

A thesis submitted in partial fulfilment for the degree of

Doctor of Clinical Psychology (DClinPsych)

Hannah Wilkinson

Harris Manchester College

University of Oxford

Trinity 2025

IN DEI NOMINE

Word Count by Component

Systematic Literature Review:	6987
Service Improvement Project:	5093
Main Research Project:	5366
Executive Summary:	544
Reflective Narrative:	895
Total word count:	18885

Table of contents

Abstracts **7**

Systematic Literature Review *7*

Service Improvement Project *8*

Theory-driven Research Project *9*

Cognitive behavioural therapy for sleep problems in psychosis: A systematic review of effectiveness and acceptability **10**

Abstract *11*

Introduction *12*

Method *14*

Search Strategy and Selection Criteria 14

Data Extraction and Synthesis 17

Quality appraisal 18

Results *19*

1. Study design, participant and intervention characteristics 19

2. Quality Appraisal 26

3. Treatment effectiveness 28

4. Treatment acceptability 35

Discussion *40*

Strengths and Limitations 44

Conclusion 46

References *47*

Improving the psychological wellbeing of parents whose babies have a prenatal diagnosis requiring surgical treatment **64**

Abstract *65*

Introduction *66*

Project background and aims 67

Methods *69*

Participants 69

Materials 69

Procedure 70

Analysis 71

Researcher Background 72

Results *72*

Survey results 72

Interview results 73

Discussion *79*

Limitations 81

Recommendations 82

Dissemination 84

Conclusion 84

References *84*

Do eating disorder cognitions play a role in Avoidant Restrictive Feeding Intake Disorder (ARFID)? A comparison between ARFID and Anorexia **94**

Abstract *95*

Introduction *96*

Method *99*
Design 99
Participants 100
Measures 100
Procedure 101
Data analysis plan 102

Results *103*
Sample characteristics 103
Internal consistency of scales 105
Primary hypotheses: prevalence of classical ED-related cognitions 105
Prevalence of weight/shape concerns 106
Cognitive predictors of restraint 108

Discussion *109*
Theoretical implications 111
Clinical implications 112
Limitations and Future Research Directions 113
Conclusion 114

References *115*

Executive Summary **130**

Reflective Narrative **133**

Acknowledgements **136**

Appendix: Systematic Literature Review **137**

Appendix 1 – PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) checklist: recommended items to address in a systematic review protocol *137*

Appendix 2 – PICO table *139*

Appendix 3 - Search terms *140*

Appendix 4 – Qualitative checklist ratings, developed from Braun and Clarke's reviewing guidelines *142*

Appendix 6 – List of included papers *156*

Appendix 7 – Quality Assessment using 'The Measurement Tool to Assess systematic Reviews 2' (AMSTAR) *158*

Appendix 8 – Author guidelines for British Journal of Psychiatry *162*

Appendix: Service Improvement Project **164**

Appendix 1 – Further detail regarding service context *164*
Appendix 2 – Further examples of illustrative quotes *166*
Appendix 3 – Survey Participant Information and Consent *170*
Appendix 4 – Interview Participant Information and Consent *173*
Appendix 5 – Survey questionnaire *177*
Appendix 6 – Interview schedule *183*
Appendix 7 – Data analysis *185*
Appendix 8 – Author guidelines *186*
Appendix: Theory-driven Research Project **188**
Appendix 1 – Ethical approval *188*
Appendix 2 – Recruitment posters *189*
Appendix 3 - Measures *192*
Appendix 4 – Scoring details *202*
Appendix 5 – Results of non-parametric (Kruskal-Wallis) test of fourth hypothesis relating to prevalence of weight/shape concerns across diagnostic groups *205*
Appendix 6 – Cronbach's reliability results *206*
Appendix 7 – Results of stepwise regression *207*
Appendix 8 – Reproduction of models of ARFID profiles from Thomas & Eddy, 2018 (p. 21-23) *208*
Appendix 9 – Author Guidelines for International Journal of Eating Disorders *210*

Abstracts

Systematic Literature Review

Background: Sleep problems are common among people with psychosis. Research suggests poor sleep is causally related to psychosis, anxiety and depression.

Aims: This review investigates the effectiveness and acceptability of cognitive behavioural therapies (CBT) in targeting sleep problems in people with and at risk of psychosis.

Method: Four databases were searched in line with PRISMA guidelines. Eligible studies either evaluated (a) CBT targeting sleep problems in people with or at risk of psychosis, or (b) subjective experiences of this treatment. Articles not published in peer-review journals were excluded. Treatment effectiveness was investigated for sleep, psychosis and other clinical outcomes. Acceptability was evaluated using qualitative data, drop-out rates, adverse events and relevant questionnaires. Adaptations to standard treatment protocols were described. Research quality was appraised using Cochrane Risk of Bias tools for randomised and non-randomised trials, and a checklist developed from Braun and Clarke's (2021) reviewing guidelines was used for qualitative papers.

Results: Of the 975 records identified, 14 were eligible. The most common CBT target was insomnia. Treatment protocols were typically adapted by omitting sleep restriction. Large effect sizes were reported for sleep outcomes, however effects for other clinical outcomes were less clear. Qualitative data and acceptability outcomes suggest that treatment was received positively by participants.

Conclusions: CBT is an effective and acceptable treatment for sleep problems in people with and at risk of psychosis. However, our conclusions are limited by few good-quality studies and small samples. Further gold-standard research is required to inform evidence-based guidelines.

Service Improvement Project

Background: Parents experience significant stress when their baby receives a prenatal diagnosis that will need surgical treatment. This study aims to explore parents' experiences of a neonatal surgical pathway and develop recommendations for support.

Methods: Twenty-four parents completed surveys and nine parents were subsequently interviewed regarding their experiences of the neonatal surgical pathway. Themes were identified using reflexive thematic analysis.

Results: Parents highlighted the importance of preparation, empathetic communication and involvement in their baby's care. Different levels of support ranged from parent-to-parent mentoring, a space to discuss feelings in medical consultations, ensuring physical health needs were met, and access to psychological services.

Conclusion: The findings show that beyond the immediate stress of having an unwell child, the way in which the wider medical system holds parents in their distress has a major impact on parents' wellbeing. Recommendations to improve support are provided.

Theory-driven Research Project

Objective: Avoidant Restrictive Feeding Intake Disorder (ARFID) is a relatively new diagnosis that featured in the 2013 edition of the Diagnostic Statistical Manual (DSM-V). There is limited evidence about whether individuals with ARFID experience unhelpful cognitions that overlap with other eating disorders and whether they play a role in driving disordered eating behaviours. This study seeks to address this gap.

Methods: Adult participants were recruited from a community sample and invited to complete an online survey. An observational design was used to compare responses by people with ARFID, anorexia nervosa, and no eating disorder who were screened using DSM-V criteria. Questionnaires were used to measure core beliefs, assumptions and automatic thoughts that are associated with other eating disorders and an assessment of concerns about weight and shape was completed.

Results: Individuals with ARFID were found to have significantly higher levels of disordered core beliefs, assumptions, automatic thoughts, and weight/shape concerns than people with no eating disorder. They showed lower levels of these cognitions relative to individuals with anorexia nervosa. Disordered assumptions and automatic thoughts explained a large proportion of variance in behaviours intended to prevent weight gain in this group.

Conclusions: The findings have implications for the diagnosis and treatment of ARFID. It challenges the DSM-V conceptualisations of ARFID as entirely separable from other eating disorders and any associated weight/shape concerns. It also highlights the need for clinicians to assess and treat unhelpful cognitions that may be maintaining disordered patterns of eating. Future research directions are discussed.

Systematic Review of the Literature

Cognitive behavioural therapy for sleep problems in psychosis: A systematic review of effectiveness and acceptability

Hannah Wilkinson

Hannah.Wilkinson@psy.ox.ac.uk

The Oxford Institute of Clinical Psychology Training and Research

University of Oxford

May 2025

Supervisors: Dr Alex Lau-Zhu & Dr Louise Johns

Word count: 6987

Accepted for publication by the British Journal of Psychiatry. This journal was chosen because its commitment to the care of mental health difficulties and focus on the clinical aspects of research.

Acknowledgements: Thank you to the Oxford Health NHS Foundation Trust librarians for their assistance with the search strategy.

Abstract

Background: Sleep problems are common among people with psychosis. Research suggests poor sleep is causally related to psychosis, anxiety and depression.

Aims: This review investigates the effectiveness and acceptability of cognitive behavioural therapies (CBT) in targeting sleep problems in people with and at risk of psychosis.

Method: Four databases were searched in line with PRISMA guidelines. Eligible studies either evaluated (a) CBT targeting sleep problems in people with or at risk of psychosis, or (b) subjective experiences of this treatment. Articles not published in peer-review journals were excluded. Treatment effectiveness was investigated for sleep, psychosis and other clinical outcomes. Acceptability was evaluated using qualitative data, drop-out rates, adverse events and relevant questionnaires. Adaptations to standard treatment protocols were described. Research quality was appraised using Cochrane Risk of Bias tools for randomised and non-randomised trials, and a checklist developed from Braun and Clarke's (2021) reviewing guidelines was used for qualitative papers.

Results: Of the 975 records identified, 14 were eligible. The most common CBT target was insomnia. Treatment protocols were typically adapted by omitting sleep restriction. Large effect sizes were reported for sleep outcomes, however effects for other clinical outcomes were less clear. Qualitative data and acceptability outcomes suggest that treatment was received positively by participants.

Conclusions: CBT is an effective and acceptable treatment for sleep problems in people with and at risk of psychosis. However, our conclusions are limited by few good-quality studies and small samples. Further gold-standard research is required to inform evidence-based guidelines.

Introduction

Sleep problems are commonly reported in psychotic spectrum disorder patients.[1] Traditionally, sleep problems have been viewed as secondary to mental distress, with troubling thoughts, feelings and physiological hyperarousal seen as responsible for keeping people awake.[2] However, recent evidence suggests that sleep difficulties and psychotic symptoms are bidirectionally related.[3] Freeman's cognitive model identifies sleep problems as a causal factor in the development and maintenance of paranoia.[4,5] Lack of sleep can increase anomalous perceptual experiences such as hearing voices or seeing figures (i.e., hallucinations) which may subsequently be misinterpreted in line with threatening delusional beliefs.[6,7] Moreover, sleep deprivation is linked to negative emotion, mood dysregulation, and poorer reasoning which may limit processing of alternative explanations and maintain a sense of threat.[8–10] Nightmares are also hypothesised to maintain persecutory beliefs by interrupting sleep and triggering feelings of threat and negative affect upon waking. Moreover, nightmares typically depict paranoid fears (e.g. physical assault) in vivid detail which may be mistaken for reality, rather than a dream. This may directly reinforce patients' worst fears.[11]

It is not only psychotic symptoms that can be exacerbated by poor sleep. There are wide-ranging consequences for people's daytime functioning, work, family and social relationships.[12] Accordingly, poor sleep is linked to an overall reduced quality of life.[13] Moreover, poor sleep increases susceptibility to developing other mental disorders such as depression and anxiety, and has strong associations with suicidal ideation.[14,15]

Sleep is increasingly identified as an important target for treatment, and patients with psychosis report desiring this.[16,17] The first-line recommended treatment for insomnia is Cognitive Behavioural Therapy (CBT).[18] Sleep is largely regulated by two processes, a homeostatic system ('sleep pressure') and a circadian timing system (circadian rhythms).[19] Sleep pressure rises with wake and falls during the sleep and correlates with alertness.

Circadian rhythms are directly influenced by environmental cues, especially light, which allow a 24-hour cycle. CBT for sleep problems draws from this theory to modify dysfunctional behaviours and beliefs, with the premise that thoughts, feelings and behaviours are linked.[20] For example, people with insomnia may believe they need to nap during the day to catch up on missed sleep. However, this makes it harder to fall asleep at night and leads to a vicious cycle. People may be encouraged to identify and challenge these unhelpful cognitions and reduce napping so that sleep pressure reaches a maximum at night-time. For nightmares, Imagery Rescripting Treatment (IRT) is advised in best practice guidelines.[21,22] IRT is a CBT technique that uses imagery-based techniques to reimagine nightmares with alternative, less threatening endings.[23]

There is growing interest in whether CBT for sleep problems can alleviate psychotic symptoms and prevent clinically high-risk individuals from developing full-blown psychosis. Several trials have shown promising results.[24–37] Strict adherence to the CBT for Insomnia (CBT-I) protocol involves elements of sleep restriction and setting a consistent sleep window, which is associated with side-effects such as sleep deprivation, fatigue and impaired vigilance.[38] There is evidence that sleep deprivation can be associated with psychotic symptoms in healthy controls.[39] Studies therefore describe making adaptations to facilitate engagement and mitigate risks of exacerbating patients' symptoms in the context of psychosis.[17,24–27,29,30,32–35,37]

Review Questions

This review seeks to address the following:

(a) What are the elements of CBT for sleep problems in people with or at high risk of psychosis-related disorders, and what adaptations are made to facilitate engagement and maximise effectiveness?

(b) What is the effectiveness in terms of sleep-related outcomes and psychotic symptoms (including delusions and hallucinations)?

(c) What are the effects on other clinical outcomes such as depression, anxiety and quality of life?

(d) What is the acceptability of treatment?

Method

Search Strategy and Selection Criteria

This review adhered to the updated version of the Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines (PRISMA). It was pre-registered on PROSPERO (ID CRD42023434097)[40,41] and was later updated to include the addition of qualitative studies. A systematic search was performed across Embase, Medline, PsycInfo and PubMed for articles published from inception to 12th December 2024. The search terms aimed to capture three main concepts: (a) sleep disorders (e.g. insomnia and nightmares), (b) CBT treatments, and (c) psychosis (e.g. symptoms and diagnoses). Keyword searching of titles, abstracts, and medical subject headings (MeSH) was undertaken following consultation with a librarian (see Appendix for search terms). Several research groups were contacted to find out if they were intending to publish any new articles that would be relevant. The decision was made to include non-randomised studies due to the low numbers of randomised-controlled trials investigating CBT for sleep problems in people with psychosis.

Table 1 *Summary of inclusion criteria*

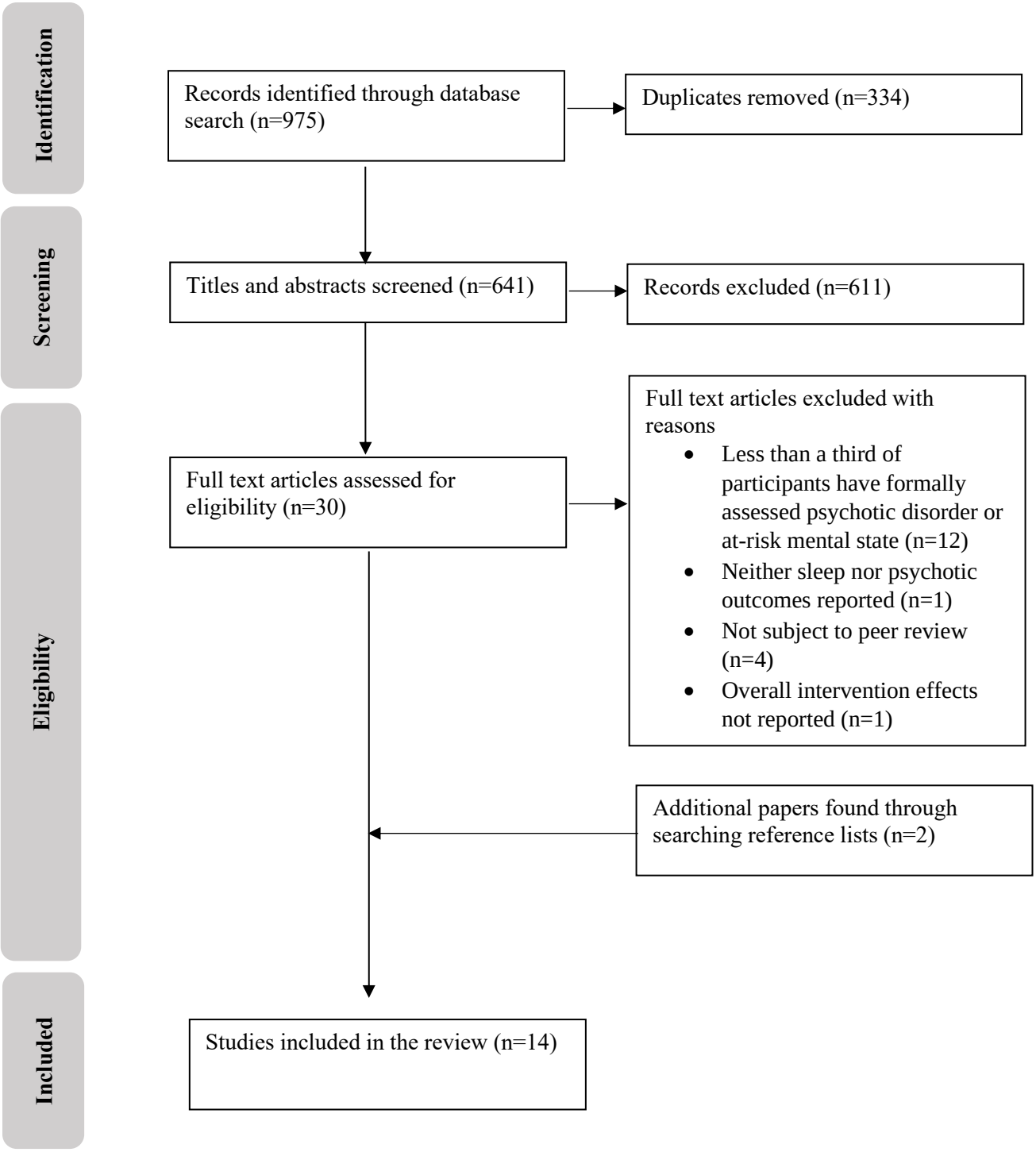
Dimension	Inclusion Criteria
General	Peer-reviewed study evaluating a CBT-based intervention, published in English. This includes controlled or non-controlled trials of an intervention, or qualitative analyses of participants' experiences of the intervention.
Population	At least a third* of participants have a formally diagnosed psychotic disorder or are formally assessed as meeting the criteria for 'at risk mental state.' Participants present with a level of sleep difficulties which are not caused by a medical sleep disorder (e.g. obstructive sleep apnoea) or neurodegenerative disease (e.g. dementia).
Intervention	CBT-based interventions target sleep problems such as insomnia (i.e., difficulties falling asleep, staying asleep, early waking), nightmares, or excessive sleeping. There are a minimum of 10 participants in experimental studies. There is no minimum number for qualitative studies.
Comparator/control	Quantitative studies compare outcomes pre- and post-treatment, and/or across a comparison group (if present). Qualitative studies do not need a control group.
Outcome	Quantitative studies assess sleep quality and/or psychotic symptoms and report overall intervention effects. Qualitative studies describe participant's subjective experiences of engaging in treatment.

**The criteria of a third of participants with psychotic disorder or at-risk mental state was chosen due to the lack of research in this area and seeking to be as inclusive as possible of papers that had the potential to be clinically relevant.*

Screening was performed by two Trainee Clinical Psychologists (HW, first reviewer; RB, second reviewer). Papers were excluded if they did not meet the criteria presented in Table 1. After searching across the databases, the studies were first screened using their titles and abstracts. The second reviewer screened 20% of these. Cohen's kappa[42] showed an almost perfect level of agreement ($\kappa = 0.85$), and disagreements were resolved by discussion.

At the next stage, the first reviewer screened eligible papers at the full-text level and the second researcher screened 25% of these. There was a perfect level of agreement ($k = 1$). Reference lists of included papers were searched to identify any further relevant studies. The study selection process is shown in Figure 1.

Figure 1 *PRISMA flowchart of the study selection process.*



Note. The reason given for the exclusion of studies reflects the first criterion for exclusion that the author identified. It is possible that some studies met several criteria for exclusion.

Data Extraction and Synthesis

Data were systematically extracted by the first reviewer starting on 30th October 2023. The second reviewer independently extracted 50% of the data to ensure accuracy and reliability. Information was gathered regarding publication date, study design, treatment focus, protocol adaptations, geographical location, participant demographics, drop-out, adverse events, means and standard deviations (SDs) of clinical outcomes and questionnaire outcomes pertaining to treatment acceptability or satisfaction. Supplementary materials were checked for additional clinical outcomes that were not reported in main papers. For trials comparing clinical outcomes across a treatment and control group, Cohen's *d* standardised between-group effect sizes were calculated.[43] For uncontrolled trials, only the direction of the effect was reported since precise estimates were unlikely to be reliable due to the lack of a control group.

Previous research has defined and assessed 'acceptability' using a variety of different means. Sekhon et al.'s[44] systematic review devises a theoretical framework of intervention acceptability comprising seven components: affective attitude, burden, perceived effectiveness, ethicality, intervention coherence, opportunity costs, and self-efficacy. This review attempted to capture acceptability using a multifaceted approach. Themes and quotes that related to these constructs were extracted from qualitative studies by the first author (HW). Since there were only two studies, no formal analytic procedure was used; instead, common ideas and quotes were organised into themes, and two supervisors (ALZ and LJ) were consulted to check their credibility. Drop-out was defined in this review as the number of participants who discontinued after having attended at least one session, during the intervention period.[45] A narrative synthesis was used to interpret these findings together with data from drop-out rates, adverse events and relevant questionnaires.

Quality appraisal

The Revised Cochrane 'Risk of Bias' (RoB)[46] tool was used to assess all randomised controlled trials (RCTs). The Cochrane 'Risk of Bias In Non-randomised Studies – of Interventions' (ROBINS-I)[47] tool was used to assess risk of bias in the remaining non-randomised trials (non-RCTs). These tools yield a risk of bias judgment across several domains (e.g. selection of participants) and for the overall paper. To appraise the quality of qualitative studies, a checklist was developed from Braun and Clarke's guidelines for reviewers to assess the rigour of thematic analysis.[48]

The second researcher independently appraised 30% of papers after calibrating their ratings with one of each kind of paper (RCT, non-RCT and qualitative). In previous research, both Cochrane tools have demonstrated slight inter-rater agreement for the overall final judgment of risk of bias and poor-to-moderate agreement for single domains.[49–51] The current review found perfect agreement ($\kappa = 1$) both overall and at the level of single domains for the RoB. Agreement was perfect ($\kappa = 1$) overall and fair ($\kappa = 0.60$) at the level of single domains for the ROBINS-I. Agreement was moderate ($\kappa = 0.76$) on the qualitative quality checklist. Differences between the reviewers were resolved through discussion.

The 'Measurement Tool to Assess systematic Reviews 2' (AMSTAR)[52] was used to evaluate the quality of the overall review (see Appendix).

Results

Fourteen studies were identified as eligible for inclusion. The results are organised by 1) Study characteristics; 2) Quality appraisal; 3) Treatment effectiveness; and 4) Treatment acceptability.

1. Study design, participant and intervention characteristics

1.1 Study design and participants

The study and participant characteristics are shown in Table 2. The majority (57.1%) of studies were conducted in the United Kingdom (UK) with the remainder being conducted in the United States of America (USA), Spain, South Korea and Australia. There were five RCTs, seven non-RCTs and a mixed-methods study combining a non-RCT and qualitative design. There was also a qualitative paper that used thematic analysis to understand participants' experiences of treatment in one of the non-RCTs.[29,53]

The total number of participants was 398 across trials (RCT and non-RCT studies) and 21 across qualitative studies. The average age of participants ranged from 18.5 to 51.6 years. Participants were mostly adults with a diagnosed psychotic disorder, however three studies recruited young people aged 14-25 years who were at risk of developing psychosis.[27,29,53] All participants received CBT for insomnia, except in one study where nightmares were the focus of treatment.[26] All studies had a relatively even split between male and females, excepting one with only male participants[25] and another study of USA veterans[30] in which most of the participants were also male.

It was challenging to summarise ethnicity and race due to a lack of consistency in the categories used. Where this information was reported, studies indicated that the highest proportion of participants identified as White, except for one study where participants

predominantly identified as African-American[31]; 35.7% of studies did not report ethnicity or race.

Table 2 *Study design and participants*

Study	Design	CBT target	Location	Population/ clinical presentation	Participants	
					Intervention	Control
Batalla-Martin et al. (2023)[37]	RCT	Insomnia	Spain	Adults with schizophrenia diagnoses and mild-to-severe insomnia	N = 20; M(age) = 50.3 (SD = 11.2); Gender: Female = 40%, Male = 60%; Ethnicity: Not reported	N = 20; M(age) = 50.4 (SD = 12.8); Gender: Female = 40%, Male = 60%; Ethnicity: Not reported
Freeman et al., (2015)[24]	RCT	Insomnia	UK	Adults with schizophrenia, schizoaffective, or delusional disorder, with persistent hallucinations or delusions, and insomnia	N = 24; M(age) = 39.6 (SD = 11.6); Gender: Female = 33%, Male = 67%; Ethnicity: White = 92%, Black = 4%, Chinese = 4%	N = 26; M(age) = 42.4 (SD = 13.5); Gender: Female = 31%, Male = 69%; Ethnicity: White = 96%, Black = 4%
Sheaves et al., (2018)[25]	RCT	Insomnia	UK	Adults admitted to an inpatient unit for acute mental health crisis with mixed psychiatric diagnoses; including schizophrenia spectrum and other psychotic disorders = 45%, with insomnia symptoms	N = 20; M(age) = 40 (SD = 12); Gender: Male = 100%; Ethnicity: Black African = 10%, Asian = 10%, Mixed ethnic group = 5%, White British = 60%, White Other = 15%	N = 20; M(age) = 40 (SD = 14); Gender: Male = 100%; Ethnicity: Asian = 5%, Mixed ethnic group = 5%, White British = 75%, White Other = 15%
Sheaves et al., (2019)[26]	RCT	Night-mares	UK	Adults with schizophrenia, schizoaffective disorder, delusional disorder, or psychosis not otherwise specified, with current persecutory delusions and nightmares	N = 12; M(age) = 43 (SD =12); Gender: Female = 42%, Male = 58%; Ethnicity: White British = 83%, Mixed/ multiple ethnic groups (White and Black African) = 8%, Black Caribbean = 8%	N = 12; M(age) = 39 (SD =13); Gender: Female = 42%, Male = 58%; Ethnicity: White British = 75%, Other: 8%, Asian/Asian British - Indian: 8%, Asian/Asian British - Pakistani: 8%
Waite et al., (2023)[27]	RCT	Insomnia	UK	Young people at ultra-high risk of psychosis with insomnia	N = 21; M(age) = 17.0 (SD =2.2);Gender: Female = 52% Male = 48%; Ethnicity: Black British = 5%, Black Caribbean = 5%, Other: 10%, White = 81%	N = 19; M(age) = 16.8 (SD = 2.8); Gender: Female = 42%, Male = 47%, Other = 11%; Ethnicity: Indian = 5%, Other: 11%, Pakistani = 5%, White = 79%
Hwang et al., (2019)[28]	Non-randomised trial with control group	Insomnia	South Korea	Adults with schizophrenia with persistent hallucinations or delusions, and insomnia	N = 31; M(age) = 45.7 (SD = 10.5); Gender: Female = 48%, Male = 52%; Ethnicity: Not reported	N = 32; M(age) = 44.2 (SD =11.0); Gender: Female: 22%, Male = 78%; Ethnicity: Not reported
Bradley et al., (2018)[29]	Case series	Insomnia	UK	Young people at ultra-high risk of psychosis with insomnia or other sleep problem	N = 12; M(age) = 18.5 (SD = 1.9); Gender: Female = 65%, Male = 45%; Ethnicity: White = 91%, Other = 9%	None
Haynes et al., (2011)[30]	Single-arm trial	Insomnia	USA	Veterans on an inpatient unit, the majority of whom met the clinical threshold for moderate-to-severe insomnia, with mixed psychiatric diagnoses; including schizophrenia = 28%, psychosis not otherwise specified = 14%	N = 19; M(age) = 51.6 (SD = 12.2); Gender: Female = 5%, Male = 95%; Ethnicity not reported	None

Holmes et al., (1995)[31]	Single-arm trial	Insomnia/ sleep quality	USA	Adults with mixed psychiatric diagnoses; including schizophrenia = 44%, schizoaffective disorder = 11%, with sleep difficulties identified using sleep logs	N = 18; M(age) = 34.3 (SD = 9.49); Gender: Female = 50%, Male = 50%; Ethnicity: African-American = 55%, Other = 45%	None
Myers et al., (2011)[32]	Case-series	Insomnia	UK	Adults with schizophrenia, psychotic disorder, psychosis, schizoaffective disorder or delusional disorder, with persistent persecutory delusions and sleep difficulties	N = 15; M(age) = 45.5 (SD = 11.3); Gender: Female = 40% , Male = 60%; Ethnicity: White = 47%, Black Caribbean = 7%, Black African = 13%, Black Other = 20%, Other = 13%	None
Novak et al., (2020)[33]	Single-arm trial	Insomnia	USA	Adults on an inpatient unit with mixed psychiatric diagnoses; either paranoid schizophrenia, psychosis or depressive disorder, and problematic insomnia	N = 13; Age, gender and ethnicity not reported	None
Waters et al., (2020)[34]	Secondary analysis of a non-randomised trial with control group	Insomnia	Austra -lia	Adults with sleep difficulties and diagnosed non-organic psychotic disorder; including schizophrenia, schizoaffective disorder, bipolar disorder, major depression with psychotic features, and psychosis not-otherwise specified	N = 50; M(age) = 41.2 (SD = 10.9); Gender: Female = 46%, Male = 54%; Ethnicity: Not reported	Not reported
Taylor et al., (2022)[35]	Single-arm trial and qualitative research (mixed methods)	Insomnia	UK	Adults with schizophrenia-spectrum diagnosis and current sleep difficulties	N = 14 for trial; N=10 for participants taking part in qualitative interviews; M(age) = 35.6 (SD = 10.9); Gender: Female: 36%, Male: 64%; Ethnicity: Asian/Asian British = 14%, Black/ African/ Caribbean/ Black British = 43%, Mixed ethnic group = 14%, Other = 7%, White = 21%.	None
Waite et al., (2018)[53]	Qualitative research	Insomnia	UK	Young people at ultra-high risk of psychosis, who had partaken Bradley et al.'s CBT-I intervention[29]	N = 11; M(age) = 18.3 (SD =2.0); Gender: Female = 55%, Male = 45%; Ethnicity: Asian = 9%, White = 91%	None

Note: This table presents all studies that are included in the review, both qualitative and quantitative. Studies are ordered by type (RCTs, non-randomised trials with a control group, single-arm trials, mixed-methods, qualitative).
N = number of participants. M(age) = mean age of participants. SD = standard deviation. RCT = Randomised Control Trial. Where possible data is reported to one decimal place.
All papers reported meeting the inclusion criteria of a third of participants having a psychotic disorder or at-risk mental state. The only exception was Novak et al.[33] where it was confirmed in an email from one of the authors that 66% of participants had a diagnosed psychotic disorder (Dr Stuart Watson, January 2025).

