Thesis submitted in partial fulfilment for the degree of Doctorate in Clinical Psychology (DClinPsych)

Emma Kinnaird

Candidate Number: 475569

Oxford Institute of Clinical Psychology Training and Research

University of Oxford

May 2023

Contents

1. Summary of Word Counts 6

2. Abstracts 6

2.1 Systematic Literature Review 6

2.2 Service Improvement Project 7

2.3 Theory Driven Research Project 8

3. Systematic Literature Review: Exploring the monitoring and reporting of unwanted events in psychotherapy trials for anorexia nervosa 10

3.1 Abstract 11

3.2 Introduction 12

3.3 Methods 16

3.3.1 Design 16

3.3.2 Eligibility Criteria 17

3.3.4 Search Strategy 17

3.3.5 Selection Process 18

3.3.6 Data Extraction 18

3.3.7 Quality Assessment 19

3.4 Results 20

3.4.1 Definitions of Unwanted Events 21

3.4.2 Safety Monitoring 22

3.4.3 Treatment Non-Completion 23

3.4.4 Treatment Non-Response 25

3.4.5 Symptom Deterioration 26

3.4.6 Other Unwanted Events 27

3.5 Discussion 28

3.6 References 35

4. Service Improvement Project: Evaluating and addressing staff training needs in relation to autism in an older adult mental health service. 60

4.1 Abstract 61

4.2 Introduction 62

4.2.1 Summary of Relevant Literature 62

4.2.2 Outline of Service Improvement Project 65

4.2.3 Aims 66

4.3 Stage 1: Audit 66

4.3.1 Methods 66

4.3.2 Findings 67

4.4 Stage 2: Needs Survey 69

4.4.1 Methods 69

4.4.2 Findings 69

4.5 Stage 3: Training and Evaluation 70

4.5.1 Methods 70

4.5.2 Findings 71

4.6 Discussion 72

4.6.1 Limitations 76

4.6.2 Service Recommendations 77

4.7 Lay Summary 80

4.8 References 81

5. Theory Driven Research Project: Exploring the relationship between clinical and personal models of recovery in anorexia nervosa 84

5.1 Abstract 85

5.2 Background 86

5.3 Methods 89

5.3.1 Design 89

5.3.2 Participants 90

5.3.3 Measures 91

5.3.4 Analysis 93

5.4 Results 95

5.4.1 How far is there agreement between self-defined and clinically defined recovery? 96

5.4.2 Is there a difference between the clinically defined recovery groups on measures of personal recovery? 97

5.4.3 Is there a relationship between measures of clinical and personal recovery across the total sample? 98

5.4.4 Participant perspectives on clinical definitions of recovery 98

5.5 Discussion 103

5.5.1 Limitations 105

5.6 References 107

6. Executive Summary 110

6.1 Systematic Literature Review: Exploring the monitoring and reporting of unwanted events in psychotherapy trials for anorexia nervosa 110

6.2 Service Improvement Project: Evaluating and addressing staff training needs in relation to autism in an older adult mental health service 110

6.3 Theory Driven Research Project: Exploring the relationship between clinical and personal models of recovery in anorexia nervosa 111

7. Connecting Narrative 114

7.1 Service Improvement Project (SIP).....114

7.2 Systematic Literature Review (SRL) 115

7.3 Theory Driven Research Project (TDRP) 116

8. Acknowledgements 117

9. Appendices.....118

9.1 Appendix 1: Published Systematic Review 118

9. 2 Appendix 2: Journal Publishing Guidelines.....138

9. 3 Appendix 3: Service Improvement Project Survey Questions.....142

9. 4 Appendix 4: Ethics Approval Letter 144

9. 5 Appendix 5: Theory Driven Research Project Questionnaires 145

1. Summary of Word Counts

- Systematic Literature Review: 6,976 words.
- Service Improvement Project: 4,952 words.
- Theory Driven Research Project: 4,996 words.
- Executive Summary: 928 words.
- Connecting Narrative: 972 words.

2. Abstracts

2.1 Systematic Literature Review

Objective: If a psychotherapeutic intervention has the capacity to produce positive change, the "deterioration effect" theory holds that it must also have the potential for negative effects. However, the definition, measurement, and reporting of unwanted events in psychotherapy is a topic of ongoing discussion. At present this area is under-explored in interventions for anorexia nervosa (AN), a severe mental illness associated with high medical and psychiatric risks. The aim of this paper was to undertake a systematic review of published randomised control trials (RCTs) evaluating psychotherapeutic interventions for AN, and to assess how unwanted events were defined, monitored and reported alongside the trial's key findings. **Method:** Using a systematic review methodology, this paper identified 23 RCTs through database searches meeting eligibility criteria. Results are presented using a narrative summary approach. **Results:** Unwanted event reporting varied widely, both regarding definitions of key unwanted events (such as non-compliance or symptom deterioration) and in the amount of detail captured in each paper. **Discussion:** The review identified two key issues: firstly, a lack of consistent definitions and unclear causality made it difficult to distinguish between unwanted events, and adverse events caused by the

interventions. Secondly, it highlighted the difficulty of defining unwanted events where different studies have different populations and goals. Recommendations are made around how the area of defining, monitoring, and reporting unwanted events in RCTs for AN can be taken forward.

2.2 Service Improvement Project

Autistic adults are at higher risk of experiencing mental health problems compared to their neurotypical peers and are more likely to report unmet healthcare needs and poor service experiences. Research suggests that this is closely related to a lack of understanding of autism by clinicians working in this area. Therefore, there is an increasing impetus in the National Health Service (NHS) for staff working in mental healthcare to receive autism training. However, there is a lack of understanding around autism in older adulthood, and whether older adult mental health teams could benefit from specialised training that reflects the specific needs of this population. This project used clinical audit and needs survey methodologies to explore whether an older adult mental health team might benefit from autism training, and to understand what clinicians might find helpful from a training intervention. Whilst the audit identified that the team did see patients on the autism spectrum, it was rare for the team to work with patients with a diagnosis of autism. Rather, clinicians were more likely to work with patients where autism is suspected but unconfirmed. Most clinicians reported having never received training in autism, and highlighted autism in older adults, and autism and mental health, as specific training priorities. A training programme based on these responses was developed, implemented, and evaluated using pre- and post-training surveys. The training was rated as useful by attendees and resulted in non-statistically significant increases in self-rated knowledge. The findings of this study suggest that clinicians working in older adult mental health can benefit from autism training. However, instead of generic autism training, clinicians may particularly benefit from training

specialised to their older adult context. Further research is required to better understand the area of older adulthood and autism, and to evaluate how to effectively enhance the capability of staff working in this area.

2.3 Theory Driven Research Project

Background: Whilst recovery from anorexia nervosa (AN) has traditionally been defined in relation to behavioural symptoms, the presence of cognitions, and weight, there is increasing research interest around understanding recovery from AN from the perspective of personal models of recovery, which emphasise factors such as personal growth and leading a meaningful life. Whilst previous research has suggested that clinical and personal recovery factors may exist in a complimentary relationship, with reduced clinical symptoms enabling growth in personal domains, to date this association has not been empirically explored. Understanding this relationship is particularly important in the context of proposals that there is a need to develop a universal definition of recovery from AN. **Methods:** The current study used a mixed methods approach to explore the relationship between personal and clinical models of recovery in AN. Eighty-one adults who self-defined as fully or partially recovered from AN were recruited into an online study and completed self-report measures of clinical and personal recovery. Participants were also asked questions about their views on the personal relevance of clinical criteria, and this data was analysed using thematic analysis. **Results:** The study found that there was good agreement between self-defined and clinically defined full recovery from AN. People who were clinically defined as fully recovered scored significantly higher on measures of personal recovery compared to those defined as not recovered, and there were strong positive associations between lower behavioural and cognitive symptoms, and personal recovery. However, there was a weaker relationship between personal recovery measures and BMI. Finally, qualitative data indicated that most participants felt that the clinical criteria were not a sufficient definition of recovery. Whilst

participants suggested potential additional criteria, and raised ideas around how these might relate to traditional symptom-based approaches, the qualitative analysis also highlighted concerns around how far a universal recovery criteria-based approach can capture individual experience. **Conclusions:** Implications for the development of future universal clinical criteria for recovery from AN are discussed, including how a clinical approach can best be combined with personal understandings of recovery.

3. Systematic Literature Review: Exploring the monitoring and reporting of unwanted events in psychotherapy trials for anorexia nervosa

Emma Kinnaird (emma.kinnaird@hmc.ox.ac.uk)

Supervised by Dr Myra Cooper.

Word count: 6,976

First submitted October 2022.

This paper has been published by the International Journal of Eating Disorders (Appendix 1) and has been formatted in line with their guidelines (Appendix 2). This journal was selected as it is the highest impact journal in the eating disorder field.

3.1 Abstract

Objective: If a psychotherapeutic intervention has the capacity to produce positive change, the "deterioration effect" theory holds that it must also have the potential for negative effects. However, the definition, measurement, and reporting of unwanted events in psychotherapy is a topic of ongoing discussion. At present this area is under-explored in interventions for anorexia nervosa (AN), a severe mental illness associated with high medical and psychiatric risks. The aim of this paper was to undertake a systematic review of published randomised control trials (RCTs) evaluating psychotherapeutic interventions for AN, and to assess how unwanted events were defined, monitored and reported alongside the trial's key findings. **Method:** Using a systematic review methodology, this paper identified 23 RCTs through database searches meeting eligibility criteria. Results are presented using a narrative summary approach. **Results:** Unwanted event reporting varied widely, both regarding definitions of key unwanted events (such as non-compliance or symptom deterioration) and in the amount of detail captured in each paper. **Discussion:** The review identified two key issues: firstly, a lack of consistent definitions and unclear causality made it difficult to distinguish between unwanted events, and adverse events caused by the interventions. Secondly, it highlighted the difficulty of defining unwanted events where different studies have different populations and goals. Recommendations are made around how the area of defining, monitoring and reporting unwanted events in RCTs for AN can be taken forward.

3.2 Introduction

There is increasing research interest in the concept of unwanted events in psychotherapy trials, with concerns that, to date, potential negative outcomes that occur during psychological interventions have been under-explored (Crawford et al., 2016; Duggan, Parry, McMurran, Davidson, & Dennis, 2014; Linden, 2013; Parry, Crawford, & Duggan, 2016). The "deterioration effect" theory holds that if therapy has the capacity to produce positive change, it must also have the potential for negative effects (Bergin, 1966). A 2014 review of National Institute for Health Research (NIHR) funded trials found that no psychological intervention trial reported on adverse event occurrence in their final report, suggesting that negative effects from these interventions may be under-reported and under-recognised (Duggan et al., 2014).

However, defining unwanted events in the psychotherapy field comes with its own complexities. On a basic level there are a wide range of definitions and terms (Parry et al., 2016). This review follows the definition of Linden (2013) in conceiving unwanted events as any unwanted event that occurs parallel or following treatment. Linden then distinguishes between types of unwanted events depending on whether there is evidence that the event was caused by the treatment, whether it was caused by incorrectly applied treatment, and specific events such as lack of improvement or deterioration. Similarly, Parry et al. (2016) have proposed three standard definitions that capture adverse effects of therapy, where it is judged that there is a causal link between the intervention and the effect: adverse events (a significant episode during or shortly after treatment that amounts to harm), clinically significant deterioration (a worsened mental state following therapy), and patient-experienced harm (negative experience of therapy). This review follows these definitions by positioning "adverse events" as a specific sub-type of unwanted events where there is an inferred causal link with the intervention.

In addition, the concept of reporting adverse events in clinical trials has been highly influenced by the medical field, where unwanted events are conceived as measurable, often physical, harms that are considered medically significant- for example, hospitalisation, or resulting in death (Duggan et al., 2014; Ioannidis, Evans, Gotzsche, et al., 2004). By contrast, in psychotherapeutic interventions unwanted events may look different and impact more psychological, social and emotional domains- for example, an intervention may lead to heightened distress in a way that impacts on the individual's relationships or performance at work (Linden, 2013). Furthermore, definitions of unwanted events in the research literature typically focus on distinguishing harms that can be identified, or inferred, as caused by the intervention itself (Duggan et al., 2014; Linden, 2013; Parry et al., 2016). However, causality can be complex to assess in this field and may not be captured in overall trial reporting. A psychotherapy trial paper may report on hospitalisation rates, but without added context it is difficult for the reader to understand whether this reflects an adverse event, or an unwanted event that occurred in parallel to the intervention for reasons unrelated to the intervention. For example, it would be unhelpful to count drug abuse as an adverse event in a trial where the intervention is aimed at individuals who already experience drug abuse (Horigian, Robbins, Dominguez, Ucha, & Rosa, 2010).

Furthermore, who defines harm? Parry et al. (2016) raise the concept of patient-experienced harms in their list of adverse events, and it is clearly evident from qualitative research that patients can and do have negative experiences of treatment (Offord, Turner, & Cooper, 2006). Linden (2013) highlights that successfully engaging in psychotherapy can create negative experiences for patients (for example, distress stemming from remembering upsetting past experiences), but that this does not necessarily mean that the psychotherapy itself is harmful. In fact, experiencing these negative effects and discomfort can be a key part of the change process. Therefore, an effect may be unwanted, and related to the treatment,

and yet not represent an adverse event from the perspective of the clinician. Indeed, it may reflect a “wanted” unwanted event in the context of the intervention.

A further complexity previously raised around the difficulties of defining and measuring unwanted events in the field of psychotherapy interventions is that of specificity (Duggan et al., 2014; Horigian et al., 2010). In their work on defining adverse events as part of a behavioural intervention trial, Horigian et al. (2010) highlight how conditions, populations and intervention types are likely to be associated with specific risks that should influence what becomes defined as an unwanted or adverse event within a trial. To date, there is a lack of research on unwanted events specific to anorexia nervosa (AN). AN is a serious eating disorder (ED) characterised by the restriction of energy intake relative to requirements resulting in low body weight, an intense fear of weight gain, disturbance in self-perceived weight and shape, and the undue influence of weight and shape on self-evaluation (American Psychiatric Association, 2022). Firstly, understanding unwanted events in the context of psychotherapeutic interventions for AN is particularly relevant as this condition is associated with treatment non-response. A range of psychotherapeutic interventions have been developed for the treatment of AN and are known to result in reductions in ED symptomatology (Murray, Quintana, Loeb, Griffiths, & Le Grange, 2019; Solmi et al., 2021). Despite these advances, research suggests that many individuals experience poor outcomes following treatment for AN: in a systematic review of the intervention literature, 46.9% of patients were found to have made a full recovery, 33.5% showed some improvement, 20.8% developed chronic AN, and 5% of patients died (Steinhausen, 2002). There is further evidence that a small proportion of patients deteriorate during treatment (Schlegl et al., 2016; Schlegl, Quadflieg, Löwe, Cuntz, & Voderholzer, 2014).

Secondly, AN is a condition associated with heightened mortality rates and other serious physical complications (Arcelus, Mitchell, Wales, & Nielsen, 2011; Birmingham, Su,

Hlynsky, Goldner, & Gao, 2005). Firstly, this creates significant complexity around inferring causality: if a trial for AN reports high hospitalisation rates, does this reflect an effect of the intervention, or pre-existing medical vulnerability? Secondly, it means that trial investigators need to simultaneously consider potential harms typically associated with psychotherapy trials, but also incorporate medical monitoring and harm reporting as seen in the medical field. Thirdly, measuring unwanted events such as treatment non-response is complicated by a lack of agreement in the field on how to define recovery from AN (Khalsa, Portnoff, McCurdy-McKinnon, & Feusner, 2017). Finally, the concept of patient-experienced harms gains additional complexity in this population. Whilst people with AN may be vulnerable to experiencing unwanted events during treatment, the egosyntonic nature of AN raises questions around whether these experiences should then be considered as unwanted in the context of treatment. Individuals with AN can experience their condition as highly valued and performing important psychosocial functions, which can contribute to experiencing treatments aimed at changing these behaviours as highly aversive and, indeed, unwanted (Gregertsen, Mandy, & Serpell, 2017).

Whilst the concept of unwanted events in psychotherapy interventions is particularly salient (and potentially complex) in AN, it is also an area which is currently under-explored. Therefore, the aim of the current review was to better understand how unwanted events are currently defined, monitored and reported alongside the findings of published randomised controlled trials (RCT) for psychological interventions for AN. It is hoped that synthesising this information is a first step in further research and discussion around unwanted events in this field.

3.3 Methods

This review was pre-registered with PROSPERO (registration number CRD42021275696) and performed according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009).

3.3.1 Design

This review originally intended to utilise a pre-existing adverse event checklist, specifically the harms extension to the CONSORT statement, an adapted version utilised in a previous review of reporting in general medical RCTs, or Linden's psychotherapy checklist, to explore this area (Ioannidis, Evans, Gotzsche, et al., 2004; Linden, 2013; Phillips, Hazell, Sauzet, & Cornelius, 2019). However, following a pilot of this approach on a smaller number of identified papers, it became evident that the definitions of adverse events and reporting standards outlined in those checklists were a poor fit with the types of information typically reported in RCTs for AN. For example, no trial reported information on absolute risk per arm (as recommended by the CONSORT extension), or included a consideration of whether the intervention had resulted in the patient having to take sick leave from their job (as recommended by Linden).

Consequently, the design of the study shifted from attempting to compare published RCTs in this area to existing standards and definitions, and instead chose to focus more on exploring what kind of information and definitions are currently being reported.

In line with this, authors EK & MC developed a data extraction checklist focused on capturing this information. This checklist was influenced both by definitions of unwanted events utilised in previous work, and the initial pilot work identifying what types of

information are typically reported in this field (Ioannidis, Evans, Gotzsche, et al., 2004; Linden, 2013; Parry et al., 2016). The final data extraction checklist is reported below.

3.3.2 Eligibility Criteria

Studies were included if they fulfilled the following criteria:

- RCT evaluating a specified psychotherapy for individuals with anorexia nervosa
- Published in English
- Published since 1980
- Article published in a peer reviewed journal

Trials evaluating medication in addition to psychotherapy were excluded as these present different risk monitoring challenges compared to psychotherapy only. As not all trials recorded consistent follow-up data, this review primarily focuses on unwanted event monitoring during and at the end of treatment (EOT), and thus follow-up studies were excluded from the analysis. Pilot, feasibility, or exploratory RCTs, or those published as brief reports, were also excluded from the analysis in case this (for example) led to reduced focus on reporting of safety monitoring procedures. Only published papers were eligible for inclusion as the intention of this review was to evaluate how unwanted events are defined in reported in final published trial papers.

3.3.4 Search Strategy

Search terms were developed by reviewing the titles of published RCTs in AN for key words. Final search terms were “anorexia nervosa" AND “rando*” AND “therapy”. Searches were performed across PsycInfo, Medline, Scopus, and Pubmed databases during June 2022. Following screening the final list of included papers was evaluated against previously published systematic reviews of RCTs in AN to identify missed studies (Brockmeyer,

Friederich, & Schmidt, 2018; Bulik, Berkman, Brownley, Sedway, & Lohr, 2007; Murray et al., 2019). No further papers were identified in this way. Searches were additionally re-run in February 2023 (using the same search terms, but results restricted to the past year) to explore if any further trials had been published during the intervening period, with no papers identified.

3.3.5 Selection Process

The selection process is detailed in Figure 1. Searches were performed by the first author, with titles and abstracts then screened against the eligibility criteria. Twenty per cent of these were also randomly selected and screened by a second researcher (Cohen's kappa 0.95). Selected full texts were then screened by the first author, with 20% screened by a second researcher (Cohen's kappa 0.8).

With the increasing trend to publish trial protocols in advance of results and final publication, it was noted during full text screening that several papers signposted readers to previous publications for further details on study design, including safety monitoring. Therefore, readers interested in unwanted event monitoring and occurrence could access this information based on signposting within the paper. Where papers indicated that further information about the study was available from other sources (for example, published protocols, available trial registration documents, or other papers relating to the trial) this was accessed and has been included in the synthesis.

3.3.6 Data Extraction

The following contextual information was extracted from papers: total sample size, mean age, mean BMI at baseline, illness duration, participant ethnicity/ socioeconomic status, diagnostic inclusion criteria, treatment setting, and interventions. The following information regarding unwanted events was extracted. For all points authors searched papers for both the

presence of this information (for example, a specified definition of adverse events), but also the absence (for example, no information on compliance rates).

- Definitions of adverse events utilised within papers
- Information about safety monitoring and procedures
- Reporting of non-compliance rates
- Treatment non-response definitions and rates
- Evidence of symptom deterioration
- Evidence of the emergence of new symptoms
- Need for additional treatment beyond interventions
- Negative feedback from patients and/or family
- Patient mortality
- Any information or analysis on causality for unwanted events

There is controversy around whether treatment non-completion should be considered an unwanted event (Duggan et al., 2014). Given the high rates of non-completion in ED treatment and its inclusion in the CONSORT recommendations for reporting RCTs, non-completion is included in this review (DeJong, Broadbent, & Schmidt, 2012; Ioannidis, Evans, Gøtzsche, et al., 2004).

Data extraction was completed by the first author, with 50% of randomly selected papers also extracted by a second researcher to ensure reliability (Cohen's kappa 0.87).

3.3.7 Quality Assessment

As this review explored the reporting of safety measures and unwanted outcomes, it was decided to assess the quality of individual studies based on the broader quality of their reporting (which also gives an indication of their methodological rigour). The quality assessment utilised a previously developed checklist consisting of eight items based on the

CONSORT reporting standards for non-pharmacological interventions (Boutron, Moher, Altman, Schulz, & Ravaud, 2008; Hoffmann, Erueti, & Glasziou, 2013). Each item has a yes/no response, and every item must have a "yes" response for the reporting to be described as "adequate". The only adjustment to Boutron et al.'s (2008) procedure was the item regarding whether intervention materials (for example, therapy manuals) were publicly available: as many manualised therapies are published as copyrighted material, this item was adjusted to reflect whether authors signposted to where readers could access the full manual (for example, a reference to a published book). Related documents such as trial registrations, pilot studies, or protocols were also included in the quality analysis. Papers were rated by the first author, with 20% randomly selected papers also evaluated by a second researcher with full agreement on adequacy ratings.

3.4 Results

Twenty-three RCTs were included in the final analysis. Study details, including proportions of participants reported as experiencing unwanted events, are summarised in Table 1. Participant details are summarised in Table 2. The majority of studies were conducted on an outpatient basis, two studies took place in an inpatient setting. The average reported mean age of participants was 20.84 years (*SD*= 5.35), with included ages ranging from 11-65 years. Six studies (26.09%) included adults only, eight (34.78%) included adolescents only, and the remainder (39.13%) either included a mix or did not report age as an inclusion criterion. The average reported mean starting BMI was 16.10 (*SD*= 1.04). The majority of studies did not report on participant ethnicity, or sociodemographic status. A wide range of different psychotherapies were included in the review, including different types of family therapy and varying individual psychotherapies such as cognitive behavioural therapy or psychodynamic interventions.

The overall quality of reporting was rated as adequate in 52.17% (*n*= 12) of the studies, with papers most commonly rated as inadequate due to not detailing the availability of treatment manuals (Hoffmann et al., 2013). Considering that it is not uncommon for psychotherapy trials to not make therapy manuals publicly available, removing this criterion meant that 73.91% (*n*= 17) papers were rated as adequate.

In the context of varying definitions and reporting standards, a narrative approach was used to synthesise data and to explore what kind of information is typically captured by RCTs for psychotherapy in AN as a first step in reviewing the relevant literature on this topic.

3.4.1 Definitions of Unwanted Events

CONSORT recommendations for reporting harm in RCTs suggest that trials should list and define addressed adverse events in the methods section (Ioannidis, Evans, Gøtzsche, et al., 2004). Only three papers (13.04%) specified definitions of unwanted events. Agras et al., (2014) included definitions of adverse events ("an unfavorable change from the participant's condition at baseline") and serious adverse events ("life-threatening events, death, hospitalizations for any cause, or other important medical or psychological events"). Fichter et al., (2012) defined "undesirable events" as serious illnesses and a loss of weight leading to a BMI below 13.5, and "serious undesirable and catastrophic events" as acute suicidality, suicide attempts, death by suicide, death by accident or by natural cause. Zipfel et al., (2014) defined serious adverse events as any fatal or life-threatening event, except for intermittent inpatient treatment of up to four weeks (due to the commonality of serious weight loss in AN).

One paper (4.35%) described intermittent inpatient/day treatment as an example of an adverse event but did not include a broader definition (Nyman-Carlsson, Birgegård, Engström, Gustafsson, & Nevonen, 2019).

3.4.2 Safety Monitoring

Papers were considered to have included information about monitoring participant safety where they detailed medical risk monitoring, psychiatric risk monitoring, procedures around hospitalisation, or procedures around monitoring and reporting adverse events. All papers included the assessment of ED and related psychiatric symptoms at specified timepoints for the purposes of trial outcome measurement, however, it was unclear if any papers utilised this information to additionally monitor patient safety or psychiatric risks (for example, if scoring above a certain threshold for depression triggered further investigation).

The majority of papers (*n*= 20, 86.96%) included some information related to monitoring participant safety, however, the amount of detail provided varied widely. Most of this information related specifically to monitoring physical risks: only one paper (4.35%) referred to the management of co-occurring psychiatric conditions (Le Grange et al., 2016). Information on safety monitoring was typically found in the methods sections of papers, with five papers (21.74%) including a specific subsection for this purpose (Agras et al., 2014; Le Grange et al., 2016; Lock et al., 2010; Schmidt et al., 2015; Touyz et al., 2013).

Three papers (13.04%) included no information relating to safety monitoring (Dalle Grave, Calugi, Conti, Doll, & Fairburn, 2013; Eisler et al., 2016; McIntosh et al., 2005). Five papers (21.74%) included a statement that their participants received regular medical monitoring but no further details (Brockmeyer et al., 2021; Channon, de Silva, Hemsley, & Perkins, 1989; Dare, Eisler, Russell, Treasure, & Dodge, 2001; Godart et al., 2012; Parling, Cernvall, Ramklint, Holmgren, & Ghaderi, 2016). Parling et al., (2016) additionally specified "we did not systematically ask for adverse events or harms". Four papers (17.39%) included information around procedure if an adverse medical event occurred, typically hospitalisation in the case of medical instability and subsequent trial withdrawal (Ball & Mitchell, 2004;

Fichter et al., 2012; Nyman-Carlsson et al., 2020; Serfaty, Turkington, Heap, Ledsham, & Jolley, 1999). Nine papers (39.13%) included a combination of detailing medical monitoring, procedures around hospitalisation (including medical criteria for admission), and subsequent impact on trial involvement (Byrne et al., 2017; Eisler et al., 2000; Geist, Heinmaa, Stephens, Davis, & Katzman, 2000; Le Grange et al., 2016; Lock, Agras, Bryson, & Kraemer, 2005; Lock et al., 2010; Robin et al., 1999; Schmidt et al., 2015; Touyz et al., 2013).

Agras et al. (2014) included the most comprehensive description of safety monitoring located in a dedicated "Participant Safety and Adverse Event" sub-section. This paper, and a previous paper detailing the trial's design, described medical monitoring, including frequency and specific medical tests, and provided detailed guidelines followed for medical hospitalisation, including specific medical criteria and procedures for trial withdrawal (Lock et al., 2012). In addition, the paper specifies the procedure for collecting and reporting data on adverse events, including describing the establishment of a dedicated board for safety monitoring. Zipfel et al. (2014) also included a higher level of detail compared to other papers around the monitoring, documentation, and reporting of adverse/serious adverse events, including an independent safety and data monitoring board, and described medical criteria, regular monitoring, and procedures around hospitalization.

3.4.3 Treatment Non-Completion

Whilst almost every trial (except Geist et al., 2000) reported treatment non-completion, not all trials differentiated between non-completion due to patient decision to drop out and non-completion due to clinician decision to withdraw, and not all trials detailed further reasons for these decisions. One trial's reporting around non-completion was unclear and has not been included in the final analysis: Robin et al., (1999) reported all participants completed the full trial including follow-up, but an earlier paper reporting on a smaller

version of the final sample stated that some participants did not complete treatment (Robin, Siegel, Koepke, Move, & Tice, 1994).

In the following summary the term “drop-out” is used to refer to discontinuation chosen by patient/family, and “withdrawal” refers to discontinuation chosen by trial clinicians (for example, due to hospitalisation). “Non-completion” is used where reasons for discontinuation are unclear. Where papers reported information using the term “drop-out” it has been assumed this refers to patient discontinuation unless otherwise specified. Non-completion rates are summarised in Table 1. In all papers withdrawal from treatment almost always related to hospitalisation or high medical risk.

Five trials (21.74%) reported non-completion rates with no further details, including not specifying if non-completion was due to patient drop-out or withdrawal by trial clinicians (Le Grange et al., 2016; Parling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013; Zipfel et al., 2014). Two trials (8.70%) reported non-completion rates and specified that a proportion of these were due to withdrawal relating to hospitalisation but did not detail reasons for other non-completers (Ball & Mitchell, 2004; McIntosh et al., 2005).

Most commonly trials reported drop-out rates only (*n*= 10, 43.48%). These trials did not report withdrawal due to hospitalisation, possibly because this did not occur during these studies. The majority (*n*= 7, 30.43%) of these studies reported drop-out rates without detailing reasons (Channon et al., 1989; Dare et al., 2001; Eisler et al., 2000; Eisler et al., 2016; Fichter et al., 2012; Lock et al., 2010; Nyman-Carlsson et al., 2020). Brockmeyer et al., (2021) and Lock et al., (2005) detailed reasons for patient drop-out, Serfaty et al., (1999) gave reasons for one treatment group only.

Two studies (8.70%) reported total non-completion and differentiated between drop-out and withdrawal, documenting reasons for withdrawal but not for drop-out (Dalle Grave et

al., 2013; Godart et al., 2012). Two studies reported total non-completion and differentiated between drop-out and withdrawal and gave reasons for both withdrawal and patient drop-out (Agras et al., 2014; Byrne et al., 2017).

Regarding further analyses of causality beyond reasons for discontinuation, most (82.61%) papers reported whether non-completion rates differed across groups (except for Godart et al., (2012), Lock et al., (2005), Schmidt et al., (2015), and Serfaty et al., (1999)). Three papers (13.04%) additionally analysed whether non-completion was related to patient characteristics, such as baseline body weight (Ball & Mitchell, 2004; Fichter et al., 2012; McIntosh et al., 2005).

3.4.4 Treatment Non-Response

No trial performed a clinical significance analysis to measure treatment non-response (Jacobson, Follette, & Revenstorf, 1984), rather, papers reported proportions of patients meeting a pre-specified categorical outcome (summarised in Table 1).

Five trials (21.74%) did not report categorical outcomes (Brockmeyer et al., 2021; Channon et al., 1989; Fichter et al., 2012; Geist et al., 2000; Touyz et al., 2013). A number of trials (*n*= 7, 30.43%) reported the proportion of patients not achieving remission, or a similar pre-defined good outcome (Agras et al., 2014; Byrne et al., 2017; Dalle Grave et al., 2013; Le Grange et al., 2016; Lock et al., 2005; Lock et al., 2010; Parling et al., 2016). This reflects participants not improving to a certain threshold, rather than a total lack of improvement, and therefore it is unclear how many of these patients were actual non-responders. One of these trials (Lock et al., 2010) additionally reported the number of patients who originally remitted then relapsed during the follow-up period. All these trials used varying criteria for a good outcome/ remission: not all included both body weight and symptom thresholds, and different body weight thresholds were utilised.

Nine trials (39.13%) reported patients not responding to treatment, typically operationalised as experiencing a poor outcome or continuing to meet diagnostic criteria for AN. Three of these trials (13.04%) used the Morgan and Russell criteria to report poor outcomes (Ball & Mitchell, 2004; Eisler et al., 2000; Eisler et al., 2016; Morgan & Russell, 1975), the remainder used outcome descriptors based around diagnostic criteria (Dare et al., 2001; McIntosh et al., 2005; Nyman-Carlsson et al., 2020; Schmidt et al., 2015; Serfaty et al., 1999; Zipfel et al., 2014).

All trials analysed whether ED outcomes differed across treatment groups.

3.4.5 Symptom Deterioration

This synthesis followed Linden's (2013) definition of "deterioration of illness" as "worsening of illness during therapy or at any other time in the course of illness". It should be noted that this definition does not capture trial follow-up periods- where reported, this information has been included. Additionally, deterioration relating to co-occurring illnesses other than AN was also included.

No trials reported the occurrence of patients deteriorating on measures of AN symptoms other than weight loss, or measures of other psychiatric pathology. Trials most commonly reported symptom deterioration as relating to a worsening of physical health resulting in hospitalisation, both during intervention and follow up periods. Almost all trials (excepting Brockmeyer et al., (2021) and Serfaty et al., (1999)) reported proportions of patients requiring hospitalisation (summarised in Table 1). Where reasons for hospitalisation were given, the majority were related to medical complications such as severe weight loss. A smaller number of trials (n=7, 30.43%) reported hospitalisations, or accessing additional treatment, relating to psychiatric risks such as suicidality, self-harm, or the occurrence of another psychiatric condition (Agras et al., 2014; Eisler et al., 2000; Eisler et al., 2016;

Godart et al., 2012; Lock et al., 2010; Parling et al., 2016; Schmidt et al., 2015) . Finally, one trial reported that a mean weight loss occurred during the study follow up period in the TAU group (Fichter et al., 2012).

In addition to reporting reasons for hospitalisation, the majority of trials (*n*=12, 52.17%) reported whether hospitalisation rates differed across groups (Agras et al., 2014; Byrne et al., 2017; Channon et al., 1989; Godart et al., 2012; Le Grange et al., 2016; Lock et al., 2005; Lock et al., 2010; McIntosh et al., 2005; Parling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013; Zipfel et al., 2014).

3.4.6 Other Unwanted Events

Papers were screened for other types of reporting that reflected unwanted events, as judged by the study authors.

11 papers (47.83%) reported proportions of participants requiring additional treatment either during the intervention or follow-up period (Agras et al., 2014; Ball & Mitchell, 2004; Brockmeyer et al., 2021; Byrne et al., 2017; Dalle Grave et al., 2013; Eisler et al., 2000; Le Grange et al., 2016; Lock et al., 2010; Parling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013). Only three of these studies (13.04%) included whether additional treatment utilisation differed across groups.

Nine papers (39.13%) reported the occurrence of patient deaths during the trial period. Four papers (17.39%) specified no deaths occurred (Agras et al., 2014; Byrne et al., 2017; Fichter et al., 2012; Zipfel et al., 2014), five papers (21.74%) reported one patient death (see Table 1) (Dare et al., 2001; McIntosh et al., 2005; Parling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013).

Four papers (17.39%) reported the emergence of new ED or psychiatric symptoms during the study. Two papers (8.70%) documented that a proportion of patients developed

additional bulimic symptoms (Eisler et al., 2000; Eisler et al., 2016). One paper (4.35%) reported the proportions of patients at end of treatment/ follow-up who now met criteria for bulimia nervosa (BN) and eating disorder not otherwise specified (EDNOS) (Nyman-Carlsson et al., 2020). One paper documented the emergence of severe obsessive-compulsive disorder and depression in some patients following weight restoration (Robin et al., 1999).

