 Age at admission: 10.7 ± 2.5 Participant related knowledge was gathered from Jichi Children’s Medical Centre, Tochigi’s records between April 2007 and March 2017.	Not specified as years, but enough knowledge was provided to estimate.	N/A	Physical: BMI > 3rd percentile (based on the Japanese Society for Paediatric Endocrinology [JSPE] BMI-for-age growth table) Behavioural and/ or Psychological: Meeting the full remission criteria (restoring behavioural and eating patterns and maintaining target weight for ≥2 weeks) or none of the ARFID or R-AN criteria according to the DSM-5. Scale: Not specified. Duration: Minimum of 6 months, after which no further treatment was provided. Other: Recovery was indicated as “full recovery”.	ARFID: 15.3 ± 9.1 months AN: 18.4 ± 15.0 months	ARFID: 10 out of 13 patients recovered (77%). AN: 34 out of 79 R-AN patients recovered (43%).
Halvorsen, I. et al. (2018) Design: Cross-sectional study Country: Oslo, Norway Conducted: England, UK Published:	AN (n): 37 Age at admission: 15.7 ± 1.9 Age at follow-up: 20.2 ± 2.6 Participants that received inpatient Family Based Treatment between May 2008 and June 2014 were recruited from the Regional Department for Eating Disorders at Oslo University Hospital.	13.2 ± 1.9	MINI 6.0 EDE 16 EDE-Q The Clinical Impairment Assessment The Beck Depression Inventory The State and Trait Anxiety Inventory	Physical: BMI ≥ 18.5 kg/m2 Behavioural and/ or Psychological: No episodes of binge eating or purging/other compensatory behaviour. Scale: EDE-Q global score ≤2.5. Duration: Minimum of 3 months. Other: Recovery was specified differently for weight and full recovery. The criteria given above emphasises “full recovery”.	4.5 ± 1.8 (1.3 - 7.1) years	12 out of 37 patients recovered (36%).

Articles Section

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
Swenne, I. et al. (2017) Design: Cohort Study Country: Uppsala, Sweden Conducted: Uppsala, Sweden Published: England, UK	AN (except weight-criteria) (n): 201 Age: 15.0 ± 1.7 Duration of symptoms (months): 9.6 ± 8.7 (2-60) Participants were assessed from the Eating Disorder Unit at the Department of Child and Adolescent Psychiatry of the Uppsala University Hospital's records (from August 2010 to July 2015).	Not specified as years, but enough knowledge was provided to estimate.	EDE-Q MADRS-S	Physical: Not specified. Behavioural and/ or Psychological: Not fulfilling criteria for an ED at a clinical interview. Scale: EDE-Q score < 2.0. Duration: Not specified.	1 year	According to the EDE-Q, 130 (65%) patients recovered at 1 year. However, the number of recovered patients were found as 106 (53%), when clinical interviews were preferred to decide recovery status.
Mustelin, L. et al. (2016a) Design: Cross-sectional (prevalence) study Country: Helsinki, Finland Conducted: Helsinki, Finland Published: England, UK	AN (n): 92 Age: 24.4 ± 0.9 Participants were recruited from the FinnTwin16 birth cohort*** (wave 4). Mean age represents all the members of wave 4 (n=2825). This studies' participants current ages weren't indicated.	17.4	SCID-I, DSM-V Checklist RAPI Mm-MAST	Physical: Restoration of weight and menstrual function (if applicable). Behavioural and/ or Psychological: Absence of binges and purges. Scale: Not specified. Duration: Minimum of a year. Other: Recovery was indicated as "clinical recovery".	5 years	72% of the patients recovered.
Mustelin, L. et al. (2016b) Design: Cross-sectional study (community-based) Country:	Total (n): 182 AN (n): 92 BN (n): 58 Participants were recruited from the FinnTwin16 birth cohort*** (wave 1,4, and 5).	18 Age-onset was reported at wave 1.	SCID-I, DSM-V Checklist	Physical: Restoration of weight and menstrual function (if applicable). Behavioural and/ or Psychological: Absence of binges and purges. Scale: Not specified. Duration: Minimum of a year.	2 follow-ups were conducted: 1. When the patients are 22-27 years	120 out of 182 women recovered (66%). (Recovered AN and BN patients were not specified.)