1.2 Elements of CBT and adaptations

Table 3 gives an overview of the interventions. Standard components of CBT-I protocols include psychoeducation, stimulus control, sleep hygiene (e.g. avoiding screen use before bed), relaxation, worry-management, and techniques that address circadian rhythm disruption such as sleep/wake cycle realignment and daily activity.[20,54] The studies which used CBT-I either explicitly described these elements, or reported following specific protocols which were known to feature them. Boosting daily activity was described as a key feature of several protocols because many psychotic patients have sedentary lifestyles. One study described greater scaffolding, for example, by first educating patients about the efficacy of CBT-I to enhance their motivation to engage, and by linking in with community resources to increase opportunities for structured activity and social connection.[37] Interventions in young people made further adjustments for the social, biological and environmental differences between adolescents and adults (e.g. involving parents, if appropriate).[27,29] Only one study investigated IRT for nightmares.[26] They also incorporated strategies for limiting worry, sleep window stabilisation, relaxation and grounding.

Interventions were delivered in a variety of settings, formats and timeframes. Most interventions took place within a community mental health setting, although three took place on a psychiatric ward.[25,30,33] The majority were delivered by clinical psychologists (53.8%) and as one-to-one sessions (53.8%). The number of intended sessions ranged from 4 to 8 and typically occurred on a weekly basis. One study differed by providing sessions more intensively within a two-week period to ensure participants received the specified dose before being discharged from acute inpatient hospitalisation.[25] Of note, one study reported outcomes after participants received only 1-2 sessions of a 4-session programme because they were limited by the duration of participants' inpatient admission.[30]

The most common adaptation was omitting formal sleep restriction from CBT-I protocols (66.7% of CBT-I studies). Other adaptations included addressing problems of delusions and hallucinations interfering with sleep, oversleeping to escape from psychotic experiences such as voices, and fears of going to bed caused by past adverse experiences.

Table 3 *Elements of CBT and adaptations*

Study	Delivery/ Format	Adaptations to standard CBT				
		Omit sleep restrict-ion	**Address delusions or hallucin -ations**	**Address oversleeping to escape psychosis**	**Address fear of bed**	**Address night-mares**
Batalla-Martin et al. (2023)[37]**	6 x 1.5hr group CBT-I sessions led by a nurse.	-	-	-	-	-
Freeman et al., (2015)[24]	8 x 1hr individual CBT-I sessions led by clinical psychologist.	X	X	X	X	X
Sheaves et al., (2018)[25]	Minimum 5 x 1hr individual CBT-I sessions led by clinical psychologist.	X	X	-	-	-
Sheaves et al., (2019)[26]**	4 x 1hr individual sessions involving CBT-for-nightmares led by clinical psychologist. Image rehearsal therapy was the key technique offered to all participants.	-	X	-	-	N/A
Waite et al., (2023)[27]	8 x 1hr individual CBT-I sessions led by clinical psychologist.	X	X	-	-	X
Hwang et al., (2019)[28]*	4 x 45min group CBT-I sessions led by mental health social worker trained in CBT.	-	-	-	-	-
Bradley et al., (2018)[29]	8 x 1hr individual CBT-I sessions led by clinical psychologist.	X	-	-	-	-
Haynes et al., (2011)[30]**	Attendance of any 1-2 x 1hr group sessions out of a 4-session rolling CBT-I programme. Led by clinical psychologist.	X	-	-	-	-
Holmes et al., (1995)[31]*	4 group sessions of a sleep management training class incorporating CBT-I principles. Length and therapist qualifications not reported.	-	-	-	-	-
Myers et al., (2011)[32]*	4 x 1hr individual CBT-I sessions led by clinical psychologist.	-	-	-	-	-
Novak et al., (2020)[33]**	Small group CBT-I; session number, length and lead therapist qualifications not reported.	X	X	-	-	-
Waters et al., (2020)[34]**	4 x 1.5hr group CBT-I sessions led by a trained psychology student.	X	X	-	-	X
Taylor et al., (2022)[35]	6 x 30min sessions delivered by mobile app. Support available from a trainee clinical psychologist.	X	X	-	-	-

Note: 'X' indicates that this adaptation was made to treatment. Papers have been marked by a single asterisk (*) where an attempt was made to contact their authors to clarify further details about adaptations, but contact was not achieved. Papers have been marked by a double asterisk (**) where the authors responded to enquiries.
Waite et al.'s paper[53] not included here since it is a qualitative analysis of participants in Bradley et al.'s trial[29], and thus reports the same intervention.
Studies are presented in order of type (RCTs, non-randomised trials with a control group, single-arm trials, mixed-methods, qualitative).

2. Quality Appraisal

Table 4 shows risk of bias ratings for RCT and non-RCT studies. Four RCTs were rated as low risk across all the Cochrane domains due to rigorous procedures in their design and analysis.[24–27] Strengths included that they were pre-registered and used blinded assessors. One RCT was evaluated as having 'some concerns' because assessment was not blinded.[37] The non-RCTs fell almost entirely into 'serious' and 'critical' categories. Their greatest area of weakness was the potential for confounding. All studies but two lacked a control group,[28,34] and of these two, only one attempted to deal with confounding by controlling for differences in baseline characteristics between the treatment and control group.[28] This study was also the only non-RCT to be use a blinded assessor and hence to be rated as 'low' on the domain of outcome measurement.[28]

Another area of weakness in the non-RCTs was the selection of participants. Without random assignment to a control, bias was likely introduced by participants self-selecting into the study or being selected by clinicians. Risk of bias due to missing data was highly variable across the non-RCTs; whilst several studies reported zero drop-out, two studies had critically high levels of drop-out, presumably reflecting the lack of control in the acute inpatient setting regarding patients' prognosis and discharge.[30,33] These two studies also suffered greater potential for confounding where it was likely that patients were receiving psychiatric stabilisation and medication changes alongside CBT-I.[30,33]

Checklist ratings for the two qualitative thematic analyses[35,53] are found in the Appendix. A strength they shared was that themes were well-evidenced and gave rise to actionable outcomes – they highlighted participants' experiences of CBT-I which led to practical implications for how treatment could be improved. However, neither paper specified what kind of thematic analysis was used or explained their theoretical orientation. The

authors also did not ‘own their perspectives’ regarding their social and personal positioning, which was arguably important where they were representing the voices of a marginalised group to which they may not have belonged.

Table 4 *Cochrane risk of bias appraisal of randomised and non-randomised trials*

RCT studies			Randomis -ation	Deviations from interventions	Missing outcome data	Measurement of outcome	Selection of result	Overall risk of bias
Batalla-Martin et al. (2023)[37]			Low	Low	Low	Some concerns	Low	Some concerns
Freeman et al., (2015)[24]			Low	Low	Low	Low	Low	Low
Sheaves et al., (2018)[25]			Low	Low	Low	Low	Low	Low
Sheaves et al., (2019)[26]			Low	Low	Low	Low	Low	Low
Waite et al., (2023)[27]			Low	Low	Low	Low	Low	Low
Non-RCT studies	**Confound -ing**	**Selection of participants**	**Intervention classification**	**Deviations from interventions**	**Missing outcome data**	**Measurement of outcome**	**Selection of result**	**Overall risk of bias**
Hwang et al., (2019)[28]	Moderate	Moderate	Low	Low	Low	Low	Low	Moderate
Bradley et al., (2018)[29]	Serious	Serious	Low	Low	Low	Moderate	Low	Serious
Haynes et al., (2011)[30]	Critical	Critical	Low	Serious	Critical	Moderate	Low	Critical
Holmes et al., (1995)[31]	Serious	Serious	Low	No information	Low	Serious	Serious	Critical
Myers et al., (2011)[32]	Serious	Serious	Low	Low	Low	Moderate	Low	Serious
Novak et al., (2020)[33]	Critical	Serious	Low	No information	Critical	Moderate	Low	Critical
Waters et al., (2020)[34]	Serious	Serious	Low	Low	Critical	Moderate	Moderate	Critical
Taylor et al., (2022)[35]	Serious	Serious	Low	Moderate	Serious	Moderate	Low	Critical

Note. In the table, studies are grouped by whether they are randomised controlled trials (RCTs) or non-randomised trials (non-RCTs). Non-RCTs are further ordered by type (controlled, single-arm, mixed-methods).
The Cochrane RoB tool was used to assess RCTs and the Cochrane ROBINS-I tool was used to assess non-RCTs. Both tools rate studies based on their risk of bias across specific domains and overall. The RoB rates each study using three categories: low, some concerns, or high risk of bias. The ROBINS-I tool rates the risk of bias for each study using five categories: low, moderate, serious, critical or no information.

3. Treatment effectiveness[1]

3.1 Sleep problems

Table 5 shows the effectiveness of targeting sleep problems using the primary sleep measure, with studies organised by their risk of bias category rating. The sign of effect sizes has been changed such that a positive effect ($d > 0$) indicates an improvement in symptoms over the course of treatment. If the study includes a control group, a positive effect size indicates a relative improvement in symptoms in the treatment group compared to the control.

[1] Of note, ranges of effect size presented in the text of this section exclude Hwang et al.'s figures.[28] This is because the effect sizes that we calculated (from their reported means and standard deviations) were extremely large and did not match what was reported in their paper. The authors were contacted about this, but no response was received. It was observed that their reported standard deviations were unusually small, and it was queried whether they were possibly reporting standard errors instead. Caution is advised in interpreting Hwang et al.'s results.

Table 5 *Effect of CBT on sleep problems*

Study	Risk of bias	Primary Outcome	Post-treatment; d, CI 95%	Follow-up; d, CI 95%	
Batalla-Martin et al. (2023)[37]	Some concerns	ISI	1.46 [0.63, 2.29]	Three months	1.98 [1.01, 2.94]
Freeman et al., (2015)[24]	Low	ISI	1.32 [0.58, 2.05]	Three months	0.82 [0.18, 1.46]
Sheaves et al., (2018)[25]	Low	ISI	0.92 [0.20, 1.63]	Two weeks Ten weeks	0.79 [0.08, 1.50] 0.75 [0.01, 1.51]
Sheaves et al., (2019)[26]	Low	DDNSI	0.95 [-0.04, 1.93]	One month	1.07 [-0.04, 2.17]
Waite et al., (2023)[27]	Low	ISI	1.96 [0.99, 2.93]	Three months	1.27 [0.47, 2.07]
Hwang et al., (2019)[28]*	Moderate	ISI	Calc: 9.42 [6.68, 12.15]* Rep: 0.5, $p < 0.001$	One month	Calc: 9.96 [7.07, 12.85]* Rep: 0.5, $p < 0.001$
Bradley et al., (2018)[29]	Serious	ISI	$d > 0$	One month	$d > 0$
Haynes et al., (2011)[30]	Serious	ISI	$d > 0$	One month	$d > 0$
Holmes et al., (1995)[31]	Critical	ISI	$d > 0$	None	
Myers et al., (2011)[32]	Critical	Subjective sleep rating	$d > 0$	None	
Novak et al., (2020)[33]	Critical	ISI	$d > 0$	None	
Waters et al., (2020)[34]	Critical	PSQI	$d > 0$	None	
Taylor et al., (2022)[35]	Critical	ISI	$d > 0$	None	

Note: This table reports the effect of CBT in alleviating sleep problems as assessed by the primary sleep measure. Studies are organised by their risk of bias category, from low to increasingly severe.
Cohen's d between-group effect sizes are shown for controlled studies, whereas the direction of the effect (pre-treatment vs post-treatment) is simply shown for uncontrolled studies. Results have been transformed such that an improvement in sleep problems in the treatment group (relative to the control group, if present) over the course of treatment is represented by a positive change score (i.e. $d > 0$). A deterioration in sleep problems in the treatment group (relative to a control group, if present) is represented by a negative score (i.e. $d < 0$).
The text in brackets refers to the 95% Confidence Interval of the effect size.
For Hwang et al.'s study,[28] 'Calc' refers to the Cohen's d effect size calculated from their reported means and SDs, whereas 'Rep' refers to the Cohen's d effect size reported in their paper. The asterisk (*) indicates to treat their findings with caution given the large discrepancy between these estimates.
Primary measures include: ISI = Insomnia Severity Index,[55] DDNSI = Disturbing Dream and Nightmare Severity Index,[56] and PSQI = Pittsburgh Sleep Quality Index.[57] Holmes et al.[31] report a subjective sleep rating that is extracted from a sleep log published elsewhere[58] but provide minimal further details. Where studies did not specify a single primary sleep measure, the most relevant measure to evaluate the intervention (e.g. an insomnia measure for a study of CBT-I) is reported.

Most studies (76.9%) chose the Insomnia Severity Index (ISI)[55] as their primary measure, whereas other studies chose to measure sleep quality or nightmares. Every study demonstrated positive effects, suggesting improvement in sleep problems following treatment. Effect sizes were calculated for studies with a control group, and ranged from medium to very large immediately post-intervention [$d = 0.50$ to 1.96]. These calculated effects were all significant (i.e. confidence intervals did not overlap zero), except for one study[26] where the lower boundary of the wide confidence interval was marginally negative. This was the smallest RCT (only 12 participants in each group) and was likely underpowered.

It was also the only study to investigate IRT, as opposed to CBT-I, and its primary outcome was the DDNSI[56] (a measure of nightmare frequency and severity) which may not be directly comparable to the other studies whose primary outcomes were sleep quality and insomnia. Over follow-up periods, effect sizes were broadly maintained [d = 0.50 to 1.98], with the longest follow-up taking place at three months.

3.2 Psychotic symptoms

The impact of CBT-I on psychosis is presented in Table 6. Psychosis was assessed in 69.2% of trials using a variety of measures relating to delusions, hallucinations and other symptoms. The effects were almost entirely positive, in the direction suggesting a relative improvement of symptoms. Indeed, most studies reported solely positive results (66.7%), however three of the studies found a mixed direction of effects across multiple measures or time-points.[24,25,27]

Calculated effect sizes ranged between *d* = -0.39 to 0.79 across post-intervention and *d* = -0.39 to 0.79 across follow-up, and confidence intervals were generally very wide which suggests high levels of uncertainty around these estimates. These effects were almost entirely non-significant since confidence intervals overlapped zero.

Table 6 *Effect of CBT on psychotic symptoms*

Study	Risk of bias	Measures	Post-treatment; d, CI 95%	Follow-up; d, CI 95%	
Batalla-Martin et al. (2023)[37]	Some concerns	None			
Freeman et al., (2015)[24]	Low	PSYRATS (delusions) PSYRATS (hallucinations) GPTS PANSS	0.16 [-0.43, 0.75] -0.33 [-0.93, 0.27] 0.21 [-0.39, 0.80] 0.40 [-0.21, 1.01]	Three months	-0.11 [-0.69, 0.47] -0.39 [-0.98, 0.20] 0.31 [-0.30, 0.92] 0.35 [-0.27, 0.96]
Sheaves et al., (2018)[25]	Low	PANSS Total PANSS (positive symptoms) PANSS (negative symptoms) PANSS (psychopathology)	0.32 [-0.32, 0.97] 0.04 [-0.60, 0.67] 0.35 [-0.30, 1.00] 0.30 [-0.34, 0.95]	Two weeks	0.06 [-0.58, 0.71] -0.02 [-0.67, 0.63] 0.50 [-0.17, 1.17] -0.16 [-0.81, 0.49]
				Ten weeks	0.34 [-0.37, 1.05] 0.19 [-0.51, 0.89] 0.69 [-0.06, 1.43] 0.06 [-0.63, 0.76]
Sheaves et al., (2019)[26]	Low	CAPS GPTS	0.03 [-0.86, 0.91] 0.72 [-0.24, 1.68]	One month	0.02 [-0.96, 0.91] 0.65 [-0.35, 1.65]
Waite et al., (2023)[27]	Low	CAARMS ČEFSA GPTS-A GPTS-B SPEQ	0.45 [-0.19, 1.10] 0.11 [-0.56, 0.79] 0.40 [-0.29, 1.08] 0.36 [-0.33, 1.03] -0.05 [-0.70, 0.61]	Three months	0.31 [-0.36, 0.97] 0.24 [-0.42, 0.91] 0.68 [-0.02, 1.37] 0.76 [0.06, 1.47] 0.19 [-0.46, 0.84]
Hwang et al., (2019)[28]*	Moder -ate	PSYRATS Total PSYRATS (hallucinations) PSYRATS (delusions)	Calc: 2.69 [1.77, 3.61]* Not reported d in paper as p > 0.1 Calc: 1.26 [0.64, 1.88]* Not reported d in paper as p > 0.1 Calc: 2.38 [1.53, 3.22]* Not reported d in paper as p > 0.1	One month	Calc. 3.08 [2.07, 4.09]* Rep. 0.1, p = 0.053 Calc. 1.63 [0.94, 2.31]* Not reported d in paper as p > 0.1 Calc: 2.58 [1.69, 3.47]* Not reported d in paper as p > 0.1
Bradley et al., (2018)[29]	Serious	GPTS SPEQ	d > 0 d > 0	One month	d > 0 d > 0
Haynes et al., (2011)[30]	Serious	GPTS-A GPTS-B PSYRATS CAPS	d > 0 d > 0 d > 0 d > 0	One month	d > 0 d > 0 d > 0 d > 0
Holmes et al., (1995)[31]	Critical	None			
Myers et al., (2011)[32]	Critical	None			
Novak et al., (2020)[33]	Critical	None			
Waters et al., (2020)[34]	Critical	MINI-p	d > 0	None	
Taylor et al., (2022)[35]	Critical	R-GPTS-A R-GPTS-B SPEQ	d > 0 d > 0 d > 0	None	

Note: This table reports the effect of CBT in alleviating symptoms of psychosis. Studies are organised by their risk of bias category, from low to increasingly severe.

Cohen's d between-group effect sizes are shown for controlled studies, whereas the direction of the effect (pre-treatment vs post-treatment) is simply shown for uncontrolled studies. Results have been transformed such that an improvement in psychotic symptoms in the treatment group (relative to the control group, if present) over the course of treatment is represented by a positive change score (i.e. d > 0). A deterioration in psychotic symptoms in the treatment group (relative to a control group, if present) is represented by a negative score (i.e. d < 0).
The text in brackets refers to the 95% Confidence Interval of the effect size.
For Hwang et al.'s study,[28] "Calc. d" refers to the Cohen's d effect size calculated from their reported means and SDs, whereas "Rep. d" refers to the Cohen's d effect size reported in their paper. The asterisk (*) indicates to treat their findings with caution given the large discrepancy between these estimates.
All reported assessment measures of psychosis in each paper are presented. Brackets denote subscales in the measures column of the table; for example, "PSYRATS (delusions)" refers to the delusions subscale of the PSYRATS.
Psychotic measures include: Cardiff Anomalous Perceptions Scale (CAPS)[59]; Černis Felt Sense of Anomaly (ČEFSA)[60]; Comprehensive Assessment of At-Risk Mental States (CAARMs)[61]; Green Paranoid Thoughts Scale (GTPS)[62] which contains subscales of 'Ideas of Reference' (GPTS-A) and 'Ideas of Persecution' (GPTS-B); Adapted Mini International Neuropsychiatric Interview – Psychosis section (MINI-p)[63]; Positive and Negative Symptoms Scale (PANSS)[64] which contains subscales relating to positive symptoms, negative symptoms, and general psychopathology; Psychotic Symptom Rating Scale (PSYRATS)[65] which contains subscales relating to delusions and hallucinations, Revised-Green Paranoid Thoughts Scale (R-GPTS)[66] which contains subscales of 'Ideas of Reference' (GTPS-A) and 'Ideas of Persecution' (GPTS-B) subscales; Specific Psychotic Experiences Questionnaire (SPEQ).[67]

3.3 Other clinical outcomes

Several trials (76.9%) evaluated whether CBT-I led to clinical improvements beyond sleep and psychosis. The most commonly assessed domains were anxiety, depression, quality of life and wellbeing. Table 7 displays effect sizes where a positive sign denotes a relative improvement in outcomes following treatment. Most effect sizes were positive suggesting that participants' outcomes improved. Nevertheless, these effects were largely non-significant. Confidence intervals were very wide and overlapped zero, suggesting that estimates were quite uncertain.

The magnitude of effect sizes ranged from around zero to very large. The largest positive effect size was d = 1.58 on a measure of functioning at one-month follow-up.[27] Of note, a large negative effect size of *d* = -0.92 was obtained on a suicidal ideation subscale at one-month follow-up.[26] This negative effect size resulted from the control group having a much higher starting mean suicidal ideation at baseline that decreased, whereas the suicidal ideation for the treatment group remained relatively stable from week 0 to week 8.

Across trials, there were a wide variety of outcomes and measures making it relatively challenging to synthesise. Depression and anxiety were measured in seven papers (53.8% of trials). A positive direction of effects was observed at all time-points immediately post-

intervention for depression (range: d = 0.31 to 0.37) and anxiety (range: d = 0.41 to 0.61), suggesting in an improvement in symptoms. These all remained positive at follow-up, except for one study[26] which found negative effects. At one month follow-up the range was d = -0.36 to 0.67 for depression, and d = -0.06 to 0.95 for anxiety.

Mental wellbeing was assessed in five papers (38.5% of trials) and similarly, all effects were positive immediately post-intervention (range: d = 0.39 to 0.56) which suggests that participants' mental wellbeing improved. They were also positive at follow-up except for one study[26] and ranged from d = -0.15 to 0.74. Finally, effect sizes for changes in quality of life (measured in five papers, representing 38.5% of trials) were all positive except in one study[26] at both post-intervention (range: d = -0.39 to 0.93) and follow-up (range: d = 0 to 1.46).

Table 7 *Effect of CBT on other clinical outcomes*

Study	Risk of bias	Measures	Post-treatment; d, CI 95%	Follow-up; d, CI 95%	
Batalla-Martin et al. (2023)[37]	Some concerns	Quality of life (ED-5Q VAS)	0.93 [0.21, 1.65]	Three months	1.38 [0.57, 2.19]
Freeman et al., (2015)[24]	Low	Psychological recovery (CHOICE)	0.57 [-0.06, 1.19]	Three months	0.26 [-0.35, 0.88]
		Quality of life (ED-5D-5L)	0.57 [-0.06, 1.19]		0.43 [-0.17, 1.04]
		Fatigue (MFI)	0.71 [0.07, 1.35]		0.49 [-0.13, 1.10]
		Mental wellbeing (WEMWBS)	0.41 [-0.20, 1.02]		0.74 [0.09, 1.38]
Sheaves et al., (2018)[25]	Low	Suicidal ideation (BSS)	0.09 [-0.55, 0.72]	Two weeks	-0.03 [-0.67, 0.62]
		Psychological distress (CORE-10)	0.70 [0.02, 1.39]		0.46 [-0.20, 1.13]
		Quality of life (ED-5Q)	-0.39 [-1.04, 0.26]		0 [-0.65, 0.65]
		Mental wellbeing (WEMWBS)	0.39 [-0.27, 1.04]		0.41 [-0.25, 1.07]
		Mania (YMRS)	0.33 [-0.31, 0.98]		0.05 [-0.59, 0.70]
					-0.19 [-0.89, 0.51]
				Ten weeks	0.28 [-0.43, 0.98]
					0.07 [-0.63, 0.76]
					0.48 [-0.25, 1.20]
					0.42 [-0.29, 1.14]
Sheaves et al., (2019)[26]	Low	Suicidal ideation (BSS)	-0.44 [-1.36, 0.47]	One month	-0.92 [-1.99, 0.14]
		Anxiety (DASS-21 anxiety)	0.41 [-0.50, 1.32]		-0.06 [-0.99, 0.88]
		Depression (DASS-21 depression)	0.31 [-0.59, 1.20]		-0.36 [-1.31, 0.60]
		Stress (DASS-21 stress)	0.75 [-0.22, 1.71]		-0.16 [-1.10, 0.78]
		Dissociation (DES-B)	0.66 [-0.29, 1.60]		0.36 [-0.60, 1.31]
		Mental wellbeing (WEMWBS)	0.56 [-0.37, 1.50]		-0.15 [-1.09, 0.79]
Waite et al., (2023)[27]	Low	Suicidal ideation (CSSRS)	0.50 [-0.22, 1.22]	Three months	0.46 [-0.26, 1.18]
		Negative schemas (BCSS-negative)	0.51 [-0.19, 1.21]		0.67 [-0.03, 1.38]
		Positive schemas (BCSS-positive)	0.75 [0.02, 1.48]		0.51 [-0.17, 1.20]
		Anxiety (DASS-21 anxiety)	0.61 [-0.10, 1.33]		0.95 [0.19, 1.71]
		Depression (DASS-21 depression)	0.37 [-0.32, 1.06]		0.67 [-0.04, 1.39]
		Stress (DASS-21 stress)	0.72 [-0.01, 1.45]		0.58 [-0.12, 1.29]
		Worry (DWQ)	0.37 [-0.32, 1.06]		0.98 [0.23, 1.73]
		Quality of life (EQ-5D-5L-Index)	0.36 [-0.33, 1.05]		0.36 [-0.31, 1.03]
		Quality of life (EQ-5D-5L-VAS)	0.43 [-0.26, 1.13]		0.62 [-0.08, 1.32]
		Agoraphobic avoidance (O-AS-A)	0.51 [-0.19, 1.22]		0.58 [-0.12, 1.29]
		Agoraphobic distress (O-AS-D)	0.51 [-0.19, 1.22]		0.84 [0.08, 1.59]
		Depression (PHQ-15)	0.42 [-0.28, 1.11]		0.46 [-0.22, 1.15]
		Recovery (QPR)	0.21 [-0.47, 0.89]		0.33 [-0.34, 1.00]
		Quality of life (ReQol)	0.81 [0.05, 1.56]		1.46 [0.61, 2.32]
		Functioning (WSAS)	0.87 [0.12, 1.63]		1.58 [0.71, 2.46]
Hwang et al., (2019)[28]	Moder -ate	Anxiety (ASI)	Calc: 1.66 [1.00, 2.35] Not reported d in paper as p > 0.1	One month	Calc: 2.78 [1.84, 3.72] Not reported d in paper as p > 0.1
		Depression (BDI)	Calc: 3.48 [2.37, 4.60] Not reported d in paper as p > 0.1		Calc: 4.54 [3.15, 5.93] Not reported d in paper as p > 0.1
Bradley et al., (2018)[29]	Serious	Anxiety (DASS-21 anxiety)	d > 0	One month	d > 0
		Depression (DASS-21 depression)	d > 0		d > 0
		Stress (DASS-21 stress)	d > 0		d > 0
		Mental wellbeing (WEMWBS)	d > 0		d > 0
		Functioning (WSAS)	d > 0		d > 0
Haynes et al., (2011)[30]	Serious	Anxiety (DASS-21 anxiety)	d > 0	One month	d > 0
		Depression (DASS-21 depression)	d > 0		d > 0

Holmes et al., (1995)[31]	Critical	None		
Myers et al., (2011)[32]	Critical	None		
Novak et al., (2020)[33]	Critical	None		
Waters et al., (2020)[34]	Critical	Impulsivity (BIS) Quality of life (MANSA) Anxiety and depression (PHQ-4)	d > 0 d > 0 d > 0	None
Taylor et al., (2022)[35]	Critical	Anxiety (DASS-21 anxiety) Depression (DASS-21 depression) Stress (DASS-21 stress) Mental wellbeing (WEMWBS) Functioning (WSAS)	d > 0 d > 0 d > 0 d > 0 d > 0	None

Note: This table reports the effect of CBT in alleviating other clinical symptoms. Studies are organised by their risk of bias category, from low to increasingly severe. Cohen's d between-group effect sizes are shown for controlled studies, whereas the direction of the effect (pre-treatment vs post-treatment) is simply shown for uncontrolled studies. Results have been transformed such that an improvement in clinical symptoms in the treatment group (relative to the control group, if present) over the course of treatment is represented by a positive change score (i.e. d > 0). A deterioration in clinical symptoms in the treatment group (relative to a control group, if present) is represented by a negative score (i.e. d < 0).
The text in brackets refers to the 95% Confidence Interval of the effect size.
For Hwang et al.'s study,[28] "Calc. d" refers to the Cohen's d effect size calculated from their reported means and SDs, whereas "Rep. d" refers to the Cohen's d effect size reported in their paper. The asterisk (*) indicates to treat their findings with caution given the large discrepancy between these estimates.
The Measures column in the table states the clinical symptom that was assessed, followed by the assessment measure in brackets. Assessment measures include: Anxiety Sensitivity Index (ASI[68]; Brief Core Schema Scale (BCSS)[69]; Beck Depression Inventory (BDI)[70]; Beck Suicide Scale (BSS)[71]; CHoice of Outcome In CBT for psychoses (CHOICE)[72]; Clinical Outcomes in Routine Evaluation 10-item scale (CORE-10)[73]; Columbia Suicide Severity Rating Scale (CSSRS)[74]; Depression Anxiety and Stress Scale–21-item (DASS-21)[75] version with subscales for anxiety (DASS-21 anxiety), depression (DASS-21 depression) and stress (DASS-21 stress); Brief Dissociative Experiences Scale (DES-B)[76]; Dunn Worry Questionnaire (DWQ)[77]; Euroqol 5D questionnaire (ED-5Q)[78]; Euroqol 5D questionnaire 5 Level version (ED-5D-5L)[79], Multidimensional Fatigue Inventory (MFI)[80]; Oxford Agoraphobic Avoidance Scale (O-AS)[81]; Patient Health Questionnaire (PHQ-9)[82]; Brief Patient Health Questionnaire (PHQ-4)[82]; Questionnaire about the Process of Recovery (QPR)[83]; Recovering Quality of Life (ReQol)[84]; Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)[85]; Work and Social Adjustment Scale (WSAS)[86]; Young Mania Rating Scale (YMRS).[87]

4. Treatment acceptability

4.1 Qualitative evidence

Treatment acceptability was evaluated through several different means. Relevant themes and quotes from two thematic analyses are organised in Table 8.[35,53] When organising themes, a realist philosophical orientation was adopted. Quotes were interpreted literally in relation to the question, 'what does the data tell us about participants' experiences of CBT-I treatment?'