One paper reported the proportion of participants who reported being dissatisfied with their treatment (Eisler et al., 2016).

Finally, a small number of papers (*n*=3, 13.04%) included statements on the occurrence of adverse events during the trial. Agras et al., (2014) was the only paper to include an "Adverse Event" sub-heading in its result section, where it documents that there were no significant differences between treatment groups on numbers of serious adverse events (and specifies types of serious adverse events that occurred). Fichter et al., (2012) and Zipfel et al., (2014) both included statements that no serious adverse events occurred relating to the trial.

3.5 Discussion

The findings of this review suggest that whilst certain key unwanted events are typically reported in AN psychotherapy RCTs (notably non-completion, hospitalisation rates, and a lack of improvement), there is significant variation in definitions which makes it difficult to compare outcomes across trials. This was particularly notable when comparing treatment non-response, or non-improvement across trials, with different papers using varying thresholds and outcomes (typically weight only, or a combination of weight and symptom measurements) to classify a lack of improvement. The need for a consistent of definition of recovery from AN in order to enable the comparison of categorical outcomes has previously been highlighted in the research literature, with suggestions that these should

include both measures of weight and symptoms (Bardone-Cone et al., 2010; Couturier & Lock, 2006; Khalsa et al., 2017). Achieving consensus around the definitions of key unwanted events, such as a lack of clinically significant improvement or the differentiation between non-completion, withdrawal, and drop-out, would be important in allowing the comparison of these outcomes across RCTs.

The second key finding of this review is that it is very difficult from the information reported in trial papers to understand whether the events described represented unwanted events, or adverse events inferred as related to the intervention. The reporting, for example, of hospitalisation rates without other context potentially reflects influences from the medical field where events like mortality and hospitalisation are reported as adverse events as standard, but does not meet the definitional requirement of adverse events where there must be an inferred causal link to the intervention itself (Duggan et al., 2014; Linden, 2013; Parry et al., 2016). This style of reporting may lack applicability to a condition like AN where patients may have medical vulnerabilities prior to starting the intervention: some RCTs may target higher risk groups of patients (such as those with lower BMIs or experiencing inpatient admissions) where high rates of hospitalisation reflect baseline illness severity rather than a consequence of the intervention. Similarly, unwanted events may be influenced by contextual factors outside of the intervention- for example, approaches to hospitalisation for AN can vary across countries, with studies based in the USA and Australia reporting higher hospital admission rates (Eisler et al., 2016). Whilst almost all papers compared key unwanted events across intervention groups (for example, comparing patient drop-out rates or hospitalisation), beyond group comparisons there was typically a lack of reflection on the causality or context of unwanted events as recommended in existing reporting standards (Ioannidis, Evans, Gotzsche, et al., 2004; Linden, 2013).

The findings of this review therefore raise the question of whether the reporting of statistics such as hospitalisation rates could be improved with added context around causality. The present typical style of reporting an uncontextualized hospitalisation statistic does give clinicians and other researchers insight into how common hospitalisation rates are during the intervention, which might be helpful in informing expectations when implementing therapy, but does not illuminate whether there are risks around the intervention causing or not effectively slowing deterioration. Where causality was considered in relation to unwanted events in this review, these typically reflected analyses of whether outcomes were related to patient factors, such as McIntosh et al. (2005) finding that treatment non-completion was related to lower starting body weight. Notable exceptions include Robin et al.'s (1999) consideration of whether the emergence of severe OCD and depression symptoms in a small number of patients may have occurred because of successful weight restoration following the intervention.

As well as highlighting difficulties in reporting and interpreting unwanted events in general when treating people with AN, this review also emphasises that understanding what counts as an unwanted versus an adverse event may depend on the nature and goals of the specific psychotherapeutic intervention. For example, not all interventions included in this review aimed for recovery as an outcome: indeed, in their RCT targeting severe and enduring AN Touyz et al., (2013) highlight that targeting recovery may be misaligned with the goals of patients in this population, and instead focused on quality of life, mood disorder symptoms and social adjustment as primary outcomes, with ED symptoms as secondary outcomes. If a standardised unwanted event checklist was applied to this trial, such as that developed by Linden (2013), this would count as treatment non-response and reflect an unwanted event; but this would not represent a helpful interpretation of this trial's goals and outcomes. This

raises further questions around the utility of standardised unwanted event checklists in this field.

Therefore, the findings of this review raise questions around how unwanted events should be measured, defined and reported within the field of psychotherapeutic interventions for AN, and how this information can then be used to inform understandings of iatrogenic harm. It does not appear that current reporting norms enable the reader to distinguish between unwanted events and negative effects. Furthermore, it was also noted that many of the unwanted events reported by trials represented a medical understanding of this concept, such as reporting hospitalisation, mortality, or treatment non-response as measured by disorder-specific symptoms. Only one paper measured patient-experienced harms in the form of patient and family feedback (Eisler et al., 2016). No paper examined unwanted events in social or work domains as recommended by Linden (2013), or broader experiences of distress or negative affect.

Therefore, the reporting norms observed in this review seem somewhat disconnected from the reality of AN as a condition, and the nature of psychotherapeutic interventions, and so limit any extrapolation into understanding how far these unwanted events translate into negative effects. In the wider field of adverse event research in psychotherapy it has been proposed that what is considered an adverse event should be defined a priori by investigators with consideration to the specific condition, population and intervention being utilised, rather than defined in relation to pre-existing checklists (Duggan et al., 2014; Horigian et al., 2010). However, AN is also a condition associated with significant medical risks, and therefore there may be benefits from also retaining some aspects of the medicalised approach noted in this paper. It is possible that an individualised approach where investigators determine trial-specific unwanted/ adverse events a priori could be combined with a core reporting framework of standards specific to AN trials, drawing on medical research standards. This

would retain the flexibility and specificity of Duggan et al's (2014) proposed approach, whilst simultaneously ensuring safety, best practice and comparability.

This is an area which clearly requires more thought and discussion. However, building on current practice and research in parallel fields, it is tentatively proposed that papers could benefit from including the following core information as a minimum. It is emphasised that this is not intended to represent a generic checklist, but rather a framework for information that should be considered and included:

- Authors should consider standards recommended by the CONSORT harms checklist (Ioannidis, Evans, Gotzsche, et al., 2004). The below points may be particularly relevant to the ED field:
    - A list of addressed adverse events within the methods section, including details on monitoring and procedures. This could include criteria for hospitalisation, pausing treatment, and withdrawal from trial, and whether regular monitoring of physical and mental health risks occurred in addition to the standard assessment schedule.
    - Reporting treatment non-completion rates across groups, differentiating between participants who dropped out (by non-mutual agreement), and proportions of participants withdrawn from the study by investigators. Reporting reasons for non-completion, including where related to the intervention (for example, patient dissatisfaction).
    - Reporting statistical comparisons of serious adverse event incidence rates across groups.
- A dedicated patient safety section within the methods section:

    - How the trial defined unwanted and adverse events, including information on how investigators inferred causality (as recommended by Duggan et al., 2014).

- A dedicated adverse event reporting section under results:
    - Reporting proportions experiencing treatment non-response/ symptom deterioration with reference to suggested standardised models of recovery and remission (Khalsa et al., 2017). Where symptom deterioration occurs, considering evidence for a causal relationship to the intervention.
    - Where hospitalisation rates and mortality rates are reported (as is currently typical in AN RCTs), contextualising this where possible with reasonable inference around whether this represents an adverse event relating to the intervention.

In the context of difficulties conceiving unwanted events within AN psychotherapy trials, it is the position of the authors that whilst the above information should be reported, it should also be contextualised with a sub-section within the discussion dedicated to reflecting on the context of unwanted events within the study, including in relation to expected harms as a result of treatment and change processes, and any potential negative effects.

Other elements that investigators could consider reporting include the use of additional treatments, whether there was a need for a prolonged intervention, and the emergence of new symptoms (for example, whether participants fulfil criteria for another ED by the end of treatment). However, the authors emphasise that investigators should consider whether these reflect adverse events and specify this within the text: for example, it may be expected that a patient would continue to access treatment following an inpatient intervention. Investigators could also consider measures that reflect patient perspectives on

treatment such as satisfaction ratings, unwanted event questionnaires, or idiosyncratic outcome ratings (Austin et al., 2020).

Finally, it was noted during data synthesis that most studies did not report data on participant ethnicity, and that the majority of individuals included in these studies were white females. The accurate recording and reporting of patient ethnicity in relation to unwanted events is important as evidence suggests that ethnic minority patients are at higher risk of adverse health outcomes, including adverse events and medical errors (Chauhan et al., 2020). Future research on unwanted events in ED treatment could benefit from exploring the potential influence of demographic variables such as race, gender, and socioeconomic status.

This review has several limitations. Firstly, by focusing on published RCT papers and excluding secondary analyses this review was unable to include papers where authors focused on negative outcomes as part of a secondary analysis: for example, whilst Lock et al., (2005) reported drop-out and remission rates in their primary analysis, the group published a later paper analysing patient factors predicting drop-out and remission in more detail (Lock, Couturier, Bryson, & Agras, 2006). Similarly, a secondary analysis of Agras et al., (2014) performed a more detailed analysis of hospitalisation rates, including relation to intervention and patient factors (Lock et al., 2016). This was a deliberate choice by the authors as the intention was to explore what kinds of information on unwanted events are presented alongside the key findings of RCTs, rather than information published at a later stage. Furthermore, this review did not include unpublished papers as its intention was to explore reporting standards in final versions of published RCTs. Therefore, this review knowingly reflects a certain degree of publication bias: the information reflected in papers, for example, might reflect details requested by journals rather than those intended for inclusion by authors. Related to this point, this review is dependent on information reported in trial papers: for

example, just because an outcome paper does not report the presence of safety monitoring does not mean that it did not occur in practice.

Secondly, it should be noted that by evaluating unwanted events as reported in RCT papers this review inevitably focuses on unwanted events deemed significant by investigators, rather than patients themselves. A benefit of future RCTs using idiosyncratic patient ratings and feedback is that they could better capture factors deemed significant by patients but not typically recorded as part of trial outcomes, such as isolation relating to inpatient admission or trauma associated with refeeding (Offord et al., 2006; Treasure, Crane, McKnight, Buchanan, & Wolfe, 2011).

In conclusion, this review highlights that at present the reporting of unwanted event monitoring and occurrence varies significantly across psychotherapy RCTs for AN in a manner that hinders interpretation and comparison. It also raises questions around how unwanted events should be conceived and defined in this field, including around whether an approach to measuring adverse events influenced by a medical understanding of this concept is an effective fit for this area. It suggests that this area could benefit from further discussion around balancing objective and comparable standards with an individualised approach that ensures that unwanted/adverse event measurement and reporting best captures the complex context of AN psychotherapy intervention studies.

3.6 References

Agras, W., Lock, J., Brandt, H., Bryson, S. W., Dodge, E., Halmi, K. A., . . . Woodside, B. (2014). Comparison of 2 family therapies for adolescent anorexia nervosa: A randomized parallel trial. *JAMA Psychiatry, 71*(11), 1279-1286.

American Psychiatric Association. (2022). Diagnostic and statistical manual of mental disorders (5th ed., text rev.).

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders. A meta-analysis of 36 studies. *Arch Gen Psychiatry, 68*(7), 724-731. doi:10.1001/archgenpsychiatry.2011.74

Ball, J., & Mitchell, P. (2004). A randomized controlled study of cognitive behavior therapy and behavioral family therapy for anorexia nervosa patients. *Eating Disorders, 12*(4), 303-314. doi:10.1080/10640260490521389

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. doi:https://doi.org/10.1016/j.brat.2009.11.001

Bergin, A. E. (1966). Some implications of psychotherapy research for therapeutic practice. *Journal of Abnormal Psychology, 71*(4), 235-246. doi:10.1037/h0023577

Birmingham, C. L., Su, J., Hlynsky, J. A., Goldner, E. M., & Gao, M. (2005). The mortality rate from anorexia nervosa. *International Journal of Eating Disorders, 38*(2), 143-146. doi:https://doi.org/10.1002/eat.20164

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., & Ravaud, P. (2008). Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration. *Annals of Internal Medicine, 148*(4), 295-309. doi:10.7326/0003-4819-148-4-200802190-00008 %m 18283207

Brockmeyer, T., Friederich, H. C., & Schmidt, U. (2018). Advances in the treatment of anorexia nervosa: a review of established and emerging interventions. *Psychological Medicine, 48*(8), 1228-1256. doi:10.1017/s0033291717002604

Brockmeyer, T., Schmidt, H., Leiteritz-Rausch, A., Zimmermann, J., Wunsch-Leiteritz, W., Leiteritz, A., & Friederich, H.-C. (2021). Cognitive remediation therapy in anorexia nervosa-A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 89*(10), 805-815. doi:10.1037/ccp0000675

Bulik, C. M., Berkman, N. D., Brownley, K. A., Sedway, J. A., & Lohr, K. N. (2007). Anorexia nervosa treatment: A systematic review of randomized controlled trials. *International Journal of Eating Disorders, 40*(4), 310-320. doi:https://doi.org/10.1002/eat.20367

Byrne, S., Wade, T., Hay, P., Touyz, S., Fairburn, C., Treasure, J., . . . Crosby, R. (2017). A randomised controlled trial of three psychological treatments for anorexia nervosa. *Psychological Medicine, 47*(16), 2823-2833.

Channon, S., de Silva, P., Hemsley, D., & Perkins, R. E. (1989). A controlled trial of cognitive-behavioural and behavioural treatment of anorexia nervosa. *Behaviour Research and Therapy, 27*(5), 529-535.

Couturier, J., & Lock, J. (2006). What is recovery in adolescent anorexia nervosa? *International Journal of Eating Disorders, 39*(7), 550-555. doi:https://doi.org/10.1002/eat.20309

Crawford, M. J., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., . . . Parry, G. D. (2016). Patient experience of negative effects of psychological treatment: results of a national survey. *British Journal of Psychiatry, 208*(3), 260-265. doi:10.1192/bjp.bp.114.162628

Dalle Grave, R., Calugi, S., Conti, M., Doll, H., & Fairburn, C. G. (2013). Inpatient cognitive behaviour therapy for anorexia nervosa: a randomized controlled trial. *Psychotherapy and Psychosomatics, 82*(6), 390-398. doi:10.1159/000350058

Dare, C., Eisler, I., Russell, G., Treasure, J., & Dodge, L. (2001). Psychological therapies for adults with anorexia nervosa. Randomised controlled trial of out-patient treatments. *British Journal of Psychiatry, 178*(MARCH.), 216-221. doi:10.1192/bjp.178.3.216

DeJong, H., Broadbent, H., & Schmidt, U. (2012). A systematic review of dropout from treatment in outpatients with anorexia nervosa. *International Journal of Eating Disorders, 45*(5), 635-647. doi:https://doi.org/10.1002/eat.20956

Duggan, C., Parry, G., McMurran, M., Davidson, K., & Dennis, J. (2014). The recording of adverse events from psychological treatments in clinical trials: evidence from a review of NIHR-funded trials. *Trials, 15*, 335. doi:10.1186/1745-6215-15-335

Eisler, I., Dare, C., Hodes, M., Russell, G., Dodge, E., & Le Grange, D. (2000). Family therapy for adolescent anorexia nervosa: the results of a controlled comparison of two family interventions. *J Child Psychol Psychiatry, 41*(6), 727-736.

Eisler, I., Simic, M., Hodsoll, J., Asen, E., Berelowitz, M., Connan, F., . . . Landau, S. (2016). A pragmatic randomised multi-centre trial of multifamily and single family therapy for adolescent anorexia nervosa. *BMC Psychiatry, 16*(1). doi:10.1186/s12888-016-1129-6

Fichter, M. M., Quadflieg, N., Nisslmüller, K., Lindner, S., Osen, B., Huber, T., & Wünsch-Leiteritz, W. (2012). Does internet-based prevention reduce the risk of relapse for anorexia nervosa? *Behaviour Research and Therapy, 50*(3), 180-190. doi:10.1016/j.brat.2011.12.003

Geist, R., Heinmaa, M., Stephens, D., Davis, R., & Katzman, D. K. (2000). Comparison of family therapy and family group psychoeducation in adolescents with anorexia nervosa. *Canadian Journal of Psychiatry, 45*(2), 173-178. doi:10.1177/070674370004500208

Godart, N., Berthoz, S., Curt, F., Perdereau, F., Rein, Z., Wallier, J., . . . Jeammet, P. (2012). A randomized controlled trial of adjunctive family therapy and treatment as usual following inpatient treatment for anorexia nervosa adolescents. *PloS one, 7*(1), e28249. doi:10.1371/journal.pone.0028249

Gregertsen, E. C., Mandy, W., & Serpell, L. (2017). The Egosyntonic Nature of Anorexia: An Impediment to Recovery in Anorexia Nervosa Treatment. *Frontiers in psychology, 8*, 2273-2273. doi:10.3389/fpsyg.2017.02273

Hoffmann, T. C., Erueti, C., & Glasziou, P. P. (2013). Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ : British Medical Journal, 347*, f3755. doi:10.1136/bmj.f3755

Horigian, V. E., Robbins, M. S., Dominguez, R., Ucha, J., & Rosa, C. L. (2010). Principles for defining adverse events in behavioral intervention research: lessons from a family-focused adolescent drug abuse trial. *Clin Trials, 7*(1), 58-68. doi:10.1177/1740774509356575

Hughes, E. K., Le Grange, D., Court, A., Yeo, M. S., Campbell, S., Allan, E., . . . Sawyer, S. M. (2014). Parent-focused treatment for adolescent anorexia nervosa: a study protocol of a randomised controlled trial. *BMC Psychiatry, 14*, 105. doi:10.1186/1471-244x-14-105

Ioannidis, J. P., Evans, S. J., Gøtzsche, P. C., O'Neill, R. T., Altman, D. G., Schulz, K., & Moher, D. (2004). Better reporting of harms in randomized trials: an extension of the

CONSORT statement. *Ann Intern Med, 141*(10), 781-788. doi:10.7326/0003-4819-141-10-200411160-00009

Ioannidis, J. P., Evans, S. J. W., Gotzsche, P. C., O'Neill, R. T., Altman, D. G., Schulz, K., & Moher, D. (2004). Better Reporting of Harms in Randomized Trials: An Extension of the CONSORT Statement. *Annals of Internal Medicine, 141*(10), 781-788. doi:10.7326/0003-4819-141-10-200411160-00009 %m 15545678

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*(4), 336-352. doi:https://doi.org/10.1016/S0005-7894(84)80002-7

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*, 20-20. doi:10.1186/s40337-017-0145-3

Le Grange, D., Eisler, I., Dare, C., & Russell, G. F. M. (1992). Evaluation of family treatments in adolescent anorexia nervosa: A pilot study. *International Journal of Eating Disorders, 12*(4), 347-357. doi:10.1002/1098-108X(199212)12:4<347::AID-EAT2260120402>3.0.CO;2-W

Le Grange, D., Hughes, E. K., Court, A., Yeo, M., Crosby, R. D., & Sawyer, S. M. (2016). Randomized Clinical Trial of Parent-Focused Treatment and Family-Based Treatment for Adolescent Anorexia Nervosa. *J Am Acad Child Adolesc Psychiatry, 55*(8), 683-692. doi:10.1016/j.jaac.2016.05.007

Linden, M. (2013). How to Define, Find and Classify Side Effects in Psychotherapy: From Unwanted Events to Adverse Treatment Reactions. *Clinical Psychology & Psychotherapy, 20*(4), 286-296. doi:https://doi.org/10.1002/cpp.1765

Lock, J., Agras, W., Bryson, S., & Kraemer, H. C. (2005). A Comparison of Short- and Long-Term Family Therapy for Adolescent Anorexia Nervosa. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(7), 632-639.

Lock, J., Agras, W. S., Bryson, S. W., Brandt, H., Halmi, K. A., Kaye, W., . . . Jo, B. (2016). Does family-based treatment reduce the need for hospitalization in adolescent anorexia nervosa? *Int J Eat Disord, 49*(9), 891-894. doi:10.1002/eat.22536

Lock, J., Brandt, H., Woodside, B., Agras, S., Halmi, W., Johnson, C., . . . Wilfley, D. (2012). Challenges in conducting a multi-site randomized clinical trial comparing treatments for adolescent anorexia nervosa. *International Journal of Eating Disorders, 45*(2), 202-213.

Lock, J., Couturier, J., Bryson, S., & Agras, S. (2006). Predictors of dropout and remission in family therapy for adolescent anorexia nervosa in a randomized clinical trial. *International Journal of Eating Disorders, 39*(8), 639-647. doi:10.1002/eat.20328

Lock, J., & Le Grange, D. (2001). Can family-based treatment of anorexia nervosa be manualized. *Journal of Psychotherapy Practice and Research, 10*(4), 253-261. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-0034798818&partnerID=40&md5=82ab4556f3ecedfbf8b970bfba2442f5

Lock, J., Le Grange, D., Agras, W., Moye, A., Bryson, S. W., & Jo, B. (2010). Randomized clinical trial comparing family-based treatment with adolescent-focused individual therapy for adolescents with anorexia nervosa. *Archives of General Psychiatry, 67*(10), 1025-1032.

McIntosh, V. V. W., Jordan, J., Carter, F. A., Luty, S. E., McKenzie, J. M., Bulik, C. M., . . . Joyce, P. R. (2005). Three psychotherapies for anorexia nervosa: a randomized, controlled trial. *The American Journal of Psychiatry, 162*(4), 741-747. doi:10.1176/appi.ajp.162.4.741

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ, 339*, b2535. doi:10.1136/bmj.b2535

Morgan, H. G., & Russell, G. F. M. (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. doi:10.1017/S0033291700056981

Murray, S. B., Quintana, D. S., Loeb, K. L., Griffiths, S., & Le Grange, D. (2019). Treatment outcomes for anorexia nervosa: a systematic review and meta-analysis of randomized controlled trials. *Psychol Med, 49*(4), 535-544. doi:10.1017/s0033291718002088

Nyman-Carlsson, E., Birgegård, A., Engström, I., Gustafsson, S. A., & Nevonen, L. (2019). Predictors of outcome among young adult patients with anorexia nervosa in a randomised controlled trial. *Eur Eat Disord Rev, 27*(1), 76-85. doi:10.1002/erv.2630

Nyman-Carlsson, E., Norring, C., Engström, I., Gustafsson, S. A., Lindberg, K., Paulson-Karlsson, G., & Nevonen, L. (2020). Individual cognitive behavioral therapy and combined family/individual therapy for young adults with Anorexia nervosa: A randomized controlled trial. *Psychother Res, 30*(8), 1011-1025. doi:10.1080/10503307.2019.1686190

Offord, A., Turner, H., & Cooper, M. (2006). Adolescent inpatient treatment for anorexia nervosa: a qualitative study exploring young adults' retrospective views of treatment and discharge. *European Eating Disorders Review, 14*(6), 377-387. doi:https://doi.org/10.1002/erv.687

Parling, T., Cernvall, M., Ramklint, M., Holmgren, S., & Ghaderi, A. (2016). A randomised trial of Acceptance and Commitment Therapy for Anorexia Nervosa after daycare treatment, including five-year follow-up. *BMC Psychiatry, 16*(1). doi:10.1186/s12888-016-0975-6

Parry, G. D., Crawford, M. J., & Duggan, C. (2016). Iatrogenic harm from psychological therapies – time to move on. *British Journal of Psychiatry, 208*(3), 210-212. doi:10.1192/bjp.bp.115.163618

Phillips, R., Hazell, L., Sauzet, O., & Cornelius, V. (2019). Analysis and reporting of adverse events in randomised controlled trials: a review. *BMJ open, 9*(2), e024537. doi:10.1136/bmjopen-2018-024537

Robin, A. L., Siegel, P. T., Koepke, T., Move, A. W., & Tice, S. (1994). Family therapy versus individual therapy for adolescent females with anorexia nervosa. *Journal of Developmental and Behavioral Pediatrics, 15*(2), 111-116. doi:10.1097/00004703-199404000-00008

Robin, A. L., Siegel, P. T., Moye, A. W., Gilroy, M., Dennis, A. B., & Sikand, A. (1999). A controlled comparison of family versus individual therapy for adolescents with anorexia nervosa. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*(12), 1482-1489. doi:10.1097/00004583-199912000-00008

Schlegl, S., Diedrich, A., Neumayr, C., Fumi, M., Naab, S., & Voderholzer, U. (2016). Inpatient Treatment for Adolescents with Anorexia Nervosa: Clinical Significance and

Predictors of Treatment Outcome. *European Eating Disorders Review, 24*(3), 214-222. doi:https://doi.org/10.1002/erv.2416

Schlegl, S., Quadflieg, N., Löwe, B., Cuntz, U., & Voderholzer, U. (2014). Specialized inpatient treatment of adult anorexia nervosa: effectiveness and clinical significance of changes. *BMC Psychiatry, 14*(1), 258. doi:10.1186/s12888-014-0258-z

Schmidt, U., Magill, N., Renwick, B., Keyes, A., Kenyon, M., Dejong, H., . . . Landau, S. (2015). The Maudsley Outpatient Study of Treatments for Anorexia Nervosa and Related Conditions (MOSAIC): Comparison of the Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA) with specialist supportive clinical management (SSCM) in outpatients with broadly defined anorexia nervosa: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 83*(4), 796-807. doi:10.1037/ccp0000019

Schmidt, U., Renwick, B., Lose, A., Kenyon, M., DeJong, H., Broadbent, H., . . . Landau, S. (2013). The MOSAIC study - comparison of the Maudsley Model of Treatment for Adults with Anorexia Nervosa (MANTRA) with Specialist Supportive Clinical Management (SSCM) in outpatients with anorexia nervosa or eating disorder not otherwise specified, anorexia nervosa type: Study protocol for a randomized controlled trial. *Trials, 14*(1). doi:10.1186/1745-6215-14-160

Serfaty, M. A., Turkington, D., Heap, M., Ledsham, L., & Jolley, E. (1999). Cognitive therapy versus dietary counselling in the outpatient treatment of anorexia nervosa: Effects of the treatment phase. *European Eating Disorders Review, 7*(5), 334-350.

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

Steinhausen, H. C. (2002). The Outcome of Anorexia Nervosa in the 20th Century. *American Journal of Psychiatry, 159*(8), 1284-1293. doi:10.1176/appi.ajp.159.8.1284

Touyz, S., Le Grange, D., Lacey, H., Hay, P., Smith, R., Maguire, S., . . . Crosby, R. (2013). Treating severe and enduring anorexia nervosa: A randomized controlled trial. *Psychological Medicine, 43*(12), 2501-2511.

Treasure, J., Crane, A., McKnight, R., Buchanan, E., & Wolfe, M. (2011). First do no harm: Iatrogenic Maintaining Factors in Anorexia Nervosa. *European Eating Disorders Review, 19*(4), 296-302. doi:https://doi.org/10.1002/erv.1056

Wild, B., Friederich, H.-C., Gross, G., Teufel, M., Herzog, W., Giel, K. E., . . . Zipfel, S. (2009). The ANTOP study: focal psychodynamic psychotherapy, cognitive-behavioural therapy, and treatment-as-usual in outpatients with anorexia nervosa--a randomized controlled trial. *Trials, 10*, 23. doi:10.1186/1745-6215-10-23

Zipfel, S., Wild, B., Groß, G., Friederich, H. C., Teufel, M., Schellberg, D., . . . Herzog, W. (2014). Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): randomised controlled trial. *Lancet, 383*(9912), 127-137. doi:10.1016/s0140-6736(13)61746-8

Figure 1: PRISMA 2020 flow diagram of study selection process

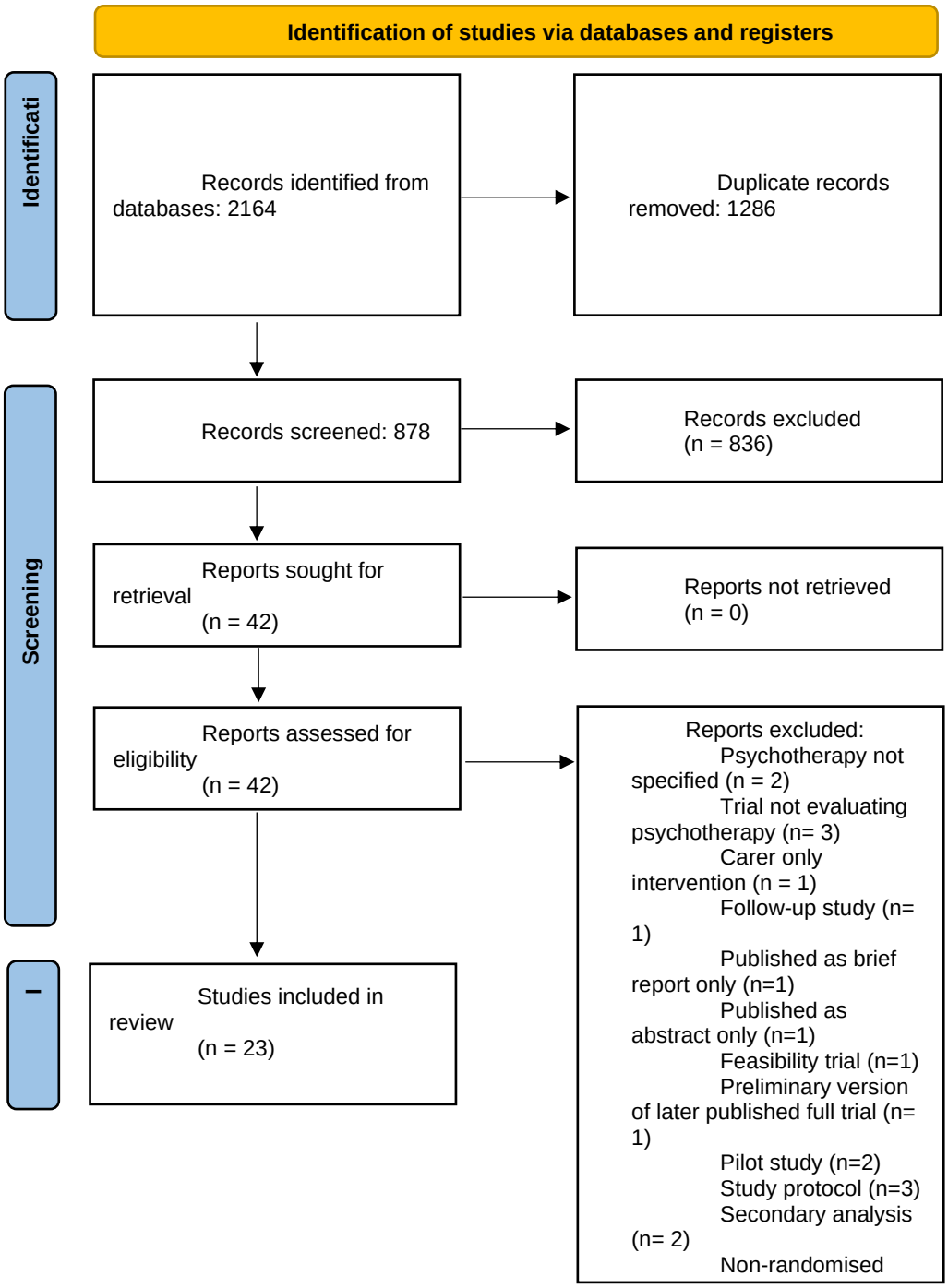

Table 1: Overview of study characteristics and unwanted event outcome reporting.