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
Conducted: Helsinki, Finland Published: New Jersey, USA	Mean age was 24.4 ± 0.9 at wave 4 and 34.8 ± 1.2 at wave 5.			Other: Recovery was indicated as “clinical recovery”.	(approximately 6-11 years after). 2. When the patients are 31-37 years (approximately 15-21 years after).	
Talbot, A. et al. (2015) Design: Cross-sectional study Country: Australia Conducted: Australia Published: USA	Total (n): 92 AN (n): 24 Age: 21.0 (18-27) Weight Recovered Group (n): 10 Age: 21.5 (19-25) Fully Recovered Group (n): 15 Age: 24.0 (21-32) Healthy Control Group (n): 43 Age: 19.0 (18-25) Patients were recruited from 2 university hospitals – no further information related to hospitals given.	AN Group: 16.5 (15–19) Weight-Recovered Group: 15.0 (13–16) Fully Recovered Group: 16.0 (14–17)	EDE-Q RCFT MFFT NART DASS-21 PI-WSUR FMPS EDQOL	Physical: BMI > 18.5 kg/m ² Behavioural and/ or Psychological: No longer meeting DSM-IV-TR and DSM-5 criteria for AN, not having ED-behaviours (binging, purging, restricting, and driven or compulsive exercise) Scale: EDE-Q within one standard deviation on all sub-scales of population norms. Duration: Minimum of 3 months. Other: Recovery was indicated as “full recovery”.	N/A	15 out of 49 AN patients recovered (30.6%).
Silen, Y. et al. (2015) Design: Cross-sectional study	Total (n): 47 Admission Age: 14.6 ± 1.2 Typical AN (n): 34 Admission Age: 14.3 ± 1.0	Not specified as years, but enough knowledge was	Morgan-Russell Scale	Physical: Weight maintained at BMI ≥ 17.5 kg/m ² or weight within 15% IBW and regular menstrual cycles Behavioural and/ or Psychological: Not specified.	N/A	60% of the patients recovered.

Articles Section

STUDY	PARTICIPANTS	AGE of ED ONSET	MEASUREMENT TOOLS	RECOVERY CRITERIA	FOLLOW UP	RECOVERY RATE
Country: Conducted: Helsinki, Finland Published: England, UK	Atypical AN (n): 13 Admission Age: 15.2 ± 1.5 Participant related knowledge was gathered from the Helsinki University Central Hospital records (from March 2012 to April 2013). Due to the fact that the data were examined retrospectively, current ages of the patients are unknown.	provided to estimate.		Scale: ‘Good outcome’ on Morgan and Russell Scale Duration: Stated as ‘over many consecutive months’		

AN: Anorexia Nervosa, ARFID: Avoidant and Restrictive Food Intake Disorder, ASD: Autism Spectrum Disorder, BED: Binge Eating Disorder, BN: Bulimia Nervosa, EDE: Eating Disorder Examination, EDE-Q: Eating Disorder Examination Questionnaire, IBW: Ideal Body Weight, N/A: Not Applicable, OSFED: Other Specified Feeding and Eating Disorders, TD: Typically Developing Control Group, UFED: Unspecified Feeding and Eating Disorders

*Gothenburg AN study covers patients who were born in 1970 and adjacent years, attending eighth grade at the time of the original study. All participants were examined for AN, and 51 AN cases were detected. In addition, nurses were asked to constitute a healthy control group – matching for age, sex, schooling and without AN diagnosis (n=51). Both groups were followed after 6 (study 2, mean age 21 years), 10 (study 3, mean age 24 years), and 18 (study 4, mean age 32 years) years. Study 5 was conducted after 30 years by Dobrescu et al. in 2020 when the mean age was 44 years.

**The FinnTwin12 Birth Cohort covers Finnish twins who were born in 1983 – 1987, followed at ages 12 (wave 1), 14 (wave 2), 17.5 (wave 3) and 22 (wave 4) years.

***The FinnTwin16 Birth Cohort covers Finnish twins who were born in 1975-1979 and their families, followed when the twins are 16 (wave 1), 17 (wave 2), 22-27 (wave 3) and 31-37 (wave 4) years.

Table 4. The Scrutinization of Recovery Criteria

	Breithaupt, L., et al. (2022)	Wentz, E., et al. (2021)	Kerr-Gaffney, J. et al. (2021)	Castro-Fornieles, J. Et al. (2021)	Silen et al. (2021)	Garriz, M. et al. (2021)	Dobrescu, S. R. et al. (2020)	Dinkler, L. et al. (2019)	Kurotori et al. (2019)	Halvorsen, I. et al. (2018)	Swenne, I. et al. (2017)	Mustelin, L. et al. (2016a)	Mustelin, L. et al. (2016b)	Talbot, A. et al. (2015)	Silen, Y. et al. (2015)
Physical	-	+	+	-	+	+	+	+	+	+	-	+	+	+	+
Behavioural and Psychological Scale	+	+	-	+	+	+	+	+	+	+	+*	+	+	+	-
Duration	-	+	+	-	-	+	+	+	+	+	-	+	+	+	+

*Behavioural and psychological criteria evaluation was based on patients’ perspectives.

Key Findings

Recovery Criteria/ Definition

The definition of recovery was mostly given in the studies that aim to distinguish any knowledge related to recovered patients from AN compared with the healthy controls and/or currently ill patients or in the studies that follow the outcomes of AN. Recovery has been defined by physical and/or behavioural and cognitive dimensions, as summarized in Table 4. While various BMI levels were referenced for explaining the physical aspects of recovery, clinical interviews that looked at the presence of symptoms and scales related to ED behaviours and cognitions were used for the behavioural and psychological ones. In addition, although some studies stipulated that the criteria they mentioned for recovery should continue for a certain period of time, some did not.