Five themes and twelve subthemes were developed from the data. The first theme was 'initial negative expectations about CBT-I.' It appears that some participants were initially apprehensive about engaging in treatment. They were concerned about being prescribed

"little methods" such as counting sheep, and it was hard to see what difference treatment could make. In addition, some participants feared being "judged," and that the content would be "depressing."

The themes 'CBT-I helped develop skills for improving sleep' and 'improved sleep had ripple effects on other areas of life' show that many participants benefitted from treatment. Participants described learning what factors influenced their sleep and acquiring strategies to enhance their sleep quality. The consequences of better sleep could be widespread. Increased energy allowed them to engage in meaningful tasks, and they noticed improvements in mood, anxiety, and psychotic experiences.

Participants also mentioned a theme of 'limitations.' Strategies were not always successful and could require trial and error. Some strategies demanded high commitment which could be difficult to maintain long-term. These issues related to the theme that 'the format of CBT-I impacted on engagement.' Some participants found booklets repetitive to complete[53]; others found that technical aspects of a mobile app could pose a barrier.[35] Across both studies, participants appreciated when treatment was collaborative and interactive. Participants in one study[53] praised how therapists tailored therapy to accommodate their preferences and lifestyle when sessions were delivered in-person. By contrast, participants in the other study[35] saw the lack of interaction with a therapist as a disadvantage of receiving treatment through a mobile-app. Participants in both studies commented on how they valued regular access to support. One study[53] achieved this by allowing participants to email therapists and access workbooks, whereas participants in the other study[35] mentioned the benefits of being able to access resources at any time via the app.

Table 8 *Qualitative themes and quotes regarding acceptability*

Theme	Subtheme	Waite et al. (2018)[53] Illustrative quotes	Taylor et al. (2022)[35] Illustrative quotes
There were initial negative expectations about CBT-I	Doubts about effectiveness	P8 (F) *"I didn't think it would work at first"* P1 (F) "*some people [I'd heard in class], they were just like, oh, just sleep late, just, you know, like hot chocolate, or just count some sheep, and I was like, "Oh, it's just going to be that all over again!"* P2 (M) "*little methods, such as like writing things down, and how actually when you think about that, that seems like it's not really going to do anything*"	P10 (M) "*I'm sceptical about this kind of stuff"*
	Fear of having a negative experience	P9 (F) "*I was scared that I was going to kind of be judged for like my sleep and stuff"* P10 (M) *"my worry was it was just going to be like some depressing. . . fest [laughing]"*	–
CBT-I helped develop skills for improving sleep	Increased awareness of factors affecting sleep	P11 (F) *"it sort of opened. . .well, not opened my eyes as such, but gave me a wider understanding of what actually affects my sleep"* P9 (F) "*I learnt quite a lot, and it was just kind of. . . some obvious stuff, but it was said kind of in a way that I hadn't thought about before and I hadn't thought like how it was really affecting my sleep*"	P7 (M) "*There wasn't like a set routine for sleep until I learned about how a routine could help"*
	Practical skills were acquired	P10 (M) *"I'm using a lot of the techniques that I was given through this"* P9 (F) *"I've been using the quarter-hour rule"*	–
	Sleep quality was enhanced	P4 (M) *"I've been falling asleep quicker and spending more time asleep"* P3 (F) "*waking up no more than twice during the night and not feeling concerned when I wake up"* P7 (F) *"I do sleep now [laughing], at like the proper time"*	P1 (F) *"think that my sleep has improved… by, I'd say, 40%"* P10 (M) "*Yeah, it's been beneficial for me"*
Improved sleep had a ripple-effect on other areas of life	Better sleep was linked to better functioning	P11 (F) "*If I get a decent night's sleep, I'm not tired during the day*" P10 (M) "*I had originally thought it was more down to stuff during the day, like how much energy you had used on this activity or this activity or, again, what you're eating, but I didn't realize how much was down to sleep*"	P7 (M) *"I can tell the impact in my everyday life… in helping take care of the girls. Yeah, I mean in my everyday interaction with people"*
	Psychotic and other clinical symptoms showed improvements	P8 (F) *"since I've been sleeping better, my, visual things have like stopped"* P3 (F) *"I used to hear things quite a lot, like walking down the street, and that's subsided, almost entirely."* P11 (F) *"the depression has been the big step. It's really eased. It's nowhere near where it used to be*" P3 (F) *"I suffer from depression so it's not going to be high all the time, but because it's so highly linked with sleep, when I sleep better, I tend to have a better day"* P9 (F) *"I find it also has helped my anxiety a lot. I'm able to cope with situations a lot better and kind of stay in control, which I think has come from me gaining control of my sleep a bit"*	–
The format of CBT-I impacted on engagement	An interactive and collaborative approach facilitated motivation to engage	P3 (F) *"[my therapist] was very happy to work through any other problems that I was having at the same time"* P3 (F) "*my therapist] would give me a technique and then I'd come back and say 'This worked for me; this didn't work for me,' and then we could kind of work around that"* P1 (F) "*It is based, like, what you do during the day, diet, social life, like taking all that into account along with your sleep*"	P7 (M) *"The fact that there's an app, and there's someone like [trial therapist] to like, keep me in tack or like keep checking up on my progress. I think those, those elements made the experience better for me"* P4 (M) [in discussing a disadvantage of the app] *"…if it's a recording of [therapist] tutoring, or [therapist] walking through the steps, I see that more interactive. Because […] I see somebody there trying to tell me what to do. I mean it will make more meaning in that sense"*

	Readily accessible support was valued	P3 (F) "*if I was struggling during the week, I could just email and say, "Having a tough day–any advice?" sort of thing. That was really helpful*" P5 (M) "*if I sort of felt like I was having sleep problems again, I can go back to this and go, well, here's things I can use*"	P1 (F) *"The advantage of using the app was it, it's there any time of the today, you can refer to it."*
	Elements were experienced as impractical and not 'user-friendly'	P8 (F) "*the booklet things, did confuse me a bit*" P3 (F) and P5 (M) "*repetitive*" P9 (F) "*tedious*"	P6 (F) "*because I'm, I'm not a technical person at all […], that's the only downside…"* P3 (M) "*the amount of work you had to do to get a quick reminder of what you wanna do […] you'd have to go through quite a bit just to get to the section that you want"*
There were limitations to treatment	Strategies could be difficult to implement and maintain	P8 (F) *"[my therapist] said if…I don't fall asleep within 15min, like to get back up again…once I'm in bed, I just don't really want to get back out of it*". P10 (M) "*every time that I try to implement some of those strategies, there is a definite benefit, but it's quite difficult to maintain*"	P6 (F) *"…it's quite a lot in a short time."* P3 (M) "*it's also a lot to keep doing everyday…"*
	Effects could be small or unsuccessful	P10 (M) *"they don't solve the situation, too much, but it does still help"* P6 (M) "*all the other things that we'd done before I felt they hadn't worked as much, and then we did the wind-down routine and it really helped*" P4 (M) "*the stuff about the nightmares. For some, it does work, but for others, it doesn't*"	–

Note: Themes and subthemes based on research in Waite et al.[53] and Taylor et al.[35] with accompanying participant quotes. Participant number and sex is indicated.

4.2 Drop-out and adverse events

Table 9 presents rates of drop-out and adverse events. Drop-out was reported in 76.9% of trial studies. Three studies observed 0% drop-out from treatment.[25,26,32] The highest rates were 25% and 27.8% in two studies which were set in acute inpatient wards and had weaker control for extraneous variables and follow-up with participants.[30,33] Reasons for drop-out were not generally stated, and it was unclear whether any studies sought to follow-up with participants who withdrew to find out the reasons for this.

Adverse events were commonly defined by papers as serious incidents that led to in-patient hospitalisation, disability or were life-threatening. The proportion of participants who had adverse events in the treatment group was slightly numerically higher (7.6% of participants) than in the control (4.8%). However, it is not possible to draw conclusions from

this as total numbers were low, and few studies reported this data. All studies deemed adverse events were unrelated to treatment.

Table 9 *Rates of drop-out and adverse events*

Study	Drop out (%)	Adverse events	
		Intervention	**Control**
Batalla-Martin et al. (2023)[37]	0%	Not reported	–
Freeman et al., (2015)[24]	8.3%	3 participants (12.5%) had 5 adverse events	0
Sheaves et al., (2018)[25]	0%	1 participant (5.0%)	0
Sheaves et al., (2019)[26]	0%	2 participants (16.7%)	1 participant (8.3%)
Waite et al., (2023)[27]	4.8%	4 participants (19.0%) had 5 adverse events	2 participants (11.0%) had 3 adverse events
Hwang et al., (2019)[28]	Not reported	Not reported	Not reported
Bradley et al., (2018)[29]	8.3%	0	–
Haynes et al., (2011)[30]	25.0%	Not reported	–
Holmes et al., (1995)[31]	Not reported	Not reported	–
Myers et al., (2011)[32]	0%	0	–
Novak et al., (2020)[33]	27.8%	None classed as related to the study but number not stated	–
Waters et al., (2020)[34]	20.0%	Not reported	Not reported
Taylor et al., (2022)[35]	15.0%	0	–

4.3 Questionnaire data

Three studies offered treatment satisfaction questionnaires. Participants in one study[26] provided therapy satisfaction ratings to an independent assessor regarding their experiences of IRT. The median score was 9 out of 10 (interquartile range IQR 6.75-10), suggesting high satisfaction. These participants were also given a client satisfaction questionnaire; 56.2% were “very satisfied” and the remainder responded that they were “mostly satisfied.” Another study[27] adopted the Abbreviated Acceptability Rating Profile (AARP)[88] to evaluate CBT-I in young people at ultra-high risk of psychosis. The average score was 45.8 (SD 4.1) and the median was 48 (IQR 46-68), out of a possible maximum of 48. Participants in the third study[35] used the ‘Experience Feedback Questionnaire’.[89] The proportion of participants who

rated the app-based CBT-I intervention as ‘enjoyable to use’ was 81.8%, ‘not disruptive to life’ was 72.7%, ‘not stopping usual activities’ was 81.1%, ‘not embarrassing’ was 100%, and ‘easy to remember to use’ was 45.4%.

Discussion

This review addresses several gaps in the research literature regarding the evidence-base for CBT interventions for sleep problems in people with, or at risk of, psychosis. We first summarise the content of these CBT interventions, before proceeding to discuss their clinical effectiveness and acceptability.

CBT interventions

Insomnia (i.e. difficulties with falling or staying asleep) was the primary focus of all CBT interventions except for one study in which nightmares were targeted. Standardised protocols were frequently adapted by omitting sleep restriction on the basis of research linking sleep loss to psychotic experiences.[7,39] Other adaptations included strategies to manage delusions and hallucinations that were preventing patients from falling asleep at night-time, problems with oversleeping to escape voices, and fears of bed caused by past adverse experiences. Several CBT-I trials also offered to address nightmares, which reflects that insomnia and nightmares can be co-occurring problems for people with psychosis.[25]

Effectiveness on sleep outcomes

The results indicate that CBT is an effective treatment for sleep problems in people with psychosis. All studies observed an improvement on their primary sleep measure, although the quality of studies was highly variable. Out of the four RCT studies that were judged to be at low risk of bias, effect sizes ranged from medium to very large post-treatment and largely persisted at three-months’ follow-up. These effects were also all significant except

for the smallest RCT, which may have been underpowered to detect this. The results are congruent with meta-analytic evidence that CBT-I and IRT are effective for healthy adults.[90,91] CBT is the recommended first-line treatment with insomnia and nightmares in clinical guidelines,[18,21] and the findings lend support to its use for people with psychosis.

Effectiveness on psychotic outcomes

Eight studies measured psychotic symptoms such as hallucinations, delusions and paranoid thinking. Study quality ranged from low to critical risk of bias. Out of the four RCTs judged as low risk, effect sizes were small-to-medium in size and mostly positive. Yet, they were not significant, and some negative effects were obtained (i.e. suggesting either a deterioration in symptoms, or that the treatment group did not improve by as much as the control group). Accordingly, the evidence is inconclusive regarding whether CBT directed at sleep problems leads to improvements in psychotic symptoms.

Of note, in the research literature, a large-scale RCT involving a non-clinical student sample of 3755 participants found that CBT-I led to significant improvements in paranoia and hallucinations.[36] Their mediation analysis supported this interpretation, as improvements in sleep accounted for almost 40-60% of the change in hallucinations and paranoia after treatment. Given that it reported small effect sizes for paranoia and hallucinations, it is possible that the studies included in this review were insufficiently powered to detect significant results (the largest RCT had a sample size of only 50 participants). Alternatively, it is also possible that the link between sleep problems and psychosis is weaker than anticipated. This would have theoretical implications for Freeman's model[4] in which sleep dysfunction is a putative cause of persecutory delusions.

Effectiveness for other clinical outcomes

In relation to other clinical outcomes, the findings are harder to interpret, partly due to the range of outcomes and measures. The most popular clinical domains were depression, anxiety, mental wellbeing and quality of life, but each domain was only captured in a small number of studies. Effect sizes were frequently in the direction suggesting an improvement in clinical symptoms, but were mostly non-significant, and thus the evidence is not conclusive.

Many studies in the wider research literature report that poor sleep is linked to depression and anxiety.[92] People who are sleep deprived often describe increased negative emotions such as anger, frustration, sadness and irritability[93,94] and reduced ability to engage in meaningful activities.[95–97] Chronic lack of sleep is also associated with a reduced immune system functioning and other physical health problems.[98] Given these findings, we hypothesised that there would be widespread ripple-effects to treating people's sleep problems. A possible explanation for why this was not clearly found is that studies were underpowered to detect effects. Another possibility is that the intervention requires time to improve people's outcomes. If people have experienced years of sleep deprivation and consequently withdrawn from important areas of their life due to fatigue and low mood, it may take time for people to regain their energy, confidence, social networks and re-engage with meaningful activities. Future research should investigate longer-term outcomes.

Acceptability

The findings of this review indicate that CBT for sleep is acceptable. None of the adverse events that occurred across studies were related to treatment. A meta-analysis has estimated that ~13% of people with schizophrenia drop-out from psychosocial treatments,[99] yet the majority of studies included in our review reported lower rates, suggesting that

engagement was reasonably good. Whilst drop-out rates ranged from 0% to 27.8%, the highest rates were found in studies with weaker methodological designs where researchers may have struggled to follow-up participants. Of note, no studies consulted participants regarding their reasons for drop-out. This would have been helpful to determine the degree to which drop-out was a valid measure of acceptability or was affected by other changes in participants' circumstances.

The feedback from questionnaires and qualitative interviews provides stronger evidence. All, or nearly all, participants rated treatment as acceptable, enjoyable and satisfying. Qualitative themes suggested that they found it useful and experienced widespread improvements in other areas of life. Participants valued an interactive approach in which support was readily accessible and user-friendly. The high levels of commitment required (e.g. in maintaining a consistent sleep window) could be a challenge when implementing strategies long-term. A process of trial and error was sometimes required to find out what strategies worked for each participant personally.

Considerations around implementation

An individualised, biopsychosocial formulation-based approach is important to clarify what factors are maintaining each patient's difficulties.[100] The clinical picture on sleep problems in psychosis is complex and the underlying causes can be varied. Erratic sleep patterns have been documented amongst patients with psychosis.[101] Often, a lack of activities and social isolation results in circadian rhythm disruption where they go to sleep too early (advanced phase), too late, (delayed phase), or they oversleep (hypersomnia).[101,102] Lifestyle factors such as napping, poor diet and nicotine can also be disruptive. Some patients experience high levels of night-time worry (including paranoid fears) which can result in insomnia by increasing hyperarousal.[103] It is also important to note that medical sleep

disorders may be present, such as sleep apnoea, which shows elevated rates amongst schizophrenia patients.[104] Hypnagogic and hypnopompic hallucinations can be caused by sleep disorders, particularly narcolepsy, or can be a part of psychosis. Patients' medications may further affect their sleep.[17] For example, certain antidepressants may specifically increase rapid-eye-movement sleep and nightmares. Sedative side-effects of neuroleptic medications can cause daytime fatigue. Hypnotics are commonly prescribed alongside neuroleptics but issues with tolerance and rebound can occur.

Unlike the full range of sleep problems described, the reviewed studies focused on insomnia and nightmares. Indeed, most trials required participants to be stabilised on medication and excluded anyone with a medical sleep disorder. No studies analysed the influence of comorbid sleep disorders and/or medication effects. Clinicians should assess for other sleep disorders and consider the impact of any medications. CBT is unlikely to be indicated or effective for all sleep problems in psychosis, and its appropriateness should be considered in conjunction with existing good clinical practice, including working within a multidisciplinary context and a wider biopsychosocial formulation. We refer to a helpful guide for treating sleep problems in schizophrenia patients[17] that advises linking in with the wider clinical team to establish if a medication review is required and collaborating with relevant organisations (e.g. supported living to implement lifestyle changes). For certain sleep disorders, clinical guidelines advise medical interventions rather than CBT (e.g. continuous positive airway pressure (CPAP) machines for sleep apnoea).[23]

Strengths and Limitations

We provide a pre-registered, rigorous, systematic review of CBT interventions directed at sleep in people with psychosis. Our scope was comprehensive as we considered the full range of psychotic symptoms from at-risk state to full-blown disorders, both

quantitative and qualitative methods, a wide range of outcomes, and both efficacy and acceptability assessments. A further strength of our study is that a rigorous methodology was used to search for and critically appraise studies, with a second reviewer duplicating several stages of the analysis to ensure interrater reliability. A mixed-methods approach allowed us to triangulate different forms of evidence and strengthen the credibility of our conclusions. We note a recent quantitative-only review with fewer included studies, which was not pre-registered[105] and likely had inflated estimates, due to an inclusion of the same dataset twice[7,106] and an error in the reporting of another paper which we highlighted (see footnote[1], page 28). The robust, converging, and accurate results of our review have clear clinical implications for healthcare professionals seeking to support people with psychosis.

There were nonetheless several limitations in the ability to draw conclusions regarding treatment effectiveness. There were few high-quality studies. Sample sizes were small, resulting in large confidence intervals which may have meant they were underpowered to detect significant results. The wide variety of measures to assess psychosis and other clinical symptoms made it difficult to generalise. As stated in our pre-registered analysis plan, we intended to conduct a meta-analysis, however, there were too few comparable studies to yield meaningful results. Low numbers of studies also meant we could not identify factors influencing treatment effectiveness including patient-status (full diagnosis versus at-risk mental state), treatment focus (insomnia versus nightmares), specific adaptations, settings, comorbidity and medication effects.

The two qualitative studies included in this review added rich insights into participants' experiences. However, they did not explain the kind of thematic analysis they were performing or its theoretical underpinnings. This is a widespread problem in thematic research.[107] The current review adopted a realist/essentialist position in seeking to 'stay close'

to participants' literal meanings when evaluating treatment acceptability. However, limitations affecting the generalisability of our findings include small numbers of participants and researcher subjectivity influencing the selection of themes. It is unknown whether the results replicate across different settings, modes of treatment, and so on.

Finally, White Western participants tended to be over-represented, and ethnicity was not reported for some studies. This reflects wider problems of under-representation of ethnic groups in clinical research.[108] Migrant and ethnic minority status is associated with increased odds of psychotic disorder.[109] It would have been useful to investigate treatment effectiveness and acceptability across cultural contexts.

Future directions

Large-scale RCTs are necessary to establish psychotic and other clinical outcomes. Further investigations of treatment effectiveness and acceptability are also needed for people at-risk of psychosis, people from a range of cultural and ethnic backgrounds, and CBT for nightmares. It would be helpful to understand reasons for treatment drop-out, and longer-term consequences of treatment (i.e. greater than three months' follow-up). Many clinical practitioners do not routinely ask about patients about their sleep problems.[110] Indeed, clinical guidelines for psychotic disorders, including those by the National Institute of Clinical Excellence,[111] do not focus on sleep. Thus it would also be useful to investigate how to bridge the research-practice gap[112] and engage practitioners in assessing and offering CBT for sleep problems in this patient group.

Conclusion

This review shows that CBT is highly effective and acceptable in treating sleep problems in people with, or at risk of, psychosis. However, results were inconclusive

regarding the wider impact of treatment on patients' psychotic and other clinical symptoms. Clinicians should be aware that adaptations to standard treatment protocols are commonly made for this population (e.g. omitting sleep restriction) and that certain adjustments may enhance acceptability (e.g. adopting a collaborative approach). High-quality, large-scale RCTs are required to understand the full benefits of treatment and explore what factors influence these.

References

1. Reeve S, Sheaves B, Freeman D. Sleep Disorders in Early Psychosis: Incidence, Severity, and Association With Clinical Symptoms. Schizophr Bull. 2019 Mar 7;45(2):287–95.

2. Kalmbach DA, Cuamatzi-Castelan AS, Tonnu CV, Tran KM, Anderson JR, Roth T, et al. Hyperarousal and sleep reactivity in insomnia: current insights. Nat Sci Sleep. 2018 Jul 17;10:193–201.

3. Waite F, Sheaves B, Isham L, Reeve S, Freeman D. Sleep and schizophrenia: From epiphenomenon to treatable causal target. Schizophr Res. 2020 Jul 1;221:44–56.

4. Freeman D. Persecutory delusions: a cognitive perspective on understanding and treatment. Lancet Psychiatry. 2016 Jul;3(7):685–92.

5. Freeman D, Garety PA, Kuipers E, Fowler D, Bebbington PE. A cognitive model of persecutory delusions. Br J Clin Psychol. 2002;41(4):331–47.

6. Reeve S, Sheaves B, Freeman D. The role of sleep dysfunction in the occurrence of delusions and hallucinations: A systematic review. Clin Psychol Rev. 2015 Dec;42:96–115.

7. Waters F, Chiu V, Atkinson A, Blom JD. Severe Sleep Deprivation Causes Hallucinations and a Gradual Progression Toward Psychosis With Increasing Time Awake. Front Psychiatry. 2018 Jul 10;9:303.

8. Dudley R, Taylor P, Wickham S, Hutton P. Psychosis, Delusions and the 'Jumping to Conclusions' Reasoning Bias: A Systematic Review and Meta-analysis. Schizophr Bull. 2016 May;42(3):652–65.

9. Marwaha S, Broome MR, Bebbington PE, Kuipers E, Freeman D. Mood instability and psychosis: analyses of British national survey data. Schizophr Bull. 2014 Mar;40(2):269–77.

10. Westermann S, Lincoln TM. Emotion regulation difficulties are relevant to persecutory ideation. Psychol Psychother. 2011 Sep;84(3):273–87.

11. Sheaves B, Rek S, Freeman D. Nightmares and psychiatric symptoms: A systematic review of longitudinal, experimental, and clinical trial studies. Clin Psychol Rev. 2023 Mar;100:102241.

12. Batalla-Martín D, Martorell-Poveda MA, Belzunegui-Eraso A, Miralles Garijo E, Del-Cuerpo Serratosa A, Valdearcos Pérez J, et al. The Experience of Insomnia in Patients With Schizophrenic Disorder: A Qualitative Study. Front Psychiatry [Internet]. 2022 [cited 2023 May 23];12. Available from: https://www.frontiersin.org/articles/10.3389/fpsyt.2021.805601

13. Ishak WW, Bagot K, Thomas S, Magakian N, Bedwani D, Larson D, et al. Quality of Life in Patients Suffering from Insomnia. Innov Clin Neurosci. 2012 Oct;9(10):13–26.

14. Li SX, Lam SP, Zhang J, Yu MWM, Chan JWY, Chan CSY, et al. Sleep Disturbances and Suicide Risk in an 8-Year Longitudinal Study of Schizophrenia-Spectrum Disorders. Sleep. 2016 Jun 1;39(6):1275–82.

15. Neckelmann D, Mykletun A, Dahl AA. Chronic Insomnia as a Risk Factor for Developing Anxiety and Depression. Sleep. 2007 Jul 1;30(7):873–80.

16. Waters F, Chiu VW, Janca A, Atkinson A, Ree M. Preferences for different insomnia treatment options in people with schizophrenia and related psychoses: a qualitative study. Front Psychol. 2015;6:990.

17. Waite F, Myers E, Harvey AG, Espie CA, Startup H, Sheaves B, et al. Treating sleep problems in patients with schizophrenia. Behav Cogn Psychother. 2016;44(3):273–87.