Paper	Linked papers/ trial documents	Population	Sample Size	Interventions	Non-Compliance	Treatment Non-Response	Symptom Deterioration	Mortality (*n*, %)
Agras et al., 2014	Trial registration documents available at clinicaltrials.gov, #NCT00610753. Trial design detailed in previous paper (Lock et al., 2012).	Adolescents (age 12-18) meeting DSM-IV criteria for AN (excluding amenorrhea criterion). Outpatient setting.	164	Family Based Therapy (FBT), Systemic Family Therapy (SyFT)	Dropout rate 26% FBT, 25% SyFT. Withdrawal due to medical instability rate 4% FBT, 9% SyFT.	Not in remission at EOT: 70% FBT, 75% SyFT. At 12 month FU: 59% FBT, 61% SyFT.	Withdrawn due to medical instability: 4% FBT, 9% SyFT. Participants experiencing an adverse event relating to hospitalisation or suicidality: 15% FBT, 25% SyFT.	0 (0%)
Ball & Mitchell, 2004	No	Female adolescents and adults (age 13-23). Included participants weighing less than 90% of their average body	25	Cognitive Behavioural Therapy (CBT), Behavioural Family Therapy (BFT)	Total non-completion 28%. 12% withdrawn due to hospitalisation.	Poor outcome at EOT: 28%. No difference at 6 month FU.	Hospitalisation: total 12%.	NR

		weight. Outpatient setting.						
Brockmeyer et al., 2021	Trial registration documents available at International Clinical Trials Registry Platform, #DRKS00006870	Adult and adolescents (aged over 17) meeting DSM-V criteria for AN. Inpatient setting.	167	Cognitive Remediation Therapy (CRT), Arts Therapy (ART)	Non-completion 13% CRT, 14% ART.	No.	No.	NR
Byrne et al., 2017	Trial registration documents available at Australian New Zealand Clinical Trials Registry, #ACTRN12611000725965.	Adult and adolescents (aged over 17) meeting DSM-IV criteria for AN, including BMIs below 18.5. Outpatient setting.	120	Specialist Supportive Clinical Management (SSCM), Maudsley Model Anorexia Nervosa Treatment for Adults (MANTRA), Enhanced Cognitive Behavioural Therapy (CBT-E)	Non-completion 42% SSCM, 43% MANTRA, 33% CBT-E. Drop-out rate 28.% SSCM, 24% MANTRA, 28% CBT-E. Withdrawal due to BMI falling below 14 8% SSCM, 20%	Not in remission at 12 month FU: SSCM 68%, MANTRA 78%, CBT-E 69%. Not achieving BMI over 18.5 at 12 month FU: 48% SSCM, 49% MANTRA,	Withdrawal due to BMI falling below 14: 8% SSCM, 20% MANTRA, 5.% CBT-E. Hospitalisation: 10% SSCM, 5% MANTRA, 5% CBT-E.	0 (0%)

					MANTRA, 5.% CBT-E. One withdrawal due to recommencing psychotropic medication.	51% CBT-E.		
Channon et al., 1989	No	Females (age range not specified) meeting Russell's (1983) criteria for AN. Outpatient setting.	24	Cognitive Behavioural Treatment (CBT), Behavioural Treatment (BT), treatment as usual (TAU)	Drop-out rate 0% CBT, 13% BT, 25% TAU.	No	Hospitalisation: 20% CBT, 12% BT, 50% TAU.	NR
Dalle Grave et al., 2013	No.	Adult and adolescents (aged 14- 65) meeting DSM-IV criteria for AN. Inpatient setting.	80	Enhanced Cognitive Behavioural Therapy focusing on AN (CBT-Ef), CBT-E including other areas of difficulty (CBT-Eb)	Non-completion 12% CBT-Ef, 8% CBT-Eb.	Not achieving BMI over 18.5 at EOT: 14% (total sample). 6 month FU: 52%, 12 month FU:	Hospitalisation: 3% (total sample).	NR

						50% (total sample). Minimal symptoms at EOT: 51% (total sample). 12 month FU: 38% (total sample).		
Dare et al., 2001	No	Adults (aged over 18) meeting DSM-IV criteria for AN. Outpatient setting.	84	Focal psychoanalytic psychotherapy (FPP), Cognitive Analytic Therapy (CAT), Family Therapy (FT), low contact routine treatment (TAU).	Drop-out rate 35% (total sample).	Poor outcome at EOT: 48% FPP, 59% FT, 68% CAT, 73% TAU.	Hospitalisation: 14% (total sample).	1 (1%)
Eisler et al., 2000	Pilot study (Le Grange, Eisler, Dare, & Russell, 1992)	Adolescents (age 11- 17) meeting DSM-IV or ICD-10 criteria for AN.	40	Conjoint Family Therapy (CFT), Separated Family Therapy (SFT)	Drop-out rate within first 3 months of treatment 10% (total sample).	Poor outcome at EOT: 53% CFT, 24% SFT.	Hospitalisation: 10% (total sample).	NR

		Outpatient setting.						
Eisler et al., 2016	Trial registration documents available at ISRCTN Registry, #ISRCTN11275465.	Adolescents (age range 13-20 years) meeting DSM-IV criteria for AN or EDNOS (restricting type), below 68% median BMI or lost 15% body weight over last 3 months. Outpatient setting.	169	Multifamily Therapy (MFT), Single Family Therapy (SFT)	Drop-out rate <11% (total sample).	Poor outcome at EOT: 42% SFT, 24% MFT. 6 month FU: 43% SFT, 22% MFT.	Hospitalisation: MFT 4%, 3% SFT.	NR
Fichter et al., 2012	No	Female adolescents and adults (minimum age 16) meeting DSM-IV criteria for AN (excluding	258	CBT-based internet relapse prevention (RP), TAU	No to low engagement with intervention: 24.2% (total sample). Drop-outs: 6% RP, 6.% TAU.	No	BMI falling below 13.5: 4% (total sample). Mean weight loss at EOT in TAU group.	0 (0%)

		amenorrhea criterion). Post-inpatient discharge setting.					Hospitalisation for inpatient treatment: 22% RP, 10% of TAU.	
Geist et al., 2000	No	Female adolescents (age range 12-17) meeting DSM-IV criteria for AN. Including participants weight <90% IBW. Outpatient treatment offered following inpatient admission.	25	Family Therapy (FT), Family Group Psychoeducation (FP).	NR	NR	Hospitalisation for inpatient treatment during intervention: 20% (total sample. FU period: 24% (total sample).	NR
Godart et al., 2012	Trial protocol provided as supplementary material but only available in French, therefore not included.	Female adolescents and adults (age 13-21) meeting DSM-IV	60	Family Therapy in addition to TAU (FT), TAU.	Non-completion: 7% TAU, 10% FT.	Poor outcome at 18 months following inclusion:	Hospitalisation: 40% FT, 55% TAU.	NR

		criteria for AN. Outpatient treatment offered following inpatient admission.				60% FT, 83% TAU.		
Le Grange et al., 2016	Trial registration documents available at Australian New Zealand Clinical Trials Registry, #ACTRN12610000216011. Study protocol available (Hughes et al., 2014).	Adolescents (age 12-18) meeting DSM-IV criteria for AN (excluding amenhorrea). Included participants weighing less than 90% of their average body weight. Outpatient setting.	107	Family Based Treatment (FBT), Parent Focused Treatment (PFT)	Non-completion: 15% (total sample).	Not in remission at EOT: 57% PFT, 78% FBT. At 6 month FU: 41% PFT, 78% FTB.	Hospitalisation: 12% PFT, 24% FBT. 83% of hospitalisations for medical reasons.	NR
Lock et al., 2005	Pilot study available (Lock & Le Grange, 2001).	Adolescents (age 12-18) meeting DSM-IV criteria for	86	Short term family therapy, Long term family therapy	Drop-out rate: 4% Short term FT, 16% Long term FT	Not in remission at EOT: 4%-33% (depending	Hospitalisation: 22% (total sample).	NR

		AN. Included some partially weight restored participants and those not meeting amenhorrea criterion. Outpatient setting.				on definitions).		
Lock et al., 2010	Trial registration documents available at clinicaltrials.gov, #NCT00149786	Adolescents (age 12-18) meeting DSM-IV criteria for AN (excluding amenorrhea criterion). Outpatient setting.	121	Family based treatment (FBT), adolescent focused individual therapy (AFT).	Drop out rate: 14% at EOT, 22% at FU (total sample).	Not in remission at EOT: 58% FBT and 77% AFT. At 6 month FU: 60% FBT, 82% AFT. Not in partial remission at EOT: 11% FBT, 33% AFT.	Hospitalisation: AFT 37%, FBT 15%. Hospitalisation during FU period: 18% AFT, 8% FBT.	NR

						Achieved full remission but relapsed by 12 month FU: 10% FBT, 40% AFT. Achieved partial remission but relapsed by 12 month FU: 18% FBT, 6% AFT.		
McIntosh et al., 2005	No	Adult females (ages 17-40) meeting DSM-IV criteria for AN (not including amenhorrea criterion). Included participants with a BMI <19.	56	Cognitive behavioural therapy (CBT), interpersonal psychotherapy (IP), control treatment (non-specific supportive clinical management; NSCM)	Non-completion: 38% (total sample).	Poor outcome at EOT: 67% IP, 53% CBT, 38% NSCM.	Hospitalisation: 7% (total sample).	1 (2%).

		Outpatient setting.						
Nyman-Carlsson et al., 2020	Trial registration documents available at ISRCTN Registry, ISRCTN25181390	Adolescents and adults (ages 17-24) meeting DSM-IV criteria for AN. Outpatient setting.	78	Cognitive Behavioural Therapy (CBT), combined individual/family therapy (FT)	Drop-out: 30% CBT, 8% FT. Withdrawal due to adverse events: 24% CBT, 22% FT.	BMI below 17.5 at EOT/ at FU: 83%-92% CBT, 83%- 95% FT. Meeting criteria for AN et EOT: 8% CBT, 8% FT. At FU: 5% CBT, 0% FT.	Proportions not reported.	NR
Parling et al., 2016	Trial registration documents available at ISRCTN Registry, ISRCTN12106530	Adults (ages 18-51) meeting DSM-IV criteria for full or partial AN. Outpatient setting following	43	Acceptance and Commitment Therapy (ACT), treatment as usual (TAU)	Non-completion: 42% (ACT).	Not achieving a good outcome at EOT: 75% ACT, 88% TAU. 6 month FU: 73% ACT, 80% TAU. 18 month	Proportions not reported.	1 (2%).

		daycare treatment.				FU: 58% ACT, 69% TAU. 24 month FU: 64% ACT, 89% TAU. 5 year FU: 44% ACT, 64% TAU.		
Robin et al., 1999	Robin et al., 1994 (earlier paper reporting on smaller sample).	Adolescents (ages 11-20) meeting DSM-III criteria for restrictive AN. Outpatient setting.	37	Behavioural family systems therapy (BFST), ego orientated individual therapy (EOIT)	Non-completion 8% (see full paper for details).	Not attaining target weight at EOT: 33% BFST, 31% EOIT. 1 year FU: 20% BFST, 31% EOIT. Not attaining 25th percentile of body mass for age at EOT: 16% BFST, 18% EOIT. At 1 year FU: 13%	Hospitalisation: 58% BFT, 28% EOIT.	NR

						BFST, 7% EOIT. Not attaining 50th percentile of body mass for age at EOT: 47% BFST, 59% EOIT. 1 year FU: 33% BFST, 53% EOIT		
Schmidt et al., 2015	Trial registration documents available at ISRCTN Registry, #ISRCTN67720902. Trial protocol (U. Schmidt et al., 2013)	Adults (age 18-60) meeting DSM-IV criteria for AN. Included those not meeting amenhorrea criterion, without fat phobia, or partial AN. Outpatient setting.	142	Maudsley Model of Treatment for Adults with Anorexia Nervosa (MANTRA), Specialist Supportive Clinical Management (SSCM).	Non-completion: 25% MANTRA, 41% SSCM.	Not recovered at 6 months: 29% SSCM, 24% MANTRA. 12 months: 18% SSCM, 16% MANTRA.	Accessed inpatient or daycare treatment for ED: 7% MANTRA, 11.42% SSCM.	1 (0.7%)

Serfaty et al., 1999	No	Adolescents and adults (minimum age 16) meeting DSM-III criteria for AN. Outpatient setting.	35	Cognitive therapy (CT), dietary advice (D)	Drop-out: 8% CT, 100% D.	Still meeting diagnostic criteria for AN at EOT: 44% CT, 100% D.	NR	NR
Touyz et al., 2013	Trial registration documents available at Australian New Zealand Clinical Trials Registry, # ACTRN12607000440426	Adults meeting DSM-IV criteria for DSM-IV (excluding amenhorrea criterion) for over 7 years). Included participants with a BMI between 17.5 and 18.5.	63	Cognitive Behavioural Therapy (CBT), Specialist Supportive Clinical Management (SSCM). Both modified for severe and enduring AN.	Non-completion: 16% CBT, 91% SSCM.	NR	Hospitalisation: 11%	1 (2%).

		Outpatient setting.						
Zipfel et al., 2014	Trial registration documents available at ISRCTN Registry, #ISRCTN72809357 Protocol (Wild et al., 2009).	Adults meeting DSM-IV criteria for DSM-IV, BMI range 15- 18.5. Outpatient setting.	242	Focal psychodynamic therapy (FPT), Enhanced Cognitive Behavioural Therapy (CBT-E), treatment as usual (TAU).	Non-completion: 22% (total sample).	Exact proportions NR.	Hospitalisation: 23% FPT, 34% CBT-E, 41% TAU.	0 (0%)

Abbreviations:
ACT: Acceptance and Commitment Therapy. AFT: Adolescent Focused Individual Therapy. ART: Arts Therapy. BFST: Behavioural Family Systems Therapy. BFT: Behavioural Family Therapy. BT: Behavioural Treatment. BMI: Body Mass Index. CAT: Cognitive Analytic Therapy. CBT: Cognitive Behavioural Therapy or Cognitive Behavioural Treatment. CBT-E: Enhanced Cognitive Behavioural Therapy. CBT-Eb: Broad Enhanced Cognitive Behavioural Therapy. CBT-Ef: Focused Enhanced Cognitive Behavioural Therapy. CFT: Conjoint Family Therapy. CRT: Cognitive Remediation Therapy. CT: Cognitive Therapy. D: Dietary Advice. DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. DSM-V: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. EOT: End of Treatment. EOIT: Ego Orientated Individual Therapy. FBT: Family Based Therapy or Family Based Treatment. FP: Family Group Psychoeducation. FPP: Focal Psychoanalytic Psychotherapy. FPT: Focal Psychodynamic Therapy. FT: Family Therapy or Combined Individual/ Family Therapy. FU: Follow up. IP: Interpersonal Psychotherapy. MANTRA: Maudsley Model Anorexia Nervosa Treatment for Adults. MFT: Multifamily Therapy. NR: Not Recorded. NSCM: Non-Specific Supportive Clinical Management. PFT: Parent Focused Treatment. RP: Relapse Prevention. SFT: Separated Family Therapy or Single Family Therapy. SSCM: Specialist Supportive Clinical Management. SyFT: Systemic Family Therapy. TAU: Treatment as Usual.

Table 2: Overview of patient clinical and sociodemographic characteristics.

Paper	Location of study	Participant mean age in years (*SD*)-	Participant sex assigned at birth	Mean baseline BMI (*SD*)	Mean illness duration (*SD*)	Participant race and ethnicity	Participant socioeconomic status
Agras et al., 2014	North America	15.3 (1.8)	89.2% female	Mean percentage of ideal body weight 81.9%	13.5 months (13.9)	79.1% of participants white. 10.1% Hispanic, 5.1% Asian.	NR
Ball & Mitchell, 2004	Australia	CBT group: 18.45 (2.57) BFT group: 17.58 (3.37)	100% female	CBT group: 16.06 (1.58) BFT group: 16.45 (0.85)	NR	NR	NR
Brockmeyer et al., 2021	Germany	CRT group: 26.06 (9.06) ART group: 25.58 (8.26)	CRT group: 97.6% female ART group: 97.6% female	CRT group: 14.90 (1.59) ART group: 14.36 (2.12)	CRT group: 50% ≤ 5 years ART group: 55.3% ≤ 5 years	NR	NR
Byrne et al., 2017	Australia	26.19 (9.47)	95.83% female	16.70 (1.22)	4 years	NR	NR
Channon et al., 1989	United Kingdom	CBT group: 21.63 (5.88) BT group: 24.13 (5.77) Control group: 25.75 (7.19)	100% female	CBT group: 14.85 (1.10) BT group: 16.06 (1.42) Control group: 14.90 (1.49)	CBT group: 5.13 years (4.85) BT group: 3.13 (1.73) Control group: 7.75 (6.09)	NR	NR
Dalle Grave et al., 2013	Italy	23.4 (6.9)	97.5% female	78.8% participants	5 years	NR	Reports occupational status

				with BMI ≤ 16.0			
Dare et al., 2001	United Kingdom	26.3 (6.7)	98% female	15.4 (1.6)	6.3 years (5.9)	NR	NR
Eisler et al., 2000	United Kingdom	15.5 (1.6)	99% female	Mean percentage of average body weight 74.3% (9.8)	12.9 months (9.4)	NR	NR
Eisler et al., 2016	United Kingdom	FT group: 15.7 (1.6) MFT group: 15.7 (1.7)	FT group: 93.9% female MFT group: 88.2% female	FT group: 15.8 (1.2) MFT group: 15.7 (1.3)	FT group: 11.4 months (12.6) MFT group: 9.6 months (10.5)	FT group: 92.7% white, 3.7% other, 3.7% missing MFT group: 88.2% white, 3.5% other, 8.2% missing	NR
Fichter et al., 2012	Germany	RP group: 23.8 (6.5) TAU group: 24.1 (3.6)	100% female	RP group: 17.8 (1.4) TAU group: 17.7 (1.2)	NR	NR	NR
Geist et al., 2000	Canada	FT group: 14.3 (1.5) FGP group: 14.9 (1.7)	100% female	Mean percentage of ideal body weight in FT group: 77.7% FGP group: 77.2%	NR	NR	NR
Godart et al., 2012	France	16.6 (1.6)	100% female	16.9 (1.1)	16.6 months (6.8)	NR	NR
Le Grange et al., 2016	Australia	15.5 (1.5)	87.7% female	16.5 (1.3)	10.5 months (8.8)	NR	Reports % of parents with

							university degrees.
Lock et al., 2005	North America	Long term FT 15.2 (1.7) Short term FT: 15.2 (1.6)	Long term FT: 91% female Short term FT: 89% female	17.1 (1.4)	Long term FT: 12.0 months (9.9) Short term FT: 11.3 months (10.4)	Long term FT: 5% Asian, 76% White, 14% Hispanic, 5% Other. Short term FT: 14% Asian, 73% White, 9% Hispanic, 2% Native American, 2% Other.	Reports family income and parental education.
Lock et al., 2010	North America	14.4 (1.6)	91% female	16.1 (1.1	11.3 months (8.6)	24% ethnic minorities	NR
McIntosh et al., 2005	New Zealand	NR	100% female	17.3 (1.1)	NR	NR	NR
Nyman-Carlsson et al., 2020	Sweden	CBT: 19.1 (1.9) FT: 18.7 (2.0)	100% female	CBT: 16.49 (0.8) FT: 16.54 (0.9)	CBT: 31.6 months (24.1) FT: 26.8 months (24.4)	NR	Reports education level
Parling et al., 2016	Sweden	ACT: 26.6 (8.1) TAU: 24.6 (5.8)	97.67% female	ACT: 17.5 (2.3) TAU: 18.1 (2.6)	NR	NR	Reports educational and occupational status
Robin et al., 1999	North America	BFST: 14.9 EOIT: 13.4	100% female	BFST: 86.5 pounds EOIT: 86.8 pounds	<12 months	94.5% White, 5.5% Middle Eastern	Reports socioeconomic status on Hollingshead index.

Schmidt et al., 2015	United Kingdom	26.7 (7.7)	97.89% female	26.6 (1.2)	8.3 years (7.3)	NR	Reports years in education
Serfaty et al., 1999	United Kingdom	CT: 22.1 (6.6) Control: 17.9 (5.6)	94.29% female	CT: 16.2 (1.6) Control: 17.0 (4.0)	CT: 5.0 years (5.5) Control: 2.2 (4.5)	NR	Reports social class
Touyz et al., 2013	Australia/ United Kingdom	33.4 (9.6)	100% female	16.2 (1.3)	16.6 (8.5)	NR	Reports educational and occupational status
Zipfel et al., 2014	Germany	FPT: 28.0 (8.6) CBT: 27.4 (7.9) TAU: 27.7 (8.1)	100% female	FPT: 16.57 (1.0) CBT: 16.82 (1.0) TAU: 16.75 (1.0)	FPT: 61% ≤ 6 years CBT: 61% ≤ 6 years TAU: 61% ≤ 6 years	NR	NR

Abbreviations:
ACT: Acceptance and Commitment Therapy. AFT: Adolescent Focused Individual Therapy. ART: Arts Therapy. BFST: Behavioural Family Systems Therapy. BFT: Behavioural Family Therapy. BT: Behavioural Treatment. CAT: Cognitive Analytic Therapy. CBT: Cognitive Behavioural Therapy or Cognitive Behavioural Treatment. CBT-E: Enhanced Cognitive Behavioural Therapy. CBT-Eb: Broad Enhanced Cognitive Behavioural Therapy. CBT-Ef: Focused Enhanced Cognitive Behavioural Therapy. CFT: Conjoint Family Therapy. CRT: Cognitive Remediation Therapy. CT: Cognitive Therapy. D: Dietary Advice. EOIT: Ego Orientated Individual Therapy. FBT: Family Based Therapy or Family Based Treatment. FP: Family Group Psychoeducation. FPP: Focal Psychoanalytic Psychotherapy. FPT: Focal Psychodynamic Therapy. FT: Family Therapy or Combined Individual/ Family Therapy. IP: Interpersonal Psychotherapy. MANTRA: Maudsley Model Anorexia Nervosa Treatment for Adults. MFT: Multifamily Therapy. NR: Not Reported. NSCM: Non-Specific Supportive Clinical Management. PFT: Parent Focused Treatment. RP: Relapse Prevention. SFT: Separated Family Therapy or Single Family Therapy. SSCM: Specialist Supportive Clinical Management. SyFT: Systemic Family Therapy. TAU: Treatment as Usual.

4. Service Improvement Project: Evaluating and addressing staff training needs in relation to autism in an older adult mental health service.

Emma Kinnaird (Emma.kinnaird@hmc.ox.ac.uk)

Supervised by Dr Kathryn Evans (internal) and Dr Anna Crabtree (external).

Word count: 4952

First submitted May 2022.

The proposed journal for this paper was Autism given its specific subject area and history of publishing papers on the topic of older adults. The paper has been formatted according to journal guidelines (Appendix 2). A shortened version of this manuscript has been accepted for publication in the British Psychological Society's Faculty for the Psychology of Older People quarterly bulletin.

4.1 Abstract

Autistic adults are at higher risk of experiencing mental health problems compared to their neurotypical peers and are more likely to report unmet healthcare needs and poor service experiences. Research suggests that this is closely related to a lack of understanding of autism by clinicians working in this area. Therefore, there is an increasing impetus in the National Health Service (NHS) for staff working in mental healthcare to receive autism training. However, there is a lack of understanding around autism in older adulthood, and whether older adult mental health teams could benefit from specialised training that reflects the specific needs of this population. This project used clinical audit and needs survey methodologies to explore whether an older adult mental health team might benefit from autism training, and to understand what clinicians might find helpful from a training intervention. Whilst the audit identified that the team did see patients on the autism spectrum, it was rare for the team to work with patients with a diagnosis of autism. Rather, clinicians were more likely to work with patients where autism is suspected but unconfirmed. Most clinicians reported never having received any training in autism, and highlighted autism in older adults, and autism and mental health, as specific training priorities. A training programme based on these responses was developed, implemented, and evaluated using pre- and post-training surveys. The training was rated as useful by attendees and resulted in non-statistically significant increases in self-rated knowledge. The findings of this study suggest that clinicians working in older adult mental health can benefit from autism training. However, instead of generic autism training, clinicians may particularly benefit from training specialised to their older adult context. Further research is required to better understand the area of older adulthood and autism, and to evaluate how to effectively enhance the capability of staff working in this area.

4.2 Introduction

4.2.1 Summary of Relevant Literature

Autism is a neurodevelopmental condition characterized by “persistent difficulties with social communication and social interaction”, and “restricted and repetitive patterns of behaviours, activities or interests” (American Psychiatric Association, 2013). It is estimated that around 0.8% of adults in the UK meet diagnostic criteria for autism (McManus, Bebbington, Jenkins, & Brugha, 2016). This paper will use the terms “autism”, “on the autism spectrum”, and “autistic” as this language is preferred by the adult autism community (Kenny et al., 2016).

Autistic adults are at a higher risk of experiencing mental and physical health conditions compared to non-autistic people and have lower life expectancies and a higher risk of dying by suicide (Croen et al., 2015; Hirvikoski et al., 2016). With 80% of autistic people experiencing at least one mental health condition across their lifetime, mental health difficulties in this population are normative rather than atypical (Lever & Geurts, 2016). In addition to experiencing heightened rates of mental health problems, autistic adults also report high levels of unmet healthcare needs (particularly relating to mental health) and low levels of satisfaction with service provision (Nicolaidis et al., 2013; Tint & Weiss, 2018).

Significantly, there appears to be a link between these poor healthcare experiences, and clinician knowledge and confidence in working with autistic people (Maddox et al., 2020). A UK-based inquiry into healthcare access for autistic people identified that this population find it difficult to access services and feel that healthcare professionals do not understand autism or associated needs, particularly in relation to the overlap between autism and mental health (The Westminster Commission on Autism, 2016). Moreover, clinicians working in mental health report a lack of knowledge, experience, competence, and

confidence in treating autistic people (Maddox et al., 2020). For this reason there is a risk that autistic patients may be rejected from services who feel that they do not have the expertise to work with this population, resulting in them falling in-between existing service provision (Department of Health, 2015).

Therefore, research suggests that although autistic adults are more likely to require support for mental health problems, there is a strong risk that existing service provision is not meeting the needs of this population. Furthermore, an important factor in services successfully providing adequate care is clinician knowledge, competence, and confidence in autism. This need is increasingly being recognized in government and NHS policies in the UK. The NHS is legally mandated to provide autistic people with equal access to mental health services, with care delivered by staff with appropriate training (Department of Health, 2015; HM Government, 2014). NHS England has additionally released improvement standards for NHS trusts to improve care for autistic people, including the stipulation that trusts must ensure that staff have "specialist knowledge and skills" in meeting the needs of autistic people, and that staff are provided with training in delivering adjusted care for autistic people that accounts for their "rights, unique needs and health vulnerabilities" (NHS Improvement, 2018). Finally, NICE guidelines clearly outline that autistic people experiencing co-occurring mental health problems should receive adapted psychological treatment delivered by clinicians with an understanding of autism (NICE, 2016).

Therefore, NHS trusts increasingly mandate that staff must attend autism training: at the time of writing, the Oxford Health NHS Foundation Trust (the setting for this project) was beginning to roll out compulsory autism training for frontline staff. However, the skills necessary for working with autistic people vary across healthcare settings, and generic autism training may not sufficiently meet the needs of clinicians working in more specialized areas (Dillenburger, McKerr, Jordan, & Keenan, 2016). It is therefore possible that an effective

intervention in increasing staff capability is one that identifies and targets their specific needs. To date, the needs of staff working in older adult mental health services in relation to autism, and how best these needs can be addressed, are under-explored.

This is reflective of a wider lack of research on autism in older adulthood, including on how aging may impact older autistic adults, and whether they have specific needs compared to their neurotypical peers (Tse, Lei, Crabtree, Mandy, & Stott, 2021). Existing knowledge and training on autism in adulthood may not fully capture the needs and experiences of autistic older adults as it is based in research literature that focuses on younger adults. Older adults with mental health problems may experience specific needs that require treatment adaptations, and the needs of autistic people within this population are not fully understood (Bennett, 2016; Hand, Angell, Harris, & Carpenter, 2020; Laidlaw & Kishita, 2015; Michael, 2016).

For example, historical changes in diagnostic criteria may mean that older adults who meet modern criteria for autism do not have an official diagnosis (Bennett, 2016). This may influence how behaviours and difficulties experienced by older adults on the autism spectrum are understood, both by them as individuals, their families, and their clinicians. Research on undiagnosed autism in other demographic groups, notably women, suggests that difficulties experienced by these individuals are at risk of being misinterpreted: for example, food refusal due to sensory problems in people with co-occurring eating disorders and undiagnosed autism being misattributed to the eating disorder, resulting in poor healthcare experiences (Kinnaird, Norton, Stewart, & Tchanturia, 2019). Furthermore, there are concerns that individuals with undiagnosed autism may be at risk of misdiagnosis of other mental health conditions, notably personality disorders (May, Pilkington, Younan, & Williams, 2021). The absence of a diagnosis may also mean that specific needs associated with autism are not considered when providing care for older adults. For example, despite evidence that transitions can cause high

stress and anxiety in autistic children and require management and planning, there is a lack of understanding around autistic people experience transitions associated with aging such as retirement or entering residential care (Stoner, Angell, House, & Bock, 2007).

Autistic older adults may also experience differing physical and cognitive needs: as autistic adults age, evidence suggests that there may be differences in cognitive functioning compared to the trajectory exhibited by their neurotypical peers (Tse et al., 2021). They may also experience a higher risk of certain physical conditions including hypertension, neurological disorders such as Parkinson's, and incontinence (Fortuna et al., 2016; Starkstein, Gellar, Parlier, Payne, & Piven, 2015). However, there is a lack of research on whether older adult community and residential services are optimised to meet potential specific needs of this population (Perkins & Berkman, 2012).

4.2.2 Outline of Service Improvement Project

In this context, the aim of this project was to assess staff training needs in relation to autism in an older adult community mental health team, to implement a bespoke staff training event, and to explore the efficacy of this staff training. The conceptualization of this project was based in the COM-B model of behavioural change (Michie, van Stralen, & West, 2011). The COM-B model posits that the behavioural changes necessary for service improvement require three elements: motivation, capability, and opportunity. The capability factor involves an individual's skill set in making the required change, and can be improved through interventions such as training. The existing project applied this model in two stages. The first stages of this study explored whether there was a need for change in the service: was the service seeing patients on the autism spectrum, and if yes, did staff report having received training in this area, and did they feel they had further training needs? The second stage of this study focused on enhancing the capability factor of the COM-B model by introducing

staff training as an intervention targeted at increasing the team's skill set in working with this population.

The service improvement project took place in the South Bucks Older Adult Mental Health Team within the Oxford Health NHS Foundation Trust. This is a multidisciplinary team (MDT) which provides care for older adults (typically aged 65 years and above) experiencing mental health problems, and additionally encompasses a memory clinic service for adults experiencing cognitive problems.

4.2.3 Aims

The first two stages of this study aimed to understand staff training needs in relation to autism by exploring the following questions:

- How often do staff see patients with either an autism diagnosis, or query autism?
- How many staff have previously received training?
- What would staff want to be included in a further training program? Are there older adult specific elements?

The final stage of this study aimed to synthesise this information and address staff training needs by delivering a bespoke training session and evaluating its efficacy. Recommendations based on these stages were then developed and delivered to the team manager.

4.3 Stage 1: Audit

4.3.1 Methods

All stages of this study received approval from the Oxford Health Clinical Audit team. The first stage used a clinical audit to explore how often the CMHT were seeing patients either with a recorded diagnosis of autism, or where the possibility of autism or

autistic traits had been mentioned in their notes. The audit was performed over a 5-year period (28th October 2016- 28th October 2021).

4.3.2 Findings

Over this time 3185 patients were referred to the CMHT. 5 of these patients (0.16%) had an autism diagnosis recorded in their notes. A further 36 patients (1.13%) had an autism spectrum condition query recorded in their notes, raised either by their treating clinicians, a family member, or by the patient themselves. Therefore, a total of 41 patients (1.29%) over the past 5 years had either an autism diagnosis, or query autism raised in their notes. Given the small sample size comparative statistics between patients in this autism spectrum sample, and patients with no diagnosis or query, were not performed. However, it should be noted that the autism spectrum sample were more likely to be male, more likely to be White British, and had a higher number of referrals and longer referral length. Demographic and referral data for the sample are summarised in Table 1.

Table 1: Summary of demographic and referral data for the audit sample.

	Autism Spectrum Sample (*n*= 41; mean (*SD*))	Non Queried Sample (*n*= 3,144; mean (*SD*))
Age (years)	76.41 (8.12)	84.34 (9.45)
Gender	27 (65.86%) male 14 (34.14%) female	1243 (39.53%) male 1900 (60.43%) female 1 (0.03%) not recorded
Ethnicity	18 (43.90%) White British 23 (56.10%) not recorded/ unknown	1137 (36.16%) White British 78 (2.48%) Other White Background 19 (0.60%) Black/ Black British 54 (1.72%) Asian/ Asian British 5 (0.20%) Mixed Background 11 (0.34%) Other Ethnic Group 1840 (58.52%) not recorded/ unknown
Number of referrals over 5 year period	1.73 (0.96)	1.40 (0.75)
Referral Length (days)	158.47 (139.62)	121.89 (169.16)

Table 2 summarises the diagnoses recorded for the patients in the autism spectrum group. In the autism spectrum sample, 12 patients (29.27%) had multiple diagnoses recorded in their notes, 25 (69.98%) had only one diagnosis, and 4 (9.75%) had no recorded diagnosis. The most common diagnosis in this sample was depression or depression symptoms, followed by psychotic conditions and memory disorders.

Table 2: Recorded diagnoses for diagnosed autism or query autism patients.

	Autism Spectrum Sample Diagnoses (*n*= 41)
Autism	5 (12.20%)
Depression	15 (36.59%)
Psychotic condition (schizophrenia, schizoaffective disorder, psychotic symptoms, delusional disorder, paranoia)	11 (26.83%)
Memory disorders	11 (26.83%)
Anxiety	5 (12.20%)
Personality disorder	5 (12.20%)
Bipolar disorder	1 (2.43%)
Adjustment disorder	1 (2.43%)
Unspecified developmental disorder	1 (2.43%)

4.4 Stage 2: Needs Survey

4.4.1 Methods

An online survey was used to evaluate whether staff subjectively reported seeing patients on the autism spectrum, whether they had previously received training, and what they would like to be included in a training programme (Appendix 3). Clinicians were also asked to self-rate their knowledge of autism in general, autism in older adults, and treatment adaptations for autism, on a scale of 0 (no knowledge) to 100 (very knowledgeable). The survey was sent to 45 clinicians across the CMHT.

4.4.2 Findings

23 clinicians completed the survey (51% response rate). On average clinicians had worked for the CMHT for 2.98 (*SD*= 4.70) years, and over that time self-reported seeing 2.18 (*SD*= 7.19) patients with diagnosed autism, and 9.04 (*SD*= 35.09) patients where they queried autism. Only one clinician (0.04%) had previously received autism training. On average

clinicians rated their knowledge of autism at 52.43 (*SD*= 20.04), and their knowledge of autism in older adults (*M*= 35.57, *SD*= 22.58) and treatment adaptations (*M*= 30.04, *SD*= 22.55) relatively lower.

Table 3 summarises what staff wanted to see included in a training programme. The most popular requests were autism in older adults, and autism and mental health, followed by treatment adaptations, and queries around if clinicians suspected autism in an undiagnosed patient.

Table 3: Staff requests for what they would want to be included in a training session on autism (participants were allowed to request multiple topics).

Topic	Frequency of requests
Autism in older adults	20
Autism and mental health	17
I think a patient might be on the spectrum but undiagnosed: what next?	14
Treatment adaptations	14
Recognising autistic traits	12
Autism diagnostic process	12
Autism and communication	11
Autism and sensory processing	11
Overlaps between autism and other conditions	10
Basic introduction to autism	9
Working with families and carers	9
Autism and physical health	8
Autism and routine/ repetition	8

4.5 Stage 3: Training and Evaluation

4.5.1 Methods

Following discussion with the team manager, it was decided not to invite members of the South Bucks CMHT working in the memory clinic only as the manager did not feel that the training would be useful for these clinicians at this time. Invitations were sent to 19 clinicians.

The training lasted an hour and was designed and delivered by EK. The content was based on a combination of the findings of the needs survey, and consultation from two autistic adults and a clinician from the Oxford Health Adult Autism Diagnostic Team. The structure included a mixture of lecturing and interactive clinical examples. The training covered the topics of introducing autism, autism in older adults, autism and mental health, and implications for treatment. The section on autism in older adults included reflections on difficulties diagnosing autism in this population. Attendees were asked to self-rate their knowledge of these four areas pre and post training using an online survey, utilising the same 0-100 ratings as the needs survey (Appendix 3). These ratings were compared using paired t-tests. Attendees were asked to rate session usefulness on a scale of 0 (not useful) to 10 (very useful), and gave qualitative responses around what they found helpful about the training, and what could be improved. Due to the small sample size these responses are summarised rather than thematically analysed.

4.5.2 Findings

A total of 11 people attended the training (attendance rate 57.89%). 8 people completed both pre and post questionnaires (response rate 72.73%). Whilst mean knowledge scores increased post training, this did not reach a statistically significant threshold (Table 4).

Table 4: Mean pre and post training knowledge ratings, and statistical comparisons.

	Pre-training (*n*= 8) Mean (Standard deviation)	Post-training (*n*= 8) Mean (Standard deviation)	T-test results
Knowledge of autism	50.63 (17.82)	59.38 (10.84)	*t*(7)= -1.62, *p*= 0.15, CI [-21.54, 4.04]
Knowledge of autistic traits in older adults	41.25 (17.27)	53.75 (20.66)	*t*(7)= -1.72, *p*= 0.13, CI [-29.66, 4.66]
Knowledge of mental health and autism	38.75 (19.59)	55.00 (12.25)	*t*(7)= -2.17, *p*= 0.06, CI [-33.95, 1.45]
Knowledge of adapting care for patients on the autism spectrum	30 (18.52)	51.25 (19.60)	*t*(7)= -1.97, *p*= 0.09, CI [-46.70, 4.20]

Attendees rated the session a mean 8.5 (*SD*= 1.22) on usefulness. Two participants reported that focusing the training specifically on older adults was helpful, and three participants found it helpful to interactively discuss case examples. Three participants suggested that the training could be improved by making it longer.