Physical Dimension

Three of the studies did not mention the physical dimensions of the recovery (Breithaupt et al., 2022; Castro-Fornieles et al., 2021; Swenne et al., 2017) whereas two of them only set physical criteria for defining the recovery (Kerr-Gaffney et al., 2021; Silén et al., 2015). According to Kerr-Gaffney et al. (2021), recovered patients should have a BMI between 18.5 - 27 kg/m² or %IBW more than 85% (for participants under 18 years). When it comes to Silen et al. (2015), the recovery definition was based on the good outcome from the Morgan-Russell Scale, which requires a BMI equal to or more than 17.5 kg/m² (BMI \geq 17.5 kg/m²) or weight within 15% of the ideal body weight together with the regular menstrual cycles.

In addition to Silen et al. (2015), only two studies required restoration of menses for applicable cases and physical criteria for recovery were defined as restoration of weight and menstrual function (Mustelin et al., 2016a; Mustelin et al., 2016b).

In three studies, the physical aspect of recovery was explained as a lack of weight deviation (Dinkler et al., 2019; Dobrescu et al., 2020; Wentz et al., 2021). Although Wentz et al. (2021) and Dobrescu et al. (2020) did not point out any specific numerical value to make that criterion more definitive, Dinkler et al. (2019) explained the absence of weight deviation with a BMI of more than 17.5 kg/m² (BMI $>$ 17.5 kg/m²). Furthermore, four studies required slightly higher BMI values. While Garriz et al. (2021) necessitated a BMI of more than 18 kg/m² (BMI $>$ 18kg/m²) for the physical aspect of recovery; Silen et al. (2021), Halvorsen et al. (2018) and Talbot et al. (2015) preferred BMI equal to or more than 18.5 kg/m² (BMI \geq 18.5 kg/m²).

Lastly, Kurotori et al. (2019) correlated BMI with percentiles rather than numerical setpoints due to the younger-aged sample characteristics and maintaining a BMI over the 3rd percentile, as present in the Japanese BMI-for-age-growth-chart that relies on age and gender, was accepted as the physical part of the recovery (Kurotori et al., 2019).

Behavioural and Psychological Dimension in conjunction with related scales

Since two of the studies emphasised only physical aspects of the recovery (Kerr-Gaffney et al., 2021; Silén et al., 2015), there are 13 studies left that examined behavioural and psychological dimensions of being recovered and 8 out of 13 studies required the absence of diagnosis as the criterion for recovery.

Talbot et al. (2015) expressed this requirement as no longer meeting the DSM-IV and DSM-5 diagnosis for AN and supported their criteria by requiring not having ED behaviours (such as restricting, bingeing, purging, driven and compulsive exercising) and EDE-Q scores within one standard deviation on all sub-scales of population norms. However, when it comes to Kurotori et al. (2019), recovery was explained as not meeting the criteria for AN-Restrictive and ARFID and achieving full remission, which means restoring behavioural eating patterns. Dinkler et al. (2019), Dobrescu et al. (2020) and Wentz et al. (2021) required being free from all diagnostic criteria of AN, BN and BED, rather than just not meeting the criteria for AN. In these three studies, exemption of diagnosis was also supported by the requirement of a lack of compensatory behaviours and deviation in weight and shape-related attitudes (including weight phobia).

In addition, Swenne, Parling and Salonen (2017), Garriz et al. (2021) and Castro-Fornieles (2021) broadened the absence of diagnosis to not fulfilling any of the EDs criteria. Thus, the lesser-mentioned subtypes of ED like Other Specified Feeding or Eating Disorder (OSFED), UFED (Unspecified Feeding or Eating Disorder) and Avoidant/Restrictive Food Intake Disorder (ARFID) were covered. Furthermore, in addition to having a lack of ED diagnosis, Swenne, Parling and Salonen (2017) required the EDE-Q global score of less than 2.0 and Garriz et al. (2021) anticipated being within 1 SD in all EDE subscales to be accepted as recovered. Moreover, Halvorsen et al. (2018) also required the combination of the absence of bingeing and compensatory behaviours with EDE-Q scores less than or equal to 2.5 SD from the recovered patients.

Unlike other studies, Breithaupt et al. (2021) elected LIFE-EAT-3 PSR Score equal to one or two in addition to the absence or residual symptoms and lack of ED-related functional impairment to define recovery criteria. In addition, in Silen et al.'s study (2021), the information related to behavioural and psychological aspects of recovery was taken by asking the patients whether they were feeling recovered or not rather than yielding on structured interviews or questionnaires.

Lastly, the physical recovery criteria were supported by only the behavioural aspects of recovery (absence of bingeing and purging), and the psychological aspects were not considered in the remaining two studies (Mustelin et al., 2016a; Mustelin et al., 2016b).

Duration of Criteria

11 studies demanded continuity of improvements (in various recovery dimensions) over a certain period – fluctuating from 3 months to a year – to accept patients as recovered.