18. National Institute of Clinical Excellence (NICE). Management | Insomnia | CKS | NICE [Internet]. Insomnia: management. 2022 [cited 2023 May 23]. Available from: https://cks.nice.org.uk/topics/insomnia/management/

19. Borbély AA. A two process model of sleep regulation. Hum Neurobiol. 1982;1:195–204.

20. Walker J, Muench A, Perlis ML, Vargas I. Cognitive Behavioral Therapy for Insomnia (CBT-I): A Primer. Klin Spetsialnaia Psikhologiia Clin Psychol Spec Educ. 2022;11(2):123–37.

21. Aurora RN, Zak RS, Auerbach SH, Casey KR, Chowdhuri S, Karippot A, et al. Best practice guide for the treatment of nightmare disorder in adults. J Clin Sleep Med JCSM Off Publ Am Acad Sleep Med. 2010 Aug 15;6(4):389–401.

22. Casement MD, Swanson LM. A meta-analysis of imagery rehearsal for post-trauma nightmares: effects on nightmare frequency, sleep quality, and posttraumatic stress. Clin Psychol Rev. 2012 Aug;32(6):566–74.

23. National Institute of Clinical Excellence (NICE). Overview | Obstructive sleep apnoea/hypopnoea syndrome and obesity hypoventilation syndrome in over 16s | Guidance | NICE [Internet]. NICE; 2021 [cited 2024 Dec 17]. Available from: https://www.nice.org.uk/guidance/ng202

24. Freeman D, Waite F, Startup H, Myers E, Lister R, McInerney J, et al. Efficacy of cognitive behavioural therapy for sleep improvement in patients with persistent delusions and hallucinations (BEST): a prospective, assessor-blind, randomised controlled pilot trial. Lancet Psychiatry. 2015 Nov;2(11):975–83.

25. Sheaves B, Isham L, Bradley J, Espie C, Barrera A, Waite F, et al. Adapted CBT to Stabilize Sleep on Psychiatric Wards: a Transdiagnostic Treatment Approach. Behav Cogn Psychother. 2018 Nov;46(6):661–75.

26. Sheaves B, Holmes EA, Rek S, Taylor KM, Nickless A, Waite F, et al. Cognitive behavioural therapy for nightmares for patients with persecutory delusions (nites): An assessor-blind, pilot randomized controlled trial. Can J Psychiatry Rev Can Psychiatr. 2019;64(10):686–96.

27. Waite F, Černis E, Kabir T, Iredale E, Johns L, Maughan D, et al. A targeted psychological treatment for sleep problems in young people at ultra-high risk of psychosis in England (SleepWell): a parallel group, single-blind, randomised controlled feasibility trial. Lancet Psychiatry. 2023 Sep 1;10(9):706–18.

28. Hwang DK, Nam M, Lee YJG. The effect of cognitive behavioral therapy for insomnia in schizophrenia patients with sleep disturbance: A non-randomized, assessor-blind trial. Psychiatry Res. 2019 Apr;274:182-188. doi: 10.1016/j.psychres.2019.02.002. Epub 2019 Feb 3. PMID: 30807969.

29. Bradley J, Freeman D, Chadwick E, Harvey AG, Mullins B, Johns L, et al. Treating sleep problems in young people at ultra-high risk of psychosis: A feasibility case series. Behav Cogn Psychother. 2018;46(3):276–91.

30. Haynes PL, Parthasarathy S, Kersh B, Bootzin RR. Examination of insomnia and insomnia treatment in psychiatric inpatients. Int J Ment Health Nurs. 2011 Apr;20(2):130–6.

31. Holmes EP, Corrigan PW, Knight S, Flaxman J. Development of a sleep management program for people with severe mental illness. Psychiatr Rehabil J. 1995;19(2):9–15.

32. Myers E, Startup H, Freeman D. Cognitive behavioural treatment of insomnia in individuals with persistent persecutory delusions: a pilot trial. J Behav Ther Exp Psychiatry. 2011;42(3):330–6.

33. Novak C., Packer E., Paterson A., Roshi A., Locke R., Keown P., et al. Feasibility and utility of enhanced sleep management on in-patient psychiatry wards. BJPsych Bull. 2020;44(6):255–60.

34. Waters F, Chiu VW, Dragovic M, Ree M. Different patterns of treatment response to Cognitive-Behavioural Therapy for Insomnia (CBT-I) in psychosis. Schizophr Res. 2020 Jul;221:57-62. doi: 10.1016/j.schres.2020.03.054.

35. Taylor KM, Bradley J, Cella M. A novel smartphone-based intervention targeting sleep difficulties in individuals experiencing psychosis: A feasibility and acceptability evaluation. Psychol Psychother Theory Res Pract. 2022;95(3):717–37.

36. Freeman D, Sheaves B, Goodwin GM, Yu LM, Nickless A, Harrison PJ, et al. The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis. Lancet Psychiatry. 2017 Oct;4(10):749–58.

37. Batalla-Martin D., Martorell-Poveda M.-A., Belzunegui-Eraso A., Marieges Gordo A., Batlle Lleal H., Pasqual Melendez R., et al. A Pilot Nurse-Administered CBT Intervention for Insomnia in Patients with Schizophrenic Disorder: A Randomized Clinical Effectiveness Trial. J Clin Med. 2023;12(19):6147.

38. Kyle SD, Miller CB, Rogers Z, Siriwardena AN, MacMahon KM, Espie CA. Sleep Restriction Therapy for Insomnia is Associated with Reduced Objective Total Sleep Time, Increased Daytime Somnolence, and Objectively Impaired Vigilance: Implications for the Clinical Management of Insomnia Disorder. Sleep. 2014 Feb 1;37(2):229–37.

39. Reeve S, Emsley R, Sheaves B, Freeman D. Disrupting Sleep: The Effects of Sleep Loss on Psychotic Experiences Tested in an Experimental Study With Mediation Analysis. Schizophr Bull. 2018 Apr 6;44(3):662–71.

40. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009 Jul 21;339:b2535.

41. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021 Mar 29;372:n160.

42. Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas. 1960 Apr 1;20(1):37–46.

43. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New York: Routledge; 1988. 567 p.

44. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res. 2017 Jan 26;17(1):88.

45. Lewis C, Roberts NP, Gibson S, Bisson JI. Dropout from psychological therapies for post-traumatic stress disorder (PTSD) in adults: systematic review and meta-analysis. Eur J Psychotraumatology. 2020;11(1):1709709.

46. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019 Aug 28;366:l4898.

47. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016 Oct 12;355:i4919.

48. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol. 2021 Jul 3;18(3):328–52.

49. Couto E, Pike E, Torkilseng EB, Klemp M. Inter-rater reliability of the Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI) | Cochrane Colloquium Abstracts [Internet]. 2015 [cited 2024 May 3]. Available from:

https://abstracts.cochrane.org/2015-vienna/inter-rater-reliability-risk-bias-assessment-tool-non-randomized-studies-interventions

50. Jerayaman M, Rabbani R, Robson R, Copstein L, Al-Yousif N, Xia J, et al. Inter-rater reliability, inter-consensus reliability and evaluator burden of ROBINS-E and ROBINS-I: a cross-sectional study | Cochrane Colloquium Abstracts [Internet]. 2019 [cited 2024 Apr 30]. Available from: https://abstracts.cochrane.org/2019-santiago/inter-rater-reliability-inter-consensus-reliability-and-evaluator-burden-robins-e-and

51. Minozzi S, Cinquini M, Gianola S, Gonzalez-Lorenzo M, Banzi R. The revised Cochrane risk of bias tool for randomized trials (RoB 2) showed low interrater reliability and challenges in its application. J Clin Epidemiol. 2020 Oct 1;126:37–44.

52. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.

53. Waite F, Bradley J, Chadwick E, Reeve S, Bird JC, Freeman D. The Experience of Sleep Problems and Their Treatment in Young People at Ultra-High Risk of Psychosis: A Thematic Analysis. Front Psychiatry. 2018;9:375.

54. Edinger JD, Carney CE, Edinger JD, Carney CE. Overcoming Insomnia: A Cognitive-Behavioral Therapy Approach, Workbook. Second Edition, Second Edition. Oxford, New York: Oxford University Press; 2014. 104 p. (Treatments That Work).

55. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med. 2001 Jul;2(4):297–307.

56. Krakow B. Nightmare complaints in treatment-seeking patients in clinical sleep medicine settings: diagnostic and treatment implications. Sleep. 2006 Oct;29(10):1313–9.

57. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res. 1989 May;28(2):193–213.

58. Hauri P, Linde S. No More Sleepless Nights. Wiley; 1990. 308 p.

59. Bell V, Halligan PW, Ellis HD. The Cardiff Anomalous Perceptions Scale (CAPS): a new validated measure of anomalous perceptual experience. Schizophr Bull. 2006 Apr;32(2):366–77.

60. Černis E, Beierl E, Molodynski A, Ehlers A, Freeman D. A new perspective and assessment measure for common dissociative experiences: 'Felt Sense of Anomaly'. PloS One. 2021;16(2):e0247037.

61. Yung AR, Yuen HP, McGorry PD, Phillips LJ, Kelly D, Dell'Olio M, et al. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. Aust N Z J Psychiatry. 2005;39(11–12):964–71.

62. Green CEL, Freeman D, Kuipers E, Bebbington P, Fowler D, Dunn G, et al. Measuring ideas of persecution and social reference: the Green et al. Paranoid Thought Scales (GPTS). Psychol Med. 2008 Jan;38(1):101–11.

63. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and

validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998;59 Suppl 20:22-33;quiz 34-57.

64. Kay SR. Positive and Negative Syndromes in Schizophrenia: Assessment and Research. Psychology Press; 1991. 306 p.

65. Haddock G, McCarron J, Tarrier N, Faragher EB. Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). Psychol Med. 1999 Jul;29(4):879–89.

66. Freeman D, Loe BS, Kingdon D, Startup H, Molodynski A, Rosebrock L, et al. The revised Green et al., Paranoid Thoughts Scale (R-GPTS): psychometric properties, severity ranges, and clinical cut-offs. Psychol Med. 2021 Jan;51(2):244–53.

67. Ronald A, Sieradzka D, Cardno AG, Haworth CMA, McGuire P, Freeman D. Characterization of Psychotic Experiences in Adolescence Using the Specific Psychotic Experiences Questionnaire: Findings From a Study of 5000 16-Year-Old Twins. Schizophr Bull. 2014 Jul;40(4):868–77.

68. Reiss S, Peterson RA, Gursky DM, McNally RJ. Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. Behav Res Ther. 1986;24(1):1–8.

69. Fowler D, Freeman D, Smith B, Kuipers E, Bebbington P, Bashforth H, et al. The Brief Core Schema Scales (BCSS): psychometric properties and associations with paranoia and grandiosity in non-clinical and psychosis samples. Psychol Med. 2006 Jun;36(6):749–59.

70. Beck AT, Steer RA, Brown GK. BDI-II Manual. San Antonio, TX: The Psychological Corporation; 1996.

71. Beck AT, Kovacs M, Weissman A. Assessment of suicidal intention: the Scale for Suicide Ideation. J Consult Clin Psychol. 1979 Apr;47(2):343–52.

72. Greenwood KE, Sweeney A, Williams S, Garety P, Kuipers E, Scott J, et al. CHoice of Outcome In Cbt for psychosEs (CHOICE): the development of a new service user-led outcome measure of CBT for psychosis. Schizophr Bull. 2010 Jan;36(1):126–35.

73. Connell J, Barkham M. CORE-10 User Manual, Version 1.1. CORE System Trust & CORE Information Management Systems Ltd.; 2007.

74. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011 Dec;168(12):1266–77.

75. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther. 1995 Mar;33(3):335–43.

76. Dalenberg C, Carlson E. Severity of Dissociative Symptoms—Adult (Brief Dissociative Experiences Scale [DES-B]—Modified) [Internet]. 2010 [cited 2024 Mar 23]. Available from: https://psycnet-apa-org.ezproxy-prd.bodleian.ox.ac.uk/record/1997-04647-022

77. Freeman D, Bird JC, Loe BS, Kingdon D, Startup H, Clark DM, et al. The Dunn Worry Questionnaire and the Paranoia Worries Questionnaire: new assessments of worry. Psychol Med. 2020 Apr;50(5):771–80.

78. Brooks R, Rabin R, de Charro F. The measurement and valuation of health status using EQ-5D: a European perspective. Dordrecht: Kluwer Academic Publishers; 2003.

79. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res Int J Qual Life Asp Treat Care Rehabil. 2011 Dec;20(10):1727–36.

80. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res. 1995 Apr;39(3):315–25.

81. Lambe S, Bird JC, Loe BS, Rosebrock L, Kabir T, Petit A, et al. The Oxford Agoraphobic Avoidance Scale. Psychol Med. 2023;53(4):1233–43.

82. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ–4. Psychosomatics. 2009 Nov 1;50(6):613–21.

83. Law H, Neil ST, Dunn G, Morrison AP. Psychometric properties of the questionnaire about the process of recovery (QPR). Schizophr Res. 2014 Jul;156(2–3):184–9.

84. Keetharuth AD, Brazier J, Connell J, Bjorner JB, Carlton J, Taylor Buck E, et al. Recovering Quality of Life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. Br J Psychiatry. 2018 Jan;212(1):42–9.

85. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes. 2007 Nov 27;5(1):63.

86. Mundt JC, Marks IM, Shear MK, Greist JH. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br J Psychiatry J Ment Sci. 2002 May;180:461–4.

87. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry J Ment Sci. 1978 Nov;133:429–35.

88. Tarnowski KJ, Simonian SJ. Assessing treatment acceptance: the Abbreviated Acceptability Rating Profile. J Behav Ther Exp Psychiatry. 1992 Jun;23(2):101–6.

89. Edwards CJ, Cella M, Tarrier N, Wykes T. The optimisation of experience sampling protocols in people with schizophrenia. Psychiatry Res. 2016 Oct 30;244:289–93.

90. van Straten A, van der Zweerde T, Kleiboer A, Cuijpers P, Morin CM, Lancee J. Cognitive and behavioral therapies in the treatment of insomnia: A meta-analysis. Sleep Med Rev. 2018 Apr;38:3–16.

91. Yücel DE, van Emmerik AAP, Souama C, Lancee J. Comparative efficacy of imagery rehearsal therapy and prazosin in the treatment of trauma-related nightmares in adults: A meta-analysis of randomized controlled trials. Sleep Med Rev. 2020 Apr;50:101248.

92. Hertenstein E, Trinca E, Wunderlin M, Schneider CL, Zust MA, Feher KD, et al. Cognitive behavioral therapy for insomnia in patients with mental disorders and comorbid insomnia: A systematic review and meta-analysis. Sleep Med Rev. 2022 Apr;62:101597. doi: 10.1016/j.smrv.2022.101597. Epub 2022 Feb 9. PMID: 35240417.

93. Killgore WDS, Weber M. Sleep Deprivation and Cognitive Performance. In: Bianchi MT, editor. Sleep Deprivation and Disease: Effects on the Body, Brain and Behavior

[Internet]. New York, NY: Springer; 2014 [cited 2024 Apr 30]. p. 209–29. Available from: https://doi.org/10.1007/978-1-4614-9087-6_16

94. Tomaso CC, Johnson AB, Nelson TD. The effect of sleep deprivation and restriction on mood, emotion, and emotion regulation: three meta-analyses in one. Sleep. 2020 Dec 26;44(6):zsaa289.

95. Ben Simon E, Walker MP. Sleep loss causes social withdrawal and loneliness. Nat Commun. 2018 Aug 14;9(1):3146.

96. Friedman J.H. Sleep disturbances in Parkinson's disease: Beyond a regional perspective. Time to do something about it. Int Psychogeriatr. 2020;32(7):793–4.

97. Wade AG. The societal costs of insomnia. Neuropsychiatr Dis Treat. 2010 Dec 20;7:1–18.

98. Colten HR, Altevogt BM, Research I of M (US) C on SM and. Extent and Health Consequences of Chronic Sleep Loss and Sleep Disorders. In: Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem [Internet]. National Academies Press (US); 2006 [cited 2024 Apr 30]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK19961/

99. Villeneuve K, Potvin S, Lesage A, Nicole L. Meta-analysis of rates of drop-out from psychosocial treatment among persons with schizophrenia spectrum disorder. Schizophr Res. 2010 Aug;121(1–3):266–70.

100. Persons JB, Brown CL, Diamond A. Case formulation-driven cognitive-behavioral therapy. In: Handbook of cognitive-behavioral therapies, 4th ed. New York, NY, US: The Guilford Press; 2019. p. 145–68.

101. Wulff K, Dijk DJ, Middleton B, Foster RG, Joyce EM. Sleep and circadian rhythm disruption in schizophrenia. Br J Psychiatry J Ment Sci. 2012 Apr;200(4):308–16.

102. Strassnig MT, Harvey PD, Miller ML, Depp CA, Granholm E. Real world sedentary behavior and activity levels in patients with schizophrenia and controls: An ecological momentary assessment study. Ment Health Phys Act. 2021 Mar;20:10.1016/j.mhpa.2020.100364.

103. Riemann D, Spiegelhalder K, Feige B, Voderholzer U, Berger M, Perlis M, et al. The hyperarousal model of insomnia: A review of the concept and its evidence. Sleep Med Rev. 2010 Feb 1;14(1):19–31.

104. Myles H, Myles N, Antic NA, Adams R, Chandratilleke M, Liu D, et al. Obstructive sleep apnea and schizophrenia: A systematic review to inform clinical practice. Schizophr Res. 2016 Jan 1;170(1):222–5.

105. Ugurlu M, Ugurlu GK, Sahin EK, Kamis GZ, Caykoylu A. Short and Long-term Effects of Cognitive Behavioral Therapy on Sleep Problems and Psychotic Symptoms in Patients with Psychotic Disorders: A Meta-Analysis Study. Rev Bras Psiquiatr Sao Paulo Braz 1999 [Internet]. 2024;(100895975). Available from: https://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=medp&DO=10.47626%2f1516-4446-2024-3623

106. Chiu, V. W., Ree, M., Janca, A., Iyyalol, R., Dragovic, M., & Waters, F. (2018). Sleep profiles and CBT-I response in schizophrenia and related psychoses. Psychiatry Research, 268, 279–287. https://doi.org/10.1016/j.psychres.2018.07.027.

107. Braun V, Clarke V. Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. Int J Transgender Health. 2023 Jan 25;24(1):1–6.

108. Redwood S, Gill PS. Under-representation of minority ethnic groups in research — call for action. Br J Gen Pract. 2013 Jul;63(612):342–3.

109. Termorshuizen F, van der Ven E, Tarricone I, Jongsma HE, Gayer-Anderson C, Lasalvia A, et al. The incidence of psychotic disorders among migrants and minority ethnic groups in Europe: findings from the multinational EU-GEI study. Psychol Med. 2022;52(7):1376–85.

110. Rehman A., Waite F., Sheaves B., Biello S., Freeman D., Gumley A. Clinician perceptions of sleep problems and their treatment in patients with non-affective psychosis. J Sleep Res. 2016;25(Supplement 1):283.

111. National Institute of Clinical Excellence (NICE). Overview | Psychosis and schizophrenia in adults: prevention and management | Guidance | NICE [Internet]. NICE; 2014 [cited 2025 Jan 17]. Available from: https://www.nice.org.uk/guidance/cg178

112. Robinson A, Moghaddam N. Psychological treatments and therapy adaptations for psychological distress in dementia and mild cognitive impairment: a systematic review and meta-analysis. Ment Health Rev J. 2022 Jan 1;27(3):295–318.

113. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015 Jan 1;4(1):1.

114. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders [Internet]. 2013 [cited 2024 Jun 1]. Available from: https://doi.org/10.1176/appi.books.9780890425596

115. World Health Organization. ICD-11: International classification of diseases (11th revision) [Internet]. 2022. Available from: https://icd.who.int/

Improving the psychological wellbeing of parents whose babies have a prenatal diagnosis requiring surgical treatment

Hannah Wilkinson

Hannah.Wilkinson@psy.ox.ac.uk

The Oxford Institute of Clinical Psychology Training and Research

University of Oxford

May 2025

Supervisors: Dr Matt Hotton and Dr Jacinta Cordwell

Word count: 4994

Submitted to The Journal of Paediatric Psychology, chosen because of its focus on promoting the health and psychological wellbeing of children and families.

Abstract

Background: Parents experience significant stress when their baby receives a prenatal diagnosis that will need surgical treatment. This study aims to explore parents' experiences of a neonatal surgical pathway and develop recommendations for support.

Methods: Twenty-four parents completed surveys and nine parents were subsequently interviewed regarding their experiences of the neonatal surgical pathway. Themes were identified using reflexive thematic analysis.

Results: Parents highlighted the importance of preparation, empathetic communication and involvement in their baby's care. Different levels of support ranged from parent-to-parent mentoring, a space to discuss feelings in medical consultations, ensuring physical health needs were met, and access to psychological services.

Conclusion: The findings show that beyond the immediate stress of having an unwell child, the way in which the wider medical system holds parents in their distress has a major impact on parents' wellbeing. Recommendations to improve support are provided.

Introduction

One in 47 babies are born with a congenital anomaly in the UK (Public Health England, 2018). Parents commonly report distress when receiving an antenatal diagnosis (Aite et al., 2004; Fonseca et al., 2013). This may be heightened when surgery is required to treat the condition and the prognosis is uncertain (Diffin et al., 2016; Dodge et al., 2022; Hearps et al., 2014; Tacy et al., 2022). It is estimated that two in five parents experience clinical levels of anxiety or post-traumatic stress when their baby requires intensive care on a neonatal unit (Malouf et al., 2022). Reasons include being separated from their baby, finding the ward environment overwhelming, and fear surrounding their baby's health and survival (Hawes et al., 2016; Williams et al., 2018). It is well documented that early stress can interfere with parent-infant attachment and have long-term consequences for children's emotional and intellectual development (Di Blasio et al., 2017; Dubois-Comtois et al., 2013; Gravener et al., 2012; Ionio et al., 2019). This highlights the need to support parents of babies whose congenital condition requires surgical treatment.

The ABC-X Model of Family Coping predicts that, in the face of a stressor, the resources available to a family and their perceived ability to cope, determine whether they experience crisis or adapt to maintain resilience (Hill, 1958). Accordingly, studies have shown how 'bad news' is delivered affects subsequent wellbeing (Luz et al., 2017; Ptacek & Eberhardt, 1996). An antenatal diagnosis can enable parents to adjust and prepare by engaging with specialist services (Aspinall, 2002; Davalbhakta & Hall, 2000; Kuttenberger et al., 2010; Smith et al., 2015). Clear communication and involving parents in care can increase a sense of empowerment (Dodge et al., 2022; Labrie et al., 2021). By contrast, when parents are confused and lack understanding about what will happen, this predicts poorer psychological outcomes (Bonnot Fazio et al., 2022). Sensitivity to parents' mental state may

be important to consider during medical consultations, as intense emotions of fear can make it difficult for parents to grasp clinical information and ask questions (Aite et al., 2004).

The British Association of Perinatal Medicine (BAPM, 2021) has in recent years published a guidance framework for neonatal units promoting 'Family Integrated Care' (FiCare). This service model empowers parents to be primary caregivers for their baby in partnership with the clinical team. It advocates for a culture of collaboration by actively seeking families' feedback to inform and improve the quality of services. This is in line with the Neonatal Critical Care Transformation Review (NCCR) which targets neonatal units with identifying and addressing gaps in parents' psychological needs (NHS England, 2019). Whilst guidance exists for neonatal pathways more broadly, no national framework currently exists on how best to communicate with parents when their babies have an antenatal diagnosis requiring surgery, and how best to support them throughout the treatment journey.

Project background and aims

This project was commissioned by the surgical team within a tertiary hospital in South-East England to evaluate how they prepare parents for their baby's surgery and improve their support. The team specialise in treating congenital gastrointestinal problems but may see any diagnosis (for example, renal and thoracic problems, and also cardiac problems where this is not the primary diagnosis). Figure 1 shows their treatment pathway. Parents usually visit their local outpatient clinic for a routine scan at 12 and 20 weeks of pregnancy. A referral is made to the hospital's Foetal Medicine Team (FMT) if the scans reveal that the baby has a medical condition that requires surgery after birth. Whilst under the FMT, further scans and genetic tests are offered to provide specialist diagnosis and monitor the baby's condition.

At 30-34 weeks, parents are given an 'antenatal consultation.' They meet with a paediatric surgeon, obstetrician, and ward nurses to discuss a plan for surgery after birth. They are also

Figure 1. Flowchart of NSP Pathway

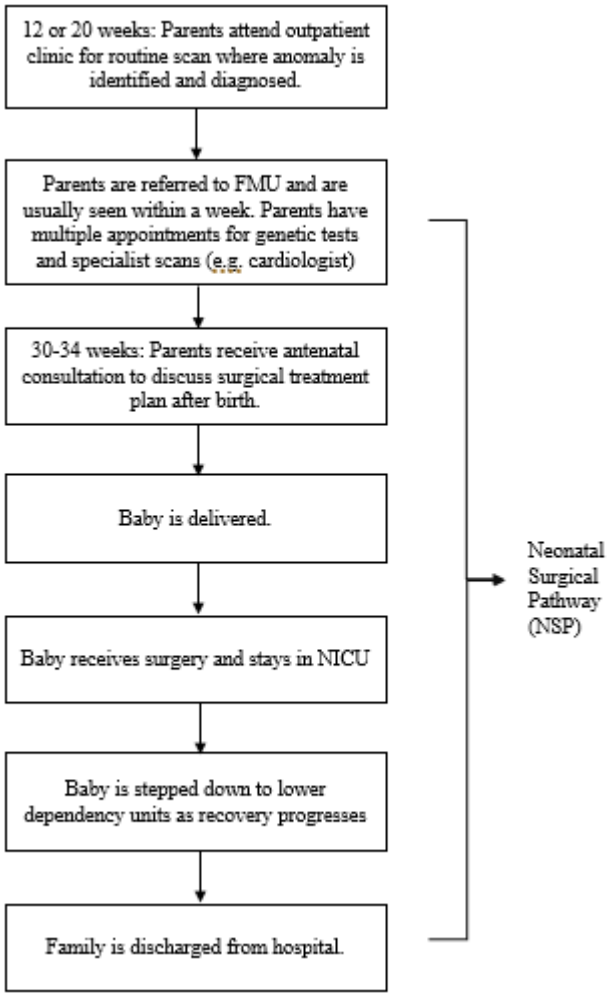


Figure 1. Of note, each family's treatment journey is different and may not follow these linear stages where, for example, a baby has multiple surgeries. Some babies are discharged home after birth and then return to the hospital several months later for surgery. Families may be discharged from the NSP pathway to less specialist hospitals that are more local to the family, community teams such as health visitors, or straight home.

sometimes offered a tour of the wards where the baby may stay. This includes the neonatal intensive care unit (NICU), which cares for babies with the highest support needs and provides 1:1 nursing support. As babies progress in their recovery, they are usually stepped down to other units where the level of support reduces. They may then be discharged home or to another hospital more local to their home should the baby have ongoing needs. Psychologists are embedded and available to speak to parents on the neonatal unit, and there is an ad-hoc on-call provision for the children's wards, however there is no provision antenatally.

We refer to the 'neonatal surgical pathway' (NSP) as beginning with treatment under the FMT through to surgery and surgical aftercare at the hospital. This study seeks to understand the stressors or unmet needs affecting parents' psychological wellbeing whilst under the NSP and the factors that helped them cope. We use this to develop recommendations to improve parent support.

Methods

Participants

This study employed a cross-sectional design. Twenty-four parents completed an online survey, and eight of these were later interviewed. The husband of an interviewee was also interviewed separately. Demographic details are presented in Table 1.

Materials

Two mothers with NSP experience (‘experts-by-experience’) were consulted when designing the project. Feedback was sought for the participant information sheet and consent form. They were also asked about their experiences of the NSP pathway including what factors (both positive and negative) had significantly impacted their wellbeing and their views on what areas the study should focus on. Members of the surgical team also provided input around the questions that would be helpful for the service to ask.

These consultations informed the development of the survey. In addition, the survey drew on research literature that had been reviewed about the key barriers and facilitators of parental wellbeing in neonatal research.The survey was first administered to understand parents’ experiences, psychological needs and sources of support. Survey responses were then used to develop a semi-structured interview schedule that aimed to provide a richer understanding of what parents found supportive or challenging across the treatment journey. One of the experts-by-experience was again consulted to check that the questions were meaningful. The interview schedule was used flexibly during interviews to ensure parents could bring what mattered to them.