4.6 Discussion

At present, there is a lack of research on the specific needs of older adults on the autism spectrum, and implications for healthcare services. This service improvement project based in an older adult mental health team identified that whilst the team did work with older adults on the autism spectrum, the small proportion of older adults either with an official diagnosis or queried as being on the spectrum compared to broader population statistics on autism prevalence is strongly suggestive that autistic older adults are under-diagnosed and under-recognised (Tromans, Chester, Kiani, Alexander, & Brugha, 2018). Furthermore, a training programme for clinicians working in this area specifically focusing on older adulthood, autism and mental health was experienced as useful by clinicians and resulted in an increase in knowledge, albeit statistically insignificant.

Both the findings from the clinical audit, and the subjective reports of clinicians in the needs survey, suggest that this team see patients on the autism spectrum. However, very few patients identified in the audit had an official diagnosis. Instead, it was more likely for patients to have been identified as potentially on the autism spectrum by clinicians, families, or the patients themselves. The finding that just 0.16% of the CMHT patients over the past 5 years had an official autism diagnosis is far lower compared to the estimated 0.8% prevalence rate in adults in the UK (McManus, Bebbington, Jenkins, & Brugha, 2016). This is consistent with previous speculation in the research literature that due to historical changes in the diagnostic criteria, and difficulties in diagnosing autism in adulthood due to the need for developmental information, older adults are less likely to be diagnosed as autistic (Bennett, 2016; Van Orden et al., 2010).

Furthermore, whilst an additional 1.13% of patients had query autism recorded in their notes, the finding that a total of 1.29% patients in the CMHT presented with either confirmed or suspected autistic traits is again lower compared to estimates that 4.8% of patients in adult mental health services in the UK meet criteria for autism (Brugha et al., 2020). This could suggest that autistic traits in older adults are underrecognized or attributed to other conditions. The overlap between autism and mental health problems, including dual or differential diagnosis and implications for presentation, is an ongoing area of discussion in the research literature, but to date there is very little information on the overlap between autism and symptoms of memory problems (Lai et al., 2019; Shulman, Rice, Morrier, & Esler, 2020).

The finding that the older adult CMHT does see individuals on the autism spectrum, and that these individuals are likely to be undiagnosed, has implications for treatment and associated staff training needs. Autistic people and older adults as two separate groups have both been identified as populations which can benefit from treatment adaptations that meet

their specific needs, however there is a lack of research on adapting care and service provision for autistic older adults (Laidlaw & Kishita, 2015; Spain & Happé, 2020). This is significant as autistic older adults may require more support from social care and healthcare services compared to their neurotypical peers (Roestorf et al., 2019). Therefore, the findings of the current study emphasise the need for further research on providing care and enhancing wellbeing in autistic older adults with mental health needs.

The prevalence of autism in older adulthood populations, and what proportion of these cases are undiagnosed, is another area identified by this study which requires further research. Effectively providing augmented treatment for autistic people requires first knowing that the individual is on the autism spectrum (Nicolaidis et al., 2013; Tint & Weiss, 2018). Whilst the current study suggests that autism in older adults is underdiagnosed and underrecognized, the conclusivity and generalisability of this finding are very limited. This was an audit of a single older adult mental health team over a five-year period, and the results relating to query autism relied on clinicians recognising and documenting possible autistic traits in patient notes. The low number of query autism cases could reflect a lack of clinician awareness of the possibility of autism in this population, or that due to the service context behaviour may be interpreted through a lens of mental health or memory problems, rather than neurodevelopmental differences (van Niekerk et al., 2011). Further research in this area would benefit from a more rigorous approach, such as using validated screening and diagnostic instruments to better estimate the prevalence of potentially undiagnosed autism in this population.

Finally, the findings of this study have implications for the provision of training programmes in this field. NHS policies and NICE guidelines clearly state that autistic people with mental health difficulties should be provided with adapted care by trained staff (Department of Health, 2015; NHS Improvement, 2018; NICE, 2016). This project identified

that this team had not been offered these training opportunities: despite seeing patients on the autism spectrum, clinicians reported a lack of training in autism, and rated their knowledge of treatment adaptations relatively low. It should be noted that the provision of autism training is an ongoing area of development across the NHS, and shortly following this survey the NHS trust introduced mandatory autism training for clinicians. However, the findings of this project suggest that for clinicians working in more specialised services, such as older adult mental health, generic training may not fully meet their training needs (Dillenburger et al., 2016). In the needs survey clinicians rated their generic understanding of autism more highly compared to their knowledge of autism in older adults, and highlighted autism in older adults as their strongest training need.

A training programme specifically addressing the area of autism and older adulthood did result in an increase in self-rated knowledge of these areas for attendees (although not statistically significant), and the staff rated the training as useful. It is possible that the non-significance reflects the small sample size: whilst 19 people were invited to the training, just over half attended and not all attendees completed pre and post surveys. There was also a low response rate to qualitative survey questions exploring what was helpful about the training and what could be improved in future, which made it difficult to do a richer analysis around how staff experienced the training. Whilst the research team did consider providing a second training session, discussion with the team manager suggested that this would be unlikely to result in more people attending. It was speculated that individuals did not attend due to time pressures which were unlikely to be resolved before a further training session. Notably, whilst there was good representation from nursing staff, support workers and occupational therapists at the training, no manager-level staff, no psychiatrists, and only one supervisor attended. The low attendance rate likely reflects high time pressure on the staff team, and highlights that upskilling staff using training as an intervention may be of limited effect in

teams where staff do not have sufficient time to attend and engage in sessions. In particular, the absence of supervisor or consultant-level staff raises concerns about whether this intervention will result in continued upskilling for the team. Ensuring that clinicians receive ongoing adequate training in this area, as mandated by the NHS, balanced with service pressures is an ongoing area for discussion.

4.6.1 Limitations

To date this is the first study exploring the prevalence of autism in older adults in a mental health setting. However, a significant limitation of the current study's audit approach is that it relies on clinical notes, and therefore likely underestimates the true prevalence of autistic patients seen by the CMHT. This area could benefit from further research exploring what proportion of older adults in mental health teams meet criteria for autism using validated diagnostic tools.

Furthermore, the conclusions drawn around the efficacy of the training are largely speculative and require further empirical investigation. Firstly, the low sample size made it difficult to quantitatively evaluate whether training was successful in increasing staff capability. Secondly, the COM-B model suggests that an increase in capability may result in a change in behaviour: it was not possible using the current design to evaluate where providing staff with training will result in long-term behavioural change and service improvements relating to the provision of care for autistic patients. Possible avenues for further evaluation could include exploring whether training results in an increase in patients being identified as exhibiting potential autistic traits, or through requesting qualitative feedback from patients on the spectrum on their care experience. In addition, this study did not explore the other aspects of the COM-B model thought to influence behaviour: opportunity and motivation. For example, the audit highlighted that at present the older adult

team sees a very small proportion of patients with diagnosed autism. It is therefore possible that managers and staff within the team lack motivation to make autism a service priority in the context of competing needs and high service pressures.

Thirdly, this paper has wondered if staff working in older adult teams could benefit from training specifically on older adulthood and autism, rather than generic autism training. This is speculative, and the most effective way of delivering autism training (for example, the outcomes of generic training compared to more specific training for different teams) requires further empirical investigation (for example, comparing outcomes from specialised older adult versus generalised autism training) to understand the best way to increase staff capability in this area.

Finally, this project evaluated the implementation of a one-off training programme and did not assess whether this resulted in long-term improvements in knowledge. Whilst it raises several ideas in relation to what staff might find helpful from training that could be beneficial to the design of future interventions, there is a risk that the current intervention will not result in long-term team upskilling, particularly given high staff turnover within the team. As meeting the needs of autistic service users continues to be a priority within the NHS, further work is required to explore the best ways to increase staff capability over the long term. For example, a recent service development project in the eating disorder field implemented a specialist autism pathway, with the provision of staff training combined with weekly MDT huddles (brief meeting spaces to support clinicians in working with autistic patients) and the improvement of staff capacity to identify potential autism cases (Tchanturia, Smith, Glennon, & Burhouse, 2020).

4.6.2 Service Recommendations

Following this project, the following service improvement recommendations for the team were developed and delivered to the team manager:

1. The older adult mental health team does see a small proportion of patients on the autism spectrum, and should be aware that NHS guidelines mandate that these patients should be provided with adapted care by trained and experienced clinicians. To ensure the team are meeting these guidelines it may be helpful to ensure that training is available to all new staff and utilised by existing staff, and to explore the best ways of disseminating relevant information (for example, in induction packs, sharing information with supervisors, developing team resources). The team could also consider how to ensure that autism is held in mind when discussing and treating patients, particularly in the context of the possibility that a significant minority of patients may be on the autism spectrum but undiagnosed.
    a. The team could consider if there is scope to create an autism working party, or designating a specialist place for where clinicians can bring queries about autism and their patients (Tchanturia et al., 2020). Whilst at present the adult autism diagnostic service in Oxford Health is experiencing high service pressure and may not have capacity to perform a consultative role, they have advised that the psychologists within the older adult CMHT may be a helpful source of support.
    b. A further service development project could explore the service experiences of patients on the autism spectrum, and whether these patients and their families do feel that the service is meeting their needs.
2. The audit findings suggest that autism is underdiagnosed and underrecognized in older adults. Clinicians should consider that they may be working with undiagnosed patients who meet the modern criteria for autism and should be supported in

considering whether their patients may present with autistic traits (potentially through supervision and training).

a. Following consultation with the Oxford Health adult autism diagnostic service, we have been advised that diagnosing autism in older adults can be difficult due to problems in obtaining developmental histories. It should also be noted that autism screening instruments such as the Autism Quotient are known to lack validity in adults with co-occurring mental health problems (Ashwood et al., 2016). Therefore, identifying patients who may be on the autism spectrum, and creating a diagnostic pathway, is likely an ongoing area of discussion and service development for the team. The adult autism diagnostic service have advised that they do accept referrals for older adults, but that they may be rejected if there are not enough clear, detailed examples and evidence that the patient meets the autism diagnostic criteria on the referral form. In this context, staff could potentially benefit from clear guidance around what happens when they query whether is on the autism spectrum, including understanding what the next steps might be in terms of treatment, at what threshold they might consider referring for a diagnostic assessment, and guidance around making referrals. The team could explore whether it is possible to work with the adult autism service in developing this pathway.
b. One possibility is that the CMHT duty staff may particularly benefit from further training in recognising autism, as these clinicians are primarily responsible for assessing new patients. The team could consider how best to support the duty team in this role, for example providing clinicians with resources or specialist supervision.

3. Our project suggests that staff feel that they particularly benefited from autism training specific to their older adult context. When considering future training delivery it may be essential to consider the benefits of generic versus older adult specific programmes.

4.7 Lay Summary

There has been a lot of research and discussion around the fact that autistic people are more likely to experience mental health problems compared to the general population, and we know autistic people benefit from adapted interventions that meet specific needs associated with autism. This has led to the NHS making training in autism for clinicians who might work with autistic people compulsory. However, there is a lack of understanding around what this means for older adult mental health teams: for example, we do not know if autistic older adults have different needs to younger autistic people, and if these clinicians might need specialist training. This study explored three areas: we looked at whether a specific older adult mental health team (CMHT) saw patients with an autism diagnosis, or where clinicians wondered if a patient might be on the autism spectrum but undiagnosed. We then explored whether staff had already received training, and how confident they felt in working with autistic people. We found that the older adult CMHT did see a very small proportion of patients with an autism diagnosis, but that there were more patients with no diagnosis where clinicians, the patient or their families wondered if they might be on the autism spectrum. This makes us think that autism in older adults might be under-diagnosed. We also found that most staff had not received training and did not feel confident in this area. We therefore designed a training programme for this staff team which resulted in small increases in staff knowledge around autism and older adults. Consequently, we think that there's a need for further research on autism and older adulthood, and how autistic older adults can best be supported by healthcare services. We also think that we need to further explore the best way

to make staff feel confident in working with this population, particularly as staff are currently under a lot of pressure and may not have time to attend training.

4.8 References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5 ed.). Washington, DC.

Ashwood, K. L., Gillan, N., Horder, J., Hayward, H., Woodhouse, E., McEwen, F. S., . . . Murphy, D. G. (2016). Predicting the diagnosis of autism in adults using the Autism-Spectrum Quotient (AQ) questionnaire. *Psychological Medicine, 46*(12), 2595-2604. doi:10.1017/S0033291716001082

Bennett, M. (2016). "What is Life Like in the Twilight Years?" A Letter About the Scant Amount of Literature on the Elderly with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 46*(5), 1883-1884. doi:10.1007/s10803-016-2710-z

Brugha, T., Tyrer, F., Leaver, A., Lewis, S., Seaton, S., Morgan, Z., . . . van Rensburg, K. (2020). Testing adults by questionnaire for social and communication disorders, including autism spectrum disorders, in an adult mental health service population. *International Journal of Methods in Psychiatric Research, 29*(1), e1814. doi:https://doi.org/10.1002/mpr.1814

Croen, L. A., Zerbo, O., Qian, Y., Massolo, M. L., Rich, S., Sidney, S., & Kripke, C. (2015). The health status of adults on the autism spectrum. *Autism, 19*(7), 814-823. doi:10.1177/1362361315577517

Department of Health. (2015). *Statutory guidance for Local Authorities and NHS organisations to support implementation of the Adult Autism Strategy*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/422338/autism-guidance.pdf

Dillenburger, K., McKerr, L., Jordan, J.-A., & Keenan, M. (2016). Staff Training in Autism: The One-Eyed Wo/Man…. *International Journal of Environmental Research and Public Health, 13*(7), 716. Retrieved from https://www.mdpi.com/1660-4601/13/7/716

Fortuna, R. J., Robinson, L., Smith, T. H., Meccarello, J., Bullen, B., Nobis, K., & Davidson, P. W. (2016). Health Conditions and Functional Status in Adults with Autism: A Cross-Sectional Evaluation. *J Gen Intern Med, 31*(1), 77-84. doi:10.1007/s11606-015-3509-x

Hand, B. N., Angell, A. M., Harris, L., & Carpenter, L. A. (2020). Prevalence of physical and mental health conditions in Medicare-enrolled, autistic older adults. *Autism, 24*(3), 755-764. doi:10.1177/1362361319890793

Hirvikoski, T., Mittendorfer-Rutz, E., Boman, M., Larsson, H., Lichtenstein, P., & Bölte, S. (2016). Premature mortality in autism spectrum disorder. *Br J Psychiatry, 208*(3), 232-238. doi:10.1192/bjp.bp.114.160192

HM Government. (2014). *Think Autism: fulfilling and rewarding lives, the strategy for adults with autism in England: an update.* Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/299866/Autism_Strategy.pdf

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism, 20*(4), 442-462. doi:10.1177/1362361315588200

Kinnaird, E., Norton, C., Stewart, C., & Tchanturia, K. (2019). Same behaviours, different reasons: what do patients with co-occurring anorexia and autism want from treatment? *Int Rev Psychiatry, 31*(4), 308-317. doi:10.1080/09540261.2018.1531831

Lai, M.-C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., . . . Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. *The Lancet Psychiatry, 6*(10), 819-829. doi:https://doi.org/10.1016/S2215-0366(19)30289-5

Laidlaw, K., & Kishita, N. (2015). Age-appropriate augmented cognitive behavior therapy to enhance treatment outcome for late-life depression and anxiety disorders. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry, 28*(2), 57-66. doi:10.1024/1662-9647/a000128

Lever, A. G., & Geurts, H. M. (2016). Psychiatric Co-occurring Symptoms and Disorders in Young, Middle-Aged, and Older Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 46*(6), 1916-1930. doi:10.1007/s10803-016-2722-8

Maddox, B. B., Crabbe, S., Beidas, R. S., Brookman-Frazee, L., Cannuscio, C. C., Miller, J. S., . . . Mandell, D. S. (2020). "I wouldn't know where to start": Perspectives from clinicians, agency leaders, and autistic adults on improving community mental health services for autistic adults. *Autism, 24*(4), 919-930. doi:10.1177/1362361319882227

May, T., Pilkington, P. D., Younan, R., & Williams, K. (2021). Overlap of autism spectrum disorder and borderline personality disorder: A systematic review and meta-analysis. *Autism Research, 14*(12), 2688-2710. doi:https://doi.org/10.1002/aur.2619

McManus, S., Bebbington, P., Jenkins, R., & Brugha, T. (2016). *Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014*. Leeds: NHS Digital

Michael, C. (2016). Why we need research about autism and ageing. *Autism, 20*(5), 515-516. doi:10.1177/1362361316647224

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science : IS, 6*, 42-42. doi:10.1186/1748-5908-6-42

NHS Improvement. (2018). *The learning disability improvement standards for NHS trusts*. Retrieved from https://www.england.nhs.uk/wp-content/uploads/2020/08/v1.17_Improvement_Standards_added_note.pdf

NICE. (2016). Autism spectrum disorder in adults: diagnosis and management. Retrieved from https://www.nice.org.uk/guidance/cg142

Nicolaidis, C., Raymaker, D., McDonald, K., Dern, S., Boisclair, W. C., Ashkenazy, E., & Baggs, A. (2013). Comparison of healthcare experiences in autistic and non-autistic adults: a cross-sectional online survey facilitated by an academic-community partnership. *J Gen Intern Med, 28*(6), 761-769. doi:10.1007/s11606-012-2262-7

Perkins, E. A., & Berkman, K. A. (2012). Into the unknown: aging with autism spectrum disorders. *Am J Intellect Dev Disabil, 117*(6), 478-496. doi:10.1352/1944-7558-117.6.478

Roestorf, A., Bowler, D. M., Deserno, M. K., Howlin, P., Klinger, L., McConachie, H., . . . Geurts, H. M. (2019). "Older Adults with ASD: The Consequences of Aging." Insights from a series of special interest group meetings held at the International Society for Autism Research 2016–2017. *Research in Autism Spectrum Disorders, 63*, 3-12. doi:https://doi.org/10.1016/j.rasd.2018.08.007

Shulman, C., Rice, C. E., Morrier, M. J., & Esler, A. (2020). The Role of Diagnostic Instruments in Dual and Differential Diagnosis in Autism Spectrum Disorder Across the Lifespan. *Child Adolesc Psychiatr Clin N Am, 29*(2), 275-299. doi:10.1016/j.chc.2020.01.002

Spain, D., & Happé, F. (2020). How to Optimise Cognitive Behaviour Therapy (CBT) for People with Autism Spectrum Disorders (ASD): A Delphi Study. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 38*(2), 184-208. doi:10.1007/s10942-019-00335-1

Starkstein, S., Gellar, S., Parlier, M., Payne, L., & Piven, J. (2015). High rates of parkinsonism in adults with autism. *Journal of Neurodevelopmental Disorders, 7*(1), 29. doi:10.1186/s11689-015-9125-6

Stoner, J. B., Angell, M. E., House, J. J., & Bock, S. J. (2007). Transitions: Perspectives from Parents of Young Children with Autism Spectrum Disorder (ASD). *Journal of Developmental and Physical Disabilities, 19*(1), 23-39. doi:10.1007/s10882-007-9034-z

The Westminster Commission on Autism. (2016). A spectrum of obstacles: an inquiry into access to healthcare for autistic people. In.

Tint, A., & Weiss, J. A. (2018). A qualitative study of the service experiences of women with autism spectrum disorder. *Autism, 22*(8), 928-937. doi:10.1177/1362361317702561

Tromans, S., Chester, V., Kiani, R., Alexander, R., & Brugha, T. (2018). The Prevalence of Autism Spectrum Disorders in Adult Psychiatric Inpatients: A Systematic Review. *Clinical practice and epidemiology in mental health : CP & EMH, 14*, 177-187. doi:10.2174/1745017901814010177

Tse, V. W. S., Lei, J., Crabtree, J., Mandy, W., & Stott, J. (2021). Characteristics of Older Autistic Adults: a Systematic Review of Literature. *Review Journal of Autism and Developmental Disorders*. doi:10.1007/s40489-021-00238-x

van Niekerk, M. E., Groen, W., Vissers, C. T., van Driel-de Jong, D., Kan, C. C., & Oude Voshaar, R. C. (2011). Diagnosing autism spectrum disorders in elderly people. *Int Psychogeriatr, 23*(5), 700-710. doi:10.1017/s1041610210002152

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. E., Jr. (2010). The interpersonal theory of suicide. *Psychological review, 117*(2), 575-600. doi:10.1037/a0018697

5. Theory Driven Research Project: Exploring the relationship between clinical and personal models of recovery in anorexia nervosa

Emma Kinnaird (emma.kinnaird@hmc.ox.ac.uk)

Supervised by Dr Myra Cooper.

Word count: 4,996

First submitted February 2023.

The proposed journal for this paper was the Journal of Eating Disorders given its specific subject area and interest in publishing qualitative and mixed methodology research. The paper has been formatted according to journal guidelines (Appendix 2).

5.1 Abstract

Background: Whilst recovery from anorexia nervosa (AN) has traditionally been defined in relation to behavioural symptoms, the presence of cognitions, and weight, there is increasing research interest around understanding recovery from AN from the perspective of personal models of recovery, which emphasise factors such as personal growth and leading a meaningful life. Whilst previous research has suggested that clinical and personal recovery factors may exist in a complimentary relationship, with reduced clinical symptoms enabling growth in personal domains, to date this association has not been empirically explored. Understanding this relationship is particularly important in the context of proposals that there is a need to develop a universal definition of recovery from AN. **Methods:** The current study used a mixed methods approach to explore the relationship between personal and clinical models of recovery in AN. Eighty-one adults who self-defined as fully or partially recovered from AN were recruited into an online study and completed self-report measures of clinical and personal recovery. Participants were also asked questions about their views on the personal relevance of clinical criteria, and this data was analysed using thematic analysis. **Results:** The study found that there was good agreement between self-defined and clinically defined full recovery from AN. People who were clinically defined as fully recovered scored significantly higher on measures of personal recovery compared to those defined as not recovered, and there were strong positive associations between lower behavioural and cognitive symptoms, and personal recovery. However, there was a weaker relationship between personal recovery measures and BMI. Finally, qualitative data indicated that most participants felt that the clinical criteria were not a sufficient definition of recovery. Whilst participants suggested potential additional criteria, and raised ideas around how these might relate to traditional symptom-based approaches, the qualitative analysis also highlighted concerns around how far a universal recovery criteria-based approach can capture individual

experience. **Conclusions:** Implications for the development of future universal clinical criteria for recovery from AN are discussed, including how a clinical approach can best be combined with personal understandings of recovery.

5.2 Background

Anorexia nervosa (AN) is an eating disorder (ED) characterised by the restriction of energy intake relative to requirements resulting in low body weight, an intense fear of weight gain, disturbance in self-perceived weight and shape, and the undue influence of weight and shape on self-evaluation (American Psychiatric Association, 2013). This condition is associated with poor treatment outcomes and a high risk of developing a severe and enduring illness course (Arkell & Robinson, 2008; Bamford & Mountford, 2012; Steinhausen, 2002).

These poor treatment outcomes reflect clinical judgements of whether an individual has recovered from AN, typically based on whether they continue to meet weight and symptom criteria for the condition (Khalsa et al., 2017; Steinhausen, 2002). This is in line with the traditional medical model of recovery, sometimes known as “clinical recovery”, which conceives recovery as cure, or symptom remission, that can be objectively measured by outcome studies (Roberts & Wolfson, 2004; Slade, 2009). Whilst there is disagreement in the literature on how clinical recovery from AN should be defined, recent papers have proposed operationalisations of recovery as a combination of weight restoration, the absence of behavioural symptoms, and a reduction in psychological symptoms, all evaluated using established and validated outcome measures (Bardone-Cone et al., 2010; Khalsa et al., 2017).

However, this approach to defining recovery has come under increasing scrutiny. There are concerns that a focus on symptoms and weight as outcome measures has had an undue influence on decisions around access to treatment and treatment content (Fogarty & Ramjan, 2016; Rance, Moller, & Clarke, 2015). Research has also highlighted that the

clinical model of recovery may not capture components of illness and recovery that are personally valuable to patients (Ashworth, Evans, & Clement, 2009; de Vos et al., 2017). A recent study found that when patients receiving treatment identified their own outcomes, over 50% of the sample identified a valued outcome not covered by standardised measures of ED symptoms or quality of life (Austin et al., 2020).

Consequently, there has been a growing interest in alternative approaches to understanding ED recovery, particularly approaches that centre the perspectives of patients themselves. A meta-analytic review of qualitative studies examining how people recovered from EDs conceive recovery found that in addition to symptom reduction, self-acceptance, positive relationships with others, personal growth, resilience, and autonomy emerged as important aspects (de Vos et al., 2017). This is reflective of wider personal models of recovery, sometimes known as the "recovery model" or "personal recovery" (Dawson, Rhodes, & Touyz, 2014). The recovery model conceives recovery as a process of personal growth, even in the context of ongoing symptoms and difficulties (Anthony, 1993; Schrank & Slade, 2007). Whilst several frameworks have been developed to operationalise this person-centred recovery approach, when this current project was designed there was only one ED-specific model (Wetzler et al., 2020). This framework highlights the core domains of supportive relationships, hope, identity, meaning and purpose, empowerment, and self-compassion, all theorised to exist in a positive relationship with a reduction in clinical symptoms. More recently a second recovery framework has been published (Kenny & Lewis, 2023). Whilst this model posits recovery domains (including physical stability, quality of life, and psychosocial functioning), its core argument is that recovery is individualised, ongoing and non-linear, and that the domains are therefore "starting points" that may or may not be relevant to individuals, rather than determinants of recovery outcome.

Clarifying how recovery from AN is defined, including how personal recovery perspectives can be incorporated with traditional clinical understandings, is particularly important in light of calls to establish a universal definition of AN recovery (Bardone-Cone, Hunt, & Watson, 2018). More broadly outside of the ED field it has been argued that personal and clinical models of recovery can be viewed as complimentary, rather than in opposition (Barber, 2012). This relationship is important to clarify: if there is poor agreement, there is a risk that an overly narrow focus on symptoms/ diagnostic criteria in treatment and research could miss distress and difficulty in other areas (Fogarty & Ramjan, 2016; Rance et al., 2015).

However, at present there is a lack of research on the relationship between clinical criteria and personally valued aspects of recovery in AN. In one study only half of people who self-defined as recovered from an ED met full clinical criteria, suggesting some disagreement, and qualitative research suggests that people with lived experience of EDs have significant concerns around the utility of clinical criteria in defining recovery (Kenny, Trottier, & Lewis, 2022; Slof-Op 't Landt, Dingemans, de la Torre Y Rivas, & van Furth, 2019). Overall, there appears to be a consensus that symptom reduction is an important element of recovery, but a lack of clarity around its relationship to personal models (Bachner-Melman, Lev-Ari, Zohar, & Lev, 2018). The Wetzler et al (2020) model theorises that reducing eating and weight behaviours and attitudes is how someone achieves personal recovery. This positions clinical recovery as a requirement for further personal recovery, rather than an end goal, and suggests a correlated relationship between the two constructs. Whilst the Kenny & Lewis (2023) model positions recovery as individualised and non-linear, noting that relationships between recovery domains will therefore vary, they do suggest that reducing ED symptoms may allow an individual to achieve greater functioning in other

domains. However, this theorised positive relationship does not appear to have been tested by quantitative research.

In this context, the aim of this study is to explore the relationship between clinical and personal recovery in AN by investigating the following research questions:

1. What is the relationship between clinical and personal models of recovery in this population?
2. How do people with a history of AN experience the relationship between their clinical symptoms and valued personal aspects of recovery?

5.3 Methods

5.3.1 Design

The study received ethical approval from the Oxford University Medical Science Ethics Board (reference R77655/RE001, Appendix 4). This study incorporated a mixed-methods survey design developed with support from individuals with lived experience of AN. Adults who self-defined as partially or fully recovered from AN were recruited online via social media (Twitter, Facebook) and asked to complete an online survey which incorporated validated questionnaire measures and survey questions about their perspectives on recovery with space for free-text responses. Responses to questionnaire measures were used to divide participants into three groups based on clinical recovery status (recovered, partially recovered, remitted, not recovered), and these three groups were statistically compared on measures of personal recovery. Finally, survey responses were analysed using qualitative analysis.

Clinical recovery from AN has been operationalised in various ways, and at present there is no definition that is universally used in this field (Bardone-Cone et al., 2018). For this study, clinical recovery was operationalised using the definition of Khalsa et al (2017).

Whilst a number of different definitions of clinical recovery from EDs exist (see Bardone-Cone et al., 2018, for a review), Khalsa et al's model has the benefit of spanning physical, behavioural and cognitive domains, and has been specifically developed for AN populations (in contrast to other ED-general models). It also includes the use of established outcome measures. Khalsa et al's definitions are summarised in Table 1.

Table 1: Khalsa et al's (2017) proposed recovery definitions, including notes on adaptation for current study.

Illness status	BMI	Symptoms	Behaviours	Scales	Duration
Full recovery	BMI ≥ 20	No significant fear of gaining weight or disturbance in body image	No restricting, bingeing, or purging	EDE within 1 SD of normal **current study used EDE-Q*	12 months
Partial recovery	BMI ≥ 19	No significant fear of gaining weight or disturbance in body image	No restricting, bingeing, or purging	EDE within 1.5 SD of normal **current study used EDE-Q*	6 months
Full remission	BMI ≥ 19	Fear of gaining weight or disturbance in body image present	No restricting, bingeing, or purging	EDE within 2 SD of normal **current study used EDE-Q*	3 months
Partial remission	BMI ≥ 18.5	Fear of gaining weight or disturbance in body image present	No restricting, bingeing, or purging	EDE within 2 SD of normal **current study used EDE-Q*	1 month

Abbreviations: body mass index (BMI), Eating Disorder Examination (EDE), Eating Disorder Examination Questionnaire (EDE-Q), standard deviation (SD).

5.3.2 Participants

Inclusion criteria for this study were that participants were aged over 18, and self-defined as fully or partially recovered from AN. Participants were excluded if they self-

reported currently experiencing another ED or other serious mental health condition (for example, psychosis or bipolar disorder). Given the lack of previous quantitative research in this area from which to generate a power analysis, sample sizes were calculated a priori using a power analysis based on Cohen's tables which suggested a minimum sample of *n*=25 in each comparator group would give the study the recommended level of 80% power (Cohen, 1992). It should be noted that the overall sample size was based on power calculations for the quantitative analysis, rather than judged upon whether data saturation had been achieved for the qualitative analysis. This is, therefore, a limitation of the qualitative analysis.

The study chose to include participants who self-defined as partially recovered in line with personal models of recovery which conceive of recovery as a process rather than a discrete end state (Anthony, 1993; Schrank & Slade, 2007). Furthermore, given the study's aim to explore the relationship between clinical symptoms and measures of personal recovery, it was felt it would be beneficial to include individuals who self-defined at different stages of this recovery process.

5.3.3 Measures

5.3.3.1 Demographics/ Clinical Information

Participants were asked to self-report demographic and clinical characteristics, including length of ED, whether they self-defined as recovered/ partially recovered, and estimated time they had not experienced ED symptoms. Participants self-reported their recovery status without any information on how clinicians currently define ED recovery.

5.3.3.2 Weight

Participants were asked to self-report their current height and weight to allow for the calculation of their BMI. A limitation of this approach is that it requires self-report, as opposed to objective measurement. However, research has found strong correlations between

self-reported and objective height and weight in people with AN, albeit this population are prone to slightly overestimating their weight (McCabe, McFarlane, Polivy, & Olmsted, 2001; Meyer, Arcelus, & Wright, 2009).

5.3.3.3 Eating Disorder Examination Questionnaire (EDE-Q)

The EDE-Q is a widely used and validated measure in the ED field, including as a determinator of outcome in treatment trials (Fairburn & Beglin, 2008). The questionnaire consists of 28 items. The EDE-Q was used to assess ED behaviours and weight/shape/eating-focused psychological symptoms. The EDE-Q showed good internal validity in this study (Cronbach's alpha 0.925).

5.3.3.4 The Process of Recovery Questionnaire (QPR)

Although an ED-specific model of personal recovery has been developed, at present there are no personal recovery measures specific to the ED field (Wetzler et al., 2020). In addition, the researchers were unable to identify a personal recovery measure which had previously been utilised in this population. In selecting measures of personal recovery, the researchers drew on Wetzler et al's (2020) research which noted that their framework closely reflected the domains of the non-ED specific personal recovery model (the CHIME (connectedness, hope/optimism, identity, meaning, empowerment) framework), with the exception of an additional domain for self-compassion (Bird et al., 2014). Although there are no measures that have specifically been developed to reflect the CHIME framework, it has been identified in previous research that the QPR is the only existing measure of personal recovery that maps onto its domains (Shanks et al., 2013). Given the lack of ED-specific options, this study chose to assess personal recovery in AN using the QPR with the additional use of a self-compassion measure in an attempt to best reflect previously identified constructs salient to this area (Wetzler et al., 2020). The QPR is a 22-item measure of personal recovery

designed in collaboration with people with experience of psychosis (Appendix 5, Neil et al., 2009). Higher scores indicate a higher degree of perceived recovery. Whilst it has been validated in psychosis populations, demonstrating good validity and reliability, this project is its first use in the ED field (Neil et al., 2009). The QPR showed high internal validity in this study (Cronbach's alpha 0.924). Whilst the QPR has two subscales gained by summing item scores (intrapersonal and interpersonal recovery), for the purposes of this study only the total QPR score (sum of all scores) was used.

5.3.3.5 Self-Compassion Scale (SCS)

The SCS (Appendix 5) is a 26-item measure of self-compassion that has been widely used and validated in previous research, including in the ED field (Braun, Park, & Gorin, 2016; Neff, 2003). The SCS was used in addition to the QPR as the Weztler et al (2020) personal recovery model emphasises the centrality of self-compassion to ED recovery. Higher scores indicate higher self-compassion. The SCS showed good internal validity in this study (Cronbach's alpha 0.947). The SCS generates six subscales through averaging item scores, with a total score achieved by averaging the subscale scores. For the purposes of this study only the total score was used.

5.3.3.6 Survey Questions

Participants were additionally presented with Khalsa et al's proposed clinical recovery criteria and asked two questions: "Do you feel this definition of recovery reflects your personal experience and understanding of recovery?", and "Do you think that you can be recovered even if you don't meet all of the above criteria?". Participants were given a yes/no response option for each question, and additionally given a free text space in which to expand on their answer.