From the physical point of view, the maintenance of weight restoration was requested for six months and a year by Kurotori et al. (2019) and Kerr-Gaffney et al. (2021), respectively. In addition, according to Kurotori et al. (2019), the countdown of these six months should be started after the treatment is terminated. On the other hand, Silen et al. (2015) required the continuity of weight restoration over many consecutive months instead of being more specific about providing duration.

Some studies emphasised the requirement of duration related to the behavioural dimension of recovery. In more detail, Garriz et al. (2021), Halvorsen et al. (2018) and Talbot et al. (2015) required the absence of ED-related behaviours (i.e. bingeing, purging, restricting, compensatory behaviours) for a minimum of 3 months, whereas Mustelin et al. (2016a) and Mustelin et al. (2016b) necessitated a year to consider patients as recovered. On the other hand, Wentz et al. (2021), Dobrescu et al. (2020) and Dinkler et al. (2019) were more stringent. They required a minimum of 6-month ED diagnosis exemption when defining the recovery criteria's behavioural and psychological aspects.

Length of Follow-Up and Recovery Rate

Due to the variances in study designs and aims, reporting follow-up years and recovery rates could not become applicable to all studies. The appropriate studies covered both short-term and long-term follow-up durations, which range from less than a year (Breithaupt et al., 2022) to 30 years (Dinkler et al., 2019; Dobrescu et al., 2020; Wentz et al., 2021).

In addition to Breithaupt et al. (2022), the studies of Swenne, Parling and Salonen (2017), Kurotori et al. (2019), Silen et al. (2021), Halvorsen et al. (2018) and Mustelin et al. (2016a) can be accepted as having a short-term follow-up with the mean follow-up years of 1, 1.5 (18 months), 4, 4.5 and 5 years, respectively. In contrast, Castro-Fornieles et al. (2021) and Garriz et al. (2019) followed the patients later on an average of 20 and 22 years, which can be considered long-term.

On the other hand, participants in Mustelin et al.'s study (2016a) were followed up two times: 1) when they were 22-27 years and 2) when they were 31-37 years. As the baseline age was given as 16 years, the follow-up years were estimated between 6-11 years for the first and 15-21 years for the second ones (Mustelin et al., 2016a).

The recovery rate was calculated manually in three studies (Castro-Fornieles et al., 2021; Gárriz et al., 2021; Talbot et al., 2015) as they gave the number of recovered patients in the AN group instead of giving a percentage. As a result of this, the recovery rate from AN was found to fluctuate from 30.6% (Talbot et al., 2015) to 72% (Mustelin et al., 2016b).

The second highest rate was seen in Mustelin et al.'s study (2016a) at 66%, while the second lowest rate was 36% (Halvorsen et al., 2018). In addition, recovery rates in Wentz et al. (2021) and Castro-Fornieles et al. (2021)'s studies were similar, at rates of 58.8% and 57.69%, respectively. However, in Silen et al. (2021) and Kurotori et al. (2019)'s studies the recovery rates were slightly lower as follows: 41.5 and 43%, unlike the studies of Garriz et al. (2021), Dobrescu et al. (2020) and Silen et al. (2015) with slightly higher recovery rates: 62.02%, 64% and 60%.

In addition, Swenne, Parling and Salonen (2017) reported that recovery rates varied depending on how they were evaluated. In more detail, while 65% of patients recovered when EDE-Q was used, this rate decreased to 53% when a clinical interview was preferred.

Discussion

The primary aim of the study was to present the criteria of recovery that were preferred by researchers while defining patients with adolescent-onset AN as recovered. As seen in the 15 reviewed papers, the researchers set their own criteria for recovery while conducting studies due to a lack of a standardised definition of it.

Two studies (Kerr-Gaffney et al., 2021; Silén et al., 2015) mentioned only the physical dimension of recovery, whereas the rest of the 13 studies had a broadened perspective and covered behavioural and psychological dimensions in addition to the restoration of weight and menses. Furthermore, the measurement tools that were used while assessing ED pathology, course and/or outcome have changed between studies. Although EDE-Q (Luce and Crowther, 1999) became prominent compared to the other measurement tools (i.e. LIFE-EAT-3 (Breithaupt et al., 2022; Keller et al., 1997), Morgan-Russell Scale (Morgan & Hayward, 1988)), it has been used by only six studies asking for different scores. While some studies required EDE-Q scores within one standard deviation (SD) on all sub-scales of population norms (Gárriz et al., 2021; Talbot et al., 2015) for some 2 (Swenne et al., 2017) and 2.5 (Halvorsen et al., 2018) SDs were thought adequate enough to accept patients as recovered. In addition, only one study emphasized the patients' perspective on recovery status by asking whether the patients thought of themselves as recovered or not (Silén et al., 2021).

In this systematic review, the recovery rate fluctuated from 30.6% to 72% as regards the applicable studies. The study with the highest rate of recovery (72%) defined recovery as an absence of binges and purges in addition to the restoration of weight and menses (for applicable cases) at least for a year and patients were followed up after an average of five years than their onset of AN (Mustelin et al., 2016b). In contrast, the lowest rate (30.6%) is seen the Talbot et al.'s study (2015), where the recovery was defined more inclusively.