Table 1
Participant Demographics

Demographic Variable	Survey		Interview	
	N	%	*N*	%
Parent status				
Mother	23	96	7	78
Father	1	4	2	22
Ethnic Background				
Asian/ Asian British	3	13	-	-
Black/African/Caribbean/Black British	1	4	1	11
Mixed/ Multiple Ethnic groups	-	-	-	-
Other ethnic group	1	4	-	-
White	18	75	8	89
First child				
Yes	17	71	8	89
No	7	29	1	11
Antenatal condition[1,2]				
Congenital diaphragmatic hernia	4	17	1	13
Congenital cystic adenomatoid malformation	2	4	1	13
Exomphalos	5	21	4	50
Gastroschisis	5	21	-	-
Other	13	54	3	38
Stage of pregnancy when first diagnosed[2]				
12-20 weeks	11	46	4	50
20-28 weeks	8	33	3	38
28-36 weeks	5	21	1	13
Stage of pregnancy when antenatal consultation occurred[2]				
28-32 weeks	7	47	2	33
32-36 weeks	8	53	4	67
Duration of hospital stay following surgery or over multiple surgeries[2]				
0-1 months	11	55	7	88
1-2 months	7	35	-	-
2-3 months	1	5	-	-
3+ months	1	5	1	12
Accessed Psychology on neonatal unit				
Yes	8	35	2	22
No	15	65	7	88

[1]Of note, some babies had more than one condition and hence percentages for this item total to over 100%. Other conditions included biliary atresia, congenital pulmonary airway malformation, craniosynostosis, cystic hygroma, duodenal atresia, jejunal atresia, dextrocardia, meconium ileus, sacrococcygeal teratoma, situs inversus, and tracheal esophageal fistula.
[2]A mother and father of same baby took part in separate interviews. For questions about the baby (i.e. about their condition and duration of hospital stay), responses have adjusted to ensure their data is not duplicated.

Procedure

This project was approved by Oxford University Health Psychological Medicine Clinical Governance committee. A nurse from the surgical team screened 102 parents by reviewing medical records for adherence to the following inclusion criteria:

- Discharged from the NSP minimum six months ago; up to three years ago
- Parent over 18 years
- No infant mortality

All eligible parents were contacted by phone and/or email about participating in the survey. The survey was anonymous and took place using the online questionnaire platform 'Qualtrics'. It took approximately 10 minutes to complete. Towards the end, parents were asked to leave their email address if they were willing to be interviewed about their experiences. Interviews were held on Microsoft Teams and were recorded and transcribed using an inbuilt function. Interviews lasted approximately one hour.

Participant Information Sheets and Consent forms were provided before the survey and interview. At the end, parents were given the opportunity to debrief with details of relevant support services.

Analysis

Survey data was summarised descriptively since its main objective was to inform the interview and there was insufficient power to conduct statistical analyses.

Reflexive Thematic Analysis was used to analyse interviews using an inductive and deductive approach (Braun & Clarke, 2006, 2021). Coded extracts of transcript data were used to create themes, which captured patterns across parents' experiences. We chose a critical realist position whereby we viewed parents' descriptions as reflecting reality, whilst acknowledging that these were shaped by the researcher's own beliefs and experiences.

The researcher discussed codes and themes regularly in supervision to develop reflexivity and ensure that they were coherent. Inter-coder reliability was not obtained because it contradicts the underlying theory by implying there is a single, objective viewpoint (O'Connor & Joffe, 2020). Instead, credibility was attended to by presenting examples of original data, contextualising information, addressing negative cases and peer reviewing with two supervisors who work across neonatal and child services within the hospital.

Researcher Background

The researcher is a Trainee Clinical Psychologist. She does not have children of her own, and interviewees may also have assumed this as she appears young and has the status of a university student. Prior to conducting interviews, the researcher visited the neonatal units and spoke with parents and staff to gain a better understanding of participants' treatment pathway. She remained curious and clarified responses where necessary. She also used her clinical skills to validate and respond sensitively when more emotional stories were shared.

Results

Survey results

Information was gathered about parents' experiences across three key time-points: after the antenatal consultation, delivery and surgery. The antenatal consultation occurred on average at 31 weeks of pregnancy (range 29-35 weeks). Most parents reported that the antenatal consultation was helpful/very helpful (83%). Some parents found it easy/very easy to understand (43%) whereas others found it hard/very hard (33%). Many parents left feeling prepared/very prepared (79%) and found it as or better than expected (80%). Just over half (52%) were given a tour of either NICU, a neonatal unit or maternity and of these, all (100%) found this helpful. The hospital NSP pathway does not offer routine psychological support antenatally however 83% of parents indicated that they might have benefitted.

After the baby was born, only 52% reported good communication between the maternity and neonatal unit. Surgery most frequently occurred on the day of birth (range 0-248 days). The mean length of recovery in hospital was 5 weeks (range 0-24 weeks). Around 35% of parents accessed a psychologist on the neonatal unit following delivery or surgery. This was generally described as helpful/very helpful (73%). Parents who were not offered a psychologist mostly said it would have been helpful/very helpful (72%).

Parents indicated that the treatment journey was highly emotional. Across the three time-points, parents commonly experienced hope, helplessness, love, protectiveness and upset or grief (see Figure 2). Overall, parents felt that their psychological wellbeing was acknowledged (55%), or partly acknowledged (20%), by the team looking after their baby. Their feelings about the diagnosis and treatment plan were considered (74%), or partly considered (16%), by the surgical team.

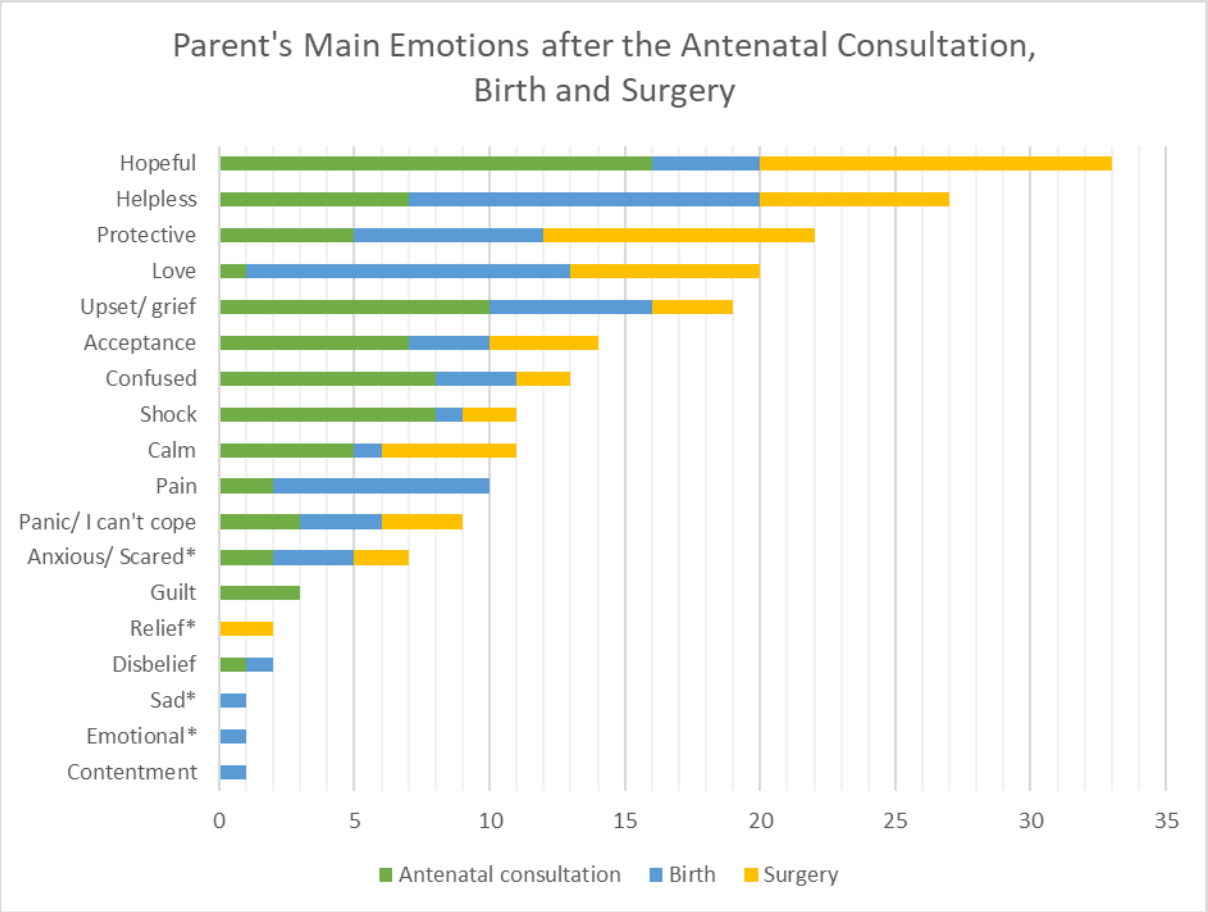


Figure 2. Parents selected from a list the three, strongest emotions they experienced after the antenatal consultation, at birth, and at surgery. Parents also had the option to state other emotions if they were not captured in the list; these entries are marked by an asterisk (*).

Interview results

Parents described various factors that challenged as well as supported their psychological wellbeing throughout their antenatal and neonatal journey (see Figure 3). All

potentially identifying information has been removed to maintain parent anonymity. Of note, although we organised subthemes into distinct themes, there were many elements of interconnection and overlap across them. For example, when former NSP parents shared their experiences with current NSP parents, this often addressed their support needs by providing preparation for next steps along the treatment pathway, more information and hope.

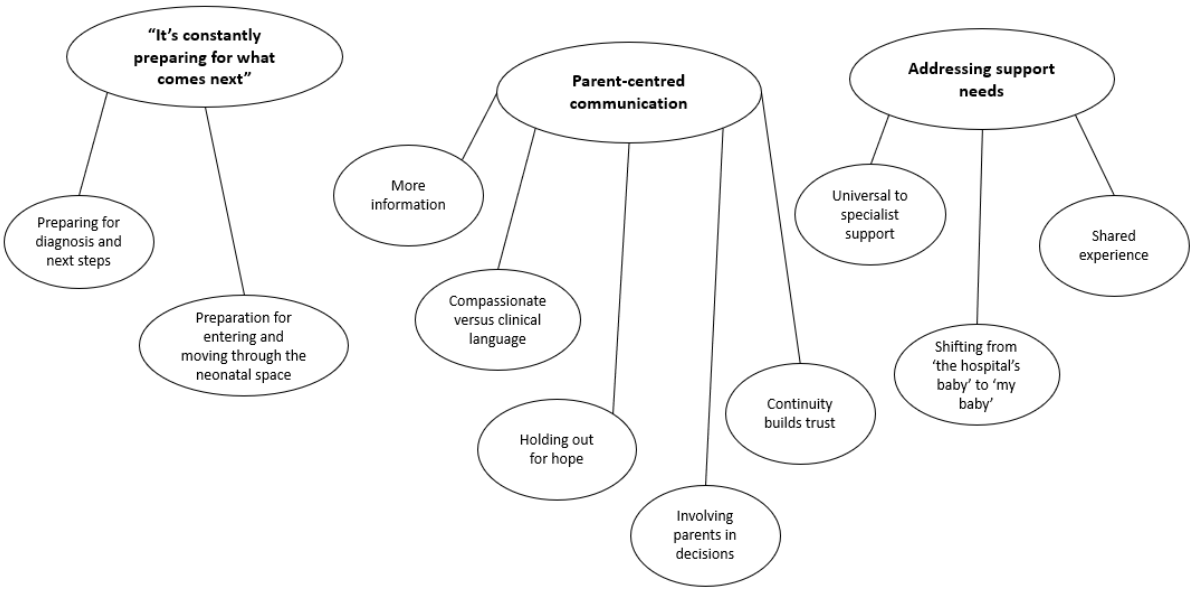


Figure 3. Map of themes and subthemes

Theme 1: "It's constantly preparing for what comes next"

1.1. Preparing for diagnosis and next steps. In this theme, we explored how parents coped with the uncertainties surrounding their baby's diagnosis and how it would be managed. Prior to birth, parents usually attended multiple appointments to diagnose genetic syndromes, scan and monitor the baby's condition. One mother explained, "*when it's all a great big unknown…it's terrifying, because how do you prepare for that psychologically?"*(P6). Parents described not knowing the purpose of an appointment which meant being faced with sudden decisions and disappointment; "*I'd suddenly have this*

question of 'do you want this test?' [when] there's a chance of miscarriage."(P5). Overall, however, the antenatal consultation was viewed positively because it enabled parents to adjust to the situation. It helped parents "*know how we would manage"*(P1), *"feel mentally prepared"*(P5) and "*process the grief about it not being a normal birth."*(P6).

1.2. Preparation for entering and moving through the neonatal space. Neonatal wards were frequently described as *"overwhelming"*(P6) due to the presence of machines and constant alarms. Parents who visited the neonatal unit before birth explained how being prepared meant they had *"one less thing to have to worry about"*(P9). Parents who did not have this opportunity said this would have lessened the initial shock.

The step-down from intensive care to other wards was often a stressful point because support was less intensive and there were fewer members of staff assigned to each baby. Parents felt anxious about this change; "*you feel your child is not completely, 100% looked after*"(P7), and often felt they needed to stay at their baby's bedside at all times which left them exhausted. Reduced space initially posed a problem for one parent (P9) whose wife was a Muslim and needed privacy to breastfeed. Parents reported that "*they didn't really prepare me for how different it is*"(P7) and identified that preparation would have helped them plan how to manage.

Theme 2: Parent-centred communication

2.1 More information. Whilst the situation was often inherently uncertain, parents shared examples of where a lack of information and containment in the system fuelled their distress; "*if you know a piece of information, however potentially shocking, I just want to know it…rather than not telling us and then we had to go through all the stress of not knowing what [it] means*"(P8). Medical jargon and confusion over staff roles implied a certain 'blindness' to what parents did or did not know, perhaps taking this for granted instead

of *"context switching"*(P8) between speaking to parents versus colleagues. There were also many positive instances of information being fully explained and adapted to parents' level of understanding, which was *"very reassuring"*(P2) and *"a comfort"*(P9).

2.2. Compassionate versus clinical language. There was a theme around how information was shared that spoke to parents' need to be 'seen,' rather than an entirely medical focus; *"every single time we went there, people would ask, 'but did you have this genetic test? Did you have this? Did you have that?' sort of reiterating the fact that there was something wrong…no one gave us reassurance…this is where I felt that no one cared, literally about us"*(P2). Clinicians sometimes gave upsetting news abruptly and reportedly lacked sympathy when parents showed distress. On the other hand, there were also many examples where clinicians were *"very patient, very kind"* (P6) and showed a lot of sensitivity at challenging moments; *"[clinician] came in and was like 'this is really stressful, have some time here, have a cup of tea' – that was all lovely."*

2.3 Holding out for hope. Parents described the foetal medicine team as being *"used to finding problems, looking for problems"*(P2) which led to a dread of upcoming appointments where parents expected bad news. When healthcare professionals believed there was a high likelihood of a genetic syndrome, *"everything was all doom and gloom…everyone was pushing for a termination*"(P4). Instead, parents expressed a need for clinicians to hold hope for them to be able to continue on their treatment journey. Examples were given of other hospitals where clinicians had gone through scans explaining *"yes there was this [problem], this and that, but also so many other things that were right…every case is different – we have this and that case where it worked okay"*(P2). Parents did not want false hope but a balanced view which included stories of families who had faced similar issues and had positive outcomes.

2.4 Involving parents in decisions. Parents described a power imbalance with doctors which could make it difficult to get their views across. They described being, *"less informed, less involved…more rushed into making decisions"*(P1), "*dismissed"*(P3) *"put under so much pressure"* (P4) and *"brushed off"*(P5), suggesting they sometimes felt like passive bystanders to the decision-making process. Sometimes, they described not being given all the options or relevant information for making the right decision, which had negative repercussions for their baby's care (e.g., risks of early discharge). Parents struggled when their concerns were not taken on board. Of note, some parents *"felt very listened to"*(P6). Some parents also preferred to let doctors lead due to their medical expertise.

2.5 Continuity builds trust. There was a theme around lack of continuity between different specialities (for example, doctors versus nurses), wards, and one hospital versus another. Conflicting advice often resulted in a breakdown in trust and raised parents' anxiety; *"you start worrying…it ruins your confidence"*(P4). Lack of consistency in staff made it harder to develop supportive relationships. Many parents identified that a named contact would have been helpful to navigate differences in opinions between medical professionals. Only one interviewee had access to this antenatally, and explained, *"psychologically, knowing that there was somebody who I could call just made me feel a bit safer*"(P6). Handovers at points of transition between wards and at discharge to other hospitals or community teams was suggested as a means of joining up care; *"you could have them all attend a briefing by the surgical team to explain the condition, the aftercare that's needed, the kind of support parents need - that could have been very helpful"(P2)*.

Theme 3: Support needs

3.1 Universal to specialist support. This theme reflects the different levels of support that parents desired. At the universal level, parents wanted professionals to

acknowledge and hold a space for their feelings. Several parents commented that no one had asked how they were coping because "*it's more like 'the baby, the baby, the baby'"*(P3) and that *"if we'd had various touch points…the chance to speak to somebody, they could have helped us to heal*"(P5). There were examples where midwives and nurses took parents aside and gave them opportunities to talk, cry and have their feelings validated, which parents greatly appreciated.

Often the stress of caring for a seriously ill baby led parents to neglect their own basic needs. Most participants were adjusting to becoming first-time parents and the remainder had to journey between the hospital and home to look after multiple children. One parent found the simple advice of having a routine *"helpful to keep balance during the day"*(P9). Another parent commented *"it's genuinely really simple things that would make a massive difference: nutrition, sleep exercise…are really vital for someone's wellbeing and generally how they cope with all the stress"*(P8). They commented it would be useful for healthcare professionals to check that parents were looking after themselves.

Specialist psychological therapies were requested by most parents. Reasons included wanting help with making difficult decisions, processing traumatic events, and managing different coping styles between partners. When accessed, psychological help was viewed positively: *"we had some really good conversations to help normalise some of what I was feeling and help me and [partner]…which had a meaningful impact in our ability to support one another through the rest of it.*"

3.2 Shifting from 'the hospital's baby' to 'my baby.' Getting involved in babies' care was seen as both a need and a mitigator of stress. Parents used terms such as *"helpless"*(P6) and *"useless"*(P7) to describe how they initially felt when they were completely reliant on other people to keep their baby alive. Inability to breastfeed compounded this; *"it was really stressful…because my solution was always to breastfeed but*

I couldn't do that"(P1). Parents praised nurses who encouraged them to get babies out of the incubator and start taking over care tasks such as nappy changing; *"you could have easily felt quite detached from your baby in that situation but [nurses] were really good in saying 'no let's get him out, you can have a cuddle'"*(P5). The confidence to perform these tasks independently also carried a sense of regaining parental identity; "*[initially] it felt like the hospital's baby that you're allowed to visit…I felt like I really needed to do [care tasks] like it was my baby"*(P7), and prepared parents for transitioning to lower dependency units and discharge.

3.3 Shared experience. Almost everyone stated that being put in contact with other parents would have been useful. They often felt *"alone"*(P3, P4, P6, P7, P9) in their treatment journey and wanted to talk to someone who could understand. Contact was particularly sought from parents with similar experience at the same hospital, because it linked to the need for preparation; *"speaking to another parent who'd done it before would be useful, even just so that you know beforehand things that I never really imagined would be difficult"*(P6). It also linked to the need for hope to get through the treatment journey; *"If I could have just spoken to someone who'd been through that before and come out the other side, that would have been really helpful…so I could ever do that, I'd love to because those are our darkest days, and actually where we are now is amazing*"(P7).

Discussion

This study aimed to understand the psychological needs of parents whose baby was undergoing assessment and surgery for a congenital anomaly. The results of a survey showed that parents experienced a wide range of emotions across their baby's treatment in hospital such as upset, helplessness, love, protectiveness and hope. However, it varied how much staff acknowledged this. Semi-structured interviews identified themes around a need for

preparation and communication that is clear and empathetic. This shows that beyond the immediate stressor of having an unwell child, a major influence on parents' wellbeing is the way in which the wider medical system responds and holds them in their distress. The final theme captured different levels of support that parents reported would help them during this time. This ranged from addressing universal needs (nutrition, exercise, sleep, routine and providing a space for parents' feelings) through to specialist psychological support. Empowerment to be involved in care tasks increased parents' sense of agency when they felt dependent on the hospital to keep their baby alive. Peer-support delivered by parents who had "*come out the other side*" had the potential to offer hope, insights and understanding.

These findings are consistent with the ABC-X model which identifies that people's perception of having adequate resources to face a stressor is linked to coping instead of being overwhelmed (Rosino, 2016). Literature reviews of giving antenatal diagnoses have consistently highlighted communication as a major theme (Kratovil & Julion, 2017; Wilpers et al., 2021). Effective communication increases parents' confidence in clinicians' care and leads to the establishment of a human connection (Chandra et al., 2018; Chaplin et al., 2005). This, in turn, has been linked to increased parental coping by widening parents' sense of the resources available to them (Kratovil & Julion, 2017; Mause et al., 2022). Shared decision-making also improves trust and partnership with professionals (Saba et al., 2006). Themes of preparation, involvement, promoting parent wellbeing and peer-support can also be understood to operate through similar mechanisms.

One of the themes identified from interviews was parents' wish for doctors to hold hope and not narrowly focus on their child's problems (Holm et al., 2021; McKechnie et al., 2015). Clinicians are legally required to be transparent about the implications of a condition and risks of treatments, which may conflict with what parents want to hear. However, our findings add to a growing literature showing that parents need holistic information about

diagnosis and prognosis which includes a child's abilities and future potential (Gokoolparsadh et al., 2024; Lemmon et al., 2022). Focusing on positives is associated with better outcomes for both parents and healthcare providers (Davies et al., 2017). Models of prognostic communication encourage clinicians to instil hope without crossing into false optimism (for example, see 'OuR-Hope' by Bracken-Roche et al., 2019; Fernandes et al., 2022; 'ALIGN' by Lemmon et al., 2023; and 'EMPATIA' Racine et al., 2017). They advise framing the prognosis in a way that acknowledges uncertainty whilst noting strengths (Turbitt et al., 2024).

Elevated rates of psychological disorder are well documented amongst parents whose babies have a congenital anomaly (Al-Maghaireh et al., 2023) and are undergoing neonatal surgery (Diffin et al., 2016). Several of the interviewees described traumatic episodes that continued to affect them months after discharge. Whilst no psychologist was employed to work with parents antenatally, both the survey and interview results highlighted the need for specialist therapies to span the breadth of the treatment journey.

Limitations

This study analyses the perspectives of a small number of parents without exploring differences across conditions. For example, all interviewees reported that antenatal genetic screens were negative, however it would be of interest to understand whether staff might respond differently if a genetic syndrome was diagnosed and what parents' support needs might be in these situations. There was a limitation in the recruitment process because the hospital tended to register mothers as the named contact for their baby. This meant that the research team could only invite partners through mothers' contact details, which likely accounts for their low uptake. Fathers are commonly a 'missing voice' in neonatal research (Burgess & Goldman, 2018), and thus it is important for future research to investigate their experiences to ensure services are also designed to meet their needs.

Recommendations

The findings above were used to develop recommendations for the service. These are presented in Figure 4, which is an adaptation of Kazak et al.'s (2003) Paediatric Psychosocial Preventative Health Model. It illustrates a stepped-care approach to intervention that matches families' level of need and risk.

The universal level proposes recommendations for all families. This includes strategies aimed to meet parents' communication and preparation needs, such as a named contact who has oversight from parents' entry to FMU through to discharge. A named contact can ensure communication is 'joined up', particularly when multiple units are involved (for example, between maternity and neonatal units) and at points of transition between units and at discharge to another care team. Training can support clinicians to think more psychologically about the work they do and build this into a framework that gives them space to deal with difficult feelings. Parents should be given time to contemplate options for appointments involving major decisions (such as amniocentesis which carries risks of miscarriage). This can be achieved by explaining in advance why each appointment is being offered and potential next steps, repeating explanations at the beginning of every appointment, and ensuring parents have plenty of opportunity to reflect without pressure. The antenatal consultation was a positive process that enabled parents to plan their next steps and should be continued.

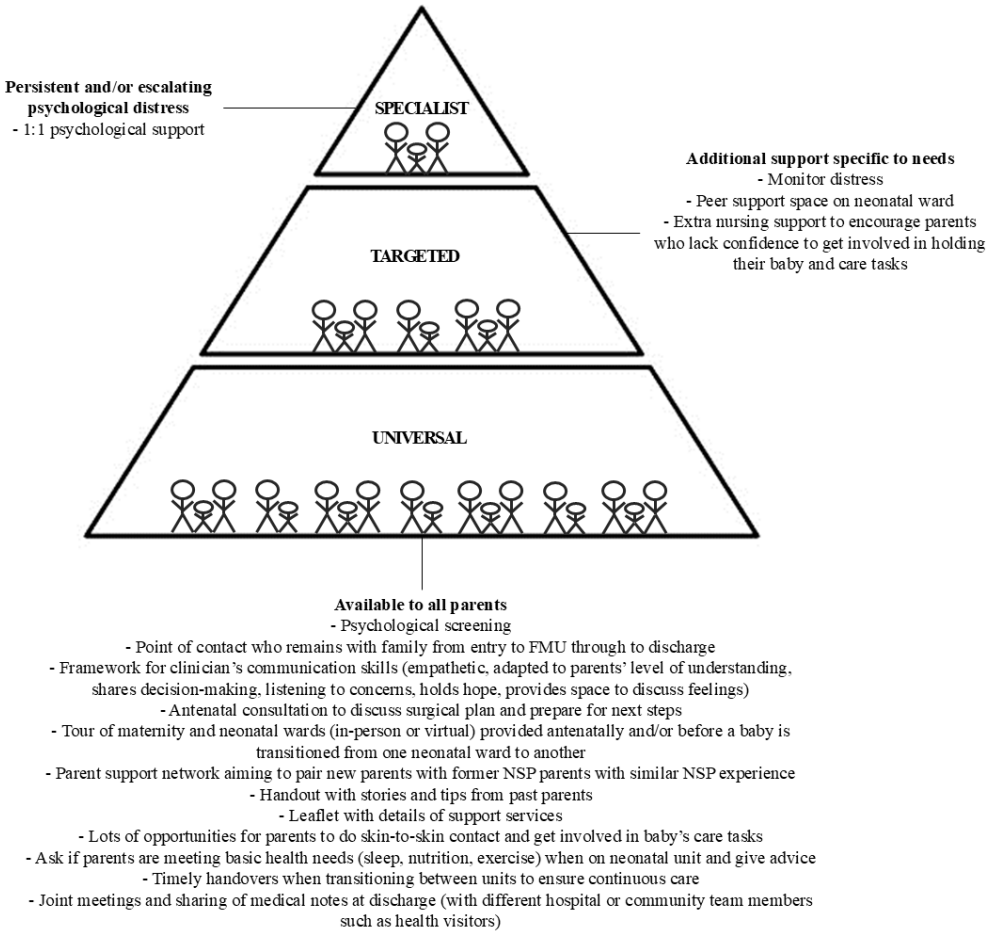


Figure 4. Recommendations for service. Adapted from Kazak et al.'s (2003) Paediatric Psychosocial Preventative Health Model

Parents should be encouraged to visit neonatal and maternity wards, however where this is not possible, a virtual tour should be offered instead. Screening for psychological distress should be provided at regular basis throughout the NSP pathway together with details of support contacts. Where possible, parents should be given the opportunity to link with former NSP families who have given consent to being contacted, ideally with similar treatment experience. A handout with stories and tips from past parents could also facilitate this.

Additional support needs may exist where parents are showing signs of struggling, for example, where they are staying on the unit for long periods to look after their baby and appear to be neglecting their own basic health needs. 'Checking in' can be a window to give

advice about self-care. A peer space on neonatal wards is recommended for parents to support each other. If it is noticed that a parent is fearful of getting involved in their baby's care tasks, they may benefit from nurses demonstrating confidence in their abilities and additional guidance.

Parents who show persistent and escalating distress may require specialist psychologist support. Based on the evidence above, dedicated and embedded psychological provision should extend across the full NSP pathway.