5.3.4 Analysis

Firstly, participant data were analysed to assess clinical recovery status on an individual level. Self-reported height and weight were used to calculate BMI. The presence of ED behaviours and symptoms were measured using the EDE-Q, with community norms derived from previous research used to calculate whether individual scores were above or below the clinical threshold defined as in Khalsa et al's clinical recovery model (Carey et al., 2019). The presence of binging or purging were based on participant responses to EDE-Q items measuring these specific behaviours, and the presence of restricting behaviours was calculated based on whether participants scored within 1 SD of community norms on the Restrictive subscale. This analysis was then used to divide participants into three groups: those meeting Khalsa et al's criteria for full recovery, a second group encompassing those meeting clinical criteria for partial recovery or remission, and a third group for participants defined as not recovered.

Statistical analyses were performed using Stata (version 17.0) software. Shapiro Wilk tests were used to evaluate normality for each variable. Group comparisons were calculated using chi square tests, *t*-tests, or Kruskal-Wallis *H* tests, depending on variable type and normality. Given the non-parametric nature of some variables, Spearman's correlations were run to assess whether clinical variables (BMI, EDE-Q Global score) were related to personal recovery variables (SCS, QPR) across the total sample.

Qualitative analyses were performed using thematic analysis from a deductive realist positioning: researchers approached the data from the theoretical positioning of clinical and personal recovery models, with a specific interest in the relationship between these constructs (Braun & Clarke, 2006). Data were initially read and reread by EK whilst making notes around potential codes. Coded data was then collated and reviewed to identify themes and sub-themes. The identification of themes was based primarily around relevance to the research question, and secondarily how frequently concepts or ideas arose in the transcripts,

including where participants disagreed on those concepts within the themes. These themes were then reviewed and refined in relation to the original dataset by EK and MC.

5.4 Results

A total of 81 people took part in the study. Thirty-seven people (45.68%) self-defined as fully recovered from AN, 44 people (54.32%) self-defined as partially recovered from AN. When clinical recovery definitions were applied, 27 (33.33%) people met criteria for the fully recovered category, 12 (14.81%) people met criteria for being partially recovered, three (3.70%) people met criteria for full remission, one (1.23%) person met criteria for partial remission, and 38 (46.91%) people were defined as not recovered. Given the power analysis suggesting that a minimum of n=25 participants in each group were required for sufficient power, and the relatively small n in the partially recovered/ remitted categories, these individuals have been grouped into a single "partial recovery/ remission" category (n= 16, 19.75% of sample). Due to this group consisting of individuals with varying clinical recovery statuses, and the group n continuing to be lower than the minimum power estimate, this group was omitted from the comparative analysis but retained for correlational calculations.

Demographic and clinical details for each of the clinical recovery groups are summarised in Table 1. As expected, there were significant differences across the three groups regarding variables relating to clinical recovery definitions (BMI, EDE-Q Global score, and self-defined recovery duration). Groups also significantly differed on gender. Further analysis suggested that this difference was due to the lower proportion of females in the clinically partially recovered/remitted group, with no significant differences between the not recovered/ fully recovered groups ($\chi^2(1) = 0.72$, p= 0.396). As the comparative analysis in this study only compared the clinically not recovered/ fully recovered groups, gender was not treated as a potential moderating variable.

Table 2: Summary of demographic and clinical details for clinical recovery groups

	Clinical recovery status (group means (*SD*))				
	Not recovered (*n*= 38)	Partial recovery/ remission (*n*= 16)	Fully recovered (*n*= 27)	Test statistic	p
Age (years)	27.17 (7.77)	29.73 (8.51)	30.00 (7.47)	$\chi^2(2) = 2.65$	0.270
Gender (*n)*	37 female (97.37%) 1 male (2.63%)	12 female (75.00%) 3 male (18.75%) 1 non-binary (6.25%)	27 female (100%)	$\chi^2(4) = 12.75$	0.013
Ethnicity (*n*)	2 Asian (5.26%) 1 Black/African/Caribbean (2.63%) 1 Mixed or multiple ethnic groups (2.63%) 4 Ethnic group not specified (10.53%) 30 White (78.95%)	0 Asian (0.00%) 0 Black/African/Caribbean (0.00%) 2 Mixed or multiple ethnic groups (12.50%) 2 Ethnic group not specified (12.50%) 12 White (0.75%)	2 Asian (7.41%) 0 Black/African/Caribbean (0.00%) 1 Mixed or multiple ethnic groups (3.70%) 0 Ethnic group not specified (0.00%) 24 White (88.89%)	$\chi^2(8) = 7.96$	0.687
Employment (*n*)	17 employed full time 3 employed part time 0 self-employed 13 students 5 unemployed	7 employed full time 2 employed part time 1 self-employed 6 students 0 unemployed	13 employed full time 4 employed part time 1 self-employed 7 students 2 unemployed	$\chi^2(8) = 5.65$	0.687
Illness duration (years)	8.54 (6.57)	11.13 (6.84)	6.59 (4.01)	$\chi^2(2) = 5.18$	0.075
Self-defined recovery duration (years)	3.93 (5.86)	6.31 (6.14)	5.69 (4.31)	$\chi^2(2) = 10.08$	0.007
BMI	19.35 (3.31)	19.46 (0.83)	23.38 (3.75)	$\chi^2(2) = 32.97$	<0.001
EDE-Q Global score	3.08 (1.27)	1.36 (0.98)	2.03 (1.45)	$\chi^2(2) = 39.20$	<0.001

Abbreviations: body mass index (BMI), Eating Disorder Examination Questionnaire (EDE-Q), number (n), standard deviation (SD).

5.4.1 How far is there agreement between self-defined and clinically defined recovery?

The majority (*n* =24, 64.86%) of people who self-defined as fully recovered met clinical criteria for full recovery. Six (16.22%) met criteria for partial recovery or remission, and seven (18.82%) were defined as not recovered. By contrast, the majority (*n* =31,70.45%) of people who self-defined as partially recovered were clinically defined as not recovered. Ten (22.73%) met criteria for partial recovery or remission, and three (6.82%) met criteria for full recovery. Therefore, in the self-defined fully recovered group there was significantly higher agreement between self-defined and clinically defined recovery in 64.86% of cases, compared to 22.73% agreement in the partially recovered group ($\chi^2(1)= 14.65$, $p <0.001$).

Of the 54 participants not meeting full clinical recovery criteria, the most common reason for not meeting criteria was a low BMI (it should be noted that most participants did not meet the full criteria for more than one reason). Thirty-eight people (70.37%) did not meet the BMI threshold of ≥ 20. Twenty-six (48.15%) reported continuing to experience restricting, binging or purging behaviours, and 21 (38.89%) self-reported a level of symptoms on the EDE-Q higher than the allowed threshold. Two (3.7%) self-reported a lower recovery duration than the 12 months specified in the criteria.

5.4.2 Is there a difference between the clinically defined recovery groups on measures of personal recovery?

Participants in the clinically fully recovered group (*M*= 71.25, *SD*= 1.91) compared to the clinically non-recovered group (*M*= 56.75, *SD*= 1.96) reported significantly higher total scores on the QPR with a large effect size ($t(58)= -5.06$, $p< 0.001$, $d= -0.97$). Participants in the clinically fully recovered group (*M*= 3.02, *SD*= 0.12) compared to the clinically non-recovered group (*M*= 2.38, *SD*= 0.13) additionally reported significantly higher mean scores on the SCS with a large effect size ($t(55)= -3.61$, $p< 0.001$, $d= -1.33$). Therefore, participants

in the clinically fully recovered group scored significantly higher than the non-recovered group on both measures of personal recovery.

5.4.3 Is there a relationship between measures of clinical and personal recovery across the total sample?

There was no relationship between BMI and the mean SCS score (r_s= 0.22, p= 0.058), and a small positive relationship between BMI and the total QPR score (r_s= 0.26, p= 0.023) across the total sample. There was a large negative relationship between the EDE-Q Global score and both the SCS (r_s= -0.53, p< 0.001) and QPR scores (r_s= -0.63, p< 0.001). In summary, there was a strong association between lower ED symptoms and higher self-rated personal recovery and self-compassion, but only a small association between a higher BMI and higher self-rated personal recovery, and no relationship between BMI and self-compassion.

5.4.4 Participant perspectives on clinical definitions of recovery

Participants were asked whether they felt Khalsa's definition of full recovery reflected their personal experience and understanding of recovery, and whether they thought that they could be recovered even if they did not meet those criteria. Most participants in the whole sample (n= 53, total responses n= 73, 72.60%) did not feel that the proposed clinical definition of recovery reflected their personal experience, and a majority (n= 52, total responses n= 74, 70.27%) believed that they could be recovered even if they did not meet those criteria. Responses to each item did not significantly vary according to clinically defined recovery status ($\chi^2(2) = 1.48$, p= 0.476, $\chi^2(2) = 1.03$, p= 0.597).

Seventy participants additionally provided qualitative responses to elaborate on their survey answers. Using a thematic analysis approach, researchers identified two themes: "Proposed Improvements to the Recovery Criteria" (sub-themes "proposed additional

personal criteria", and "relationship between existing clinical and proposed personal criteria"), and "Problematising the Concept of Recovery Criteria" (sub-themes "universal versus individualised", "threshold versus process"). Themes with supporting quotes are summarised in Table 3.

Table 3: Summary of identified thematic structure with supporting quotes.

Main theme	Sub-theme	Supporting Quotes
Proposed Improvements to the Recovery Criteria	Proposed Additional Criteria	"Recovery is broader and includes things like living a meaningful life" (female participant, age 29, self-identifying as fully recovered).
	Relationship Between Existing Clinical and Proposed Personal Criteria	"I think that's the minimum towards recovery… for me it has been all of the above but far more than that" (female participant, age 26, self-identifying as fully recovered). "I believe recovery is being able to live with anorexia. I still sometimes restrict my eating and I do feel unhappy about my body and weight. But it doesn't take up my entire life and I am able to live a fulfilled life and do things I enjoy despite it" (female participant, age 27, self-identifying as fully recovered).
Problematising the Concept of Recovery Criteria	Universal Versus Individualised	"It needs to be so much broader than just physical and food related issues. Recovery is holistic and also disability needs to be factored in as well. Recovery is unique to the individual. For disabled people their "recovery" may look different" (male participant, age 25, self-identifying as partially recovered).

	"I think everyone has a certain level of body image concerns and their own way of eating, and if we held everyone to this criteria then everyone would still have an eating disorder. To overcome your own body image and fight diet culture takes an extra level of commitment that shouldn't have to be the marker for recovery" (female participant, age 23, self-identifying as fully recovered).
Threshold versus Process	"ED recovery is not all or none. That perspective reinforces the black and white thinking" (female participant, age 46, self-identifying as partially recovered). "I consider recovery to mean that the eating disorder no longer rules your life. I believe recovery (even in the long term) is not linear and that occasional, short-lived symptoms will arise in times of crisis" (female participant, age 31, self-identifying as fully recovered).

5.4.4.1 Theme 1: Proposed Improvements to the Recovery Criteria

5.4.4.1.1 Sub-Theme: Proposed Additional Criteria

Firstly, there was a strong sense that the proposed clinical criteria did not sufficiently capture aspects of recovery which felt meaningful to participants. Participants identified several additional concepts which they felt were important in their understanding of recovery. Some ideas linked into broader aspects of mental health not addressed by the ED clinical

criteria, including self-esteem, developing alternative coping mechanisms for stress, lowering anxiety, lessening self-harm, or dealing with past trauma. Participants also highlighted the importance of social and occupational aspects, such as being able to engage in relationships, employment, education, and hobbies.

5.4.4.1.2 Sub-Theme: Relationship Between Existing Clinical and Proposed Personal Criteria

Whilst there was a consensus that the idea of recovery needed to be extended to include other personal aspects, there was disagreement around the role of symptoms within this broadened conceptualisation: if recovery encompasses both personal and clinical domains, which has primacy? A minority of participants accepted the primacy of clinical criteria, positioning clinical thresholds as a basic requirement from which the individual can then build towards broader well-being in other domains.

In contrast, most of the sample felt that it was possible to be recovered in a broader sense whilst still meeting some clinical criteria. In this conceptualisation of recovery, the presence of enduring symptoms was positioned as secondary in relation to a primary criterion of living a meaningful life. Recovery was centred around reducing symptoms to the point where one can fully engage in life, rather than complete symptom extinction. From this perspective, it would be possible to experience enduring symptoms and still consider oneself fully recovered. It should be noted that within this group there was significant disagreement around how each of the different clinical criteria related to a broader understanding of recovery. Some participants felt that BMI was essential, others disagreed with its inclusion, and there were similar disagreements across both the ED-related thoughts and behaviours criterion.

5.4.4.2 Theme 2: Problematising the Concept of Recovery Criteria

Although the first two themes highlight participant recommendations for how definitions of recovery from AN could be improved (for example, including additional criterion, or changing acceptable symptom levels), this final theme captures concerns raised by participants around the basic concept of creating a universal definition recovery from AN.

5.4.4.2.1 Sub-Theme: Universal Versus Individualised

Firstly, in a continuation of ideas raised in the previous two themes, the concept of recovery reflecting an individual achieving a meaningful life (and judging whether enduring symptoms interferes with this process) raised tensions around whether this state can ever be objectively measured. Whilst common ideas around what a meaningful life might reflect were raised across the sample (for example, self-esteem, or engagement in social domains), there also was a sense that recovery should be understood as an experience which varies across individuals, including intersections between the experience of recovery and other co-occurring conditions (such as experiencing ongoing atypical eating due to physical disability), and cultural context. Participants raised concerns around how far proposed clinical criterion reflected a degree of disordered eating that they felt was normalised in their societies, suggesting that recovering to this standard would render them atypical in their cultural contexts. Other participants similarly highlighted concerns around the validity of BMI as a measure of individual health, including relating to cultural and racial differences in healthy BMI ranges.

5.4.4.2.2 Sub-Theme: Threshold Versus Process

Finally, there was discomfort around the positioning of recovery as a threshold state. Participants criticised the idea of clinical criteria as overly concrete in a way that resonated with thinking styles and perspectives characteristic of the ED. These concerns were also

linked to the concept of recovery as a process. Recovery was conceived by several participants as an active, ongoing, non-linear individual experience, problematising a criteria-based approach.

5.5 Discussion

The findings of this study suggest that there is some agreement between personal and clinical models of recovery in AN, in line with previous work hypothesising that there is a positive relationship between symptom reduction and greater functioning in personal recovery domains (Kenny & Lewis, 2023, Wetzler et al., 2020). Most participants who self-defined as fully recovered met clinical criteria for full recovery (albeit around a fifth of this group were clinically defined as not recovered), and those who met criteria for full recovery scored significantly higher on personal recovery measures compared to the not recovered group. Furthermore, there were strong correlations between lower ED behavioural and psychological symptoms, and higher personal recovery ratings.

However, the results also suggest limitations to this complimentary relationship. Firstly, there was poor agreement between self-defined partial recovery and clinical definitions, with the majority of those self-defining as partially recovered falling into the clinically not recovered category. This suggests that “partial recovery” is a harder to define state than full recovery and reflects the qualitative findings of this study that it may be particularly difficult to capture a process model of recovery using objective criterion. Secondly, around a fifth of participants self-defining as fully recovered were not recovered according to clinical criteria, suggesting that these criteria may lack relevance to a significant minority of this population.

Moreover, most participants in this study did not feel that the clinical criteria fully reflected their personal experience or understanding of recovery, and the majority felt that

they could still be recovered despite not fully meeting criteria. It should be noted that Khalsa et al's (2017) model is only one example of several different clinical approaches to defining recovery from AN in the absence of a universally agreed or widely used operationalisation, and exact criteria used across previous literature vary (Bardone-Cone et al., 2018). Given that this paper's focus is on the broader relationship between clinical and personal models of recovery, and that Khalsa et al's (2017) definitions were chosen as a pragmatic and particularly detailed operationalisation of clinical recovery in the absence of broader consensus in the field, the current paper is not intended to represent an evaluation or critique of the specific criteria. Rather, it makes three broader points relating to the ongoing conceptualisation of recovery in this field.

Firstly, researchers have called for establishing a uniform consensus on defining recovery from AN, and have highlighted that including the perspectives of those with experience of EDs is vital in ensuring these definitions are meaningful (Bardone-Cone et al., 2018; Couturier & Lock, 2006). The findings of this study suggest that a clinical perspective is not at odds with more personal models of recovery which highlight the salience of domains beyond clinical symptoms, with the significant caveat that these personal models may better co-exist with a clinical focus on symptom reduction compared to symptom extinction. Results suggest that clinical criteria which focus on the absence of symptoms may ultimately lack relevance for some people with lived experience of AN who feel that recovery is possible in the context of some enduring symptoms. An approach in line with these findings would have implications for the future design of clinical thresholds, with a focus on defining low levels of cognitions or behaviours as indicators of recovery rather than a complete absence.

Secondly, a point of controversy among participants in this study was around the use of BMI as a recovery indicator, with participants disagreeing on its usefulness. Moreover, the

quantitative results suggested that weight was the only aspect of recovery which did not have a strong relationship with measures of personal recovery. This is in line with previous research suggesting that psychological aspects of EDs were viewed by patients as more central to recovery than weight (Emanuelli, Waller, Jones-Chester, & Ostuzzi, 2012). Within the AN recovery construct, BMI may be the one element which does not have strong associations with personal aspects of recovery. This reinforces the suggestion that BMI should not be used in isolation to measure recovery, and should be combined with behavioural and psychological indices as a minimum to ensure that clinical criteria remain meaningful to lived experience (Bardone-Cone et al., 2010).

Thirdly, the findings of this study, particularly the diversity of opinions expressed in the qualitative section, emphasise that whilst a universal criterion-based approach to defining recovery from AN is vital for ongoing research and clinical work in this field, it will inevitably have limitations on an individual level. This is in line with concerns raised by researchers in this field who highlight that qualitative research suggests that recovery may best be captured using a flexible, individualised, and non-linear approach (Kenny & Lewis, 2023). For example, the presence of co-occurring conditions may complicate the measurement of ongoing behaviours: previous research on the overlap between autism and AN has highlighted that autistic individuals may experience disordered eating unrelated to AN which persists following recovery (Kinnaird et al., 2019). Clinicians could consider using personal recovery tools to support a more holistic exploration of what recovery means to an individual, such as the worksheet developed as part of Kenny & Lewis' (2023) proposed recovery model.

5.5.1 Limitations

This study has several limitations. Firstly, it was an online study wholly utilising self-report measures. Whilst this approach benefited the paper's aim of exploring how individuals self-define and conceptualise their own recovery, a drawback is this could not be combined with objective data on clinical variables such as height, weight, and illness duration. Therefore, the findings on how far self-defined recovery agreed with clinically defined recovery will inevitably lack accuracy, particularly around the BMI and recovery duration criteria. However, it should be noted that using self-report measures (i.e. the EDE-Q) as a measure of ED symptoms is widely practiced in the field, including in measuring recovery status (Bardone-Cone et al., 2018).

Secondly, the operationalisation of personal recovery in this study was limited due to the absence of personal recovery measures that are either specific to the ED field, or that have been validated in this population. Whilst this study's use of the QPR and SCS were in line with domains identified in Weztler et al's (2020) proposed model of ED personal recovery, further research in this area would significantly benefit from the development of ED specific personal recovery measures that reflect areas of functioning previously identified as salient to this population.

Thirdly, the paper had to make similarly difficult choices around the operationalisation of clinical recovery from AN given the lack of consensus in this field (Bardone-Cone et al., 2018; Couturier & Lock, 2006). In exploring the relationship between clinical and personal recovery the study ultimately chose to utilise the AN-specific criteria developed by Khalsa et al (2017). However, as highlighted in previous research if this paper had chosen a different set of definitions it would have found different results on agreement between self-defined and clinically defined recovery (Couturier & Lock, 2006). Nonetheless, it is notable that ideas raised by the qualitative analysis in this current paper (for example, concerns around the "black and white" nature of diagnostic criteria, recovery as

individualised) align with the findings of a recent study exploring the perspectives of people with lived experience of AN on alternative recovery criteria proposed by Bardone-Cone et al (2018), suggesting that there may be consistent concerns from a personal recovery perspective across different operationalisations of clinical recovery (Kenny et al., 2022).

Finally, this study's intention was not to explore how participants defined recovery from AN as this field has already been well-explored in the qualitative literature (see de Vos et al., 2017 for a meta-synthesis). Therefore, survey questions were specifically targetted to explore participant perspectives on clinical recovery criteria. The qualitative findings around how participants understand recovery should therefore be understood as being generated from this specific and limited context, rather than a broader exploration of how people recovered from AN define recovery.

5.6 References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5 ed.). Washington, DC.

Anthony, W. A. (1993). Recovery from mental illness: The guiding vision of the mental health service system in the 1990s. *Psychosocial Rehabilitation Journal, 16*(4), 11-23. doi:10.1037/h0095655

Arkell, J., & Robinson, P. (2008). A pilot case series using qualitative and quantitative methods: biological, psychological and social outcome in severe and enduring eating disorder (anorexia nervosa). *Int J Eat Disord, 41*(7), 650-656. doi:10.1002/eat.20546

Ashworth, M., Evans, C., & Clement, S. (2009). Measuring psychological outcomes after cognitive behaviour therapy in primary care: a comparison between a new patient-generated measure "PSYCHLOPS" (Psychological Outcome Profiles) and "HADS" (Hospital Anxiety and Depression Scale). *Journal of Mental Health, 18*(2), 169-177. doi:10.1080/09638230701879144

Austin, A., Potterton, R., Flynn, M., Richards, K., Allen, K., Grant, N., . . . Schmidt, U. (2020). Exploring the use of individualised patient-reported outcome measures in eating disorders: Validation of the Psychological Outcome Profiles. *European Eating Disorders Review, n/a*(n/a). doi:https://doi.org/10.1002/erv.2819

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*, 2456-2456. doi:10.3389/fpsyg.2018.02456

Bamford, B. H., & Mountford, V. A. (2012). Cognitive behavioural therapy for individuals with longstanding anorexia nervosa: Adaptations, clinician survival and system issues. *European Eating Disorders Review, 20*(1), 49-59. doi:https://doi.org/10.1002/erv.1080

Barber, M. E. (2012). Recovery as the new medical model for psychiatry. *Psychiatr Serv, 63*(3), 277-279. doi:10.1176/appi.ps.201100248

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. doi:https://doi.org/10.1016/j.brat.2009.11.001

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. doi:10.1007/s11920-018-0932-9

Bird, V., Leamy, M., Tew, J., Le Boutillier, C., Williams, J., & Slade, M. (2014). Fit for purpose? Validation of a conceptual framework for personal recovery with current mental health consumers. *Aust N Z J Psychiatry, 48*(7), 644-653. doi:10.1177/0004867413520046

Braun, T. D., Park, C. L., & Gorin, A. (2016). Self-compassion, body image, and disordered eating: A review of the literature. *Body Image, 17*, 117-131. doi:https://doi.org/10.1016/j.bodyim.2016.03.003

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa

Carey, M., Kupeli, N., Knight, R., Troop, N. A., Jenkinson, P. M., & Preston, C. (2019). Eating Disorder Examination Questionnaire (EDE-Q): Norms and psychometric properties in U.K. females and males. *Psychol Assess, 31*(7), 839-850. doi:10.1037/pas0000703

Cohen, J. (1992). A power primer. *Psychol Bull, 112*(1), 155-159. doi:10.1037//0033-2909.112.1.155

Couturier, J., & Lock, J. (2006). What is recovery in adolescent anorexia nervosa? *International Journal of Eating Disorders, 39*(7), 550-555. doi:https://doi.org/10.1002/eat.20309

Davidson, L., O'Connell, M. J., Tondora, J., Lawless, M., & Evans, A. C. (2005). Recovery in Serious Mental Illness: A New Wine or Just a New Bottle? *Professional Psychology: Research and Practice, 36*(5), 480-487. doi:10.1037/0735-7028.36.5.480

Dawson, L., Rhodes, P., & Touyz, S. (2014). The recovery model and anorexia nervosa. *Australian & New Zealand Journal of Psychiatry, 48*(11), 1009-1016. doi:10.1177/0004867414539398

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), 34. doi:10.1186/s40337-017-0164-0

Emanuelli, F., Waller, G., Jones-Chester, M., & Ostuzzi, R. (2012). Recovery from Disordered Eating: Sufferers' and Clinicians' Perspectives. *European Eating Disorders Review, 20*(5), 363-372. doi:https://doi.org/10.1002/erv.2159

Fairburn, C. G., & Beglin, S. J. (2008). Eating Disorder Examination Questionnaire (EDE-Q 6.0). In C. G. Fairburn (Ed.), *Cognitive Behavior Therapy and Eating Disorders*. New York: Guilford Press.

Fogarty, S., & Ramjan, L. M. (2016). Factors impacting treatment and recovery in Anorexia Nervosa: qualitative findings from an online questionnaire. *Journal of Eating Disorders, 4*(1), 18. doi:10.1186/s40337-016-0107-1

Kenny, T. E., Trottier, K., & Lewis, S. P. (2022). Lived experience perspectives on a definition of eating disorder recovery in a sample of predominantly white women: a

mixed method study. *Journal of Eating Disorders, 10*(1), 149. doi:10.1186/s40337-022-00670-2

Kenny, T. E. & Lewis, S. P. (2023). More than an outcome: a person-centred, ecological framework for eating disorder recovery. *Journal of Eating Disorders, 11*. doi: 10.1186/s40337-023-00768-1

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*, 20-20. doi:10.1186/s40337-017-0145-3

Kinnaird, E., Norton, C., Stewart, C., & Tchanturia, K. (2019). Same behaviours, different reasons: what do patients with co-occurring anorexia and autism want from treatment? *Int Rev Psychiatry, 31*(4), 308-317. doi:10.1080/09540261.2018.1531831

McCabe, R. E., McFarlane, T., Polivy, J., & Olmsted, M. P. (2001). Eating disorders, dieting, and the accuracy of self-reported weight. *International Journal of Eating Disorders, 29*(1), 59-64. doi:https://doi.org/10.1002/1098-108X(200101)29:1<59::AID-EAT9>3.0.CO;2-#

Meyer, C., Arcelus, J., & Wright, S. (2009). Accuracy of self-reported weight and height among women with eating disorders: A replication and extension study. *European Eating Disorders Review, 17*(5), 366-370. doi:https://doi.org/10.1002/erv.950

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity, 2*(3), 223-250. doi:10.1080/15298860309027

Neil, S. T., Kilbride, M., Pitt, L., Nothard, S., Welford, M., Sellwood, W., & Morrison, A. P. (2009). The questionnaire about the process of recovery (QPR): A measurement tool developed in collaboration with service users. *Psychosis, 1*(2), 145-155. doi:10.1080/17522430902913450

Rance, N., Moller, N. P., & Clarke, V. (2015). 'Eating disorders are not about food, they're about life': Client perspectives on anorexia nervosa treatment. *Journal of Health Psychology, 22*(5), 582-594. doi:10.1177/1359105315609088

Roberts, G., & Wolfson, P. (2004). The rediscovery of recovery: open to all. *Advances in Psychiatric Treatment, 10*(1), 37-48. doi:10.1192/apt.10.1.37

Schrank, B., & Slade, M. (2007). Recovery in psychiatry. *Psychiatric Bulletin, 31*(9), 321-325. doi:10.1192/pb.bp.106.013425

Shanks, V., Williams, J., Leamy, M., Bird, V., Le Boutillier, C., & Slade, M. (2013). Measures of Personal Recovery: A Systematic Review. *Psychiatric Services, 64*(10), 974-980. doi:10.1176/appi.ps.005012012

Slade, M. (2009). What is recovery? In M. Slade (Ed.), *Personal Recovery and Mental Illness: A Guide for Mental Health Professionals* (pp. 35-44). Cambridge: Cambridge University Press.

Slof-Op 't Landt, M. C. T., Dingemans, A. E., de la Torre Y Rivas, J., & van Furth, E. F. (2019). Self-assessment of eating disorder recovery: Absence of eating disorder psychopathology is not essential. *International Journal of Eating Disorders, 52*(8), 956-961. doi:https://doi.org/10.1002/eat.23091

Steinhausen, H. C. (2002). The Outcome of Anorexia Nervosa in the 20th Century. *American Journal of Psychiatry, 159*(8), 1284-1293. doi:10.1176/appi.ajp.159.8.1284

Wetzler, S., Hackmann, C., Peryer, G., Clayman, K., Friedman, D., Saffran, K., . . . Pike, K. M. (2020). A framework to conceptualize personal recovery from eating disorders: A systematic review and qualitative meta-synthesis of perspectives from individuals with lived experience. *International Journal of Eating Disorders, 53*(8), 1188-1203. doi:https://doi.org/10.1002/eat.23260

6. Executive Summary

6.1 Systematic Literature Review: Exploring the monitoring and reporting of unwanted events in psychotherapy trials for anorexia nervosa

Anorexia nervosa (AN) is a condition associated with poor treatment outcomes and heightened physical and mental risks. In the broader field of psychotherapy there is increasing research interest in the occurrence of unwanted events during treatment, however, despite the heightened risks of unwanted events in AN, at present this area is underexplored in the research literature. This project undertook a systematic review to explore the definition, monitoring and reporting of unwanted events in published randomised control trials (RCTs) evaluating psychotherapeutic interventions for AN. Findings suggested that unwanted event reporting in RCTs for AN varied widely, with papers using different definitions of key terms (for example, recovery or non-compliance), and often not exploring whether the unwanted event was potentially causally related to the intervention. The review also highlighted difficulties in operationalising the concept of unwanted events in AN: for example, not achieving recovery or remission is widely understood as an unwanted event, but would not hold relevance in trials targeting stabilisation instead of recovery. The review concludes with suggestions for how unwanted events in AN could be defined, monitored and reported in RCT protocols and publications in the future.

6.2 Service Improvement Project: Evaluating and addressing staff training needs in relation to autism in an older adult mental health service

There is a growing awareness that autistic people with co-occurring mental health problems may experience specific needs that require consideration and adaptation by mental health service providers. For this reason, NHS improvement standards and NICE guidelines state that clinicians working in this area should be provided with autism training. However,

there is a lack of research on how best this requirement can be operationalised in older adult mental health services. Autistic older adults may experience specific needs compared to their younger peers, and there is a risk that generic autism training would not fully meet the requirements of staff working in this area.

This current project used a mixture of audit and survey methodologies to explore staff training needs relating to autism in older adults in an older adult community mental health team. The project identified that whilst staff did see patients on the autism spectrum, in most cases autism was suspected by clinicians rather than officially diagnosed. Staff reported not having previously received autism training, and the most popular request for training was that the programme focus on autism in older adults. Subsequently the project designed and delivered a single session autism training for staff, evaluated using pre- and post-session surveys. The training was rated as useful by attendees and resulted in non-statistically significant increases in self-rated knowledge.

These findings suggest that clinicians working in this area may particularly benefit from autism training specialised to their older adult context. However, further consideration is needed to better understand whether training is effective in long term up-skilling in this area, and whether staff could benefit from further initiatives to further support skill development. Further research is also needed to better understand the prevalence of undiagnosed autistic older adults accessing mental health services.

6.3 Theory Driven Research Project: Exploring the relationship between clinical and personal models of recovery in anorexia nervosa

At present, there is no universal definition of recovery from anorexia nervosa (AN) that is consistently used across eating disorder (ED) research and service providers. Whilst there have been calls from researchers to develop a universal definition of recovery, this

raises broader conceptual questions about how recovery should be defined. Does recovery from AN represent attaining a healthy body weight and no longer experiencing symptoms (an approach sometimes known as clinical recovery), or does it represent a broader process of self-acceptance and personal growth (known as personal recovery)? Whilst researchers have theorised that these two approaches may exist in a complimentary relationship (with lower ED symptoms leading to recovery in other personal domains), to date this has not been explored by empirical research.

This research study used a mixed methods approach to explore the relationship between clinical and personal models of recovery in AN. Adult participants self-defining as partially or fully recovered from AN were recruited online and completed measures of clinical symptoms, measures of personal recovery, and answered questions about their views on the clinical/personal recovery relationship. The study found that there was good agreement between self-defined and clinically defined full recovery from AN. People who were clinically defined as fully recovered scored significantly higher on measures of personal recovery compared to those defined as not recovered, and there were strong positive associations between lower behavioural and cognitive symptoms, and personal recovery. However, there was a weaker relationship between personal recovery measures and BMI. Finally, the majority of participants suggested that they did not feel that clinical criteria were a sufficient definition of recovery: they felt that recovery was broader, including domains such as self-esteem and addressing co-occurring mental health problems, and also questioned whether objective criteria could ever fully capture individual experience.

These findings suggest that clinical approaches to defining recovery are not wholly at odds with personal models and may exist in a complimentary relationship. Researchers or service providers in this area could benefit from combining clinical outcome measures with measures that explore other domains, such as personal recovery measures or personalised

outcome measures. Further research is needed to explore how best a personal recovery approach can be synthesised with clinical criteria in creating a universal definition of recovery from AN in a way that feels meaningful to both patients, clinicians and researchers working in this area.

7. Connecting Narrative

I came onto the course having completed a PhD, and so I looked forward to the opportunity to continue to do research. For me, the biggest challenge has been learning how to balance and time-manage research whilst simultaneously doing clinical work. I quickly learned during the project development stage that I need a lot of mental space when thinking reflectively about concepts and research ideas, and that I could not fit trying to think about how to empirically test models of recovery in anorexia nervosa into my lunch break on placement. This has been a valuable lesson as my hope entering the course was to practice post-qualification as a psychologist balancing clinical work with research: I now have a clearer idea of what I would need for this to be a realistic prospect. In the future I want to further develop my research alongside working clinically: as raised below, I feel that my best research has been closely aligned with an awareness of the needs of clinicians, patients and families.

7.1 Service Improvement Project (SIP)

My SIP was the first project I developed. I had previously contributed towards a service development project developing an autism pathway within an eating disorder (ED) service, and when I started my older adult placement I was struck by the extent to which this was not a topic of conversation within the team: it seemed like no patients were diagnosed as autistic (despite the fact I knew that autism prevalence was significantly higher in mental health services compared to the general population), and the possibility of autism as a potential explanation never seemed to be raised in team meetings.

At this time my supervisor, aware of my interest in this area, gave me a case involving an older adult woman in a care home presenting with challenging behaviour, where her children suspected she might be on the autistic spectrum. The focus of the work became

providing her staff team with autism training. This led to the development of my SIP: whilst I initially wanted to conduct interviews with staff about their experiences working with autism (or suspected autism) in an older adult context, my supervisor felt that she wanted to create a project which would give something back to the staff team (rather than asking them to give up their time for interviews with no clear benefit to them). This led to the final design of a staff needs survey leading to a bespoke training session. This experience taught me a valuable lesson around how much easier it is to get people to engage in research when there is a clear buy-in and perceived benefit, and the value of aligning research designs with the needs of the target population. For me, a point of pride in the final research product was that it was experienced as valuable and helpful by the staff team, and that whilst I lack confidence in its long-term effects or generalisability it has at least started a conversation about autism in older adults within that specific team.