Physical Dimension of Recovery

Like many of the studies included in this review, previous studies (Bardone-Cone et al., 2010; Dawson et al., 2013); suggest incorporating restoration of weight and menstrual cycle into a definition of recovery – but not solely, only as a part of a physical dimension of a more comprehensive definition. Although the reliability and generalisability of BMI for a physical recovery decision are highly questionable (Dawson et al., 2013), BMI (for adults) and percentage weight (for children and adolescents) were preferred mostly to express physical recovery (Bardone-Cone et al., 2010).

As seen in the eligible studies of this review, BMI was required between 17.5 kg/m² to 18.5 kg/m². These cut-offs are similar to the ones that were recommended in Bardone-Cone, Hunt and Watson's overview (2018) and Khalsa et al's systematic review (2017), except for some included studies that recommended a BMI of more than 19 kg/m² or 20 kg/m². In Dawson et al.'s study (2013), the experts' thoughts were divided when they asked whether the BMI should be equal to 18.5 kg/m² or 20 kg/m² for defining the physical aspect of recovery from AN. However, at the end of that study, a BMI equal to or higher than 18.5 kg/m² was agreed upon by most (92%) experts in this field (Dawson et al., 2013).

When it comes to the restoration of menses, although most of the studies did not mention the return of menses, it was stipulated in 3 (out of 15) studies in this review, raising doubts on whether it is necessary. The return of menses was perceived as an unreliable factor for recovery due to the existence of women with amenorrhea without substantial weight loss and also emaciated ones without amenorrhea (due to personal differences or oral contraceptives) (Couturier & Lock, 2006; Dawson et al., 2013; Watson & Andersen, 2003).

Dawson, Rhodes and Touyz (2013) suggest evaluating BMI cut-offs and the requirement for regular menstruation individually. Considering that patients with AN symptomatology experience severe distress regardless of their weight status (Watson & Andersen, 2003), showing regard to the AED's biologically appropriate weight criteria: a) absence of restricting or dieting, bingeing, obsessive and compulsive exercising, b) supporting normal functioning and growth and c) being consistent with pre-morbid weight, gender, ethnicity, family history) instead of specified BMI cut-offs can be useful (Academy for Eating Disorders Nutrition Working Group).

Behavioural and Psychological Dimension of Recovery

Recovery was explained based on physical improvements previously; however, due to the continuity of symptoms and dysfunctions after physical recovery, studies agreed on supporting recovery criteria by including behavioural and psychological dimensions (Bardone-Cone et al., 2010; Couturier & Lock, 2006; Dawson et al., 2013; Khalsa et al., 2017). In this review, although recovery was

defined based on weight solely in two out of 15 studies, the rest of the 13 studies supported the idea of including behavioural and/or psychological improvements in the recovery criteria.

When it comes to the absence of a diagnosis criterion, a lack of consensus can also be seen in how it was defined. For instance, some studies required being free from criteria from AN only (Talbot et al., 2015); whereas, some preferred not meeting any criteria for AN, BN and BED (Dinkler et al., 2019; Dobrescu et al., 2020; Wentz et al., 2021) or EDs in general (Bardone-Cone et al., 2010; Castro-Fornieles et al., 2021; Gárriz et al., 2021; Swenne et al., 2017). This difference may be explained as related to participants' diagnosis of the studies. For example, the studies that research AN-BP may need to emphasize the lack of bingeing, in contrast to ones with AN-R that need to focus on restricting. On the other hand, there is also one questionable point on whether the criteria of recovery should be transdiagnostic or AN-specific. Bardone-Cone et al. (2010) intentionally proposed a transdiagnostic definition, due to the frequency of the transition between ED subtypes. When the transition/ relapse of AN is considered (Andrés-Pepiñá et al., 2020; Dobrescu et al., 2020; Wentz et al., 2021), it would be more useful to yield transdiagnostic criteria than AN-specific ones.

Assessment

Another point that requires clarity is the way of assessment. The studies mostly – but not restrictively – use EDE and its' questionnaire version, EDE-Q; however, no agreement was achieved on cut-offs that may present recovery. Some researchers focused on the global score of the EDE-Q (Halvorsen et al., 2018; Swenne et al., 2017), and some required cut-offs for each subscale (Gárriz et al., 2021; Talbot et al., 2015). Also, although this review presented that the required cut-offs ranged from being within 1 SD to 2.5 SD (Gárriz et al., 2021; Halvorsen et al., 2018; Swenne et al., 2017; Talbot et al., 2015), having scores of 2 SDs away from societies' norms was criticized for allowing continuity of pathologies (Wade & Lock, 2020). Apart from these, Swenne, Parling and Salonen (2017) emphasized the differences in recovery rates regarding the preference for clinical interviews (53%) or questionnaires (65%). To avoid misinterpretation, in-person examinations and questionnaires should be combined during recovery assessment (Dawson et al., 2013).