Dissemination

These outcomes were shared with the surgical team and the Trust's Quality Improvement Lead. They were also submitted to the British Association of Perinatal Medicine 2025 Conference.

Conclusion

Ultimately, these findings add to previous literature by highlighting the importance of how medical professionals accompany parents in enhancing their wellbeing. Health services should also consider parent-to-parent mentoring and psychological support that is available across the full treatment journey.

References

Aite, L., Trucchi, A., Nahom, A., Casaccia, G., Zaccara, A., Giorlandino, C., & Bagolan, P. (2004). Antenatal diagnosis of diaphragmatic hernia: Parents' emotional and cognitive reactions. *Journal of Pediatric Surgery, 39*(2), 174–178. https://doi.org/10.1016/j.jpedsurg.2003.10.010

Al-Maghaireh, D. F., Kawafha, M. M., Abdullah, K. L., Shawish, N. S., Abu Kamel, A. M., & Basyouni, N. R. (2023). Psychological problems among parents of children with

congenital anomalies. *Journal of Neonatal Nursing*, *29*(6), 846–850. https://doi.org/10.1016/j.jnn.2023.07.005

Aspinall, C. L. (2002). Dealing with the Prenatal Diagnosis of Clefting: A Parent's Perspective. *The Cleft Palate Craniofacial Journal*, *39*(2), 183–187. https://doi.org/10.1597/1545-1569_2002_039_0183_dwtpdo_2.0.co_2

BAPM. (2021). *Family Integrated Care*. British Association of Perinatal Medicine. https://www.bapm.org/resources/ficare-framework-for-practice

Bonnot Fazio, S., Dany, L., Dahan, S., & Tosello, B. (2022). Communication, information, and the parent–caregiver relationship in neonatal intensive care units: A review of the literature. *Archives de Pédiatrie*, *29*(5), 331–339. https://doi.org/10.1016/j.arcped.2022.05.013

Bracken-Roche, D., Shevell, M., & Racine, E. (2019). Chapter 24 - Understanding and addressing barriers to communication in the context of neonatal neurologic injury: Exploring the ouR-HOPE approach. In L. S. de Vries & H. C. Glass (Eds.), *Handbook of Clinical Neurology* (Vol. 162, pp. 511–528). Elsevier. https://doi.org/10.1016/B978-0-444-64029-1.00024-2

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

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qualitative Research in Psychology, 18(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Burgess, A., & Goldman, R. (2018). Who's the bloke in the room? *Fathers during Pregnancy and at the Birth in the United Kingdom (Ful Report). Contemporary Fathers in the UK Series. Marlborough: Fatherhood Institute.*

Chandra, S., Mohammadnezhad, M., & Ward, P. (2018). Trust and Communication in a Doctor- Patient Relationship: A Literature Review. *Journal of Healthcare Communications*, *03*. https://doi.org/10.4172/2472-1654.100146

Chaplin, J., Schweitzer, R., & Perkoulidis, S. (2005). Experiences of Prenatal Diagnosis of Spina Bifida or Hydrocephalus in Parents Who Decide to Continue with Their Pregnancy. *Journal of Genetic Counseling*, *14*(2), 151–162. https://doi.org/10.1007/s10897-005-0488-9

Davalbhakta, A., & Hall, P. N. (2000). The impact of antenatal diagnosis on the effectiveness and timing of counselling for cleft lip and palate. *British Journal of Plastic Surgery*, *53*(4), 298–301. https://doi.org/10.1054/bjps.2000.3330

Davies, B., Steele, R., Krueger, G., Albersheim, S., Baird, J., Bifirie, M., Cadell, S., Doane, G., Garga, D., Siden, H., Strahlendorf, C., & Zhao, Y. (2017). Best Practice in Provider/Parent Interaction. *Qualitative Health Research*, *27*(3), 406–420. https://doi.org/10.1177/1049732316664712

Di Blasio, P., Camisasca, E., Miragoli, S., Ionio, C., & Milani, L. (2017). Does Maternal Parenting Stress Mediate the Association Between Postpartum PTS Symptoms and Children's Internalizing and Externalizing Problems? A Longitudinal Perspective. *Child & Youth Care Forum*, *46*(5), 685–701. https://doi.org/10.1007/s10566-017-9400-7

Diffin, J., Spence, K., Naranian, T., Badawi, N., & Johnston, L. (2016). Stress and distress in parents of neonates admitted to the neonatal intensive care unit for cardiac surgery. *Early Human Development*, *103*, 101–107. https://doi.org/10.1016/j.earlhumdev.2016.08.002

Dodge, A., Gibson, C., Williams, M., & Ross, K. (2022). Exploring the needs and coping strategies of New Zealand parents in the neonatal environment. *Journal of Paediatrics and Child Health*, *58*(6), 1060–1065. https://doi.org/10.1111/jpc.15908

Dubois-Comtois, K., Moss, E., Cyr, C., & Pascuzzo, K. (2013). Behavior problems in middle childhood: The predictive role of maternal distress, child attachment, and mother-child interactions. *Journal of Abnormal Child Psychology*, *41*(8), 1311–1324. https://doi.org/10.1007/s10802-013-9764-6

Feldman, R., Granat, A., Pariente, C., Kanety, H., Kuint, J., & Gilboa-Schechtman, E. (2009). Maternal Depression and Anxiety Across the Postpartum Year and Infant Social Engagement, Fear Regulation, and Stress Reactivity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(9), 919–927. https://doi.org/10.1097/CHI.0b013e3181b21651

Fernandes, I. A. M., Menezes, R. O. A., & Rego, G. (2022). Communication of the Diagnosis of Spinal Muscular Atrophy in the Views of Patients and Family Members, a Qualitative Analysis. *International Journal of Environmental Research and Public Health*, *19*(24), Article 24. https://doi.org/10.3390/ijerph192416935

Fonseca, A., Nazaré, B., & Canavarro, M. C. (2013). Clinical Determinants of Parents' Emotional Reactions to the Disclosure of a Diagnosis of Congenital Anomaly. *Journal*

of Obstetric, Gynecologic & Neonatal Nursing, *42*(2), 178–190. https://doi.org/10.1111/1552-6909.12010

Gokoolparsadh, A., Bourne, M., McEwen, A., Amor, D. J., & Turbitt, E. (2024). Parents' perspectives on conversations about prognosis and an assessment of prognostic information available online: A mixed-methods study. *Disability and Health Journal*, 101718. https://doi.org/10.1016/j.dhjo.2024.101718

Gravener, J. A., Rogosch, F. A., Oshri, A., Narayan, A. J., Cicchetti, D., & Toth, S. L. (2012). The relations among maternal depressive disorder, maternal Expressed Emotion, and toddler behavior problems and attachment. *Journal of Abnormal Child Psychology*, *40*(5), 803–813. https://doi.org/10.1007/s10802-011-9598-z

Hassan, H., Williams, F., Cordwell, J., & Mann, J. (n.d.). Ethnic minority fathers' experiences of the Neonatal Care Unit: Barriers to accessing psychological support. *Journal of Reproductive and Infant Psychology*, *0*(0), 1–13. https://doi.org/10.1080/02646838.2023.2287079

Hawes, K., McGowan, E., O'Donnell, M., Tucker, R., & Vohr, B. (2016). Social Emotional Factors Increase Risk of Postpartum Depression in Mothers of Preterm Infants. *The Journal of Pediatrics*, *179*, 61–67. https://doi.org/10.1016/j.jpeds.2016.07.008

Hearps, S. J., McCarthy, M. C., Muscara, F., Hearps, S. J. C., Burke, K., Jones, B., & Anderson, V. A. (2014). Psychosocial risk in families of infants undergoing surgery for a serious congenital heart disease. *Cardiology in the Young*, *24*(4), 632–639. https://doi.org/10.1017/S1047951113000760

Hill, R. (1958). 1. Generic Features of Families under Stress. *Social Casework*, *39*(2–3), 139–150. https://doi.org/10.1177/1044389458039002-318

Holm, K. G., Neville, A. J., Pierini, A., Latos Bielenska, A., Jamry-Dziurla, A., Cavero-Carbonell, C., Garne, E., & Clemensen, J. (2021). The Voice of Parents of Children With a Congenital Anomaly – A EUROlinkCAT Study. *Frontiers in Pediatrics*, *9*, 654883. https://doi.org/10.3389/fped.2021.654883

Ionio, C., Mascheroni, E., Colombo, C., Castoldi, F., & Lista, G. (2019). Stress and feelings in mothers and fathers in NICU: Identifying risk factors for early interventions. *Primary Health Care Research & Development*, *20*, e81. https://doi.org/10.1017/S1463423619000021

Kazak, A. E., Rourke, M. T., & Crump, T. A. (2003). Families and Other Systems in Pediatric Psychology. In *Handbook of pediatric psychology, 3rd ed* (pp. 159–175). The Guilford Press.

Kratovil, A. L., & Julion, W. A. (2017). Health-care provider communication with expectant parents during a prenatal diagnosis: An integrative review. *Journal of Perinatology*, *37*(1), 2–12. https://doi.org/10.1038/jp.2016.123

Kuttenberger, J., Ohmer, J. N., & Polska, E. (2010). Initial counselling for cleft lip and palate: Parents' evaluation, needs and expectations. *International Journal of Oral and Maxillofacial Surgery*, *39*(3), 214–220. https://doi.org/10.1016/j.ijom.2009.12.013

Labrie, N. H. M., van Veenendaal, N. R., Ludolph, R. A., Ket, J. C. F., van der Schoor, S. R. D., & van Kempen, A. A. M. W. (2021). Effects of parent-provider communication during infant hospitalization in the NICU on parents: A systematic review with meta-synthesis and narrative synthesis. *Patient Education and Counseling*, *104*(7), 1526–1552. https://doi.org/10.1016/j.pec.2021.04.023

Lemmon, M. E., Barks, M. C., Bansal, S., Bernstein, S., Kaye, E. C., Glass, H. C., Ubel, P. A., Brandon, D., & Pollak, K. I. (2023). The ALIGN Framework. *Neurology*, *100*(8), e800–e807. https://doi.org/10.1212/WNL.0000000000201600

Lemmon, M. E., Barks, M. C., Bernstein, S., Davis, J. K., Jiao, M. G., Kaye, E. C., Glass, H. C., Brandon, D., & Ubel, P. A. (2022). Prognostic Discussion for Infants with Neurologic Conditions: Qualitative Analysis of Family Conferences. *Annals of Neurology*, *92*(4), 699–709. https://doi.org/10.1002/ana.26457

Luz, R., George, A., Spitz, E., & Vieux, R. (2017). Breaking bad news in prenatal medicine: A literature review. *Journal of Reproductive and Infant Psychology*, *35*(1), 14–31. https://doi.org/10.1080/02646838.2016.1253052

Malouf, R., Harrison, S., Burton, H. A. L., Gale, C., Stein, A., Franck, L. S., & Alderdice, F. (2022). Prevalence of anxiety and post-traumatic stress (PTS) among the parents of babies admitted to neonatal units: A systematic review and meta-analysis. *EClinicalMedicine*, *43*, 101233. https://doi.org/10.1016/j.eclinm.2021.101233

Mause, L., Hoffmann, J., Reimer, A., Dresbach, T., Horenkamp-Sonntag, D., Klein, M., Scholten, N., & Neo-CamCare. (2022). Trust in medical professionals and its influence on the stress experience of parents of premature infants. *Acta Paediatrica*, *111*(3), 527–535. https://doi.org/10.1111/apa.16187

McKechnie, A. C., Pridham, K., & Tluczek, A. (2015). Preparing Heart and Mind for Becoming a Parent Following a Diagnosis of Fetal Anomaly. *Qualitative Health Research*, *25*(9), 1182–1198. https://doi.org/10.1177/1049732314553852

NHS England. (2019). *NHS England » Implementing the Recommendations of the Neonatal Critical Care Transformation Review*.

https://www.england.nhs.uk/publication/implementing-the-recommendations-of-the-neonatal-critical-care-transformation-review/

O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. International Journal of Qualitative Methods, 19. https://doi.org/10.1177/1609406919899220

Ptacek, J. T., & Eberhardt, T. L. (1996). Breaking Bad News: A Review of the Literature. *JAMA, 276*(6), 496–502. https://doi.org/10.1001/jama.1996.03540060072041

Public Health England. (2018). *NCARDRS statistics 2018: Summary report*. GOV.UK. https://www.gov.uk/government/publications/ncardrs-congenital-anomaly-annual-data/ncardrs-statistics-2018-summary-report

Racine, E., Bell, E., Farlow, B., Miller, S., Payot, A., Rasmussen, L. A., Shevell, M. I., Thomson, D., & Wintermark, P. (2017). The 'ouR-HOPE' approach for ethics and communication about neonatal neurological injury. *Developmental Medicine & Child Neurology*, *59*(2), 125–135. https://doi.org/10.1111/dmcn.13343

Robbins, J. M., Damiano, P., Druschel, C. M., Hobbs, C. A., Romitti, P. A., Austin, A. A., Tyler, M., Reading, J. A., & Burnett, W. (2010). Prenatal Diagnosis of Orofacial Clefts: Association with Maternal Satisfaction, Team Care, and Treatment Outcomes. *The Cleft Palate Craniofacial Journal*, *47*(5), 476–481. https://doi.org/10.1597/08-177

Rosino, M. (2016). ABC-X Model of Family Stress and Coping. In *Encyclopedia of Family Studies* (pp. 1–6). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119085621.wbefs313

Saba, G. W., Wong, S. T., Schillinger, D., Fernandez, A., Somkin, C. P., Wilson, C. C., & Grumbach, K. (2006). Shared Decision Making and the Experience of Partnership in Primary Care. *Annals of Family Medicine*, *4*(1), 54. https://doi.org/10.1370/afm.393

Smith, K., Christensen, B., & Sylvester, D. (2015). Parental perspective of cleft lip/palate: Experience in prenatal counseling for parents of cleft patients. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, *119*(5), e254. https://doi.org/10.1016/j.oooo.2015.02.009

Tacy, T. A., Kasparian, N. A., Karnik, R., Geiger, M., & Sood, E. (2022). Opportunities to enhance parental well-being during prenatal counseling for congenital heart disease. *Seminars in Perinatology*, *46*(4), 151587. https://doi.org/10.1016/j.semperi.2022.151587

Turbitt, E., Callinan, E., Shakes, P., McEwen, A., & Amor, D. J. (2024). Fostering Hope and Acknowledging Uncertainty: Meeting Parents' Needs and Preferences When Communicating Prognosis in Genetic Neurodevelopmental Conditions. *Current Developmental Disorders Reports*, *11*(1), 21–31. https://doi.org/10.1007/s40474-024-00291-1

Williams, K. G., Patel, K. T., Stausmire, J. M., Bridges, C., Mathis, M. W., & Barkin, J. L. (2018). The Neonatal Intensive Care Unit: Environmental Stressors and Supports. *International Journal of Environmental Research and Public Health*, *15*(1), 60. https://doi.org/10.3390/ijerph15010060

Wilpers, A., Bahtiyar, M. O., Stitelman, D., Batten, J., Calix, R. X., Chase, V., Yung, N., Maassel, N., & Novick, G. (2021). The parental journey of fetal care: A systematic

review and metasynthesis. *American Journal of Obstetrics & Gynecology MFM*, *3*(3), 100320. https://doi.org/10.1016/j.ajogmf.2021.100320

Do eating disorder cognitions play a role in Avoidant Restrictive Feeding Intake Disorder (ARFID)? A comparison between ARFID and Anorexia

Hannah Wilkinson

Hannah.Wilkinson@psy.ox.ac.uk

The Oxford Institute of Clinical Psychology Training and Research

University of Oxford

May 2025

Supervisors: Dr Myra Cooper and Dr Anne Carbonnier

Word count: 4929

Submitted to International Journal of Eating Disorders, chosen because of its focus on scientific knowledge needed for understanding, treating, and preventing eating disorders, and it has a Special Issue call for papers relating to Avoidant Restrictive Feeding Intake Disorder.

Abstract

Objective: Avoidant Restrictive Feeding Intake Disorder (ARFID) is a relatively new diagnosis that featured in the 2013 edition of the Diagnostic Statistical Manual (DSM-V). There is limited evidence about whether individuals with ARFID experience unhelpful cognitions that overlap with other eating disorders and whether they play a role in driving disordered eating behaviours. This study seeks to address this gap.

Methods: Adult participants were recruited from a community sample and invited to complete an online survey. An observational design was used to compare responses by people with ARFID, anorexia nervosa, and no eating disorder who were screened using DSM-V criteria. Questionnaires were used to measure core beliefs, assumptions and automatic thoughts that are associated with other eating disorders and an assessment of concerns about weight and shape was completed.

Results: Individuals with ARFID were found to have significantly higher levels of disordered core beliefs, assumptions, automatic thoughts, and weight/shape concerns than people with no eating disorder. They showed lower levels of these cognitions relative to individuals with anorexia nervosa. Disordered assumptions and automatic thoughts explained a large proportion of variance in behaviours intended to prevent weight gain in this group.

Discussion: The findings have implications for the diagnosis and treatment of ARFID. It challenges the DSM-V conceptualisations of ARFID as entirely separable from other eating disorders and any associated weight/shape concerns. It also highlights the need for clinicians to assess and treat unhelpful cognitions that may be maintaining disordered patterns of eating. Future research directions are discussed.

Introduction

Avoidant Restrictive Feeding Intake Disorder (ARFID) is a relatively new diagnosis that was introduced in the 2013 edition of the Diagnostic Statistical Manual (DSM-V, 2013). It represents restrictive eating caused by sensitivity to sensory characteristics (e.g. texture), lack of interest, or fear of aversive consequences (e.g. choking). It is estimated to affect 0.3-15.5% of the general population (Sanchez-Cerezo et al., 2023). ARFID can have major physical health consequences by causing nutritional deficiencies (Nitsch et al., 2021) and can severely impact people's quality of life (Hay et al., 2017; Zickgraf et al., 2016) .

Research has focused on how ARFID might differ from classical eating disorders (EDs) particularly anorexia nervosa (AN) which similarly involves restrictive eating. The DSM-V states that AN and ARFID are distinct conditions that are distinguished by the presence of body image concerns (DSM-V, 2013). In AN, weight or shape concerns represent the main driver of restriction. By contrast, much has been made of the absence of these concerns in ARFID and indeed this is one of the diagnostic criteria. Yet, this account has been challenged. Increased rates of autism are well documented amongst AN patients with estimates suggesting a prevalence of 25-30% (Inal-Kaleli et al., 2025). One of the core features of autism is atypical sensory sensitivity which can lead to food selectivity and avoidance, and indeed, these behaviours may be diagnosed as ARFID when the impact is clinically significant (Baraskewich et al., 2021; Keski-Rahkonen & Ruusunen, 2023; Sader et al., 2025). One might therefore expect that some autistic AN patients may show signs of ARFID. Becker et al. (2020) have explained that their eating disorder team has observed several adolescent girls with ARFID (e.g. with choking phobia, or high selectivity for certain textures) who have simultaneously had AN or subsequently gone on to develop this disorder. Forbush (2021) has co-authored a chapter with Becker, Eddy and Thomas stating that approximately 10% of patients in their ARFID clinic present with clinically significant

dissatisfaction with their weight or shape (see also Barney et al., 2022 and Norris et al., 2020). It seems notable that these concerns are raised by some of the leading experts in the field with extensive research and treatment experience, and also by the authors of one of the main treatment manuals (see Thomas et al., 2020; Thomas & Eddy, 2018).

The presence of body image concerns has implications for diagnosis. It also raises the question of how far people with ARFID have unhelpful cognitions associated with other EDs, which may be a target for treatment. The first-line treatment for classical EDs is cognitive behavioural therapy (CBT). It works by addressing unhelpful behaviours and cognitions that are maintaining the disorder. In CBT, cognitions are conceptualised at three levels: core beliefs, assumptions and automatic thoughts (Beck, 1976). Core beliefs are deep-seated beliefs people have about themselves and the world (Greenberger & Padesky, 1995). They give rise to assumptions, demands and 'rules for living', which in turn, shape people's automatic thoughts in a specific situation. This is illustrated in Figure 1. Negative core self-beliefs play a causal role in classical EDs (Segal et al., 2002; Teasdale et al., 2000), and it has been argued that lasting recovery requires modifying them (Cooper, 2012; 2005). Cognitive CBT techniques involve identifying distorted thinking patterns (e.g. catastrophising) and gathering evidence to challenge and reframe them.



Figure 1. An illustration of Beck's cognitive model of emotional disorders and an example of this for a patient. Adapted from Willson & Branch (2019)

Psychological treatments for ARFID are still in the phase of being piloted and tested and the evidence-base is very sparse (Bryant-Waugh et al., 2021; Davey et al., 2023). However, they also tend to favour a CBT approach, albeit with more of a behavioural focus employing strategies such as regular eating, graded exposure to novel foods, relaxation and behavioural experiments, rather than any thought-challenging techniques (Kambanis & Thomas, 2023; Thomas et al., 2018). This perhaps reflects that there has been limited research on whether unhelpful cognitions play a role in ARFID and on their specific content.

In this study, we aim to address this gap by determining the degree to which people with ARFID overlap in experiencing core beliefs, assumptions and automatic thoughts that are commonly found in patients with classical EDs (referred to as 'classical ED-related cognitions'). In addition, we explore whether these cognitions play a role in maintaining behaviours to prevent weight gain (referred to as 'restraint', see definition under Measures section). There are several grounds for our predictions. As discussed above, body image concerns appear to be present in some individuals with ARFID. Moreover, ARFID is associated with several factors which are prominent in classical EDs, including fear of negative evaluation (Trompeter et al., 2019), emotional dysregulation (Zanna et al., 2021), high levels of adverse childhood events (Rienecke et al., 2022), and some parental behaviours (Wagner et al., 2020). In classical EDs, these events have been linked to unhelpful assumptions and core beliefs (Cooper, 2005; Fairburn et al., 1999; Lock et al., 2008; Loeb et al., 2012), and thus this raises the question whether these cognitions are also present in ARFID. We might not expect people with ARFID to show as high levels of classical ED-related cognitions as those with AN given that ARFID is not a classical ED, the DSM-V mandates that weight/shape concerns are not the main driver of restriction, and we recognise that many people with ARFID show no such concerns. Nevertheless, we predict that they will show higher levels than the general population (Kambanis et al., 2024; Zickgraf et al., 2016).

This study compares people with ARFID, AN and healthy controls (HC) in order to examine the hypotheses below:

Primary hypotheses

1. It is predicted that the highest prevalence of classical ED-related core beliefs will be found in people with AN, followed by those with ARFID, followed by HC (i.e., people with AN will show significantly more classical ED-related core beliefs than both people with ARFID and HC, and ARFID patients will show significantly more of these classical ED core beliefs than HC).
2. Individuals with AN will show the highest prevalence of classical ED-related assumptions, followed by individuals with ARFID, followed by HC.
3. Individuals with AN will show the highest prevalence of classical ED-related automatic thoughts, followed by individuals with ARFID, followed by HC.

If the data does not falsify these hypotheses, then further predictions can be made.

Secondary Hypotheses

4. Individuals with AN will show the greatest number of weight/shape concerns, followed by those with ARFID, followed by HC.
5. Restraint will be positively related to classical ED-related cognitions in people with ARFID.

Method

Design

An observational design was used to compare questionnaire responses across three independent groups: people with ARFID, AN and HC. Ethical approval was obtained from the Central University Research Ethics Committee at the University of Oxford (ref: R88894/RE001). Three experts-by-experience were involved in designing advertising material, information and consent forms, the survey, and debrief page.

Participants

ARFID pathways in national healthcare services are few/non-existent and therefore participants were recruited from a community sample. Between March and August 2024, advertisements were shared through social media (Facebook and Instagram) and relevant mental health charities (Student Minds, Participate-MQ, ARFID Awareness UK). Participants had to be aged over 18 years and living in the UK, USA, Canada, Australia or New Zealand. They were excluded if they did not belong to one of the above three groups, and/or reported losing weight because of medical illness or a reason other than avoidance or restriction.

Measures

The survey consisted of various self-report questionnaires assessing ED diagnosis, classical ED-related cognitions, and sample characteristics (see Table 1). Two of the Eating Disorder Beliefs Questionnaire (EDBQ) subscales ('weight and shape as a means to acceptance by others' and 'weight and shape as a means to self-acceptance') were used to assess the concept of 'weight/shape concerns.' Four items of the Eating Disorder Diagnostic Scale (EDDS-V) were summed to yield the total number of times that people engaged in various disordered eating behaviours over the past three months, as an index of 'restraint' (see Abber et al., 2024). These behaviours were fasting, laxative/diuretic use, purging, and more intensive exercising specifically with the aim of counteracting the effects of eating or preventing weight gain. Full scales and details of scoring are provided in the Appendix.

Table 1
Outcome Measures Used Within the Study

Domain	Measure
Grouping Variables	
Classical eating disorder diagnosis	The DSM-V version of the Eating Disorder Diagnostic Scale (EDDS-V) is a 23-item self-report measure that diagnoses anorexia, bulimia nervosa and binge eating disorder from healthy controls (Stice, 2014). It represents an update of the original EDDS (Stice et al., 2000), but has not yet been validated. However, the original version demonstrates good test-retest reliability, internal consistency, criterion validity and convergent validity (Krabbenborg et al., 2012; Stice et al., 2004).
ARFID diagnosis	Pica, ARFID and Rumination Disorder Questionnaire (PARDI-AR-Q; Bryant-Waugh et al., 2019) which is a 32 item self-report questionnaire. It contains a 7-point

	Likert scale to measure common ARFID features including sensory-based avoidance, lack of interest, and fear of aversive consequences of food or eating. It has been validated and shows high specificity (Bryant-Waugh et al., 2022).
Primary Outcome Variables	
Classical ED-related core beliefs	Eating Disorder Core Beliefs Questionnaire - Revised (ED-CBQ-R; Fairchild & Cooper, 2010; Hatoum et al., 2022a) is designed to capture common core beliefs in patients with classical eating disorders, such as anorexia and bulimia. This 15-item scale assesses four dimensions of eating disorder core beliefs of self-loathing, unassertive, demanding, and abandoned. The subscales possess adequate internal consistency, and convergent and divergent validity (Hatoum et al., 2022b).
Classical ED-related assumptions	Eating Disorder Beliefs Questionnaire (EDBQ; Cooper et al., 1997). It contains four subscales measuring dysfunctional assumptions in classical eating disorder patients (22 items). These subscales pertain to negative beliefs, weight and shape as a means of acceptance by others, weight or shape as a means of self-acceptance, and control over eating. Results indicate good internal consistency, convergent, divergent and criterion-related validity.
Classical ED-related automatic thoughts	Thoughts Questionnaire (TQ; Cooper et al., 2006). This is a 26-item scale that captures automatic thoughts about eating in classical eating disorder patients. It comprises three subscales relating to positive, negative and permissive thoughts about eating. The subscales demonstrate good internal consistency, construct validity and discriminant validity.
Sample characteristics	
Demographic details	Demographic questions relating to age, sex, gender identification, race, ethnicity and occupation were derived from the Office for National Statistics census, which was developed through research testing and user-engagement (ONS, 2021).
Depression	Patient Health Questionnaire-8 (PHQ-8; Kroenke et al, 2009). This is an 8-item assessment of depression symptoms. It is derived from the PHQ-9 but omits an item that asks about suicide and suicidal ideation.[2] It demonstrates adequate reliability, sensitivity, specificity, internal consistency, and cross-country equivalence (Wu et al., 2019; de la Torre., 2023)
Generalised Anxiety	The Generalised Anxiety Disorder Assessment (GAD-7; Spitzer et al., 2006) is a 7-item screening tool and severity measure. Even though it was originally developed to assess generalised anxiety disorder, it is increasingly used as a general measure of anxiety (Beard & Björgvinsson, 2014) and anxiety disorder (Dear et al., 2011). It displays excellent convergent validity and internal consistency (Johnson et al., 2019) and is reliable and valid in general population samples (Löwe et al., 2008; Hinz et al., 2017).

Procedure

Participants were invited to follow a link which presented them with an information sheet and consent form. Eligible participants were then able to complete the study which involved an anonymous online survey. Participants who declined consent were not able to

[2] *Removing this item was required by ethical research governance since the study was anonymous and hence it would not be possible to provide support to individuals flagged as at risk of suicide.*

proceed. The final page of the survey contained a statement reminding participants of their self-worth and provided an optional set of items where participants could select their values, identify motivational quotes, and engage in guided imagery. There was also a list of support services.