7.2 Systematic Literature Review (SRL)

I found this project to be the most straightforward in terms of design and execution. I had previously done research into the (largely negative) treatment experiences of autistic people with anorexia nervosa (AN), and I was aware from personal relationships that many people with EDs have negative experiences of treatment. I was therefore interested to explore the idea of iatrogenic harm in ED treatment, which led me to the broader concept of unwanted events in psychotherapy. The most challenging part of this project for me was the write-up stage, including going through the peer review journal publication process. I have sometimes felt that I have lacked confidence in thinking about how my research contributes to broader conceptual discussions in the field, I think because I feel that my research projects are not significant enough to have an impact on bigger concepts. In the past I have felt more confident using my research to test empirical models created by other people rather than going into a new area myself. However, a peer reviewer encouraged me to be more forward

in reflecting on how the findings of my SRL related to the broader concept of defining unwanted events in psychotherapy for AN, an area of research which is currently novel and under-explored in the ED field. I was grateful for their encouragement as this is the project that I feel most proud of, and I feel reflects my maturation and growing confidence as a researcher.

7.3 Theory Driven Research Project (TDRP)

Similarly to my SRL, my previous research in autism and EDs had made me aware that many people with EDs have significant disagreements with current clinical approaches, including how we conceptualise severity and recovery. I went into this project with a very clear idea of wanting to research personal models of recovery but found it unusually hard to come up with a clear research question or design: I found myself emailing my supervisor one week with a qualitative grounded theory proposal, and then the following day emailing to say "no, changed my mind, terrible idea". When I finally came up with a clear research question, I instantly knew it was the right one because it felt simple, straightforward, and easy to explain. It also resonated with the experts by experience consulted during development. It taught me a lesson about getting the research question right, and the design will follow. Compared to the difficult process of coming up with a research design I found the rest of the research project straightforward: I felt it played to my strengths (previous experience in this area, familiarity with methods used, existing relationships within the online ED community).

8. Acknowledgements

I would like to thank my supervisors Dr Myra Cooper, Dr Kathryn Evans, and Dr Anna Crabtree for their support during this process. I would particularly like to thank Myra for her help and encouragement in her dual role as my research supervisor and course tutor. I would also like to thank the other people who have helped my research, including the managers at the older adult CMHT for supporting the staff team in attending the SIP training, Meghan and Amelia for their help with my systematic review, and the experts by experience who joined with the development of my SIP and TDRP projects (Lyn Roberts, Mark Wood, and Cheryl Russell). Finally, I would like to thank my parents, partner and friends for all of their emotional support throughout the course.

9. Appendices

9.1 Appendix 1: Published Systematic Review

Kinnaird, E. et al. (2023). Systematic Review: Exploring the monitoring and reporting of unwanted events in psychotherapy trials for anorexia nervosa. *International Journal of Eating Disorders.*

Check for updates

Received: 11 November 2022 | Revised: 14 February 2023 | Accepted: 18 February 2023
DOI: 10.1002/eat.23926

REVIEW

INTERNATIONAL JOURNAL OF EATING DISORDERS WILEY

Systematic review: Exploring the monitoring and reporting of unwanted events in psychotherapy trials for anorexia nervosa

Emma Kinnaird PhD[1] | Meghan Nicholson BSc[1] | Amelia Thomas[2] | Myra Cooper DPhil[1]

[1]Oxford Institute of Clinical Psychology Training and Research, Oxford University, Oxford, UK
[2]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK

Correspondence
Emma Kinnaird, Oxford Institute of Clinical Psychology Training and Research, Oxford University, Oxford, UK.
Email: emma.kinnaird@psc.ox.ac.uk

Action Editor: Kelly L. Klump

Abstract

Objective: If a psychotherapeutic intervention has the capacity to produce positive change, the "deterioration effect" theory holds that it must also have the potential for negative effects. However, the definition, measurement, and reporting of unwanted events in psychotherapy is a topic of ongoing discussion. At present this area is under-explored in interventions for anorexia nervosa (AN), a severe mental illness associated with high medical and psychiatric risks. The aim of this article was to undertake a systematic review of published randomized controlled trials (RCTs) evaluating psychotherapeutic interventions for AN, and to assess how unwanted events were defined, monitored, and reported alongside the trial's key findings.

Method: Using a systematic review methodology, this article identified 23 RCTs through database searches meeting eligibility criteria. Results are presented using a narrative summary approach.

Results: Unwanted event reporting varied widely, both regarding definitions of key unwanted events (such as non-compliance or symptom deterioration) and in the amount of detail captured in each paper.

Discussion: The review identified two key issues: firstly, a lack of consistent definitions and unclear causality made it difficult to distinguish between unwanted events, and adverse events caused by the interventions. Secondly, it highlighted the difficulty of defining unwanted events where different studies have different populations and goals. Recommendations are made around how the area of defining, monitoring, and reporting unwanted events in RCTs for AN can be taken forward.

Public Significance: Whilst psychotherapies can be effective in treating mental health conditions, negative or unwanted events can sometimes occur. This review explored how RCTs examining psychotherapy for AN report how they monitor the safety of participants, and how they report unwanted events. We found that the reporting was often inconsistent or hard to interpret, and we have made recommendations for how this can be improved in the future.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. International Journal of Eating Disorders published by Wiley Periodicals LLC.

Resumen

Objetivo: Si una intervención psicoterapéutica tiene la capacidad de producir un cambio positivo, la teoría del "efecto de deterioro" sostiene que también debe tener el potencial de efectos negativos. Sin embargo, la definición, medición y reporte de eventos no deseados en psicoterapia es un tema de discusión continua. En la actualidad, esta área está poco explorada en las intervenciones para la anorexia nerviosa (AN), una enfermedad mental grave asociada con altos riesgos médicos y psiquiátricos. El objetivo de este trabajo fue realizar una revisión sistemática de los ensayos controlados aleatorios (ECA) publicados que evalúan las intervenciones psicoterapéuticas para la AN, y evaluar cómo se definieron, monitorearon e informaron los eventos no deseados junto con los hallazgos clave del ensayo.

Método: Mediante una metodología de revisión sistemática, este documento identificó 23 ECA mediante búsquedas en bases de datos que cumplían los criterios de elegibilidad. Los resultados se presentan utilizando un enfoque de resumen narrativo.

Resultados: El informe de eventos no deseados varió ampliamente, tanto con respecto a las definiciones de eventos no deseados clave (como el incumplimiento o el deterioro de los síntomas) como en la cantidad de detalles capturados en cada artículo.

Discusión: La revisión identificó dos cuestiones clave: en primer lugar, la falta de definiciones consistentes y la causalidad poco clara dificultaron la distinción entre eventos no deseados y eventos adversos causados por las intervenciones. En segundo lugar, destacó la dificultad de definir eventos no deseados donde diferentes estudios tienen diferentes poblaciones y objetivos. Se hacen recomendaciones sobre cómo se puede llevar adelante el área de definición, monitoreo e informe de eventos no deseados en ECA para anorexia nerviosa.

KEYWORDS

adverse events, anorexia nervosa, eating disorders, psychotherapy, randomized controlled trials, systematic review, unwanted events

1 | INTRODUCTION

There is increasing research interest in the concept of unwanted events in psychotherapy trials, with concerns that, to date, potential negative outcomes that occur during psychological interventions have been under-explored (Crawford et al., 2016; Duggan et al., 2014; Linden, 2013; Parry et al., 2016). The "deterioration effect" theory holds that if therapy has the capacity to produce positive change, it must also have the potential for negative effects (Bergin, 1966). A 2014 review of National Institute for Health Research (NIHR) funded trials found that no psychological intervention trial reported on adverse event occurrence in their final report, suggesting that negative effects from these interventions may be under-reported and under-recognized (Duggan et al., 2014).

However, defining unwanted events in the psychotherapy field comes with its own complexities. On a basic level, there is a wide range of definitions and terms (Parry et al., 2016). This review follows the definition of Linden (2013) in conceiving unwanted events as any unwanted event that occurs parallel or following treatment. Linden then distinguishes between types of unwanted events depending on whether there is evidence that the event was caused by the treatment, whether it was caused by incorrectly applied treatment, and specific events such as lack of improvement or deterioration. Similarly, Parry et al. (2016) have proposed three standard definitions that capture adverse effects of therapy, where it is judged that there is a causal link between the intervention and the effect: adverse events (a significant episode during or shortly after treatment that amounts to harm), clinically significant deterioration (a worsened mental state following therapy), and patient-experienced harm (negative experiences of therapy). This review follows these definitions by positioning "adverse events" as a specific sub-type of unwanted events where there is an inferred causal link with the intervention.

In addition, the concept of reporting adverse events in clinical trials has been highly influenced by the medical field, where unwanted events are conceived as measurable, often physical, harms that are considered medically significant—for example, hospitalization, or resulting in death (Duggan et al., 2014; Ioannidis et al., 2004). By contrast, in psychotherapeutic interventions unwanted events may look different

and impact: more psychological, social, and emotional domains—for example, an intervention may lead to heightened distress in a way that impacts on the individuals relationships or performance at work (Linden, 2013). Furthermore, definitions of unwanted events in the research literature typically focus on distinguishing harms that can be identified, or inferred, as caused by the intervention itself (Duggan et al., 2014; Linden, 2013; Perry et al., 2016). However, causality can be complex to assess in this field and may not be captured in overall trial reporting. A psychotherapy trial paper may report on hospitalization rates, but without added context it is difficult for the reader to understand whether this reflects an adverse event, or an unwanted event that occurred in parallel to the intervention for reasons unrelated to the intervention. For example, it would be unhelpful to count drug abuse as an adverse event in a trial where the intervention is aimed at individuals who already experience drug abuse (Horigian et al., 2010).

Furthermore, who defines harm? Perry et al. (2016) helpfully raise the concept of patient-experienced harms in their list of adverse events, and it is clearly evident from qualitative research that patients can and do have negative experiences of treatment (Clifford et al., 2008). Linden (2013) highlights that successfully engaging in psychotherapy can create negative experiences for patients (e.g., distress stemming from remembering upsetting past experiences), but that this does not necessarily mean that the psychotherapy itself is harmful. In fact, experiencing these negative effects and discomfort can be a key part of the change process. Therefore, an effect may be unwanted, and related to the treatment, and yet not represent an adverse event from the perspective of the clinician. Indeed, it may reflect a "wanted" unwanted event in the context of the intervention.

A further complexity previously raised around the difficulties of defining and measuring unwanted events in the field of psychotherapy interventions is that of specificity (Duggan et al., 2014; Horigian et al., 2010). In their work on defining adverse events as part of a behavioral intervention trial, Horigian et al. (2010) highlight how conditions, populations, and intervention types are likely to be associated with specific risks that should influence what becomes defined as an unwanted or adverse event within a trial. To date, there is a lack of research on unwanted events specific to anorexia nervosa (AN). AN is a serious eating disorder (ED) characterized by the restriction of energy intake relative to requirements resulting in low body weight, an intense fear of weight gain, disturbance in self-perceived weight and shape, and the undue influence of weight and shape on self-evaluation (American Psychiatric Association, 2022). Firstly, understanding unwanted events in the context of psychotherapeutic interventions for AN is particularly relevant as this condition is associated with treatment non-response. A range of psychotherapeutic interventions have been developed for the treatment of AN and are known to result in reductions in ED symptomatology (Murray et al., 2019; Solmi et al., 2021). Despite these advances, research suggests that many individuals experience poor outcomes following treatment for AN: In a systematic review of the intervention literature, 46.9% of patients were found to have made a full recovery, 33.5% showed some improvement, 20.8% developed chronic AN, and 5% of patients died (Steinhausen, 2002). There is further evidence from more recent research that a small proportion of patients may indeed deteriorate during treatment (Schlegl et al., 2014, 2016).

Secondly, AN is a condition associated with heightened mortality rates and other serious physical complications (Arcelus et al., 2011; Birmingham et al., 2020). Firstly, this creates significant complexity around inferring causality: If a trial for AN reports high hospitalization rates, does this reflect an effect of the intervention, or pre-existing medical vulnerability? Secondly, it means that trial investigators need to simultaneously consider potential harms typically associated with psychotherapy trials, but also incorporate medical monitoring and harm reporting as seen in the medical field. Thirdly, measuring unwanted events such as treatment non-response is complicated by a lack of agreement in the field on how to define recovery from AN (Khalsa et al., 2017). Finally, the concept of patient-experienced harms gains additional complexity in this population. Whilst people with AN may be vulnerable to experiencing unwanted events during treatment, the egosyntonic nature of AN raises questions around whether these experiences should then be considered as unwanted in the context of treatment. Individuals with AN can experience their condition as highly valued and performing important psychosocial functions, which can contribute to experiencing treatments aimed at changing these behaviors as highly aversive and, indeed, unwanted (Gregertsen et al., 2017).

Whilst the concept of unwanted events in psychotherapy interventions is particularly salient (and potentially complex) in AN, it is also an area which is currently under-explored. Therefore, the aim of the current review was to better understand how unwanted events are currently defined, monitored, and reported alongside the findings of published randomized controlled trials (RCTs) for psychological interventions for AN. It is hoped that synthesizing this information is a first step in further research and discussion around unwanted events in this field.

2 | METHODS

This review was pre-registered with PROSPERO (registration number CRD42021275696) and performed according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Moher et al., 2009).

2.1 | Design

This review originally intended to utilize a pre-existing adverse event checklist, specifically the harms extension to the CONSORT statement, an adapted version utilized in a previous review of reporting in general medical RCTs, or Linden's psychotherapy checklist, to explore this area (Ioannidis et al., 2004; Linden, 2013; Phillips et al., 2019). However, following a pilot of this approach on a smaller number of identified papers, it became evident that the definitions of adverse events and reporting standards outlined in these checklists were a poor fit with the types of information typically reported in RCTs for AN. For example, no trial reported information on absolute risk per arm (as recommended by the CONSORT extension), or included a

10990886, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/eat.23769 by Test, Wiley Online Library on [07/06/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

consideration of whether the intervention had resulted in the patient having to take sick leave from their job (as recommended by Linden).

Consequently, the design of the study shifted from attempting to compare published RCTs in this area to existing standards and definitions, and instead chose to focus more on exploring what kind of information and definitions are currently being reported.

In line with this, authors EK and MC developed a data extraction checklist focused on capturing this information. This checklist was influenced both by definitions of unwanted events utilized in previous work, and the initial pilot work identifying what types of information are typically reported in this field (Ioannidis et al., 2004; Linden, 2013; Perry et al., 2016). The final data extraction checklist is reported below.

2.2 | Eligibility criteria

Studies were included if they fulfilled the following criteria:

- RCT evaluating a specified psychotherapy for individuals with AN.
- Published in English.
- Published since 1980.
- Article published in a peer-reviewed journal.

Trials evaluating medication in addition to psychotherapy were excluded as these present different risk monitoring challenges compared to psychotherapy only. As not all trials recorded consistent follow-up data, this review primarily focuses on unwanted event monitoring during and at the end of treatment (EOT), and thus follow-up studies were excluded from the analysis. Pilot, feasibility, or exploratory RCTs, or those published as brief reports, were also excluded from the analysis in case this (for example) led to reduced focus on reporting of safety monitoring procedures. Only published papers were eligible for inclusion as the intention of this review was to evaluate how unwanted events are defined in reported in final published trial papers.

2.3 | Search strategy

Search terms were developed by reviewing the titles of published RCTs in AN for key words. Final search terms were "anorexia nervosa" AND "rando*" AND "therapy." Searches were performed across PsycInfo, Medline, Scopus, and PubMed databases during June 2022. Following screening, the final list of included papers was evaluated against previously published systematic reviews of RCTs in AN to identify missed studies (Brockmeyer et al., 2018; Bulik et al., 2007; Murray et al., 2019). No further papers were identified in this way.

2.4 | Selection process

The selection process is detailed in Figure 1. Searches were performed by the first author, with titles and abstracts then screened against the eligibility criteria. Twenty percent of these were also randomly selected and screened by a second researcher (Cohen's kappa .95). Selected full texts were then screened by the first author, with 20% screened by a second researcher (Cohen's kappa .8).

With the increasing trend to publish trial protocols in advance of results and final publication, it was noted during full text screening that several papers signposted readers to previous publications for further details on study design, including safety monitoring. Therefore, readers interested in unwanted event monitoring and occurrence could access this information based on signposting within the paper. Where papers indicated that further information about the study was available from other sources (e.g., published protocols, available trial registration documents, or other papers relating to the trial) this was accessed and has been included in the synthesis.

2.5 | Data extraction

The following contextual information was extracted from papers: total sample size, mean age, mean BMI at baseline, illness duration, participant ethnicity/socioeconomic status, diagnostic inclusion criteria, treatment setting, and interventions. The following information regarding unwanted events was extracted. For all points, authors searched papers for both the presence of this information (e.g., a specified definition of adverse events), but also the absence (e.g., no information on compliance rates).

- Definitions of adverse events utilized within papers.
- Information about safety monitoring and procedures.
- Reporting of non-compliance rates.
- Treatment non-response definitions and rates.
- Evidence of symptom deterioration.
- Evidence of the emergence of new symptoms.
- Need for additional treatment beyond interventions.
- Negative feedback from patients and/or family.
- Patient mortality.
- Any information or analysis on causality for unwanted events. This included where causality was hypothesized by authors, or potentially indicated by statistical analyses.

There is controversy around whether treatment non-completion should be considered an unwanted event (Duggan et al., 2014). Given the high rates of non-completion in ED treatment and its inclusion in the CONSORT recommendations for reporting RCTs, non-completion is included in this review (DeJong et al., 2012; Ioannidis et al., 2004).

Data extraction was completed by the first author, with 90% of randomly selected papers also extracted by a second researcher to ensure reliability (Cohen's kappa .87).

2.6 | Quality assessment

As this review explored the reporting of safety measures and unwanted outcomes, it was decided to assess the quality of individual studies based on the broader quality of their reporting (which also gives an indication of their methodological rigor). The quality assessment utilized a

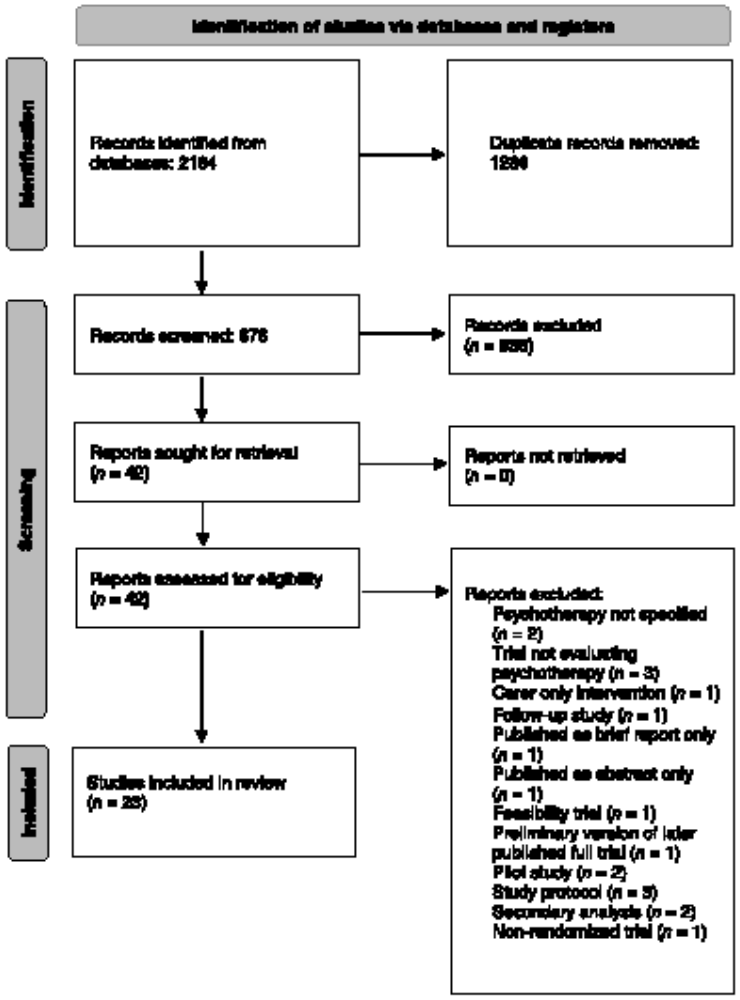

FIGURE 1 PRISMA 2020 flow diagram of study selection process.

previously developed checklist consisting of eight items based on the CONSORT reporting standards for non-pharmacological interventions (Boutron et al., 2008; Hoffmann et al., 2013). Each item has a yes/no response, and every item must have a "yes" response for the reporting to be described as "adequate." The only adjustment to Boutron et al.'s (2008) procedure was the item regarding whether intervention materials (e.g., therapy manuals) were publicly available as many manualized therapies are published as copyrighted material, this item was adjusted to reflect whether authors signposted to where readers could access the full manual (e.g., a reference to a published book). Related documents such as trial registrations, pilot studies, or protocols were also included in the quality analysis. Papers were rated by the first author, with 20% randomly selected papers also evaluated by a second researcher with full agreement on adequacy ratings.

3 | RESULTS

Twenty-three RCTs were included in the final analysis. Study details, including proportions of participants reported as experiencing unwanted events, are summarized in Table 1. Participant details are summarized in Table 2. The majority of studies were conducted on an outpatient basis;

TABLE 1 Overview of study characteristics and unwanted event outcome reporting.

Paper	Linked papers/trial documents	Population	Sample size	Intervention	Non-completers	Treatment non-response	Symptom deterioration	Mortality (n, %)
Agras et al. (2014)	Trial registration documents available at clinicaltrials.gov, #NCT00431978. Trial design detailed in previous paper (Lock et al., 2010).	Adolescents (age 12–18) meeting DSM-IV criteria for AN (excluding amenorrhea criterion). Outpatient setting.	144	FBT, BeFT	Dropout rate 20% FBT, 20% BeFT. Withdrawal due to medical instability: n/a FBT, n/a BeFT.	Not in remission at EOT: 70% FBT, 75% SyFT. At 12-month FU: 89% FBT, 61% BeFT.	Withdrawn due to medical instability: 4% FBT, 7% SyFT. Participants reporting an adverse event relating to hospitalization or suicidality: 18% FBT, 25% SyFT.	0 (0%)
Ball and Mitchell (2004)	No	Female adolescents and adults (age 13–39). Inclusion criteria: participants weighing less than 90% of their average body weight. Outpatient setting.	35	CBT, SFT	Total non-completers 34%; 16% withdrawn due to hospitalization.	Poor outcome at EOT: 86%. No difference at 6-month FU.	Hospitalization total 16%.	NR
Brockhoper et al. (2010)	Trial registration documents available at International Clinical Trials Registry Platform, (ICRS00000607).	Adult and adolescents (aged over 17) meeting DSM-IV criteria for AN. Inpatient setting.	167	CRT, ART	Non-completers 13% CRT; 14% ART.	No	No	NR
Byrne et al. (2017)	Trial registration documents available at Australian New Zealand Clinical Trials Registry, (ACTRN12610000256055).	Adult and adolescents (aged over 17) meeting DSM-IV criteria for AN, including BMI cut-off below 14.5. Outpatient setting.	150	SSCM, MANTRA, CBT-E	Non-completers 45% SSCM, 45% MANTRA 30% CBT-E. Dropout rate 30% SSCM, 26% MANTRA, 26% CBT-E. Withdrawal due to BMI falling below 14 kg/m² SSCM, 0% MANTRA, 3% CBT-E. One withdrawal due to recommencing psychotropic medication.	Not in remission at 12-month FU: SSCM 45%, MANTRA 76%, CBT-E 69%. Not still losing BMI over 0.5 kg/m² at 12-month FU: n/a SSCM, 43% MANTRA, 31% CBT-E.	Withdrawal due to BMI falling below 14 kg/m²: SSCM, 0%, MANTRA, 6%, CBT-E. Hospitalization: 30% SSCM, 3% MANTRA, 0% CBT-E.	0 (0%)
Channon et al. (1989)	No	Female (age range not specified) meeting Russell's (1970) criteria for AN. Outpatient setting.	24	CBT, BT, TAU	Drop-out rate 0% CBT, 33% BT, 25% TAU.	No	Hospitalization: 20% CBT, 33% BT, 50% TAU.	NR
Dalle Grave et al. (2013)	No	Adult and adolescents (aged 16–65) meeting DSM-IV criteria for AN. Inpatient setting.	60	CBT-IE, CBT-ID	Non-completers 13% CBT-IE, 0% CBT-ID.	Not still losing BMI over 0.5 at EOT: 34% (total sample), 6-month FU: 65%, 12-month FU: 50% (total sample). Mutual agreement at EOT: 31% (total sample), 12-month FU: 39% (total sample).	Hospitalization: 0% (total sample).	NR
Dare et al. (2001)	No	Adults (aged over 18) meeting DSM-IV criteria for AN. Outpatient setting.	84	FPP, CBT, FT, low contact control treatment (TAU)	Drop-out rate: 29% (total sample).	Poor outcome at EOT: 49% FPP, 50% FT, 69% CAT, 79% TAU.	Hospitalization: 14% (total sample).	1 (1%)
Eiber et al. (2020)	Pilot study to Lo Coco et al., (2018).	Adolescents (age 13–17) meeting DSM-IV or ICD-10 criteria for AN. Outpatient setting.	40	CFT, SFT	Dropout rate within first 5 months of treatment: 10% (total sample).	Poor outcome at EOT: 53% CFT, 34% SFT.	Hospitalization: 10% (total sample).	NR
Eiber et al. (2019)	Trial registration documents available at ISRCTN Registry, #ISRCTN11570468.	Adolescents (age range 12–20 years) meeting DSM-IV criteria for AN or BN-NOS.	149	MFT, SFT	Drop-out rate =13% (total sample).	Poor outcome at EOT: 42% SFT, 34% MFT. 6-month FU: 42% SFT, 33% MFT.	Hospitalization: MFT 4%, SFT 8%.	NR

www.interscience.wiley.com/journal/eat by Aristotle University of Thessaloniki, Wiley Online Library on [06/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 (Continued)

Paper	Linked papers/trial documents	Population	Sample size	Intervention	Non-completion	Treatment non-response	Symptom deterioration	Mortality (n, %)
		(restricting type), below 85% median BMI or just 15% body weight over last 3 months. Outpatient setting.						
Richter et al. (2012)	No	Female adolescents and adults (median age 24) meeting DSM-IV criteria for AN (excluding amenorrhea criterion). Postrecovery during weight gain.	226	CBT-based telemedicine, TAU	No-to-low engagement with interventions: 24.3% (total sample). Drop-outs: 6% (9), 8% TAU.	No	BMI falling below 13.5–6% (total sample). Mean weight loss at EOT: in TAU group. Hospitalization for inpatient treatment: 22% (9), 33% of TAU.	0 (0%)
Gelin et al. (2009)	No	Female adolescents (age range 13–17) meeting DSM-IV criteria for AN, including participants weight <90% IBW. Outpatient treatment offered following inpatient admission.	20	FT, FP	NR	NR	Hospitalization for inpatient treatment during intervention: 30% (total sample, FU period: 34% (total sample).	NR
Calvert et al. (2019)	Trial protocol provided as supplementary material but only available in French, therefore not included.	Female adolescents and adults (age 13–21) meeting DSM-IV criteria for AN. Outpatient treatment offered following inpatient admission.	40	Family therapy in addition to TAU (FT), TAU	Non-completion: 5% TAU, 40% FT.	Poor response at 18 months following inclusion: 40% FT, 80% TAU.	Hospitalization: 40% FT, 20% TAU.	NR
Le Grange et al. (2016)	Trial registration documents available at Australian New Zealand Clinical Trials Registry (ACTRN12608000001411. Study protocol available (Le Grange et al., 2014).	Adolescents (age 12–18) meeting DSM-IV criteria for AN (excluding amenorrhea), included participants weighing less than 85% of their average body weight. Outpatient setting.	167	FBT, SFT	Non-completion: 15% (total sample).	Not in remission at EOT: 37% FFT, 76% FBT. At 6-month FU: 43% FFT, 70% FTB.	Hospitalization: 23% FBT, 34% FBT. 50% of hospitalizations for medical reasons.	NR
Lock et al. (2010)	Pilot study available (Lock & Le Grange, 2001).	Adolescents (age 12–18) meeting DSM-IV criteria for AN, included some partially weight restored participants and those not meeting amenorrhea criterion. Outpatient setting.	86	Short-term FT, Long-term FT	Drop-out rate: 4% Short-term FT; 16% Long-term FT	Not in remission at EOT: 46–65% (depending on definition).	Hospitalization: 23% (total sample).	NR
Lock et al. (2019)	Trial registration documents available at clinicaltrials.gov (NCT02044978).	Adolescents (age 12–18) meeting DSM-IV criteria for AN (excluding amenorrhea criterion). Outpatient setting.	121	FBT, AFT	Drop-out rate: 14% at EOT, 22% at FU (total sample).	Not in remission at EOT: 58% FBT and 77% AFT. At 6-month FU: 43% FBT, 63% AFT. Not in partial remission at EOT: 15% FBT, 49% AFT. Achieved full remission, but relapsed by 12-month FU: 33% FBT, 40% AFT.	Hospitalization: AFT 30%, FBT 15%. Hospitalization during FU period: 18% AFT, 8% FBT.	NR

[Continued]

TABLE 1 (Continued)

Paper	Linked papers/trial documents	Population	Sample size	Interventions	Non-completers	Treatment non-response	Symptom deterioration	Mortality (n, %)
						Achieved partial remission N/I: relapsed by 12-month FU: 18% FBT, 45% APT.		
Hubbard et al. (2003)	No	Adult females (ages 17–40) meeting DSM-IV criteria for AN (not including assessment criterion) included participants with a BMI < 18. Outpatient setting.	46	CBT, IPt (partial treatment (NRCH))	Non-completers 38% (total sample).	Poor outcome at EOT: 40% IPt, 39% CBT, 39% NRCH.	Hospitalisations: 7% (total sample).	1 (6%)
Nyman-Carlsson et al. (2015)	Trial registration documents available at: ISRCTN Registry, ISRCTN92853690.	Adolescents and adults (ages 12–24) meeting DSM-IV criteria for AN. Outpatient setting.	79	CBT, FT	Drop-out: 30% CBT, 8% FT. Withdrawal due to adverse events: 24% CBT, 22% FT.	BMI below 17.5 at EOT: 6% FU: 18% CBT, 8% FT. Meeting criteria for AN at EOT: 8% CBT, 8% FT. At FU: 5% CBT, 0% FT.	Proportions not reported.	NR
Pierluigi et al. (2016)	Trial registration documents available at: ISRCTN Registry, ISRCTN53250530.	Adults (ages 18–51) meeting DSM-IV criteria for full or partial AN. Outpatient setting following daycare treatment.	43	ACT, TAU	Non-completers: 42% (ACT).	Not achieving a good outcome at: EOT: 79% ACT, 69% TAU. 6-month FU: 79% ACT, 85% TAU. 18-month FU: 88% ACT, 69% TAU. 24-month FU: 64% ACT, 85% TAU. 5-year FU: 44% ACT, 44% TAU.	Proportions not reported.	1 (5%)
Eddy et al. (1999)	Robin et al. (1994) (earlier paper reporting on smaller sample).	Adolescents (ages 11–18) meeting DSM-III criteria for restrictive AN. Outpatient setting.	37	BFST, EOIT	Non-completers: 8% (see full paper for details).	Not attaining target weight at: EOT: 25% BFST, 33% EOIT. 1 year FU: 29% BFST, 31% EOIT. Not attaining 25th percentile of body mass for age at EOT: 18% BFST, 19% EOIT. At 1 year FU: 19% BFST, 7% EOIT. Not attaining 50th percentile of body mass for age at EOT: 47% BFST, 69% EOIT. 1 year FU: 50% BFST, 25% EOIT	Hospitalisations: 25% BFT, 25% EOIT.	NR
Schmidt et al. (2016)	Trial registration documents available at: ISRCTN Registry, ISRCTN57735660. Trial protocol (G. Schmidt et al., 2010).	Adults (ages 16–45) meeting DSM-IV criteria for AN, included those not meeting amenorrhea criterion, with less restrictive, or partial AN. Outpatient setting.	142	MANTRA, SSCM	Non-completers: 29% MANTRA, 41% SSCM.	Not recovered at 6 months: 69% SSCM, 84% MANTRA. 12 months: 58% SSCM, 58% MANTRA.	Assessed hyperactive-impulsive treatment for 85: 7% MANTRA, 11.43% SSCM.	1 (1%)
Dennhag et al. (1999)	No	Adolescents and adults (adolescents age 14) meeting DSM-III criteria for AN. Outpatient setting.	35	CT, D	Dropout: 6% CT, 10.5% D.	Still meeting diagnostic criteria for AN at EOT: 44% CT, 100% D.	N/I	NR

www.interscience.wiley.com/journal/eat by the Applicable Creative Commons License

TABLE 1 (Continued)

Paper	Linked papers/trial documents	Population	Sample size	Interventions	Non-completers	Treatment non-response	Symptom deterioration	Mortality (N, %)
Thege et al. (2018)	Trial registration documents available at Australian New Zealand Clinical Trials Registry: ACTRN12612000404820.	Adults meeting DSM-IV criteria for DSM-IV [anorexia nervosa criteria] for over 7 years, included participants with a BMI between 13.5 and 18.5. Outpatient setting.	49	CBT, SSCM, BAS modified for severe and enduring AN.	Non-completion: 18% CBT, 29% SSCM.	NR	Hospitalisation: 11%	1 (2%)
Zipfel et al. (2014)	Trial registration documents available at ISRCTN Registry, reference ISRCTN72886057; Protocol (Wild et al., 2009).	Adults meeting DSM-IV criteria for DSM-IV, BMI range 13–18.5. Outpatient setting.	242	FPT, CBT-E, TAU	Non-completers: 22% total sample.	Basic properties NR.	Hospitalisation: 25% FPT, 34% CBT-E, 45% TAU.	0 (0%)

Abbreviations: ACT, acceptance and commitment therapy; AFT, adolescent focused individual therapy; AN, anorexia nervosa; ANT, ante therapy; BFST, behavioural family systems therapy; BFT, behavioural family therapy; BMA, body mass index; BT, behavioural treatment; CAU, cognitive analytic therapy; CBT, cognitive behavioural therapy or cognitive behavioural treatment; CBT-E, enhanced cognitive behavioural therapy; CBT-ED, focused enhanced cognitive behavioural therapy; CFT, cognitive family therapy; CRT, cognitive remediation therapy; CT, cognitive therapy; D, dietary advice; DSM-IV, diagnostic and statistical manual of mental disorders, fourth edition; DSM-V, diagnostic and statistical manual of mental disorders, fifth edition; EDNOS, eating disorder not otherwise specified; EOET, age orientated individual therapy; EOT, end of treatment; FBT, family-based therapy or family-based treatment; FP, family group psychoeducation; FPP, focal psychoanalytic psychotherapy; FPT, focal psychotherapy; FT, family therapy or caretaker individual/family therapy; FU, follow up; IP, interpersonal psychotherapy; MANTRA, Maudsley model emotion regulation treatment for adults; MFT, multifamily therapy; NR, not reported; NSCP, non-specific supportive clinical management; PFT, parent focused treatment; RP, relapse prevention; SFT, supported family therapy or single family therapy; SSCM, specialist supportive clinical management; SyFT, systemic family therapy; TAU, treatment as usual.

two studies took place in an inpatient setting. The average reported mean age of participants was 20.64 years (*SD* = 5.35), with included ages ranging from 11 to 65 years. Six studies (26.09%) included adults only, eight (34.78%) included adolescents only, and the remainder (39.13%) either included a mix or did not report age as an inclusion criteria. The average reported mean starting BMI was 16.10 (*SD* = 1.04). The majority of studies did not report on participant ethnicity, or socio-demographic status. A wide range of different psychotherapies were included in the review, including different types of family therapy and varying individual psychotherapies such as cognitive behavioural therapy or psychodynamic interventions.