Duration

While defining recovery, it is suggested to cover a specific time frame in which improvements are sustained (Dawson et al., 2013). When the experts were asked to define the difference between remission and recovery, time was defined as the main difference (Dawson et al., 2015). But, how long duration is required to recover? Apart from the ones that do not specify a time frame, the suggestions ranged

from some consecutive months or at least three months to a year of the reviewed articles. According to Dawson et al.'s study (2015), three years were advised by 15% of experts (fellows of AED and members of EDRS); however, more than a year was suggested by the EDRS, more recently (Wade & Lock, 2020).

Beyond the Physical, Behavioural and Psychological Dimensions

Lack of recovery criteria related to general psychology

The studies that were included in this review did not emphasize the importance of general psychology or social functionality for assessing recovery. However, Dawson et al. (2015) conducted a study that sought an answer to the question about the best method to measure recovery. At the end of their study, experts from all around the world (fellows of the AED and members of the EDRS) agreed on the importance of incorporating psychosocial functioning and quality of life into measurements in addition to physical, behavioural and psychological dimensions (Dawson et al., 2013). More recently, the definition of recovery was discussed in the EDRS annual meeting, and experts also agreed to incorporate quality of life assessments (in terms of psychological, social and emotional domains) into a recovery definition (Wade & Lock, 2020). It might be advisable to reconsider the possible pros and cons of including general psychology. Currently, both the routine (66.4%) and urgent cases (59%) rates of receiving timely access to ED treatment are under national standards (95%) (Nuffieldtrust, 2023). Although including improvements in general psychology may decrease relapse rates (Rodriguez et al., 2005), prolonged therapy durations could cause further reduced access to treatment for the cases on waiting lists. At this point, it may be considered as a suggestion to perform two separate evaluations, one for ED pathology and the other for general psychology, and to plan the referral of patients with problematic scores related to general psychology but without ED pathology from ED-focused units to other relevant treatment units.

Limitations

There are some limitations to the methodology of this study. This review included only the studies with patients that are diagnosed with AN during their adolescent period. Future studies can reach more definitions of recovery used by the researchers by removing this criterion. Also, eligibility criteria remained incapable of restricting other factors that may affect recovery rates, such as ages at treatment and/ or follow-up, the type of received treatments and admission settings, the level of BMI at admission and discharge, symptom severity, personal traits, AN subtypes like AN-BP or AN-R (Andrés-Pepiñá et al., 2020; Breithaupt et al., 2022; Couturier & Lock, 2006; Fichter et al., 2017; Glasofer et al., 2020). Further studies interested in the recovery rate should also take into account the variations that may occur due

to these factors. Lastly, there might be unintentionally unnoticed studies during the screening process. Although each part of the review was carried out under the supervision of experts in the field, the study selection processes were conducted by only one author.

Conclusion

In the current circumstances, there is no consensus on defining recovery from AN. As seen in the reviewed articles, different definitions were used to accept patients as recovered such as evaluating only physical and/ or behavioural dimensions, setting varied cut-off points for scale-related measurements or preferring different expressions due to the differences in assessment tools. Perhaps the absence of clear indicators is a natural outcome of the myriad perspectives available to us in approaching the process of recovery. Although the researchers have conducted studies to build a consensus to define recovery and agreed on a multidimensional definition, none of their suggestions was placed in ED-related guidelines, resulting in a continuous lack of consensus. Until reaching an agreement, researchers interested in recovered patients or the outcomes of AN should be aware that inconsistencies in recovery definition can affect the results of their research.

Authors' note

Funding information: There are no funders to report for this submission.

Acknowledgements: No acknowledgement.

Conflict of interest: The authors declare no potential conflict of interest.

References

- Academy for Eating Disorders Nutrition Working Group. *Guidebook for Nutrition Treatment of Eating Disorders*. Academy for Eating Disorders https://higher.logicdownload.s3.amazonaws.com/AEDWEB/27a3b69a-8aae-45b2-a04c-2a078d02145d/UploadedImages/Publications_Slider/FINAL_AED_Purple_Nutrition_Book.pdf
- Academy of Nutrition and Dietetics. (2022, 11/09/2022). *EVIDENCE ANALYSIS MANUAL*. Academy of Nutrition and Dietetics. Retrieved 17/11/2023 from <https://www.andeal.org/evidence-analysis-manual>
- American Psychiatric Association, D., & Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Vol. 5). American psychiatric association Washington, DC.