Data analysis plan

Power. An a-priori power analysis was performed using Cohen's Power Primer (1992). This calculated a minimum required sample size of $N = 156$ to detect a medium-sized difference ($d=.50$) across three independent groups with $.80$ power.

Data preparation. Prior to conducting any analyses, participants were classified into groups of AN, ARFID and HC using DSM-V diagnostic criteria. The EDDS-V was used to identify participants who had a clinical diagnosis of AN and screen out participants with other classical ED diagnoses (e.g. bulimia). The PARDI-AR-Q identified participants with ARFID. People with no clinical eating disorder were HC. In line with DSM-V criteria, it was assumed that people cannot have a dual diagnosis of AN and ARFID, with AN functioning as the primary diagnosis.

Data checking. Participants who were missing responses on questionnaire items were excluded from the dataset. Next, all data was checked for adherence to statistical assumptions for analyses. The majority of data for each group was normally distributed.[3] However, there was an issue of a floor effect in the healthy group where, expectedly, they showed low levels of classical ED-related cognitions and weight/shape concerns. Following guidance from a statistician, the decision was made to continue with parametric testing since Multivariate Analysis of Variance (MANOVA) is robust to violations of normality particularly when reporting Pillai's Trace statistic (Olson, 1974), and the remaining data were suitable.

[3] *When investigating the fourth hypothesis regarding weight/shape concerns across the three groups, the equivalent Kruskal-Wallis's test was also conducted since concerns about the violations of normality were stronger; however, the pattern of results remained consistent. This analysis is reported in the Appendix.*

Identification of covariates. Demographic variables that differed significantly between AN, ARFID and HC participants were intended to be included as covariates in analyses of group differences (hypotheses 1-4). A one-way Analysis of Variance (ANOVA) showed that age was a significant covariate. However, categorical demographic variables (sex, gender identification, race, ethnicity, occupation) were unsuitable for running Chi-Square tests since there were not enough data in each cell (N < 5).

Statistical tests. For the primary hypotheses, a one-way multivariate analysis of covariance (MANCOVA) aimed to analyse whether diagnostic groups showed different levels of classical ED-related cognitions (core beliefs, assumptions and thoughts), whilst controlling for age. Following this, post-hoc one-way analysis of covariance (ANCOVA) and Bonferroni corrected t-tests would explore any significant effects. Plans for secondary analyses included a one-way MANCOVA to compare group differences in weight/shape concerns, with age as a covariate. Linear regression analyses aimed to test whether classical ED-related cognitions predicted 'restraint' behaviours designed to prevent weight gain in people with ARFID.

Results

Sample characteristics

The final dataset included 184 participants assigned to groups of AN (*N = 68*), ARFID (*N = 61*) and HC (*N = 55*). Table 2 shows sample characteristics organised by group. Overall, most participants were White (87%) and reported female sex (94%) and gender identification (79%). Approximately two thirds were in some form of paid work (66%). The overall mean age of participants was 28.50 years, however age varied significantly across the three groups, $F_{(2, 154)} = 9.80$, $p = <.001$. HC participants were significantly older than AN ($p <.001$) and ARFID participants ($p = .001$), whereas there was no significant difference

between AN and ARFID participants (p =1.000). Body mass index (BMI) also varied significantly across the three groups ($F_{(2, 181)} = 36.38$, p = <.001) where, consistent with diagnosis, AN participants had a significantly lower BMI than the other two groups (p <.001 for both). BMI was in the underweight range for AN participants and in the normal range for ARFID and HC participants.

Group-level differences also existed in self-reported symptoms of anxiety, $F_{(2, 181)}$ = 47.89, p = <.001, and depression, $F_{(2, 181)} = 52.01$, p <.001. The AN group reported the highest levels of anxiety compared to the ARFID group (p = .042), who in turn, had higher levels of anxiety than HC (p < .001). The same trend was seen for depressive symptoms whereby AN participants had significantly more symptoms than ARFID participants (p = .006), who in turn, had more symptoms than HC (p <.001).

Table 2. Sample characteristics organised by group

Characteristics	**Anorexia**	**ARFID**	**Healthy**	**Total**
Ethnicity (N, %)				
Asian/Asian British	6 (8.8)	2 (3.3)	2 (3.6)	10 (5.4)
Black/African/Caribbean/Black British	1 (1.5)	-	3 (5.5)	4 (2.2)
Mixed/Multiple ethnic groups	2 (2.9)	3 (4.9)	2 (3.6)	7 (3.8)
Other ethnic group	-	2 (3.3)	1 (1.8)	3 (1.6)
White	59 (86.8)	54 (88.5)	47 (85.5)	160 (87.0)
Sex (N, %)				
Female	66 (97.1)	59 (96.7)	48 (87.3)	173 (94.0)
Male	2 (2.9)	2 (3.3)	7 (12.7)	11 (6.0)
Gender identification (N, %)				
Female	58 (85.3)	41 (67.2)	46 (83.6)	145 (78.8)
Genderfluid	-	2 (3.3)	-	2 (1.1)
Male	-	2 (3.3)	7 (12.7)	9 (4.9)
Non-binary	8 (11.8)	13 (21.3)	2 (3.6)	23 (12.5)
Not specified	1 (1.5)	1 (1.6)	-	2 (1.1)
Transgender	1 (1.5)	2 (3.3)	-	3 (1.6)
Occupation (N, %)				
In paid work (employee or self-employed)	36 (52.9)	26 (42.6)	46 (83.6)	108 (58.7)
In paid work taking temporary leave	8 (11.8)	4 (6.6)	1 (1.8)	13 (7.1)
Long term sick or disabled	7 (10.3)	5 (8.2)	2 (3.6)	14 (7.6)
Looking after family or home	-	2 (3.3)	2 (3.6)	4 (2.2)
Other	3 (4.4)	2 (3.3)	-	5 (2.7)
Retired	-	2 (3.3)	-	2 (1.1)
Student	14 (20.6)	20 (32.8)	4 (7.3)	38 (7.1)
Age (M, SD)	25.66 (7.80)	26.58 (9.93)	34.26 (2.02)	28.50 (11.13)
BMI (M,SD)	16.25 (1.81)	23.47 (7.80)	22.86 (4.89)	20.62 (6.29)
Total anxiety scores (M, SD)	15.21 (5.13)	12.85 (5.58)	5.91 (5.42)	11.64 (6.60)

Total depressive scores (M, SD)	16.37 (5.65)	13.15 (6.16)	5.69 (5.75)	12.11 (7.31)

Sample characteristics are shown for the Anorexia, ARFID and healthy control groups and at an overall level. Number (N) and percentage (%) of participants are reported for categorical data. Means (M) and standard deviations (SD) are reported for continuous data.
Of note, on the GAD-7 measure of anxiety symptoms, a total score of 0-4 indicates minimal anxiety, 5-9 indicates mild anxiety,10-14 indicates moderate anxiety, and greater than 15 indicates severe anxiety (Spitzer et al., 2011). On the PHQ-8 measure of depressive symptoms, a score of 0-4 is rated as normal, 5-9 as mild, 10-14 as moderate, and 15-19 as moderately severe, and 20-24 as severe (Kroenke et al., 2009).

Internal consistency of scales

Cronbach's reliability showed good to excellent consistency for the complete ED-CBQ-R, EDBQ, TQ questionnaires and the two EDBQ subscales assessing weight/shape concerns (α = 0.82-0.98; see Appendix). This is important as previous research has only validated individual subscales of the ED-CBQ-R, EDBQ and TQ, whereas the tests of the primary hypotheses were not adequately powered to detect differences at this level.

Primary hypotheses: prevalence of classical ED-related cognitions

A one-way MANCOVA was conducted to investigate whether people with AN, ARFID and HC showed different levels of classical ED-related core beliefs, assumptions and thoughts, whilst controlling for age. The analysis revealed a significant overall effect of diagnostic group on classical ED-related cognitions, F (6, 304) = 12.61, $p < .001$, $\eta^2 = .20$. Univariate tests confirmed that there were group differences in classical ED-related core beliefs, F (2, 153) = 23.51, $p < .001$, $\eta^2 = .26$; assumptions, F (2, 153) = 47.64, $p < .001$, $\eta^2 = .38$; and thoughts, F (2, 153) = 26.99, $p < .001$, $\eta^2 = .26$. Pairwise comparisons revealed a pattern whereby people with AN had significantly more classical ED-related core beliefs, assumptions and thoughts than people with ARFID, who in turn, had significantly more of these cognitions than HC. These results are summarised in Table 3 and estimated marginal means are shown graphically in Figure 2.

Table 3. MANCOVA results comparing levels of ED-related cognitions across groups

Dependent variable	(I) Diagnosis	Mean	Std. error	(J) Diagnosis	Mean diff. (I-J)	Std. error	Sig.
Thoughts	Anorexia	35.19	1.88	ARFID	12.24	2.64	<.001
				Healthy	20.99	2.92	<.001

	ARFID	22.95	1.89	Anorexia	-12.24	2.64	<.001
				Healthy	8.78	2.90	.009
	Healthy	14.20	2.14	Anorexia	-20.99	2.92	<.001
				ARFID	-8.78	2.90	.009
Assumptions	Anorexia	62.28	3.03	ARFID	23.39	4.25	<.001
				Healthy	45.45	4.69	<.001
	ARFID	38.89	3.04	Anorexia	-23.39	4.25	<.001
				Healthy	22.06	4.66	<.001
	Healthy	16.83	3.44	Anorexia	-45.45	4.69	<.001
				ARFID	-22.06	4.66	<.001
Core beliefs	Anorexia	4.54	0.15	ARFID	.76	.22	.002
				Healthy	1.64	.24	<.001
	ARFID	3.78	0.16	Anorexia	-.77	.22	.002
				Healthy	.87	.24	.001
	Healthy	2.91	0.18	Anorexia	-1.67	.24	<.001
				ARFID	-.87	.24	.001

'Mean' refers to the estimated marginal means with the covariate of age being evaluated in the model at 28.50 years, and 'std. error' refers to their standard error. 'Mean diff.' represents the mean difference between two groups of estimated marginal means.

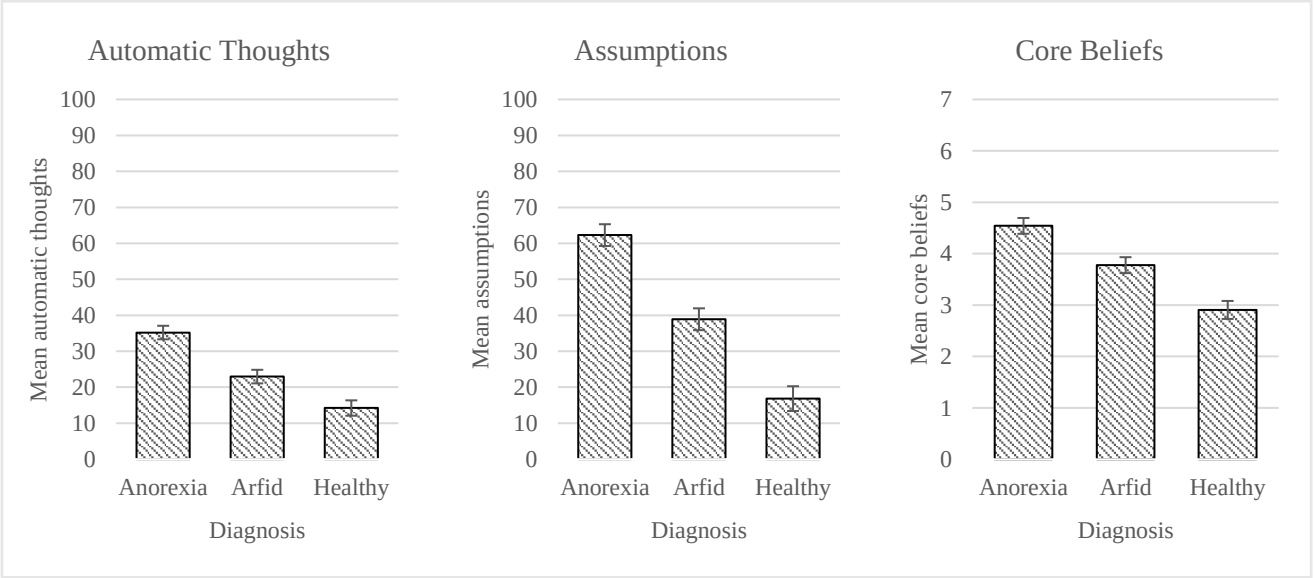


Figure 2. Graphs depict estimated marginal means and standard errors for scores on the TQ (automatic thoughts), EDBQ (assumptions) and ED-CBQ-R (core beliefs). This is shown for each group (Anorexia, ARFID and healthy controls).

Prevalence of weight/shape concerns

To investigate whether there were differences in weight/shape concerns among the diagnostic groups, a one-way MANCOVA was performed. The analysis revealed a significant overall effect of diagnostic group on weight/shape concerns, $F(4, 306) = 13.40$, $p < .001$, $\eta^2 = .15$. Univariate tests confirmed that there were group differences in both EDBQ subscales relating to weight/shape as a means of acceptance from others, $F(2, 153) = 21.41$, $p < .001$, η^2

= .22; and weight/shape as a means of self-acceptance, $F(2, 153) = 27.90$, $p < .001$, $\eta^2 = .27$. Pairwise comparisons revealed the same trend as previously shown, whereby people with AN had significantly more ED-related weight/shape concerns, followed by people with ARFID, and then HC. Results of these analyses are summarised in Table 4 with estimated marginal means shown in Figure 3.

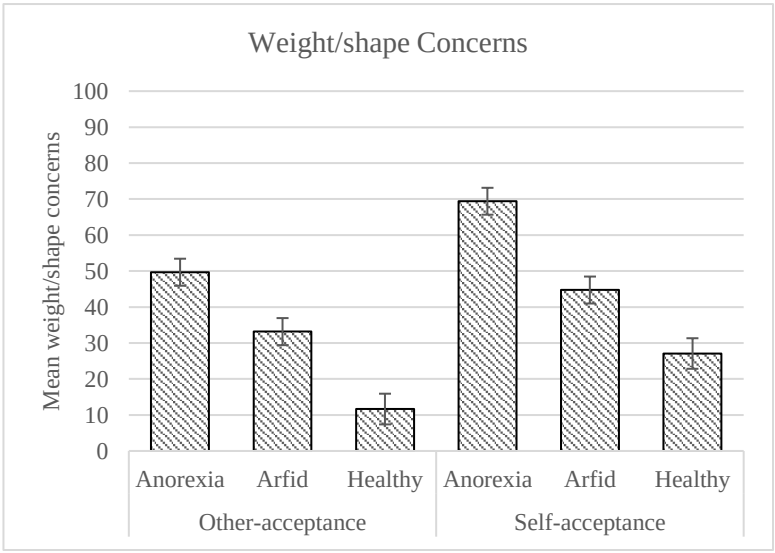


Figure 3. Graphs depict estimated marginal means and standard errors for scores on the two EDBQ subscales of weight/shape concerns as a means of acceptance by others (left graph) and weight/shape concerns as a means of self-acceptance (right graph). This is shown for each group (Anorexia, ARFID and healthy controls).

Table 4. MANCOVA results comparing levels of weight/shape concerns across groups

Dependent variable	(I) Diagnosis	Mean	Std. error	(J) Diagnosis	Mean diff. (I-J)	Std. error	Sig.
Weight/shape as a means of acceptance from others	Anorexia	49.68	3.76	ARFID	16.51	5.27	.006
				Healthy	38.03	5.81	<.001
	ARFID	33.17	3.76	Anorexia	-16.51	5.27	.006
				Healthy	21.52	5.78	<.001
	Healthy	11.65	4.27	Anorexia	-38.03	5.81	<.001
				ARFID	-21.52	5.78	<.001
Weight/shape as a means of self-acceptance	Anorexia	69.40	3.74	ARFID	24.67	5.24	<.001
				Healthy	42.32	5.78	<.001
	ARFID	44.73	3.74	Anorexia	-24.67	5.24	<.001
				Healthy	17.64	5.75	.008
	Healthy	27.09	4.25	Anorexia	-42.32	5.78	<.001
				ARFID	-17.64	5.75	.008

'Mean' refers to the estimated marginal means with the covariate of age being evaluated in the model at 28.50 years, and 'std. error' refers to their standard error. 'Mean diff.' represents the mean difference between two groups of estimated marginal means.

Cognitive predictors of restraint

Table 6 shows a correlation matrix between the predictor variables and restraint. Classical ED-related assumptions and automatic thoughts had significant positive correlations with restraint. They were also significantly correlated with each other. Initially, a stepwise multiple regression had been planned to identify the most influential cognitive predictors of restraint in the ARFID group. However, this test was found to be unsuitable because suppression effects were observed, including a reversal paradox, whereby the association between one of the predictors and the outcome variable became negative when other predictors were added to the analysis (see Appendix for details). This artefact was likely due to the multicollinearity between the three predictor variables resulting in unstable regression estimates. Hence, three separate regression models were run to address this.

Table 5. Pearson's correlation between variables for the ARFID group

Variable	Core beliefs	Assumptions	Thoughts	Restraint
Core beliefs	-			
Assumptions	.56 (<.001)	-		
Thoughts	.31 (.014)	.79 (<.001)	-	
Restraint	.15 (.235)	.67 (<.001)	.51 (<.001)	-

The linear regression model with core beliefs as a predictor was not found to be significant, $F(1, 59) = 1.44$, $p = .235$. However, the regression model with assumptions as a predictor was statistically significant, $F(1, 59) = 46.91$, $p < .001$. This model explained 44.3% of the variance in restraint (adjusted R^2). Greater levels of these assumptions predicted higher numbers of restraint behaviours in ARFID participants: $\beta = 0.67$, $SE = .04$, $p < .001$. The regression model with automatic thoughts as a predictor was also significant, $F(1, 59) = 20.84$, $p < .001$. This model explained 24.9% of the variance in restraint (adjusted R^2), indicating that these automatic thoughts also resulted in increased restraint behaviours: $\beta = .51$, $SE = .10$, $t = 4.57$, $p < .001$.

Discussion

This study investigated whether people with ARFID experience unhelpful cognitions associated with classical eating disorders, and whether these cognitions predict behaviours to prevent weight gain. In line with our hypotheses, people with ARFID showed significantly higher levels of classical ED-related core beliefs, assumptions and automatic thoughts than people with no eating disorder. They also showed significantly more weight/shape concerns than people with no eating disorder. Individuals with AN showed the very highest levels of classical ED-related cognitions and weight/shape concerns. A further finding was that, in the ARFID group, these classical ED-related assumptions and automatic thoughts explained a large proportion of the variance in restraint behaviours. To summarise, these results demonstrate that individuals with ARFID and classical ED patients have some shared experiences of unhelpful cognitions and weight/shape concerns. They also point to the significant role that cognitions may play in maintaining patterns of disordered eating in people with ARFID.

Of note, classical ED-related core beliefs did not significantly predict restraint. A critical difference could lie in automatic thoughts and assumptions being conditional (e.g. "I am unworthy because I am fat"), whereas core beliefs are deeply-entrenched, unconditional beliefs about the self (e.g. "I am unworthy"). It has been argued that eating disorders occur through the "fusion" of negative core beliefs and assumptions about needing to control weight and shape, whereby people restrict food and engage in these behaviours to compensate from these highly damaging beliefs (e.g. striving for thinness to feel less unworthy) (Cooper & Turner, 2000; Jones et al., 2007; Sarin & Abela, 2003). Thus, core beliefs alone may not be sufficient to predict restraint in ARFID. Alternatively, it may be that different core beliefs are related to restraint behaviours in ARFID and classical EDs. Indeed, it is important to say that ED-related cognitions predicted classical ED-related behaviours.

We were not investigating ARFID-specific eating behaviours which may have other relationships to these cognitions.

A strength of our study is that these findings were obtained despite a highly stringent application of the diagnostic criteria. Following the DSM-V, AN was treated as the primary diagnosis, such that people with ARFID who also met the threshold for AN were designated as having AN. In other words, our results likely underestimate the presence of classical-ED related cognitions and weight/shape concerns in people with ARFID. It is also perhaps remarkable that the ARFID group showed this pattern of findings despite not being underweight (like the AN group). Validated measures were used to diagnose individuals with classical EDs and ARFID and to ensure that healthy controls did not have either of these conditions according to DSM-V criteria. We highlight the novelty of our findings since this is the first-known study to investigate classical ED-related core beliefs, assumptions and automatic thoughts in people with ARFID, and to further determine their relationship to actual behaviours of disordered eating.

Our findings add to a small but growing literature suggesting ARFID is not so distinct from classical EDs as the DSM-V would suggest. There are several reports that ARFID and weight/shape concerns can co-occur (Barney et al., 2022; Becker et al., 2020; Burton Murray et al., 2021; Forbush, 2021; Santonicola et al., 2012; Sharpe, 2024; Zucker & Bulik, 2020). Indeed, a study by Jhe et al. (2024) identified weight/shape concerns in over 250 youth with ARFID. A number of clinicians and researchers have proposed that there are three ARFID profiles depending on whether food restriction is driven by either sensory sensitivity, low appetite, or fear of aversive consequences (e.g. see Sharp & Stubbs, 2019). Whilst Jhe et al.'s (2024) study did not find that weight/shape concerns varied between different ARFID profiles or age, higher concerns were found in youth who were gender diverse and youth with overweight/obesity status. A recent study in adults also found a degree of overlap between

ARFID and classical ED patients in their motivations to restrict food (Abber et al., 2024). A subgroup of ED patients showed both ARFID-based (i.e., driven by sensory sensitivity, low appetite, or fear) and weight/shape-based motivations to eat less food, which suggests that these motivations are not clearly separable. Indeed, several patients with ARFID diagnoses showed weight/shape motivations, and several patients with classical ED diagnoses showed ARFID-based motivations. Our findings present a further challenge to DSM-V criteria by demonstrating that people with ARFID and AN show overlapping cognitions, and further, that these cognitions predict actual engagement in behaviours to prevent weight gain.

Theoretical implications

Thus, taken together with other research, our results raise interesting questions about whether diagnostic refinement is needed. Currently, the DSM-V treats ARFID as an exception by presenting it as the only ED for which it is not possible to have a dual diagnosis with other EDs such as AN and bulimia. However, it is unclear from a theoretical standpoint, why weight/shape concerns should be mutually exclusive to any sensory sensitivities, low appetite, or fear of aversive consequences around food. Given that autism frequently co-occurs with AN, one might expect a substantial proportion of AN patients to show sensory sensitivity related to their autism (Adams et al., 2024). This sensory sensitivity may result in clinically significant food restriction or avoidance, and accordingly, there are elevated rates of ARFID among autistic groups (Sader et al., 2025). Overall, there appear to be increasing counterexamples to the DSM-V definition of ARFID in the literature.

Our study was theoretically-informed in applying Beck's (1976) cognitive model of emotional disorders. This framework of core beliefs, assumptions and automatic thoughts underpins much of CBT. Indeed, CBT models of a wide range of clinical presentations are derived from this framework, including models of AN and bulimia (Cooper, 2005, 2012;

Cooper et al., 2004; Fairburn, 2008; Fairburn et al., 1999, 2003). Current models of ARFID appear to be relatively limited in specifying the connections between disordered cognitions and behaviours. For example, Thomas and Eddy's CBT manual depicts three models relating to each ARFID profile (Thomas & Eddy, 2018, p. 21-23, see Appendix). These models propose that "negative feelings and predictions about consequences of eating" mediate relationships from biological predispositions and/or food-related trauma, to food restriction, but they do not describe the specific content of these predictions, nor address how weight/shape concerns may also come into play. It is possible that individuals with ARFID experience cognitions specific to this disorder, or cognitions that are transdiagnostic with classical EDs, or both of these. It may be useful for future models to separate the different levels of cognition according to Beck's (1976) cognitive framework rather than overlook this and treat all cognitions as the same. More validation is needed for cognitive models of ARFID including demonstrations of their relevance.

Clinical implications

Based on our findings, we recommend that clinicians routinely assess ARFID patients' cognitions in line with Beck's (1974) framework and screen for any weight/shape concerns. Clinicians may need to target disordered cognitions within a patient's formulation and treatment plan. The evidence that cognitions overlap between ARFID and classical EDs suggests that a transdiagnostic approach may be suitable for some ARFID patients, and that it may be useful to consider drawing on cognitive CBT techniques found in other ED treatments.

Of note, we observed that the ARFID group showed clinical levels of anxiety and depression, which is consistent with other research showing increased rates of mood and neurodevelopmental disorders in ARFID (Manwaring et al., 2024; Nitsch et al., 2021; Thomas et al., 2021) which are broadly comparable to other EDs (MacDonald et al., 2024;

Vanzhula et al., 2023; Zickgraf et al., 2016). Bryant-Waugh et al. (2021) have put forward an evidence-informed ARFID care pathway for children and adolescents, which recommends screening for other mental health disorders as part of clinical assessments. Our findings are consistent with this and suggest that this screening should also extend to symptoms of other EDs including weight/shape concerns.

Limitations and Future Research Directions

Our study was not sufficiently powered to analyse subscale differences on the questionnaires measuring classical ED-related cognitions. It is possible that some subscales may be particularly characteristic of certain diagnoses, for example, AN patients have been found to score more highly on 'negative self-beliefs' than dieters and healthy controls (Cooper & Turner, 2000). It would be useful for future research to analyse which subscales relate most strongly to ARFID, which may support diagnosis and enable cognitive interventions to be more targeted.

It would also be useful to determine how patients' pattern of responses may vary according to their ARFID profile, severity, and other clinical features such as comorbid depression andanxiety. Indeed, great heterogeneity has been documented in clinical ARFID presentations (Bourne et al., 2020; Sanchez-Cerezo et al., 2024; Thomas et al., 2017) and it has been suggested that this may reflect different underlying mechanisms responsible for the onset and maintenance of ARFID (Bryant-Waugh et al., 2021). Whilst autism was not assessed in the current study, it would be of interest to assess and explore whether it accounts for any of the observed overlap between AN and ARFID. This study was limited to recruiting a community sample because there are few specialist ARFID services for adults in the UK restricting patients' opportunities for diagnosis and treatment (Harrison, 2011). Published accounts of treatment-seeking families show that seeking help does not always result in

appropriate referral (Bartels, 2024; Cardona Cano et al., 2015; Young et al., 2024). Nevertheless, it would be beneficial for future research to replicate this study in patients with a formal diagnosis and determine whether the findings broadly generalise or only apply to certain patient subgroups.

Finally, it would be useful for future research to characterise and determine what drives the cognitions and behaviours that are specific to ARFID. The measures we used were developed with classical ED patients, and thus it is possible that ARFID patients demonstrate other core beliefs, assumptions and automatic thoughts that have not been captured. Surveys, momentary sampling analysis, and phenomenological and thematic qualitative studies may reveal the content of ARFID patients' cognitions. Interviews in which a qualified clinician used the CBT questioning technique of 'downward arrowing' (see Corstorphine et al., 2007) could enable further exploration of patients' deeper meanings around food and eating, and determine assumptions and core beliefs linked to their eating difficulties. It would also be useful to explore how cognitions relate to behaviours characteristic of ARFID eating difficulties (e.g. restriction based on texture) since our study focused on behaviours typical of classical EDs with the specific intention of preventing weight gain.

Conceptualising these factors in models of ARFID will be useful to understanding the mechanisms by which patients' cognitions may be maintaining disordered eating and how this can be addressed through evidence-based treatment.

Conclusion

This study presents novel findings that individuals with ARFIDdemonstrate core beliefs, assumptions, automatic thoughts and weight/shape concerns that are characteristic of classical eating disorders such as anorexia nervosa, and further, that some of these cognitions predict disordered eating behaviours aimed to prevent weight gain. These findings add to

growing literature that challenges the DSM-V conceptualisation of ARFID as distinct from classical eating disorders. These results also highlight a need for clinicians to assess whether unhelpful cognitions may be contributing to a patient's eating difficulties and attend to this by using cognitive treatment techniques.