The overall quality of reporting was rated as adequate in 52.17% (*n* = 12) of the studies, with papers most commonly rated as inadequate due to not detailing the availability of treatment manuals (Hoffmann et al., 2013). Considering that it is not uncommon for psychotherapy trials to not make therapy manuals publicly available, removing this criterion meant that 73.91% (*n* = 17) papers were rated as adequate.

In the context of varying definitions and reporting standards, a narrative approach was used to synthesise data and to explore what kind of information is typically captured by RCTs for psychotherapy in AN as a first step in reviewing the relevant literature on this topic.

3.1 | Definitions of unwanted events

CONSORT recommendations for reporting harm in RCTs suggest that trials should list and define addressed adverse events in Section 2 (Ioannidis et al., 2004). Only three papers (13.04%) specified definitions of unwanted events. Agras et al. (2014) included definitions of adverse events ("an unfavorable change from the participant's condition at baseline") and serious adverse events ("life-threatening events, death, hospitalisations for any cause, or other important medical or psychological events"). Fichter et al. (2012) defined "undesirable events" as serious illnesses and a loss of weight leading to a BMI below 13.5, and "serious undesirable and catastrophic events" as acute suicidality, suicide attempts, death by suicide, death by accident or by natural cause. Zipfel et al. (2014) defined serious adverse events as any fatal or life-threatening event, except for intermittent inpatient treatment of up to 4 weeks (due to the commonality of serious weight loss in AN).

One paper (4.35%) described intermittent inpatient/day treatment as an example of an adverse event but did not include a broader definition (Nyman-Carlsson et al., 2019).

3.2 | Safety monitoring

Papers were considered to have included information about monitoring participant safety where they detailed medical risk monitoring, psychiatric risk monitoring, procedures around hospitalisation, or procedures around monitoring and reporting adverse events. All papers included the assessment of ED and related psychiatric symptoms at specified timepoints for the purposes of trial outcome measurement; however, it was unclear if any papers utilised this information to additionally

...by guidelines for the purpose of text and data mining, or training artificial intelligence technologies, or similar technologies. Protected by Copyright. Use of this article is governed by the Wiley Online Library Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Overview of patient clinical and sociodemographic characteristics.

Paper	Location of study	Participants' mean age in years (SD)	Participants sex assigned at birth	Mean baseline BMI (SD)	Mean illness duration (SD)	Participant race and ethnicity	Participant socioeconomic status
Agras et al. (2014)	North America	15.3 (1.4)	88.2% female	Mean percentage of ideal body weight: 81.9%	13.6 months (13.5)	79.4% of participants White, 10.1% Hispanic, 5.1% Asian	NR
Bell and Mitchell (2004)	Australia	CBT group: 18.45 (2.57) IPT group: 17.58 (3.07)	100% female	CBT group: 16.06 (1.58) IPT group: 16.43 (1.85)	NR	NR	NR
Brockmeyer et al. (2021)	Germany	CBT group: 26.56 (9.06) ART group: 29.96 (9.26)	CBT group: 97.4% female ART group: 97.4% female	CBT group: 14.90 (1.59) ART group: 14.86 (2.12)	CBT group: 50% ≤ 5 years ART group: 35.9% ≤ 5 years	NR	NR
Byrne et al. (2017)	Australia	26.19 (6.47)	95.83% female	14.79 (1.22)	4 years	NR	NR
Channon et al. (1989)	United Kingdom	CBT group: 21.43 (5.86) BT group: 24.13 (5.77) Control group: 25.79 (7.19)	100% female	CBT group: 14.85 (1.10) BT group: 16.08 (1.42) Control group: 14.80 (1.44)	CBT group: 5.13 years (4.69) BT group: 8.13 (1.79) Control group: 7.25 (6.09)	NR	NR
Dalle Grave et al. (2013)	Italy	22.4 (6.9)	97.2% female	78.8% participants with BMI ≤ 16.0	5 years	NR	Reports occupational status
Dare et al. (2001)	United Kingdom	26.3 (6.7)	98% female	13.4 (1.4)	6.3 years (5.9)	NR	NR
Eisler et al. (2000)	United Kingdom	15.5 (1.6)	99% female	Mean percentage of average body weight 74.36% (9.8)	12.9 months (16.4)	NR	NR
Espie et al. (2014)	United Kingdom	FT group: 15.7 (1.8) MFT group: 15.7 (1.7)	FT group: 93.8% female MFT group: 88.2% female	FT group: 15.8 (1.5) MFT group: 15.7 (1.3)	FT group: 11.4 months (12.4) MFT group: 9.6 months (10.3)	FT group: 92.3% White, 3.7% other, 3.7% missing MFT group: 88.2% White, 5.9% other, 5.9% missing	NR
Fisher et al. (2013)	Germany	IBT group: 23.8 (4.5) TAU group: 24.1 (5.4)	100% female	IBT group: 17.6 (1.4) TAU group: 17.7 (1.3)	NR	NR	NR
Gelin et al. (2009)	Canada	FT group: 14.3 (1.5) FGP group: 14.9 (1.7)	100% female	Mean percentage of ideal body weight in FT group: 77.7% FGP group: 77.2%	NR	NR	NR
Gowers et al. (2015)	France	16.6 (1.4)	100% female	16.9 (1.1)	16.6 months (6.8)	NR	NR
Le Grange et al. (2016)	Australia	15.5 (1.5)	87.7% female	16.5 (1.5)	10.5 months (8.6)	NR	Reports 76 of parents with university degree
Lock et al. (2005)	North America	Long-term FT: 15.3 (1.7) Short-term FT: 15.3 (1.6)	Long-term FT: 91% female	17.1 (1.4)	Long-term FT: 13.0 months (9.5)	Long-term FT: 5% Asian, 76% White, 14% Hispanic, 5% Other.	Reports family income and

www.interscience.wiley.com/journal/eat by Applewell on [redacted] See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 (Continued)

Paper	Location of study	Participant mean age in years (SD)	Participant sex assigned at birth	Mean baseline BMI (SD)	Mean illness duration (SD)	Participant race and ethnicity	Participant socioeconomic status
			Short-term FT: 89% female		Short-term FT: 11.3 months (10.4)	Short-term FT: 14% Asian, 73% White, 9% Hispanic, 2% Native American, 2% Other	parental education
Lock et al. (2010)	North America	14.4 (1.6)	91% female	16.1 (1.1	11.8 months (8.6)	24% ethnic minorities	NR
McIntosh et al. (2005)	New Zealand	NR	100% female	17.3 (1.1)	NR	NR	NR
Nyman-Carlsson et al. (2020)	Sweden	CBT: 19.1 (1.9) FT: 18.7 (2.0)	100% female	CBT: 16.49 (.8) FT: 16.54 (.9)	CBT: 31.6 months (24.1) FT: 26.8 months (24.4)	NR	Reports education level
Perling et al. (2016)	Sweden	ACT: 26.6 (8.1) TAU: 24.6 (5.8)	97.67% female	ACT: 17.5 (2.3) TAU: 18.1 (2.6)	NR	NR	Reports educational and occupational status
Robin et al. (1999)	North America	BFST: 14.9 EDIT: 13.4	100% female	BFST: 86.5 pounds EDIT: 86.8 pounds	<12 months	94.5% White, 5.5% Middle Eastern	Reports socioeconomic status on Hollingshead index
Schmidt et al. (2015)	United Kingdom	26.7 (7.7)	97.89% female	26.6 (1.2)	8.3 years (7.3)	NR	Reports years in education
Serfaty et al. (1999)	United Kingdom	CT: 22.1 (6.6) Control: 17.9 (5.6)	94.29% female	CT: 16.2 (1.6) Control: 17.0 (4.0)	CT: 5.0 years (5.5) Control: 2.2 (4.5)	NR	Reports social class
Tosyz et al. (2013)	Australia/United Kingdom	33.4 (9.6)	100% female	16.2 (1.3)	16.6 (8.5)	NR	Reports educational and occupational status
Zipfel et al. (2014)	Germany	FPT: 28.0 (8.6) CBT: 27.4 (7.9) TAU: 27.7 (8.1)	100% female	FPT: 16.57 (1.0) CBT: 16.82 (1.0) TAU: 16.79 (1.0)	FPT: 61% ≤ 6 years CBT: 61% ≤ 6 years TAU: 61% ≤ 6 years	NR	NR

Abbreviations: ACT, acceptance and commitment therapy; AFT, adolescent focused individual therapy; ART, arts therapy; BFST, behavioral family systems therapy; BFT, behavioral family therapy; BT, behavioral treatment; CAT, cognitive analytic therapy; CBT, cognitive behavioral therapy or cognitive behavioral treatment; CBT-E, enhanced cognitive behavioral therapy; CBT-Eb, broad enhanced cognitive behavioral therapy; CBT-Ef, focused enhanced cognitive behavioral therapy; CFT, conjoint family therapy; CRT, cognitive remediation therapy; CT, cognitive therapy; D, dietary advice; EDIT, ego orientated individual therapy; FBT, family-based therapy or family-based treatment; FP, family group psychoeducation; FPP, focal psychoanalytic psychotherapy; FPT, focal psychodynamic therapy; FT, family therapy or combined individual/family therapy; IP, interpersonal psychotherapy; MANTRA, Maudsley model anorexia nervosa treatment for adults; MFT, multifamily therapy; NR, not reported; NSCM, non-specific supportive clinical management; PFT, parent focused treatment; RP, relapse prevention; SFT, separated family therapy or single family therapy; SSCM, specialist supportive clinical management; SyFT, systemic family therapy; TAU, treatment as usual.

monitor patient safety or psychiatric risks (e.g., if scoring above a certain threshold for depression triggered further investigation).

The majority of papers (*n* = 20, 86.96%) included some information related to monitoring participant safety; however, the amount of detail provided varied widely. Most of this information related specifically to monitoring physical risks; only one paper (4.35%) referred to the management of co-occurring psychiatric conditions (Le Grange et al., 2016). Information on safety monitoring was typically found in the methods sections of papers, with five papers (21.74%) including a specific subsection for this purpose (Agras et al., 2014; Le Grange et al., 2016; Lock et al., 2010; Schmidt et al., 2015; Touyz et al., 2013).

Three papers (13.04%) included no information relating to safety monitoring (Dalle Grave et al., 2013; Ekeler et al., 2016; McIntosh et al., 2005). Five papers (21.74%) included a statement that their participants received regular medical monitoring but no further details (Brockmeyer et al., 2021; Channon et al., 1989; Dare et al., 2001; Godart et al., 2012; Perling et al., 2016). Perling et al. (2016) additionally specified "we did not systematically ask for adverse events or harms." Four papers (17.39%) included information around procedure if an adverse medical event occurred, typically hospitalization in the case of medical instability and subsequent trial withdrawal (Ball & Mitchell, 2004; Fichter et al., 2012; Nyman-Carlsson et al., 2020; Serfaty et al., 1999). Nine papers (39.13%) included a combination of detailing medical monitoring procedures around hospitalization (including medical criteria for admission), and subsequent impact on trial involvement (Byrne et al., 2017; Ekeler et al., 2000; Geist et al., 2000; Le Grange et al., 2016; Lock et al., 2005, 2010; Robin et al., 1999; Schmidt et al., 2015; Touyz et al., 2013).

Agras et al. (2014) included the most comprehensive description of safety monitoring located in a dedicated "Participant Safety and Adverse Event" subsection. This paper, and a previous paper detailing the trial's design, described medical monitoring, including frequency and specific medical tests, and provided detailed guidelines followed for medical hospitalization, including specific medical criteria and procedures for trial withdrawal (Lock et al., 2012). In addition, the paper specifies the procedure for collecting and reporting data on adverse events, including describing the establishment of a dedicated board for safety monitoring. Zipfel et al. (2014) also included a higher level of detail compared to other papers around the monitoring, documentation, and reporting of adverse/serious adverse events, including an independent safety and data monitoring board, and described medical criteria, regular monitoring, and procedures around hospitalization.

3.3 | Treatment non-completion

Whilst almost every trial (except Geist et al., 2000) reported treatment non-completion, not all trials differentiated between non-completion due to patient decision to drop out and non-completion due to clinician decision to withdraw, and not all trials detailed further reasons for these decisions. One trial's reporting around non-completion was unclear and has not been included in the final analysis; Robin et al. (1999) reported all participants completed the full trial including follow-up, but an earlier paper reporting on a smaller version of the final sample stated that some participants did not complete treatment (Robin et al., 1994).

In the following summary, the term "drop-out" is used to refer to discontinuation chosen by patient/family, and "withdrawal" refers to discontinuation chosen by trial clinicians (e.g., due to hospitalization). "Non-completion" is used where reasons for discontinuation are unclear. Where papers reported information using the term "drop-out" it has been assumed this refers to patient discontinuation unless otherwise specified. Non-completion rates are summarized in Table 1. In all papers, withdrawal from treatment almost always related to hospitalization or high medical risk.

Five trials (21.74%) reported non-completion rates with no further details, including not specifying if non-completion was due to patient drop-out or withdrawal by trial clinicians (Le Grange et al., 2016; Perling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013; Zipfel et al., 2014). Two trials (8.70%) reported non-completion rates and specified that a proportion of these were due to withdrawal relating to hospitalization but did not detail reasons for other non-completers (Ball & Mitchell, 2004; McIntosh et al., 2005).

Most commonly trials reported drop-out rates only (*n* = 10, 43.48%). These trials did not report withdrawal due to hospitalization, possibly because this did not occur during these studies. The majority (*n* = 7, 30.43%) of these studies reported drop-out rates without detailing reasons (Channon et al., 1989; Dare et al., 2001; Ekeler et al., 2000, 2016; Fichter et al., 2012; Lock et al., 2010; Nyman-Carlsson et al., 2020). Brockmeyer et al. (2021) and Lock et al. (2005) detailed reasons for patient drop-out; Serfaty et al. (1999) gave reasons for one treatment group only.

Two studies (8.70%) reported total non-completion and differentiated between drop-out and withdrawal, documenting reasons for withdrawal but not for drop-out (Dalle Grave et al., 2013; Godart et al., 2012). Two studies reported total non-completion and differentiated between drop-out and withdrawal and gave reasons for both withdrawal and patient drop-out (Agras et al., 2014; Byrne et al., 2017).

Regarding further analyses of causality beyond reasons for discontinuation, most (82.61%) papers reported whether non-completion rates differed across groups (except for Godart et al., 2012; Lock et al., 2005; Schmidt et al., 2015; Serfaty et al., 1999). Three papers (13.04%) additionally analyzed whether non-completion was related to patient characteristics, such as baseline body weight (Ball & Mitchell, 2004; Fichter et al., 2012; McIntosh et al., 2005).

3.4 | Treatment non-response

No trial performed a clinical significance analysis to measure treatment non-response (Jacobson et al., 1984), rather, papers reported proportions of patients meeting a pre-specified categorical outcome (summarized in Table 1).

Five trials (21.74%) did not report categorical outcomes (Brockmeyer et al., 2021; Channon et al., 1989; Fichter et al., 2012; Geist et al., 2000; Touyz et al., 2013). A number of trials (*n* = 7, 30.43%) reported the proportion of patients not achieving remission,

or a similar pre-defined good outcome (Agras et al., 2014; Byrne et al., 2017; Dalle Grave et al., 2013; Le Grange et al., 2016; Lock et al., 2005, 2010; Parling et al., 2016). This reflects participants not improving to a certain threshold, rather than a total lack of improvement, and therefore it is unclear how many of these patients were actual non-responders. One of these trials (Lock et al., 2010) additionally reported the number of patients who originally remitted then relapsed during the follow-up period. All these trials used varying criteria for a good outcome/remission: not all included both body weight and symptom thresholds, and different body weight thresholds were utilised.

Nine trials (29.13%) reported patients not responding to treatment, typically operationalised as experiencing a poor outcome or continuing to meet diagnostic criteria for AN. Three of these trials (13.04%) used the Morgan and Russell criteria to report poor outcomes (Ball & Mitchell, 2004; Elder et al., 2000, 2016; Morgan & Russell, 1975), the remainder used outcome descriptors based around diagnostic criteria (Dare et al., 2001; McIntosh et al., 2005; Nyman-Carlsson et al., 2020; Schmidt et al., 2015; Serfaty et al., 1999; Zipfel et al., 2014).

All trials analysed whether ED outcomes differed across treatment groups.

3.5 | Symptom deterioration

This synthesis followed Linden's (2013) definition of "deterioration of illness" as "worsening of illness during therapy or at any other time in the course of illness." It should be noted that this definition does not capture trial follow-up periods—where reported, this information has been included. Additionally, deterioration relating to co-occurring illnesses other than AN was also included.

No trials reported the occurrence of patients deteriorating on measures of AN symptoms other than weight loss, or measures of other psychiatric pathology. Trials most commonly reported symptom deterioration as relating to a worsening of physical health resulting in hospitalization, both during intervention and follow-up periods. Almost all trials (except Brockmeyer et al., 2021 and Serfaty et al., 1999) reported proportions of patients requiring hospitalisation (summarised in Table 1). Where reasons for hospitalisation were given, the majority were related to medical complications such as severe weight loss. A smaller number of trials ($n = 7$, 30.43%) reported hospitalisations, or accessing additional treatment, relating to psychiatric risks such as suicidality, self-harm, or the occurrence of another psychiatric condition (Agras et al., 2014; Elder et al., 2000, 2016; Godart et al., 2012; Lock et al., 2010; Parling et al., 2016; Schmidt et al., 2015). Finally, one trial reported that a mean weight loss occurred during the study follow-up period in the TAU group (Fichter et al., 2012).

In addition to reporting reasons for hospitalisation, the majority of trials ($n = 12$, 52.17%) reported whether hospitalisation rates differed across groups (Agras et al., 2014; Byrne et al., 2017; Channon et al., 1989; Godart et al., 2012; Le Grange et al., 2016; Lock et al., 2005, 2010; McIntosh et al., 2005; Parling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013; Zipfel et al., 2014).

3.6 | Other unwanted events

Papers were screened for other types of reporting that reflected unwanted events, as judged by the study authors.

Eleven papers (47.83%) reported proportions of participants requiring additional treatment either during the intervention or follow-up period (Agras et al., 2014; Ball & Mitchell, 2004; Brockmeyer et al., 2021; Byrne et al., 2017; Dalle Grave et al., 2013; Elder et al., 2000; Le Grange et al., 2016; Lock et al., 2010; Parling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013). Only three of these studies (13.04%) included whether additional treatment utilisation differed across groups.

Nine papers (39.13%) reported the occurrence of patient deaths during the trial period. Four papers (17.39%) specified no deaths occurred (Agras et al., 2014; Byrne et al., 2017; Fichter et al., 2012; Zipfel et al., 2014), five papers (21.74%) reported one patient death (see Table 1) (Dare et al., 2001; McIntosh et al., 2005; Parling et al., 2016; Schmidt et al., 2015; Touyz et al., 2013).

Four papers (17.39%) reported the emergence of new ED or psychiatric symptoms during the study. Two papers (8.70%) documented that a proportion of patients developed additional bulimic symptoms (Elder et al., 2000, 2016). One paper (4.35%) reported the proportions of patients at EOT/follow-up who now met criteria for bulimia nervosa (BN) and eating disorder not otherwise specified (EDNOS) (Nyman-Carlsson et al., 2020). One paper documented the emergence of severe OCD and depression in some patients following weight restoration (Rosin et al., 1999).

One paper reported the proportion of participants who reported being dissatisfied with their treatment (Elder et al., 2016).

Finally, a small number of papers ($n = 3$, 13.04%) included statements on the occurrence of adverse events during the trial. Agras et al. (2014) was the only paper to include an "Adverse Event" subheading in its result section, where it documents that there were no significant differences between treatment groups on numbers of serious adverse events (and specifies types of serious adverse events that occurred). Fichter et al. (2012) and Zipfel et al. (2014) both included statements that no serious adverse events occurred relating to the trial.

4 | DISCUSSION

The findings of this review suggest that whilst certain key unwanted events are typically reported in AN psychotherapy RCTs (notably non-completion, hospitalisation rates, and a lack of improvement), there is significant variation in definitions which makes it difficult to compare outcomes across trials. This was particularly notable when comparing treatment non-response, or non-improvement across trials, with different papers using varying thresholds and outcomes (typically weight

21982394, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/eat.23778 by Test, Wiley Online Library on [06/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

only, or a combination of weight and symptom measurements) to clarify a lack of improvement. The need for a consistent definition of recovery from AN in order to enable the comparison of categorical outcomes has previously been highlighted in the research literature, with suggestions that these should include both measures of weight and symptoms (Bardone-Cone et al., 2010; Couturier & Lock, 2006; Kinkes et al., 2017). Achieving consensus around the definitions of key unwanted events, such as a lack of clinically significant improvement or the differentiation between non-completion, withdrawal, and drop-out, would be important in allowing the comparison of these outcomes across RCTs.

The second key finding of this review is that it is very difficult from the information reported in trial papers to understand whether the events described represented unwanted events, or adverse events inferred as related to the intervention. The reporting, for example, of hospitalization rates without other context potentially reflects influences from the medical field where events like mortality and hospitalization are reported as adverse events as standard, but does not meet the definitional requirement of adverse events where there must be an inferred causal link to the intervention itself (Duggan et al., 2014; Linden, 2013; Perry et al., 2016). This style of reporting may lack applicability to a condition like AN where patients may have medical vulnerabilities prior to starting the intervention; some RCTs may target higher risk groups of patients (such as those with lower BMIs or experiencing inpatient admissions) where high rates of hospitalization reflect baseline illness severity rather than a consequence of the intervention. Similarly, unwanted events may be influenced by contextual factors outside of the intervention—for example, approaches to hospitalization for AN can vary across countries, with studies based in the United States and Australia reporting higher hospital admission rates (Eisler et al., 2016). Whilst almost all papers compared key unwanted events across intervention groups (e.g., comparing patient drop-out rates or hospitalization), beyond group comparisons there was typically a lack of reflection on the causality or context of unwanted events as recommended in existing reporting standards (Ioannidis et al., 2004; Linden, 2013).

The findings of this review therefore raise the question of whether the reporting of statistics such as hospitalization rates could be improved with added context around causality. The present typical style of reporting an uncontextualised hospitalization statistic does give clinicians and other researchers insight into how common hospitalization rates are during the intervention, which might be helpful in informing expectations when implementing therapy, but does not illuminate whether there are risks around the intervention causing or not effectively slowing deterioration. Where causality was considered in relation to unwanted events in this review, these typically reflected analyses of whether outcomes were related to patient factors, such as McIntosh et al. (2005) finding that treatment non-completion was related to lower starting body weight. Notable exceptions include Rosin et al.'s (1999) consideration of whether the emergence of severe OCD and depression symptoms in a small number of patients may have occurred as a consequence of successful weight restoration following the intervention.

As well as highlighting difficulties in reporting and interpreting unwanted events in general when treating people with AN, this review also emphasizes that understanding what counts as an unwanted versus an adverse event may depend on the nature and goals of the specific psychotherapeutic intervention. For example, not all interventions included in this review aimed for recovery as an outcome; indeed, in their RCT targeting severe and enduring AN, Touyz et al. (2013) highlight that targeting recovery may be misaligned with the goals of patients in this population, and instead focused on quality of life, mood disorder symptoms and social adjustment as primary outcomes, with ED symptoms as secondary outcomes. If a standardized unwanted event checklist was applied to this trial, such as that developed by Linden (2013), this would count as treatment non-response and reflect an unwanted event; but this would not represent a helpful interpretation of this trial's goals and outcomes. This raises further questions around the utility of standardized unwanted event checklists in this field.

Therefore, the findings of this review raise questions around how unwanted events should be measured, defined, and reported within the field of psychotherapeutic interventions for AN, and how this information can then be used to inform understandings of iatrogenic harm. It does not appear that current reporting norms enable the reader to distinguish between unwanted events and negative effects. Furthermore, it was also noted that many of the unwanted events reported by trials represented a medical understanding of this concept, such as reporting hospitalization, mortality, or treatment non-response as measured by disorder-specific symptoms. Only one paper measured patient-experienced harms in the form of patient and family feedback (Eisler et al., 2016). No paper examined unwanted events in social or work domains as recommended by Linden (2013), or broader experiences of distress or negative affect.

Therefore, the reporting norms observed in this review seem somewhat disconnected from the reality of AN as a condition, and the nature of psychotherapeutic interventions, and so limit any extrapolation into understanding how far these unwanted events translate into negative effects. In the wider field of adverse event research in psychotherapy, it has been proposed that what is considered an adverse event should be defined a priori by investigators with consideration to the specific condition, population, and intervention being utilized, rather than defined in relation to pre-existing checklists (Duggan et al., 2014; Hardgin et al., 2010). However, AN is also a condition associated with significant medical risks, and therefore there may be benefits from also retaining some aspects of the medicalized approach noted in this paper. It is possible that an individualized approach where investigators determine trial-specific unwanted/ adverse events a priori could be combined with a core reporting framework of standards specific to AN trials, drawing on medical research standards. This would retain the flexibility and specificity of Duggan et al.'s (2014) proposed approach, whilst simultaneously ensuring safety, best practice, and comparability.

This is an area that clearly requires more thought and discussion. However, building on current practice and research in parallel fields, it is tentatively proposed that papers could benefit from including the following core information as a minimum. It is emphasized that this is

not intended to represent a generic checklist, but rather a framework for information that should be considered and included:

- Authors should consider standards recommended by the CONSORT harms checklist (Ioannidis et al., 2004). The below points may be particularly relevant to the ED field:
  - A list of addressed adverse events within Section 2, including details on monitoring and procedures. This could include criteria for hospitalization, pausing treatment, and withdrawal from trial, and whether regular monitoring of physical and mental health risks occurred in addition to the standard assessment schedule.
  - Reporting treatment non-completion rates across groups, differentiating between participants who dropped out (by non-mutual agreement), and proportions of participants withdrawn from the study by investigators. Reporting reasons for non-completion, including where related to the intervention (e.g., patient dissatisfaction).
  - Reporting statistical comparisons of serious adverse event incidence rates across groups.
- A dedicated patient safety section within Section 2:
  - How the trial defined unwanted and adverse events, including information on how investigators inferred causality (as recommended by Duggan et al., 2014).
- A dedicated adverse event reporting section under results:
  - Reporting proportions experiencing treatment non-response/symptom deterioration with reference to suggested standardized models of recovery and remission (Khalsa et al., 2017). Where symptom deterioration occurs, considering evidence for a causal relationship to the intervention.
  - Where hospitalization rates and mortality rates are reported (as is currently typical in AN RCTs), contextualizing this where possible with reasonable inference around whether this represents an adverse event relating to the intervention.

In the context of difficulties conceiving unwanted events within AN psychotherapeutic interventions, it is the position of the authors that whilst the above information should be reported, it should also be contextualized with a subsection within the discussion dedicated to reflecting on the context of unwanted events within the study, including in relation to expected harms as a result of treatment and change processes, and any potential negative effects.

Finally, it was noted during data synthesis that a majority of studies did not report data on participant ethnicity, and that the majority of individuals included in these studies were White females. The accurate recording and reporting of patient ethnicity in relation to unwanted events is important as evidence suggests that ethnic minority patients are at higher risk of adverse health outcomes, including adverse events and medical errors (Chauhan et al., 2020). Future research on unwanted events in ED treatment could benefit from exploring the potential influence of demographic variables such as race, gender, and socioeconomic status.

This review has several limitations. Firstly, by focusing on published RCT papers and excluding secondary analyses this review was unable to include papers where authors focused on negative outcomes as part of a secondary analysis: for example, whilst Lock et al. (2005) reported drop-out and remission rates in their primary analysis, the group published a later paper analyzing patient factors predicting drop-out and remission in more detail (Lock et al., 2006). Similarly, a secondary analysis of Agras et al. (2014) performed a more detailed analysis of hospitalization rates, including relation to the interventions and patient factors (Lock et al., 2016). This was a deliberate choice by the authors as the intention was to explore what kinds of information on unwanted events are presented alongside the key findings of RCTs, rather than information published at a later stage. Furthermore, this review did not include unpublished papers as its intention was to explore reporting standards in final versions of published RCTs. Therefore, this review knowingly reflects a certain degree of publication bias: the information reflected in papers, for example, might reflect details requested by journals rather than those intended for inclusion by authors. Related to this point, this review is dependent on information reported in trial papers: for example, just because an outcome paper does not report the presence of safety monitoring does not mean that it did not occur in practice.

Secondly, it should be noted that by evaluating unwanted events as reported in RCT papers this review inevitably focuses on unwanted events deemed significant by investigators, rather than patients themselves. A benefit of future RCTs using idiosyncratic patient ratings and feedback is that they could better capture factors deemed significant by patients but not typically recorded as part of trial outcomes, such as isolation relating to inpatient admission or trauma associated with refeeding (Offord et al., 2006; Treasure et al., 2011).

In conclusion, this review highlights that at present the reporting of unwanted event monitoring and occurrence varies significantly across psychotherapy RCTs for AN in a manner that hinders interpretation and comparison. It also raises questions around how unwanted events should be conceived and defined in this field, including around whether an approach to measuring adverse events influenced by a medical understanding of this concept is an effective fit for this area. It suggests that this area could benefit from further discussion around balancing objective and comparable standards with an individualized approach that ensures that unwanted/adverse event measurement and reporting best captures the complex context of AN psychotherapy intervention studies.

AUTHOR CONTRIBUTIONS

Emma Kinnaird: Conceptualization; formal analysis; investigation; project administration; writing – original draft; writing – review and editing. **Meghan Nicholson:** Formal analysis; writing – original draft. **Amelia Thomas:** Formal analysis; writing – original draft. **Myra Cooper:** Conceptualization; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Amy Spray for her assistance with the abstract screening stage of this review.

CONFLICT OF INTEREST

The authors have no funding or conflicts of interest to declare.

not intended to represent a generic checklist, but rather a framework for information that should be considered and included:

- Authors should consider standards recommended by the CONSORT harms checklist (Ioannidis et al., 2004). The below points may be particularly relevant to the ED field:
  - A list of addressed adverse events within Section 2, including details on monitoring and procedures. This could include criteria for hospitalization, pausing treatment, and withdrawal from trial, and whether regular monitoring of physical and mental health risks occurred in addition to the standard assessment schedule.
  - Reporting treatment non-completion rates across groups, differentiating between participants who dropped out (by non-mutual agreement), and proportions of participants withdrawn from the study by investigators. Reporting reasons for non-completion, including where related to the intervention (e.g., patient dissatisfaction).
  - Reporting statistical comparisons of serious adverse event incidence rates across groups.
- A dedicated patient safety section within Section 2:
  - How the trial defined unwanted and adverse events, including information on how investigators inferred causality (as recommended by Duggan et al., 2014).
- A dedicated adverse event reporting section under results:
  - Reporting proportions experiencing treatment non-response/symptom deterioration with reference to suggested standardized models of recovery and remission (Khalsa et al., 2017). Where symptom deterioration occurs, considering evidence for a causal relationship to the intervention.
  - Where hospitalization rates and mortality rates are reported (as is currently typical in AN RCTs), contextualizing this where possible with reasonable inference around whether this represents an adverse event relating to the intervention.

In the context of difficulties conceiving unwanted events within AN psychotherapeutic interventions, it is the position of the authors that whilst the above information should be reported, it should also be contextualized with a subsection within the discussion dedicated to reflecting on the context of unwanted events within the study, including in relation to expected harms as a result of treatment and change processes, and any potential negative effects.

Finally, it was noted during data synthesis that a majority of studies did not report data on participant ethnicity, and that the majority of individuals included in these studies were White females. The accurate recording and reporting of patient ethnicity in relation to unwanted events is important as evidence suggests that ethnic minority patients are at higher risk of adverse health outcomes, including adverse events and medical errors (Chauhan et al., 2020). Future research on unwanted events in ED treatment could benefit from exploring the potential influence of demographic variables such as race, gender, and socioeconomic status.

This review has several limitations. Firstly, by focusing on published RCT papers and excluding secondary analyses this review was unable to include papers where authors focused on negative outcomes as part of a secondary analysis: for example, whilst Lock et al. (2005) reported drop-out and remission rates in their primary analysis, the group published a later paper analysing patient factors predicting drop-out and remission in more detail (Lock et al., 2006). Similarly, a secondary analysis of Agras et al. (2014) performed a more detailed analysis of hospitalization rates, including relation to the interventions and patient factors (Lock et al., 2016). This was a deliberate choice by the authors as the intention was to explore what kinds of information on unwanted events are presented alongside the key findings of RCTs, rather than information published at a later stage. Furthermore, this review did not include unpublished papers as its intention was to explore reporting standards in final versions of published RCTs. Therefore, this review knowingly reflects a certain degree of publication bias: the information reflected in papers, for example, might reflect details requested by journals rather than those intended for inclusion by authors. Related to this point, this review is dependent on information reported in trial papers: for example, just because an outcome paper does not report the presence of safety monitoring does not mean that it did not occur in practice.

Secondly, it should be noted that by evaluating unwanted events as reported in RCT papers this review inevitably focuses on unwanted events deemed significant by investigators, rather than patients themselves. A benefit of future RCTs using idiosyncratic patient ratings and feedback is that they could better capture factors deemed significant by patients but not typically recorded as part of trial outcomes, such as isolation relating to inpatient admission or trauma associated with refeeding (Clifford et al., 2006; Treasure et al., 2011).

In conclusion, this review highlights that at present the reporting of unwanted event monitoring and occurrence varies significantly across psychotherapy RCTs for AN in a manner that hinders interpretation and comparison. It also raises questions around how unwanted events should be conceived and defined in this field, including around whether an approach to measuring adverse events influenced by a medical understanding of this concept is an effective fit for this area. It suggests that this area could benefit from further discussion around balancing objective and comparable standards with an individualized approach that ensures that unwanted/adverse event measurement and reporting best captures the complex context of AN psychotherapy intervention studies.

AUTHOR CONTRIBUTIONS

Emma Kinnaird: Conceptualization; formal analysis; investigation; project administration; writing – original draft; writing – review and editing. **Meghan Nicholson:** Formal analysis; writing – original draft. **Amelia Thomas:** Formal analysis; writing – original draft. **Myra Cooper:** Conceptualization; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Amy Spray for her assistance with the abstract screening stage of this review.