- Andrés-Pepiñá, S., Plana, M. T., Flamarique, I., Romero, S., Borràs, R., Julià, L., Gárriz, M., & Castro-Fornieles, J. (2020). Long-term outcome and psychiatric comorbidity of adolescent-onset anorexia nervosa. *Clinical Child Psychology and Psychiatry*, 25(1), 33-44.
- Bachner-Melman, R., Lev-Ari, L., Zohar, A. H., & Lev, S. L. (2018). Can recovery from an eating disorder be measured? Toward a standardized questionnaire. *Frontiers in psychology*, 9, 2456.
- Bachner-Melman, R., Zohar, A. H., & Ebstein, R. P. (2006). An examination of cognitive versus behavioral components of recovery from anorexia nervosa. *The Journal of nervous and mental disease*, 194(9), 697-703.
- 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: Conceptualization, validation, and examination of psychosocial functioning and psychiatric comorbidity. *Behaviour Research and Therapy*, 48(3), 194-202.
- 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, 1-18.
- Breithaupt, L., Kahn, D. L., Slattery, M., Plessow, F., Mancuso, C., Izquierdo, A., Dreier, M. J., Becker, K., Franko, D. L., & Thomas, J. J. (2022). Eighteen-month course and outcome of adolescent restrictive eating disorders: Persistence, crossover, and recovery. *Journal of Clinical Child & Adolescent Psychology*, 51(5), 715-725.
- Castro-Fornieles, J., de la Serna, E., Calvo, A., Pariente, J., Andrés-Perpiña, S., Plana, M. T., Romero, S., Flamarique, I., Gárriz, M., & Bargalló, N. (2021). Cortical thickness 20 years after diagnosis of anorexia nervosa during adolescence. *European Archives of Psychiatry and Clinical Neuroscience*, 271, 1133-1139.
- Accurso, E. C., Ciao, A. C., Fitzsimmons-Craft, E. E., Lock, J. D., & Le Grange, D. (2014). Is weight gain really a catalyst for broader recovery?: The impact of weight gain on psychological symptoms in the treatment of adolescent anorexia nervosa. *Behav Res Ther*, 56, 1-6. <https://doi.org/10.1016/j.brat.2014.02.006>
- Couturier, J., & Lock, J. (2006). What is recovery in adolescent anorexia nervosa? *International Journal of Eating Disorders*, 39(7), 550-555.
- Dawson, L., Rhodes, P., & Touyz, S. (2013). Defining recovery from anorexia nervosa: A Delphi study to explore practitioners' views. *Journal of Eating Disorders*, 1(1), 1-1.
- de Vos, J. A., LaMarre, A., Radstaak, M., Bijkerk, C. A., Bohlmeijer, E. T., & Westerhof, G. J. (2017). Identifying fundamental criteria for eating disorder recovery: a systematic review and qualitative meta-analysis. *Journal of Eating Disorders*, 5(1), 1-14.
- Dinkler, L., Rydberg Dobrescu, S., Råstam, M., Gillberg, I. C., Gillberg, C., Wentz, E., & Hadjikhani, N. (2019). Visual scanning during emotion recognition in

- long-term recovered anorexia nervosa: An eye-tracking study. *International Journal of Eating Disorders*, 52(6), 691-700.
- Dobrescu, S. R., Dinkler, L., Gillberg, C., Råstam, M., Gillberg, C., & Wentz, E. (2020). Anorexia nervosa: 30-year outcome. *The British Journal of Psychiatry*, 216(2), 97-104.
- Fichter, M. M., Quadflieg, N., Crosby, R. D., & Koch, S. (2017). Long-term outcome of anorexia nervosa: Results from a large clinical longitudinal study. *International Journal of Eating Disorders*, 50(9), 1018-1030.
- Gárriz, M., Andres-Perpina, S., Plana, M. T., Flamarique, I., Romero, S., Julià, L., & Castro-Fornieles, J. (2021). Personality disorder traits, obsessive ideation and perfectionism 20 years after adolescent-onset anorexia nervosa: a recovered study. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 26, 667-677.
- Glasofer, D. R., Muratore, A. F., Attia, E., Wu, P., Wang, Y., Minkoff, H., Rufin, T., Walsh, B. T., & Steinglass, J. E. (2020). Predictors of illness course and health maintenance following inpatient treatment among patients with anorexia nervosa. *Journal of Eating Disorders*, 8, 1-10.
- Graell, M., de Andrés, P., Sepúlveda, A. R., Moreno, A., Villaseñor, Á., Faya, M., Martínez-Cantarero, C., Gómez-Martínez, S., Marcos, A., Morandé, G., & Nova, E. (2018). The adolescent onset anorexia nervosa study (ANABEL): Design and baseline results. *Int J Methods Psychiatr Res*, 27(3), e1739. <https://doi.org/10.1002/mpr.1739>
- Halvorsen, I., Reas, D. L., Nilsen, J. V., & Rø, Ø. (2018). Naturalistic outcome of family-based inpatient treatment for adolescents with anorexia nervosa. *European eating disorders review*, 26(2), 141-145.
- Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P., & Andreasen, N. C. (1987). The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry*, 44(6), 540-548. <https://doi.org/10.1001/archpsyc.1987.01800180050009>
- Kerr-Gaffney, J., Hayward, H., Jones, E. J., Halls, D., Murphy, D., & Tchanturia, K. (2021). Autism symptoms in anorexia nervosa: A comparative study with females with autism spectrum disorder. *Molecular Autism*, 12(1), 47.
- 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(1), 1-12.
- Kordy, H., Krämer, B., Palmer, R. L., Papezova, H., Pellet, J., Richard, M., Treasure, J., & Action, C. (2002). Remission, recovery, relapse, and recurrence in eating disorders: Conceptualization and illustration of a validation strategy. *Journal of clinical psychology*, 58(7), 833-846.
- Kurotori, I., Shioda, K., Abe, T., Kato, R., Ishikawa, S., & Suda, S. (2019). An inpatient observational study: characteristics and outcomes of