CONFLICT OF INTEREST

The authors have no funding or conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Emma Kinnaird https://orcid.org/0000-0002-6831-6386
Myra Cooper https://orcid.org/0000-0002-2676-4354

REFERENCES

Agras, W., Lock, J., Brandt, H., Bryson, S. W., Dodge, E., Halmi, K. A., Jo, B., Johnson, C., Kaye, W., Wilfley, D., & Woodside, B. (2014). Comparison of 2 family therapies for adolescent anorexia nervosa: A randomized parallel trial. *JAMA Psychiatry*, *71*(11), 1279–1286.

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders. A meta-analysis of 36 studies. *Archives of General Psychiatry*, *68*(7), 724–731. https://doi.org/10.1001/archgenpsychiatry.2011.74

Ball, J., & Mitchell, P. (2004). A randomized controlled study of cognitive behavior therapy and behavioral family therapy for anorexia nervosa patients. *Eating Disorders*, *12*(4), 303–314. https://doi.org/10.1080/10640260490521389

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. https://doi.org/10.1016/j.brat.2009.11.001

Bergin, A. E. (1966). Some implications of psychotherapy research for therapeutic practice. *Journal of Abnormal Psychology*, *71*(4), 235–246. https://doi.org/10.1037/h0023577

Birmingham, C. L., Su, J., Hlynsky, J. A., Goldner, E. M., & Gao, M. (2005). The mortality rate from anorexia nervosa. *International Journal of Eating Disorders*, *38*(2), 143–146. https://doi.org/10.1002/eat.20164

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., & Ravaud, P. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: Explanation and elaboration. *Annals of Internal Medicine*, *148*(4), 295–309. https://doi.org/10.7326/0003-4819-148-4-200802190-00008

Brockmeyer, T., Friederich, H. C., & Schmidt, U. (2018). Advances in the treatment of anorexia nervosa: A review of established and emerging interventions. *Psychological Medicine*, *48*(8), 1228–1256. https://doi.org/10.1017/s0033291717003604

Brockmeyer, T., Schmidt, H., Leferink-Bausch, A., Zimmermann, J., Wünsch-Leiteritz, W., Leiteritz, A., & Friederich, H.-C. (2021). Cognitive remediation therapy in anorexia nervosa—A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *89*(10), 805–815. https://doi.org/10.1037/ccp0000675

Bulik, C. M., Berkman, N. D., Brownley, K. A., Sedway, J. A., & Lohr, K. N. (2007). Anorexia nervosa treatment: A systematic review of randomized controlled trials. *International Journal of Eating Disorders*, *40*(4), 310–320. https://doi.org/10.1002/eat.20367

Byrne, S., Wade, T., Hay, P., Touyz, S., Fairburn, C., Treasure, J., Schmidt, U., McIntosh, V., Allen, K., Fursland, A., & Crosby, R. (2017). A randomised controlled trial of three psychological treatments for anorexia nervosa. *Psychological Medicine*, *47*(14), 2522–2533.

Channon, S., de Silva, P., Hemsley, D., & Perkins, R. E. (1989). A controlled trial of cognitive-behavioural and behavioural treatment of anorexia nervosa. *Behaviour Research and Therapy*, *27*(5), 529–535.

Chouhan, A., Walton, M., Manias, E., Walpola, R. L., Seale, H., Lalanli, M., Leane, D., Moore, S. & Harrison, R. (2020). The safety of health care for ethnic minority patients: a systematic review. *International Journal for Equity in Health*, *19*(1). https://doi.org/10.1186/s12939-020-01223-2

Couturier, J., & Lock, J. (2006). What is recovery in adolescent anorexia nervosa? *International Journal of Eating Disorders*, *39*(7), 550–555. https://doi.org/10.1002/eat.20309

Crawford, M. J., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., Clarke, J., & Parry, G. D. (2016). Patient experience of negative effects of psychological treatment: Results of a national survey. *British Journal of Psychiatry*, *208*(3), 260–265. https://doi.org/10.1192/bjp.bp.114.162628

Dalle Grave, R., Calugi, S., Conti, M., Doll, H., & Fairburn, C. G. (2013). Inpatient cognitive behaviour therapy for anorexia nervosa: A randomized controlled trial. *Psychotherapy and Psychosomatics*, *82*(6), 390–398. https://doi.org/10.1159/000350058

Dare, C., Eisler, I., Russell, G., Treasure, J., & Dodge, L. (2001). Psychological therapies for adults with anorexia nervosa. Randomised controlled trial of out-patient treatments. *British Journal of Psychiatry*, *178*, 216–221. https://doi.org/10.1192/bjp.178.3.216

DeJong, H., Broadbent, H., & Schmidt, U. (2012). A systematic review of dropout from treatment in outpatients with anorexia nervosa. *International Journal of Eating Disorders*, *45*(5), 635–647. https://doi.org/10.1002/eat.20956

Duggan, C., Parry, G., McMurran, M., Davidson, K., & Dennis, J. (2014). The recording of adverse events from psychological treatments in clinical trials: Evidence from a review of NIHR-funded trials. *Trials*, *15*, 335. https://doi.org/10.1186/1745-6215-15-335

Eisler, I., Dare, C., Hodes, M., Russell, G., Dodge, E., & Le Grange, D. (2000). Family therapy for adolescent anorexia nervosa: The results of a controlled comparison of two family interventions. *Journal of Child Psychology and Psychiatry*, *41*(6), 727–736.

Eisler, I., Simic, M., Hodsoll, J., Asen, E., Berelowitz, M., Connan, F., Ellis, G., Hugo, P., Schmidt, U., Treasure, J., Yi, I., & Landau, S. (2016). A pragmatic randomised multi-centre trial of multifamily and single family therapy for adolescent anorexia nervosa. *BMC Psychiatry*, *16*(1), 422. https://doi.org/10.1186/s12888-016-1129-6

Fichter, M. M., Quadflieg, N., Nisslmüller, K., Lindner, S., Osen, B., Huber, T., & Wünsch-Leiteritz, W. (2012). Does Internet-based prevention reduce the risk of relapse for anorexia nervosa? *Behaviour Research and Therapy*, *50*(3), 180–190. https://doi.org/10.1016/j.brat.2011.12.003

Gelet, R., Hehrmann, M., Stephens, D., Davis, R. & Katzman, D. K. (2000). Comparison of family therapy and family group psychoeducation in adolescents with anorexia nervosa. *Canadian Journal of Psychiatry*, *45*(2), 173–178. https://doi.org/10.1177/070674370004500208

Godart, N., Berthoz, S., Curt, F., Perdereau, F., Rein, Z., Wallier, J., Horrard, A. S., Kaganski, I., Lucet, R., Atger, F., Cervas, M., Fermanian, J., Falissard, B., Flament, M., Elsler, I., & Jeammet, P. (2012). A randomized controlled trial of adjunctive family therapy and treatment as usual following inpatient treatment for anorexia nervosa adolescents. *PLoS One*, *7*(1), e28469. https://doi.org/10.1371/journal.pone.0028469

Gregertsen, E. C., Mandy, W., & Serpell, L. (2017). The Egosyntonic nature of anorexia: An impediment to recovery in anorexia nervosa treatment. *Frontiers in Psychology*, *8*, 2273. https://doi.org/10.3389/fpsyg.2017.02273

Hoffmann, T. C., Erueti, C., & Glasziou, P. P. (2013). Poor description of non-pharmacological interventions: Analysis of consecutive sample of randomised trials. *British Medical Journal*, *347*, f3755. https://doi.org/10.1136/bmj.f3755

Horigian, V. E., Robbins, M. S., Dominguez, R., Ucha, J. & Rosa, C. L. (2010). Principles for defining adverse events in behavioral intervention research: Lessons from a family-focused adolescent drug abuse trial. *Clinical Trials*, *7*(1), 59–68. https://doi.org/10.1177/1740774509356573

Hughes, E. K., Le Grange, D., Court, A., Yeo, M. S., Campbell, S., Allen, E., Crosby, R. D., Louk, K. L., & Sawyer, S. M. (2014). Parent-focused

treatment for adolescent anorexia nervosa: A study protocol of a randomized controlled trial. *BMC Psychiatry, 14*, 109. https://doi.org/10.1186/1471-244x-14-109

Ioannidis, J. P., Evans, S. J., Gøtzsche, P. C., O'Neill, R. T., Altman, D. G., Schulz, K., & Moher, D. (2004). Better reporting of harms in randomized trials: An extension of the CONSORT statement. *Annals of Internal Medicine, 141*(10), 781–788. https://doi.org/10.7326/0003-4819-141-10-200411160-00009

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7

Kinasz, K. S., Pontoff, L. C., McCurdy-McKinnon, D., & Fournier, 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/s40337-017-0145-8

La Grange, D., Eisler, I., Dare, C., & Russell, G. F. M. (1992). Evaluation of family treatments in adolescent anorexia nervosa: A pilot study. *International Journal of Eating Disorders, 12*(4), 347–357. https://doi.org/10.1002/1098-108x(199212)12:4<347::AID-EAT2260120402>3.0.CO;2-W

La Grange, D., Hughes, E. K., Court, A., Yao, M., Crosby, R. D., & Sawyer, S. M. (2016). Randomized clinical trial of parent-focused treatment and family-based treatment for adolescent anorexia nervosa. *Journal of the American Academy of Child and Adolescent Psychiatry, 55*(8), 683–692. https://doi.org/10.1016/j.jaac.2016.05.007

Linden, M. (2013). How to define, find and classify side effects in psychotherapy: From unwanted events to adverse treatment reactions. *Clinical Psychology & Psychotherapy, 20*(4), 286–296. https://doi.org/10.1002/cpp.1768

Lock, J., Agras, W., Bryson, S., & Kraemer, H. C. (2005). A comparison of short- and long-term family therapy for adolescent anorexia nervosa. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(7), 632–639.

Lock, J., Agras, W. S., Bryson, S. W., Brandt, H., Halmi, K. A., Kaye, W., Wilfley, D., Woodside, B., Pajerito, S., & Jo, B. (2016). Does family-based treatment reduce the need for hospitalization in adolescent anorexia nervosa? *The International Journal of Eating Disorders, 49*(9), 891–894. https://doi.org/10.1002/eat.22536

Lock, J., Brandt, H., Woodside, B., Agras, S., Halmi, W., Johnson, C., Kaye, W., & Wilfley, D. (2012). Challenges in conducting a multi-site randomized clinical trial comparing treatments for adolescent anorexia nervosa. *International Journal of Eating Disorders, 45*(2), 202–213.

Lock, J., Couturier, J., Bryson, S., & Agras, S. (2006). Predictors of dropout and remission in family therapy for adolescent anorexia nervosa in a randomized clinical trial. *International Journal of Eating Disorders, 39*(8), 639–647. https://doi.org/10.1002/eat.20328

Lock, J., & La Grange, D. (2001). Can family-based treatment of anorexia nervosa be manualized. *Journal of Psychotherapy Practice and Research, 10*(4), 253–261.

Lock, J., La Grange, D., Agras, W., Moye, A., Bryson, S. W., & Jo, B. (2010). Randomized clinical trial comparing family-based treatment with adolescent-focused individual therapy for adolescents with anorexia nervosa. *Archives of General Psychiatry, 67*(10), 1025–1032.

McIntosh, V. V., Jordan, J., Carter, F. A., Luty, S. E., McKenzie, J. M., Bulik, C. M., Frampton, C. M. A., & Joyce, P. R. (2005). Three psychotherapies for anorexia nervosa: A randomized, controlled trial. *The American Journal of Psychiatry, 162*(4), 741–747. https://doi.org/10.1176/appi.ajp.162.4.741

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ, 339*, b2535. https://doi.org/10.1136/bmj.b2535

Morgan, H. G., & Russell, G. F. M. (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/S0033291700058901

Murray, S. B., Quintana, D. S., Loeb, K. L., Griffiths, S., & Le Grange, D. (2019). Treatment outcomes for anorexia nervosa: A systematic review and meta-analysis of randomized controlled trials. *Psychological Medicine, 49*(4), 535–544. https://doi.org/10.1017/s0033291718002088

Nyman-Carlsson, E., Engstrand, A., Engström, I., Gustafsson, S. A., & Norring, C. (2019). Predictors of outcome among young adult patients with anorexia nervosa in a randomized controlled trial. *European Eating Disorders Review, 27*(1), 76–85. https://doi.org/10.1002/erv.2630

Nyman-Carlsson, E., Norring, C., Engström, I., Gustafsson, S. A., Lindberg, K., Paulson-Karlsson, G., & Nevonen, L. (2020). Individual cognitive behavioral therapy and combined family/individual therapy for young adults with anorexia nervosa: A randomized controlled trial. *Psychotherapy Research, 30*(8), 1011–1025. https://doi.org/10.1080/10503307.2019.1686190

Offord, A., Turner, H., & Cooper, M. (2006). Adolescent inpatient treatment for anorexia nervosa: A qualitative study exploring young adults' retrospective views of treatment and discharge. *European Eating Disorders Review, 14*(6), 377–387. https://doi.org/10.1002/erv.697

Poling, T., Cornwall, M., Rashkind, M., Holmgren, S., & Glandel, A. (2016). A randomized trial of acceptance and commitment therapy for anorexia nervosa after daycare treatment, including five-year follow-up. *BMC Psychiatry, 16*(1), 272. https://doi.org/10.1186/s12888-016-0975-6

Parry, G. D., Crawford, M. J., & Duggan, C. (2016). Iatrogenic harm from psychological therapies – Time to move on. *British Journal of Psychiatry, 209*(3), 233–238. https://doi.org/10.1192/bjp.bp.115.168518

Phillips, R., Hazell, L., Sauzet, O., & Cornelius, V. (2019). Analysis and reporting of adverse events in randomised controlled trials: A review. *BMJ Open, 9*(2), e024537. https://doi.org/10.1136/bmjopen-2018-024537

Robin, A. L., Siegel, P. T., Koepke, T., Moye, A. W., & Tice, S. (1994). Family therapy versus individual therapy for adolescent females with anorexia nervosa. *Journal of Developmental and Behavioral Pediatrics, 15*(2), 111–116. https://doi.org/10.1097/00004703-199404000-00008

Robin, A. L., Siegel, P. T., Moye, A. W., Gilroy, M., Dennis, A. B., & Sikand, A. (1999). A controlled comparison of family versus individual therapy for adolescents with anorexia nervosa. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*(12), 1482–1489. https://doi.org/10.1097/00004583-199912000-00008

Russell G. F. M. (1985). Anorexia nervosa and bulimia nervosa. In *Handbook of Psychiatry 4: The Neurose and Personality Disorders*. Edited by Russell G. F. M. and Hersov L. A. Cambridge University Press, Cambridge.

Schlegl, S., Diedrich, A., Neumayr, C., Fussl, M., Naab, S., & Voderholzer, U. (2016). Inpatient treatment for adolescents with anorexia nervosa: Clinical significance and predictors of treatment outcome. *European Eating Disorders Review, 24*(3), 214–222. https://doi.org/10.1002/erv.2416

Schlegl, S., Quadflieg, N., Löwe, B., Cuntz, U., & Voderholzer, U. (2014). Specialized inpatient treatment of adult anorexia nervosa: Effectiveness and clinical significance of changes. *BMC Psychiatry, 14*(1), 258. https://doi.org/10.1186/s12888-014-0258-z

Schmidt, U., Magill, N., Renwick, B., Keyes, A., Kenyon, M., Dejong, H., Lose, A., Broadbent, H., Loomes, R., Yasin, H., Watson, C., Ghaleni, S., Bonin, E. M., Serpell, L., Richards, L., Johnson-Sabine, E., Boughton, N., Whitehead, L., Beecham, J., … Landau, S. (2015). The Maudsley outpatient study of treatments for anorexia nervosa and related conditions (MOSAIC): Comparison of the Maudsley model of anorexia nervosa treatment for adults (MANTRA) with specialist supportive clinical management (SSCM) in outpatients with broadly defined anorexia nervosa: A randomized controlled trial. *Journal of*

Consulting and Clinical Psychology, *83*(4), 796–807. https://doi.org/10.1037/ccp0000019

Schmidt, U., Renwick, B., Lose, A., Kenyon, M., DeJong, H., Broadbent, H., Loomes, R., Watson, C., Ghelani, S., Serpell, L., Richards, L., Johnson-Sabine, E., Boughton, N., Whitehead, L., Beecham, J., Treasure, J., & Landau, S. (2013). The MOSAIC study – Comparison of the Maudsley model of treatment for adults with anorexia nervosa (MANTRA) with specialist supportive clinical management (SSCM) in outpatients with anorexia nervosa or eating disorder not otherwise specified, anorexia nervosa type: Study protocol for a randomized controlled trial. *Trials*, *14*(1), 160. https://doi.org/10.1186/1745-6215-14-160

Serfaty, M. A., Turkington, D., Heap, M., Ledsham, L., & Jolley, E. (1999). Cognitive therapy versus dietary counselling in the outpatient treatment of anorexia nervosa: Effects of the treatment phase. *European Eating Disorders Review*, *7*(5), 334–350.

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

Steinhausen, H. C. (2002). The outcome of anorexia nervosa in the 20th century. *American Journal of Psychiatry*, *159*(8), 1284–1293. https://doi.org/10.1176/appi.ajp.159.8.1284

Touyz, S., Le Grange, D., Lacey, H., Hay, P., Smith, R., Maguire, S., Stanford, S., Pike, K. M., & Crosby, R. (2013). Treating severe and enduring anorexia nervosa: A randomized controlled trial. *Psychological Medicine*, *43*(12), 2501–2511.

Treasure, J., Crane, A., McKnight, R., Buchanan, E., & Wolfe, M. (2011). First do no harm: Iatrogenic maintaining factors in anorexia nervosa. *European Eating Disorders Review*, *19*(4), 296–302. https://doi.org/10.1002/erv.1056

Wild, B., Friederich, H.-C., Gross, G., Teufel, M., Herzog, W., Giel, K. E., de Zwaan, M., Schauenburg, H., Schade-Brittinger, C., Schäfer, H., & Zipfel, S. (2009). The ANTOP study: Focal psychodynamic psychotherapy, cognitive-behavioural therapy, and treatment-as-usual in outpatients with anorexia nervosa—A randomised controlled trial. *Trials*, *10*, 28. https://doi.org/10.1186/1745-6215-10-28

Zipfel, S., Wild, B., Groß, G., Friederich, H. C., Teufel, M., Schellberg, D., Giel, K. E., de Zwaan, M., Dinkel, A., Herpertz, S., Burgmer, M., Löwe, B., Tagay, S., von Wietersheim, J., Zeeck, A., Schade-Brittinger, C., Schauenburg, H., & Herzog, W. (2014). Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): Randomised controlled trial. *Lancet*, *383*(9912), 127–137. https://doi.org/10.1016/S0140-6736(13)61746-8

How to cite this article: Kinnaird, E., Nicholson, M., Thomas, A., & Cooper, M. (2023). Systematic review: Exploring the monitoring and reporting of unwanted events in psychotherapy trials for anorexia nervosa. *International Journal of Eating Disorders*, 1–18. https://doi.org/10.1002/eat.23926

9. 2 Appendix 2: Journal Publishing Guidelines

International Journal of Eating Disorders (Systematic Literature Review)

Title Page

The Title Page of the manuscript should comprise:

A brief informative title containing the major keywords. The title should not contain abbreviations (see Wiley's best practice SEO tips).

All co-author details, including affiliation and email address.

Up to ten keywords.

If published already as a preprint, a link to the preprint server.

An author contributions statement that succinctly indicates how each author contributed to the piece of work, using the CRediT "Contributor Roles Taxonomy". Author contributions are also required within the submission form of both original and revised submissions.

Any applicable statements relating to our ethics and integrity policies, such as:

- data, materials and code availability statement
- funding statement or other acknowledgements of support
- conflict of interest disclosure
- permission to reproduce material from other sources

Abstract

The Abstract provides a succinct summary of the article content. The recommended format and word limit vary by article type.

Structured abstracts have a recommended maximum of 250 words and should be organized into: Objective: state the primary purpose of the article, or major question addressed in the study. Method: indicate the sources of data, give brief overview of methodology, or, if it is a review article, how the literature was searched and articles were selected for discussion. For research-based articles, briefly note study design, how participants were selected, and major study measures. If your data are based on a preregistered study, provide the preregistration number or link. Results: summarize the key findings. Discussion: indicate main clinical, theoretical, or research applications/implications.

Main Text File

The main text file should be in MS Word and include the following content and recommended formatting:

Main body, formatted as Introduction, Method, Results, and Discussion, as recommended by the International Committee of Medical Journal Editors (ICMJE) (J. Pharmacol. Pharmacother. 2010, 1, 42–58). Exceptions to these formatting recommendations include Commentaries, Forum articles, and Perspective articles.

A Public Significance statement (< 70 words) that explains why this research is important and is written in plain English for a general, educated public.

Figure titles should be supplied as a complete list in the text.

References

Please refer to article types regarding the number of permissible references.

This journal offers Free Format submission and authors may submit using their preferred referencing style, as long as consistency is applied throughout the manuscript.

The typesetter will apply the American Psychological Association reference style on manuscripts accepted for publication. If authors wish, they may review reference style guidelines prior to submission.

Tables

Tables should include a descriptive title and, if needed, footnotes defining abbreviations and any other information critical to interpreting the data shown.

Figures

Figures should have legends (and if needed, notes) that succinctly describe the information being displayed. Figures should be uploaded in the highest resolution possible.

Autism (Service Improvement Project)

Formatting

Autism asks that authors use the APA style for formatting. The APA Guide for New Authors can be found on the APA website, as can more general advice for authors.

Artwork, figures and other graphics

For guidance on the preparation of illustrations, pictures and graphs in electronic format, please visit SAGE's Manuscript Submission Guidelines.

Figures supplied in colour will appear in colour online regardless of whether or not these illustrations are reproduced in colour in the printed version. For specifically requested colour reproduction in print, you will receive information regarding the costs from SAGE after receipt of your accepted article.

Supplementary material

This journal is able to host additional materials online (e.g. datasets, podcasts, videos, images etc) alongside the full-text of the article. For more information please refer to our guidelines on submitting supplementary files.

Terminology

Autism has researched and produced its own guidance on terminology and language used in autism research. Please consult the guide here: autism terminology guidelines.

Likewise, Autism has also produced the following guidance to be considered when writing about race and ethnicity. Please consult the guide here: race and ethnicity language guidelines.

Reference style

Autism adheres to the APA reference style. View the APA guidelines to ensure your manuscript conforms to this reference style.

Journal of Eating Disorders (Theory Driven Research Project)

Title page

The title page should:

- present a title that includes, if appropriate, the study design e.g.:

"A versus B in the treatment of C: a randomized controlled trial", "X is a risk factor for Y: a case control study", "What is the impact of factor X on subject Y: A systematic review"

or for non-clinical or non-research studies a description of what the article reports

- list the full names and institutional addresses for all authors
- if a collaboration group should be listed as an author, please list the Group name as an author. If you would like the names of the individual members of the Group to be searchable through their individual PubMed records, please include this information in the “Acknowledgements” section in accordance with the instructions below

Large Language Models (LLMs), such as ChatGPT, do not currently satisfy our authorship criteria. Notably an attribution of authorship carries with it accountability for the work, which cannot be effectively applied to LLMs. Use of an LLM should be properly documented in the Methods section (and if a Methods section is not available, in a suitable alternative part) of the manuscript.

- Indicate the corresponding author

Abstract

The Abstract should not exceed 350 words. Please minimize the use of abbreviations and do not cite references in the abstract. Reports of randomized controlled trials should follow the CONSORT extension for abstracts. The abstract must include the following separate sections:

Background: the context and purpose of the study

Methods: how the study was performed and statistical tests used

Results: the main findings

Conclusions: brief summary and potential implications

Trial registration: If your article reports the results of a health care intervention on human participants, it must be registered in an appropriate registry and the registration number and date of registration should be stated in this section. If it was not registered prospectively (before enrollment of the first participant), you should include the words 'retrospectively registered'. See our editorial policies for more information on trial registration

Keywords

Three to ten keywords representing the main content of the article.

Background

The Background section should explain the background to the study, its aims, a summary of the existing literature and why this study was necessary or its contribution to the field.

Methods

The methods section should include:

- the aim, design and setting of the study
- the characteristics of participants or description of materials
- a clear description of all processes, interventions and comparisons. Generic drug names should generally be used. When proprietary brands are used in research, include the brand names in parentheses
- the type of statistical analysis used, including a power calculation if appropriate

Results

This should include the findings of the study including, if appropriate, results of statistical analysis which must be included either in the text or as tables and figures.

Discussion

This section should discuss the implications of the findings in context of existing research and highlight limitations of the study.

Conclusions

This should state clearly the main conclusions and provide an explanation of the importance and relevance of the study reported.

List of abbreviations

If abbreviations are used in the text they should be defined in the text at first use, and a list of abbreviations should be provided.

9. 3 Appendix 3: Service Improvement Project Survey Questions

Table 1: Training needs Survey

What is your role?	
How long have you worked with this team (in years)?	
How many clients have you seen whilst working with the OA CMHT who have a diagnosis of autism?	
How many clients have you seen whilst working with the OA CMHT who did not have a diagnosis, but you thought might have autistic traits/ had a query autism raised?	
Have you completed any autism training whilst working with the OA CMHT?	
On a scale of 0 (no knowledge) -100 (very knowledgeable), where would you rate your knowledge of autism?	
On a scale of 0-100 (no knowledge/ very knowledgable), where would you rate your knowledge of adapting standard care for patients on the autism spectrum?	
On a scale of 0 (no knowledge) -100 (very knowledgeable), where would you rate your knowledge of autism in older adults?	
What would you like to see included in a training session on autism in older adulthood? (You can circle/tick multiple options).	Basic introduction to autism Autism and communication Autism and sensory processing Autism and routine/ repetition Recognising autistic traits I think a patient might be on the spectrum but undiagnosed: what next? Treatment adaptations Autism and mental health Autism and physical health Autism in older adults Autism diagnostic process Overlaps between autism and other conditions Working with families and carers Other:

Table 2: Pre/Post Training Surveys

Question	**Asked pre training, post training, or both**
What is your job role?	Both
How long have you worked with this team (in years)?	Both
On a scale of 0 (no knowledge)- 100 (very knowledgeable, where would you rate your knowledge of autism?	Both
On a scale of 0 (no knowledge)- 100 (very knowledgeable, where would you rate your knowledge of autistic traits in older adults?	Both
On a scale of 0 (no knowledge)- 100 (very knowledgeable, where would you rate your knowledge of mental health and autism?	Both
On a scale of 0 (no knowledge)- 100 (very knowledgeable, where would you rate your knowledge of adapting standard care for patients on the autism spectrum?	Both
How would you rate the usefulness of today's training on a scale of 1 (not useful) to 10 (very useful)?	Post only
What did you find helpful about today's training?	Post only
What could we do to improve today's training?	Post only

9. 4 Appendix 4: Ethics Approval Letter

MEDICAL SCIENCES INTERDIVISIONAL RESEARCH ETHICS COMMITTEE
Research Services, Boundary Brook House, Churchill Drive, Headington, Oxford, OX3 7GB
Tel: +44(0)1865 616575
ethics@medsci.ox.ac.uk

CONFIDENTIAL

Emma Kinnaird
Oxford Institute of Clinical Psychology Training and Research
Isis Education Centre
Warneford Hospital
Headington
Oxford

17 November 2021

Dear Emma,

Research Ethics Approval - CUREC 1

Ethics Approval Reference: R77655/RE001

Study title: Exploring the relationship between clinical and personal models of recovery in anorexia nervosa

Short title: Exploring perceptions of recovery from anorexia nervosa

The above application has been considered on behalf of the Medical Sciences Interdivisional Research Ethics Committee (MS IDREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the IDREC, the proposed research has been judged as meeting appropriate ethical standards, and approval has been granted for a period of **18 months**, commencing on **17th November 2021.**

Amendments

Should there be any subsequent changes to the study, you should submit details to the MS IDREC for consideration and approval. Details of changes must be listed on an amendment form.

Yours Sincerely

Dr Helen Barnby-Porritt
Research Ethics Manager

9. 5 Appendix 5: Theory Driven Research Project Questionnaires

The Process of Recovery Questionnaire (QPR; Neil et al., 2009).

The Process of Recovery Questionnaire (the QPR): Guidelines for Clinicians, Researchers and Service Users for the uses, administration and scoring of the QPR

Developed by: Sandra Neil, Liz Pitt, Martina Kilbride, Anthony Morrison, Sarah Nothard, Mary Welford and William Sellwood in collaboration with The Bolton Salford and Trafford Service User Steering Committee

What is the QPR? The QPR is a 22-item measure developed from service users' accounts of recovery from psychosis in collaboration with local service users. The idea of the QPR is to ask people about aspects of recovery that are meaningful to them. The QPR is reliable and valid and is strongly associated with general psychological wellbeing, quality of life and empowerment all of which are crucial in recovery from psychosis. There are two subscales: 1) intrapersonal tasks involved in recovery and 2) interpersonal factors that facilitate recovery (see below).

What are the applications of the QPR?

- **Clinical practice:** Because the QPR asks about aspects of recovery that are important to service users this measure could help to facilitate communication and engagement. The QPR may be used to illustrate to people that other individuals progressed to achieve similar goals and this positive message might instil hope, which is crucial to recovery.

    - The QPR could be used both as a tool for setting goals for individual outcomes and as a measure of achievement of these individual goals. For example, the QPR could be used to help people open up, give them structure and offer a focus for individual goals they could work towards and then be used to track progress and provide evidence of this.

    - The sensitivity of QPR is currently being evaluated, to assess the QPR's use as a measure of service effectiveness and as a routine outcome measure.

- **Research**: It is suggested that researchers could use the QPR to expand and add to the evidence base in the area of recovery

How do I administer the QPR? The service users involved in the development of the QPR suggest that before administering the QPR clinicians or researchers using this measure should ensure that:

- All service users who are asked to complete the QPR are given general information (as above) about the measure and are provided with an explanation as to why they are being asked to complete this questionnaire, such as "It is hoped through asking you to complete this measure that we can identify the areas in your life where things are going well and also any areas where you might be having difficulties."

- All service users must give their written or verbal consent to complete the QPR

- The QPR should where possible be completed with another professional or person with whom they can discuss any issues raised.

- The QPR should be used judiciously and responsibly by clinicians, and service users who are in crisis and / or very distressed **should not** be asked to complete the QPR

- The QPR should not be used in a sterile manner, but rather as a vehicle to facilitate discussion about individual goals.

How do I score the QPR? The QPR has 22 items each scored on a 4-point scale (0= disagree strongly, 1=disagree, 2=neither agree nor disagree, 3=agree, 4=agree strongly). The intrapersonal subscale includes items; 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 18, 19, 21, 22. The interpersonal subscale includes items: 14, 15, 16, 17, 20 (highlighted in grey on the QPR). Higher scores are indicative of recovery. However, those involved in developing this measure suggest that subscale or total scores should not only be added to give total recovery scores, but the QPR should be used as described above e.g. as a tool for engagement, setting goals relative to the individual and as a measure of outcome for these.

PLEASE TURN OVER THE PAGE AND CONTINUE OVERLEAF

The Process of Recovery Questionnaire (QPR)

[15/10/2007- Version 1]

We developed this questionnaire in order to understand more about the process of recovery; what's helpful and what's not so helpful.
Everyone is different and there will be differences for everyone. The items on this questionnaire were developed through a process of interviewing service users about their recovery journeys. We hope that by filing in this questionnaire you will help us find out information that is important to you and your own recovery. Not all factors will be important to you, since everyone is different. This questionnaire is not intended to be used to impose anything against your wishes.

If you would like to fill in the questionnaire, please take a moment to consider and sum up how things stand for you at the present time, in particular over the last 7 days, with regards to your mental health and recovery. Please respond to the following statements by putting a tick in the box which best describes your experience.

	Disagree strongly	Disagree	Neither agree nor disagree	Agree	Agree Strongly
I feel better about myself					
I feel able to take chances in life					
I am able to develop positive relationships with other people					
I feel part of society rather than isolated					
I am able to assert myself					
I feel that my life has a purpose					

	My experiences have changed me for the better					
	I have been able to come to terms with things that have happened to me in the past and move on with my life					
	I am basically strongly motivated to get better					
	I can recognise the positive things I have done					
	I am able to understand myself better					
	I can take charge of my life					
	I am able to access independent support					
	I can weigh up the pros and cons of psychiatric treatment					
	I feel my experiences have made me more sensitive towards others					
	Meeting people who have had similar experiences makes me feel better					
	My recovery has helped challenge other peoples views about getting better					
	I am able to make sense of my distressing experiences					
	I can actively engage with life					
	I realise that the views of some mental health professionals is not the only way of looking at things					
	I can take control of aspects of my life					
	I can find the time to do the things I enjoy					

Self-Compassion Scale (SCS; Neff, 2003).

Self-Compassion Scale (SCS)
HOW I TYPICALLY ACT TOWARDS MYSELF IN DIFFICULT TIMES

Please read each statement carefully before answering. For each item, indicate how often you behave in the stated manner, using the following 1-5 scale. Please answer according to what really reflects your experience rather than what you think your experience should be.

Almost never				Almost always
1	2	3	4	5

1. I'm disapproving and judgmental about my own flaws and inadequacies.
2. When I'm feeling down I tend to obsess and fixate on everything that's wrong.
3. When things are going badly for me, I see the difficulties as part of life that everyone goes through.
4. When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world.
5. I try to be loving towards myself when I'm feeling emotional pain.
6. When I fail at something important to me I become consumed by feelings of inadequacy.
7. When I'm down, I remind myself that there are lots of other people in the world feeling like I am.
8. When times are really difficult, I tend to be tough on myself.
9. When something upsets me I try to keep my emotions in balance.
10. When I feel inadequate in some way, I try to remind myself that feelings of inadequacy are shared by most people.
11. I'm intolerant and impatient towards those aspects of my personality I don't like.
12. When I'm going through a very hard time, I give myself the caring and tenderness I need.
13. When I'm feeling down, I tend to feel like most other people are probably happier than I am.
14. When something painful happens I try to take a balanced view of the situation.
15. I try to see my failings as part of the human condition
16. When I see aspects of myself that I don't like, I get down on myself.
17. When I fail at something important to me I try to keep things in perspective.
18. When I'm really struggling, I tend to feel like other people must be having an easier time of it.
19. I'm kind to myself when I'm experiencing suffering.
20. When something upsets me I get carried away with my feelings.
21. I can be a bit cold-hearted towards myself when I'm experiencing suffering.
22. When I'm feeling down I try to approach my feelings with curiosity and openness.
23. I'm tolerant of my own flaws and inadequacies.
24. When something painful happens I tend to blow the incident out of proportion.
25. When I fail at something that's important to me, I tend to feel alone in my failure.
26. I try to be understanding and patient towards those aspects of my personality I don't like.