- avoidant/restrictive food intake disorder (ARFID) in children and adolescents in Japan. *Neuropsychiatric Disease and Treatment*, 3313-3321.
- Lock, J., Agras, W. S., Le Grange, D., Couturier, J., Safer, D., & Bryson, S. W. (2013). Do end of treatment assessments predict outcome at follow-up in eating disorders? *International Journal of Eating Disorders*, 46(8), 771-778.
- Luce, K. H., & Crowther, J. H. (1999). The reliability of the Eating Disorder Examination-Self-Report Questionnaire Version (EDE-Q). *Int J Eat Disord*, 25(3), 349-351. [https://doi.org/10.1002/\(sici\)1098-108x\(199904\)25:3<349::aid-eat15>3.0.co;2-m](https://doi.org/10.1002/(sici)1098-108x(199904)25:3<349::aid-eat15>3.0.co;2-m)
- Moore, C. A., & Bokor, B. R. (2023). Anorexia Nervosa. Retrieved 12/11/2023, from <https://fxflare.com/sample-of-differential-diagnosis-on-anorexia-nervosa-pdf#/word-and-definision-worksheets>
- Morgan, H., & Hayward, A. (1988). Clinical assessment of anorexia nervosa: the Morgan-Russell outcome assessment schedule. *The British Journal of Psychiatry*, 152(3), 367-371.
- Morgan, H., & 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.
- Mustelin, L., Latvala, A., Raevuori, A., Rose, R. J., Kaprio, J., & Keski-Rahkonen, A. (2016a). Risky drinking behaviors among women with eating disorders—A longitudinal community-based study. *International Journal of Eating Disorders*, 49(6), 563-571.
- Mustelin, L., Silén, Y., Raevuori, A., Hoek, H. W., Kaprio, J., & Keski-Rahkonen, A. (2016b). The DSM-5 diagnostic criteria for anorexia nervosa may change its population prevalence and prognostic value. *Journal of psychiatric research*, 77, 85-91.
- Nuffieldtrust. (2023). *Children and young people with an eating disorder waiting times*. Nuffieldtrust. Retrieved 12/11/2023 from <https://www.nuffieldtrust.org.uk/resource/children-and-young-people-with-an-eating-disorder-waiting-times#background>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery*, 88, 105906.
- Rodriguez, B. F., Bruce, S. E., Pagano, M. E., & Keller, M. B. (2005). Relationships among psychosocial functioning, diagnostic comorbidity, and the recurrence of generalized anxiety disorder, panic disorder, and major depression. *Journal of anxiety disorders*, 19(7), 752-766.
- Silén, Y., Raevuori, A., Jüriloo, E., Tainio, V. M., Marttunen, M., & Keski-Rahkonen, A. (2015). Typical versus atypical anorexia nervosa among adolescents: Clinical characteristics and implications for ICD-11. *European eating disorders review*, 23(5), 345-351.

- Silén, Y., Sipilä, P. N., Raevuori, A., Mustelin, L., Marttunen, M., Kaprio, J., & Keski-Rahkonen, A. (2021). Detection, treatment, and course of eating disorders in Finland: A population-based study of adolescent and young adult females and males. *European eating disorders review*, 29(5), 720-732.
- Steinhausen, H.-C. (2002). The outcome of anorexia nervosa in the 20th century. *American Journal of Psychiatry*, 159(8), 1284-1293.
- Swenne, I., Parling, T., & Salonen Ros, H. (2017). Family-based intervention in adolescent restrictive eating disorders: early treatment response and low weight suppression is associated with favourable one-year outcome. *BMC psychiatry*, 17, 1-10.
- Talbot, A., Hay, P., Buckett, G., & Touyz, S. (2015). Cognitive deficits as an endophenotype for anorexia nervosa: An accepted fact or a need for re-examination? *International Journal of Eating Disorders*, 48(1), 15-25.
- Wade, T. D., & Lock, J. (2020). Developing consensus on the definition of remission and recovery for research. *International Journal of Eating Disorders*, 53(8), 1204-1208.
- Watson, T., & Andersen, A. E. (2003). A critical examination of the amenorrhea and weight criteria for diagnosing anorexia nervosa. *Acta Psychiatrica Scandinavica*, 108(3), 175-182.
- Wentz, E., Gillberg, I. C., Anckarsäter, H., Gillberg, C., & Råstam, M. (2009). Adolescent-onset anorexia nervosa: 18-year outcome. *Br J Psychiatry*, 194(2), 168-174. <https://doi.org/10.1192/bjp.bp.107.048686>
- Wentz, E., Dobrescu, S. R., Dinkler, L., Gillberg, C., Gillberg, C., Blennow, K., Råstam, M., & Zetterberg, H. (2021). Thirty years after anorexia nervosa onset, serum neurofilament light chain protein concentration indicates neuronal injury. *European Child & Adolescent Psychiatry*, 30, 1907-1915.