References

Abber, S. R., Becker, K. R., Stern, C. M., Palmer, L. P., Joiner, T. E., Breithaupt, L., Kambanis, P. E., Eddy, K. T., Thomas, J. J., & Burton-Murray, H. (2024). Latent profile analysis reveals overlapping ARFID and shape/weight motivations for restriction in eating disorders. *Psychological Medicine, 54*(11), 2956–2966. https://doi.org/10.1017/S003329172400103X

Adams, K. L., Mandy, W., Catmur, C., & Bird, G. (2024). Potential mechanisms underlying the association between feeding and eating disorders and autism. Neuroscience & Biobehavioral Reviews, 162, 105717. https://doi.org/10.1016/j.neubiorev.2024.105717

Baraskewich, J., von Ranson, K. M., McCrimmon, A., & McMorris, C. A. (2021). Feeding and eating problems in children and adolescents with autism: A scoping review. Autism : the international journal of research and practice, 25(6), 1505–1519. https://doi.org/10.1177/1362361321995631

Barney, A., Bruett, L. D., Forsberg, S., & Nagata, J. M. (2022). Avoidant Restrictive Food Intake Disorder (ARFID) and Body Image: A case report. *Journal of Eating Disorders*, *10*(1), 61. https://doi.org/10.1186/s40337-022-00583-0

Bartels, N. (2024). *Improving Healthcare Provider Knowledge on Recognition and Treatment of Avoidant/Restrictive Food Intake Disorder (ARFID).* https://hdl.handle.net/10919/120883

Beard, C., & Björgvinsson, T. (2014). Beyond generalized anxiety disorder: Psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. Journal of Anxiety Disorders, 28(6), 547–552. https://doi.org/10.1016/j.janxdis.2014.06.002

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders* (p. 356). International Universities Press.

Becker, K. R., Breithaupt, L., Lawson, E. A., Eddy, K. T., & Thomas, J. J. (2020). Co-occurrence of avoidant/restrictive food intake disorder and traditional eating psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(2), 209–212. https://doi.org/10.1016/j.jaac.2019.09.037

Bourne, L., Bryant-Waugh, R., Cook, J., & Mandy, W. (2020). Avoidant/restrictive food intake disorder: A systematic scoping review of the current literature. *Psychiatry Research*, *288*, 112961. https://doi.org/10.1016/j.psychres.2020.112961

Bryant-Waugh, R., Loomes, R., Munuve, A., & Rhind, C. (2021). Towards an evidence-based out-patient care pathway for children and young people with avoidant restrictive food intake disorder. *Journal of Behavioral and Cognitive Therapy*, *31*(1), 15–26. https://doi.org/10.1016/j.jbct.2020.11.001

Bryant-Waugh, R., Micali, N., Cooke, L., Lawson, E. A., Eddy, K. T., & Thomas, J. J. (2019). Development of the Pica, ARFID, and Rumination Disorder Interview, a multi-informant, semi-structured interview of feeding disorders across the lifespan: A pilot

study for ages 10 to 22. *The International Journal of Eating Disorders*, *52*(4), 378–387. https://doi.org/10.1002/eat.22958

Bryant-Waugh, R., Stern, C. M., Dreier, M. J., Micali, N., Cooke, L. J., Kuhnle, M. C., Burton Murray, H., Wang, S. B., Breithaupt, L., Becker, K. R., Misra, M., Lawson, E. A., Eddy, K. T., & Thomas, J. J. (2022). Preliminary validation of the pica, ARFID and rumination disorder interview ARFID questionnaire (PARDI-AR-Q). *Journal of Eating Disorders*, *10*(1), 179. https://doi.org/10.1186/s40337-022-00706-7

Burton Murray, H., Dreier, M. J., Zickgraf, H. F., Becker, K. R., Breithaupt, L., Eddy, K. T., & Thomas, J. J. (2021). Validation of the nine item ARFID screen (NIAS) subscales for distinguishing ARFID presentations and screening for ARFID. *The International Journal of Eating Disorders*, *54*(10), 1782–1792. https://doi.org/10.1002/eat.23520

Cardona Cano, S., Tiemeier, H., Van Hoeken, D., Tharner, A., Jaddoe, V. W. V., Hofman, A., Verhulst, F. C., & Hoek, H. W. (2015). Trajectories of picky eating during childhood: A general population study. *International Journal of Eating Disorders*, *48*(6), 570–579. https://doi.org/10.1002/eat.22384

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Cooper, M. (2005). Cognitive theory in anorexia nervosa and bulimia nervosa: Progress, development and future directions. *Clinical Psychology Review*, *25*(4), 511–531.

Cooper, M. (2012). Cognitive Behavioural Models in Eating Disorders. In *Eating and its Disorders* (pp. 204–224). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118328910.ch15

Cooper, M., Cohen-Tovée, E., Todd, G., Wells, A., & Tovée, M. (1997). The Eating Disorder Belief Questionnaire: Preliminary development. *Behaviour Research and Therapy*, *35*(4), 381–388. https://doi.org/10.1016/S0005-7967(96)00115-5

Cooper, M. J., Todd, G., Woolrich, R., Somerville, K., & Wells, A. (2006). Assessing Eating Disorder Thoughts and Behaviors: The Development and Preliminary Evaluation of Two Questionnaires. *Cognitive Therapy and Research*, *30*(5), 551–570. https://doi.org/10.1007/s10608-006-9003-3

Cooper, M. J., Wells, A., & Todd, G. (2004). A cognitive model of bulimia nervosa. *British Journal of Clinical Psychology*, *43*(1), 1–16. https://doi.org/10.1348/014466504772812931

Cooper, M., & Turner, H. (2000). Underlying assumptions and core beliefs in anorexia nervosa and dieting. *The British Journal of Clinical Psychology*, *39*(2), 215–218. https://doi.org/10.1348/014466500163239

Corstorphine, E., Waller, G., Cordery, H., Hinrichsen, H., Russell, K., Lawson, R., & Mountford, V. (Eds.). (2007). Downward arrowing. In *Cognitive Behavioral Therapy for Eating Disorders: A Comprehensive Treatment Guide* (pp. 179–182). Cambridge University Press. https://doi.org/10.1017/CBO9781139644204.023

Davey, E., Bennett, S. D., Bryant-Waugh, R., Micali, N., Takeda, A., Alexandrou, A., & Shafran, R. (2023). Low intensity psychological interventions for the treatment of feeding and eating disorders: A systematic review and meta-analysis. *Journal of Eating Disorders*, *11*(1), 56. https://doi.org/10.1186/s40337-023-00775-2

de la Torre, J. A. de la, Vilagut, G., Ronaldson, A., Valderas, J. M., Bakolis, I., Dregan, A., Molina, A. J., Navarro-Mateu, F., Pérez, K., Bartoll-Roca, X., Elices, M., Pérez-Sola,

V., Serrano-Blanco, A., Martín, V., & Alonso, J. (2023). Reliability and cross-country equivalence of the 8-item version of the Patient Health Questionnaire (PHQ-8) for the assessment of depression: Results from 27 countries in Europe. The Lancet Regional Health – Europe, 31. https://doi.org/10.1016/j.lanepe.2023.100659

Dear, B. F., Titov, N., Sunderland, M., McMillan, D., Anderson, T., Lorian, C., & Robinson, E. (2011). Psychometric comparison of the generalized anxiety disorder scale-7 and the Penn State Worry Questionnaire for measuring response during treatment of generalised anxiety disorder. Cognitive Behaviour Therapy, 40(3), 216–227. https://doi.org/10.1080/16506073.2011.582138

DSM-V. (2013). *Diagnostic and statistical manual of mental disorders, 5th edn.* American Psychiatric Association.

Fairburn, C. G. (2008). *Cognitive Behavior Therapy and Eating Disorders*. Guilford Press.

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

Fairburn, C. G., Shafran, R., & Cooper, Z. (1999). A cognitive behavioural theory of anorexia nervosa. *Behaviour Research and Therapy*, *37*(1), 1–13. https://doi.org/10.1016/S0005-7967(98)00102-8

Fairchild, H., & Cooper, M. (2010). A multidimensional measure of core beliefs relevant to eating disorders: Preliminary development and validation. *Eating Behaviors*, *11*(4), 239–246. https://doi.org/10.1016/j.eatbeh.2010.05.004

Forbush, K. (2021). Shape and Weight Concerns in the Context of ARFID Symptoms. In J. J. Thomas, K. T. Eddy, & K. R. Becker (Eds.), *The Picky Eater's Recovery Book: Overcoming Avoidant/Restrictive Food Intake Disorder* (pp. 180–206). Cambridge University Press. https://doi.org/10.1017/9781108854603.013

Greenberger, D., & Padesky, C. A. (1995). *Mind over mood: A cognitive therapy treatment manual for clients* (pp. xi, 215). Guilford Press.

Harrison, A. (2011). Falling Through the Cracks: UK Health Professionals' Perspective of Diagnosis and Treatment for Children and Adolescents with Avoidant/Restrictive Food Intake Disorder. *Child Care in Practice*, *0*(0), 1–11. https://doi.org/10.1080/13575279.2021.1958751

Hatoum, A. H., Burton, A. L., & Abbott, M. J. (2022a). Assessing negative core beliefs in eating disorders: Revision of the Eating Disorder Core Beliefs Questionnaire. *Journal of Eating Disorders*, *10*(1), 18. https://doi.org/10.1186/s40337-022-00542-9

Hatoum, A. H., Burton, A. L., & Abbott, M. J. (2022b). Validation of the revised eating disorder core beliefs questionnaire (ED-CBQ-R) in an Australian sample. *Clinical Psychologist*, *0*(0), 1–9. https://doi.org/10.1080/13284207.2022.2144717

Hay, P., Mitchison, D., Collado, A. E. L., González-Chica, D. A., Stocks, N., & Touyz, S. (2017). Burden and health-related quality of life of eating disorders, including Avoidant/Restrictive Food Intake Disorder (ARFID), in the Australian population. *Journal of Eating Disorders*, *5*(1), 21. https://doi.org/10.1186/s40337-017-0149-z

Hinz, A., Klein, A. M., Brähler, E., Glaesmer, H., Luck, T., Riedel-Heller, S. G., Wirkner, K., & Hilbert, A. (2017). Psychometric evaluation of the Generalized Anxiety Disorder

Screener GAD-7, based on a large German general population sample. Journal of Affective Disorders, 210, 338–344. https://doi.org/10.1016/j.jad.2016.12.012

Inal-Kaleli, I., Dogan, N., Kose, S., & Bora, E. (2025). Investigating the Presence of Autistic Traits and Prevalence of Autism Spectrum Disorder Symptoms in Anorexia Nervosa: A Systematic Review and Meta-Analysis. International Journal of Eating Disorders, 58(1), 66-90.Jhe, G. B., Chad-Friedman, E., Eldredge, O., Milliren, C. E., Lin, J., Carmody, J., Freizinger, M., Gearhart, M., Bern, E., & Richmond, T. (2024). Weight/Shape concerns in youth with Avoidant/Restrictive Food Intake Disorder (ARFID). *Eating Disorders*, *0*(0), 1–15. https://doi.org/10.1080/10640266.2024.2404788

Johnson, S. U., Ulvenes, P. G., Øktedalen, T., & Hoffart, A. (2019). Psychometric Properties of the General Anxiety Disorder 7-Item (GAD-7) Scale in a Heterogeneous Psychiatric Sample. Frontiers in Psychology, 10, 1713. https://doi.org/10.3389/fpsyg.2019.01713

Jones, C., Leung, N., & Greville-Harris, G. (2007). Dysfunctional core beliefs in eating disorders: A review. Journal of Cognitive Psychotherapy, 21(2).

Kambanis, P. E., Palmer, L. P., & Zayas, L. (2024). Differential Diagnosis Between Avoidant/Restrictive Food Intake Disorder and Anorexia Nervosa. Psychiatric Annals, 54(2), e37–e41. https://doi.org/10.3928/00485713-20240110-01

Kambanis, P. E., & Thomas, J. J. (2023). Assessment and Treatment of Avoidant/Restrictive Food Intake Disorder. *Current Psychiatry Reports*, *25*(2), 53–64. https://doi.org/10.1007/s11920-022-01404-6

Keski-Rahkonen, A., & Ruusunen, A. (2023). Avoidant-restrictive food intake disorder and autism: epidemiology, etiology, complications, treatment, and outcome. Current

opinion in psychiatry, 36(6), 438–442. https://doi.org/10.1097/YCO.0000000000000896

Krabbenborg, M. A. M., Danner, U. N., Larsen, J. K., van der Veer, N., van Elburg, A. A., de Ridder, D. T. D., Evers, C., Stice, E., & Engels, R. C. M. E. (2012). The Eating Disorder Diagnostic Scale: Psychometric Features Within a Clinical Population and a Cut-off Point to Differentiate Clinical Patients from Healthy Controls. *European Eating Disorders Review*, *20*(4), 315–320. https://doi.org/10.1002/erv.1144

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, *114*(1), 163–173. https://doi.org/10.1016/j.jad.2008.06.026

Lock, J., Le Grange, D., & Crosby, R. (2008). Exploring possible mechanisms of change in family-based treatment for adolescent bulimia nervosa. *Journal of Family Therapy*, *30*(3), 260–271. https://doi.org/10.1111/j.1467-6427.2008.00430.x

Loeb, K. L., Lock, J., Greif, R., & Le Grange, D. (2012). Transdiagnostic Theory and Application of Family-Based Treatment for Youth With Eating Disorders. *Cognitive and Behavioral Practice*, *19*(1), 17–30. https://doi.org/10.1016/j.cbpra.2010.04.005

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and Standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the General Population. Medical Care, 46(3), 266. https://doi.org/10.1097/MLR.0b013e318160d093

MacDonald, D. E., Liebman, R., & Trottier, K. (2024). Clinical characteristics, treatment course and outcome of adults treated for avoidant/restrictive food intake disorder

(ARFID) at a tertiary care eating disorders program. *Journal of Eating Disorders*, *12*(1), 15. https://doi.org/10.1186/s40337-024-00973-6

Manwaring, J. L., Blalock, D. V., Rienecke, R. D., Le Grange, D., & Mehler, P. S. (2024). A descriptive study of treatment-seeking adults with avoidant/restrictive food intake disorder at residential and inpatient levels of care. *Eating Disorders*, *32*(1), 13–28. https://doi.org/10.1080/10640266.2023.2241266

Nickerson, C. A., & Brown, N. J. L. (2019). Simpson's Paradox is suppression, but Lord's Paradox is neither: Clarification of and correction to Tu, Gunnell, and Gilthorpe (2008). *Emerging Themes in Epidemiology*, *16*(1), 5. https://doi.org/10.1186/s12982-019-0087-0

Nitsch, A., Knopf, E., Manwaring, J., & Mehler, P. S. (2021). Avoidant/Restrictive Food Intake Disorder (ARFID): Its Medical Complications and Their Treatment—an Emerging Area. *Current Pediatrics Reports*, *9*(2), 21–29. https://doi.org/10.1007/s40124-021-00239-8

Norris, M. L., Santos, A., Obeid, N., Hammond, N. G., Valois, D. D., Isserlin, L., & Spettigue, W. (2020). Characteristics and clinical trajectories of patients meeting criteria for avoidant/restrictive food intake disorder that are subsequently reclassified as anorexia nervosa. *European Eating Disorders Review*, *28*(1), 26–33. https://doi.org/10.1002/erv.2710

Olson, C. L. (1974). Comparative Robustness of Six Tests in Multivariate Analysis of Variance. *Journal of the American Statistical Association*, *69*(348), 894–908. https://doi.org/10.1080/01621459.1974.10480224

ONS. (2021). *Question development and research—Office for National Statistics*. Census 2021 Question Development and Research. https://www.ons.gov.uk/census/planningforcensus2021/questiondevelopment/questiondevelopmentandresearch

Rienecke, R. D., Johnson, C., Le Grange, D., Manwaring, J., Mehler, P. S., Duffy, A., McClanahan, S., & Blalock, D. V. (2022). Adverse childhood experiences among adults with eating disorders: Comparison to a nationally representative sample and identification of trauma profiles. *Journal of Eating Disorders*, *10*(1), 72. https://doi.org/10.1186/s40337-022-00594-x

Sader, M., Weston, A., Buchan, K., Kerr-Gaffney, J., Gillespie-Smith, K., Sharpe, H., & Duffy, F. (2025). The Co-Occurrence of Autism and Avoidant/Restrictive Food Intake Disorder (ARFID): A Prevalence-Based Meta-Analysis. The International Journal of Eating Disorders, 58(3), 473–488. https://doi.org/10.1002/eat.24369

Sanchez-Cerezo, J., Nagularaj, L., Gledhill, J., & Nicholls, D. (2023). What do we know about the epidemiology of avoidant/restrictive food intake disorder in children and adolescents? A systematic review of the literature. *European Eating Disorders Review*, *31*(2), 226–246. https://doi.org/10.1002/erv.2964

Sanchez-Cerezo, J., Neale, J., Julius, N., Croudace, T., Lynn, R. M., Hudson, L. D., & Nicholls, D. (2024). Subtypes of avoidant/restrictive food intake disorder in children and adolescents: A latent class analysis. *eClinicalMedicine*, *68*. https://doi.org/10.1016/j.eclinm.2024.102440

Santonicola, A., Siniscalchi, M., Capone, P., Gallotta, S., Ciacci, C., & Iovino, P. (2012). Prevalence of functional dyspepsia and its subgroups in patients with eating disorders.

World Journal of Gastroenterology, *18*(32), 4379–4385. https://doi.org/10.3748/wjg.v18.i32.4379

Sarin, S., & Abela, J. R. Z. (2003). The Relationship Between Core Beliefs and a History of Eating Disorders: An Examination of the Life Stories of University Students. *Journal of Cognitive Psychotherapy*, *17*(4), 359–374. https://doi.org/10.1891/jcop.17.4.359.52540

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse* (pp. xiv, 351). Guilford Press.

Sharp, W. G., & Stubbs, K. H. (2019). Avoidant/restrictive food intake disorder: A diagnosis at the intersection of feeding and eating disorders necessitating subtype differentiation. International Journal of Eating Disorders, 52(4), 398–401. https://doi.org/10.1002/eat.22987

Sharpe, S. L. (2024). A living experience proposal for the co-occurring diagnosis of avoidant/restrictive food intake disorder and other eating disorders. *Journal of Eating Disorders*, *12*(1), 110. https://doi.org/10.1186/s40337-024-01073-1

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine, 166(10), Article 10. https://doi.org/10.1001/archinte.166.10.1092

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2011). *Generalized Anxiety Disorder 7* [Dataset]. https://doi.org/10.1037/t02591-000

Stice, E. (2014). Eating Disorder Diagnostic Scale for DSM-5. Personal Communication. The scale and coding instructions can be found at http://www.ori.org/sticemeasures or available upon request from estice@stanford.edu.

Stice, E., Fisher, M., & Martinez, E. (2004). Eating Disorder Diagnostic Scale: Additional Evidence of Reliability and Validity. *Psychological Assessment*, *16*, 60–71. https://doi.org/10.1037/1040-3590.16.1.60

Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological Assessment*, *12*, 123–131. https://doi.org/10.1037/1040-3590.12.2.123

Teasdale, J. D., Segal, Z. V., Williams, J. M., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, *68*(4), 615–623. https://doi.org/10.1037//0022-006x.68.4.615

Thomas, J. J., Becker, K. R., Breithaupt, L., Murray, H. B., Jo, J. H., Kuhnle, M. C., Dreier, M. J., Harshman, S., Kahn, D. L., Hauser, K., Slattery, M., Misra, M., Lawson, E. A., & Eddy, K. T. (2021). Cognitive-behavioral therapy for adults with avoidant/restrictive food intake disorder. *Journal of Behavioral and Cognitive Therapy*, *31*(1), 47–55. https://doi.org/10.1016/j.jbct.2020.10.004

Thomas, J. J., Becker, K. R., Kuhnle, M. C., Jo, J. H., Harshman, S. G., Wons, O. B., Keshishian, A. C., Hauser, K., Breithaupt, L., Liebman, R. E., Misra, M., Wilhelm, S., Lawson, E. A., & Eddy, K. T. (2020). Cognitive-behavioral therapy for avoidant/restrictive food intake disorder: Feasibility, acceptability, and proof-of-

concept for children and adolescents. *The International Journal of Eating Disorders*, *53*(10), 1636–1646. https://doi.org/10.1002/eat.23355

Thomas, J. J., & Eddy, K. E. (2018). *Cognitive-Behavioral Therapy for Avoidant/Restrictive Food Intake Disorder: Children, Adolescents, and Adults*. Cambridge University Press.

Thomas, J. J., Lawson, E. A., Micali, N., Misra, M., Deckersbach, T., & Eddy, K. T. (2017). Avoidant/Restrictive Food Intake Disorder: A Three-Dimensional Model of Neurobiology with Implications for Etiology and Treatment. *Current Psychiatry Reports*, *19*(8), 54. https://doi.org/10.1007/s11920-017-0795-5

Thomas, J. J., Wons, O., & Eddy, K. (2018). Cognitive-Behavioral Treatment of Avoidant/Restrictive Food Intake Disorder. *Current Opinion in Psychiatry*, *31*(6), 425–430. https://doi.org/10.1097/YCO.0000000000000454

Trompeter, N., Bussey, K., Hay, P., Griffiths, S., Murray, S. B., Mond, J., Lonergan, A., Pike, K. M., & Mitchison, D. (2019). Fear of negative evaluation among eating disorders: Examining the association with weight/shape concerns in adolescence. *International Journal of Eating Disorders*, *52*(3), 261–269. https://doi.org/10.1002/eat.23018

Vanzhula, I. A., Wang, E., Martinelli, M. K., Schreyer, C., & Guarda, A. S. (2023). Inpatient hospital course and self-reported symptomatology in underweight adults with ARFID compared to age- and sex-matched controls with anorexia nervosa. *Journal of Eating Disorders*, *11*(1), 206. https://doi.org/10.1186/s40337-023-00912-x

Wagner, A. F., Zickgraf, H. F., & Lane-Loney, S. (2020). Caregiver accommodation in adolescents with avoidant/restrictive food intake disorder and anorexia nervosa: Relationships with distress, eating disorder psychopathology, and symptom change. *European Eating Disorders Review*, *28*(6), 657–670. https://doi.org/10.1002/erv.2786

Willson, R., & Branch, R. (2019). Cognitive Behavioural Therapy For Dummies. John Wiley & Sons.

Wu, Y., Levis, B., Riehm, K. E., Saadat, N., Levis, A. W., Azar, M., Rice, D. B., Boruff, J., Cuijpers, P., Gilbody, S., Ioannidis, J. P. A., Kloda, L. A., McMillan, D., Patten, S. B., Shrier, I., Ziegelstein, R. C., Akena, D. H., Arroll, B., Ayalon, L., … Thombs, B. D. (2019). Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: a systematic review and individual participant data meta-analysis. Psychological Medicine, 50(8), 1368–1380. https://doi.org/10.1017/S0033291719001314Young, M. A., Boerner, K. E., Marshall, S., Dhariwal, A., & Coelho, J. S. (2024). Exploring Family Care Journeys to Inform Cognitive-Behavioral Therapy for Avoidant/Restrictive Food Intake Disorder and Somatic Symptom Disorders. *Cognitive and Behavioral Practice*, *31*(3), 356–366. https://doi.org/10.1016/j.cbpra.2024.01.002

Zanna, V., Criscuolo, M., Mereu, A., Cinelli, G., Marchetto, C., Pasqualetti, P., Tozzi, A. E., Castiglioni, M. C., Chianello, I., & Vicari, S. (2021). Restrictive eating disorders in children and adolescents: A comparison between clinical and psychopathological profiles. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, *26*(5), 1491–1501. https://doi.org/10.1007/s40519-020-00962-z

Zickgraf, H. F., Franklin, M. E., & Rozin, P. (2016). Adult picky eaters with symptoms of avoidant/restrictive food intake disorder: Comparable distress and comorbidity but different eating behaviors compared to those with disordered eating symptoms. *Journal of Eating Disorders*, *4*(1), 26. https://doi.org/10.1186/s40337-016-0110-6

Zucker, N. L., & Bulik, C. M. (2020). On bells, saliva, and abdominal pain or discomfort: Early aversive visceral conditioning and vulnerability for anorexia nervosa. *The*

International Journal of Eating Disorders, *53*(4), 508–512.

https://doi.org/10.1002/eat.23255

Executive Summary

Avoidant Restrictive Feeding Intake Disorder (ARFID) is a relatively new eating disorder diagnosis. It describes eating difficulties that are caused by sensory sensitivity to food (e.g. texture), fear of an aversive outcome (e.g. choking), or lack of appetite. There is limited research into whether ARFID patients experience unhelpful cognitions that might be an important target for treatment. Evidence-based treatments for other eating disorders such as Anorexia Nervosa (AN) focus on addressing unhelpful cognitions and behaviours that are maintaining restrictive eating patterns. Currently, interventions for ARFID are in the early stages of being piloted and developed and so their evidence base is limited. However, most of the interventions described in the research literature for ARFID prioritise targeting problematic behaviours and neglect to focus on unhelpful beliefs.

Cognitive-behavioural models of cognitions often conceptualise them as occurring at three levels of depth:

- Automatic thoughts (these are immediate, rapid, surface-level thoughts that often arise in response to a specific situation, for example, “I have made a mistake in my exam”, “I will never succeed at life”)
- Assumptions (these contain ‘rules for living’ or assumptions about the world such as “I should never make mistakes”, “If I am not top of the class, people will reject me”)
- Core beliefs (these are deep-seated beliefs about oneself, others or the world, such as “I am a failure”, “I am unworthy”).

This study aimed to investigate whether ARFID patients experience unhelpful cognitions that are usually be associated with other eating disorders. It recruited 184

participants (68 AN patients, 62 ARFID patients, and 55 people with no eating disorder). Participants were an adult community sample, mostly recruited through social media sites, who completed an anonymous online survey and were screened using Diagnostic Statistical Manual (DSM-5) criteria. They were given questionnaires that measured the presence of core beliefs, assumptions and automatic thoughts that are associated with other eating disorders. They were also assessed for the presence of body image concerns relating to dissatisfaction with their weight or shape. An observational design was then used to compare how participants' responses varied according to their clinical status.

The results revealed that ARFID patients did indeed experience significant levels of unhelpful cognitions associated with other eating disorders. They experienced more of these cognitions (core beliefs, assumptions and automatic thoughts) than people with no eating disorder, but less of these cognitions compared to AN patients. ARFID patients also experienced more weight/shape concerns than people with no eating disorder, but less of these concerns than AN patients. Another finding was that disordered assumptions and core beliefs in patients with ARFID were linked to higher frequencies of behaviours aimed at preventing weight gain.

These findings have implications for the clinical assessment and diagnosis of patients. Diagnostic manuals (such as the DSM-5) currently state that ARFID and AN are mutually exclusive diagnoses, and that ARFID can be distinguished from AN on the basis of a lack of weight/shape concerns. However, the findings above add to a growing literature that challenges this view by showing that ARFID patients may experience weight/shape concerns and that there are overlapping features between ARFID and AN. The findings also have

implications for clinical treatments for ARFID. They suggest that unhelpful beliefs might be important to address since they may be maintaining disordered patterns of eating.

Reflective Narrative

Growing a Thesis

The following is a reflective summary of the steps taken to complete the three projects that constitute this thesis. Although these projects appear very different to each other, they are united by a theme of challenging prevailing models or attitudes to achieve better care for people's mental health.

Planting the seeds: developing research interests

Prior to the doctorate, I had done lots of voluntary work with children and families and gained clinical experience as an Assistant Psychologist in a CAMHS specialist community team. This paved the foundations of my Service Improvement Project (SIP) focused on improving parents' wellbeing when their child was acutely sick. I was interested in what helped families to maintain resilience in the face of a stressor. Translating this to a highly medicalised setting where psychologists were 'outnumbered' by medical professions was something I was curious to learn more about – how do psychologists challenge dominant perspectives and work alongside other professions to influence their practice? This project involved supporting healthcare staff to not just focus on the baby's medical needs but to care for the family as a whole.

The seeds of other interests were also laid by my previous work experiences. As an Assistant Psychologist, I witnessed debates arise between different teams as to who was going to provide interventions for people with ARFID. During the pandemic, there could be a tribalism between teams in their efforts to limit mounting waiting lists by tightening their referral criteria and gatekeeping. It seemed the eating disorders team was the most natural place in terms of their skill mix and knowledge to monitor physical health risks, yet they argued that they were only "commissioned to see eating disorders but not disordered eating." I questioned this distinction between 'eating disorder vs disordered eating' – was it that clear

cut? From clinical experience, it seemed that ARFID patients frequently had eating habits and motivations that overlapped with Anorexia, and vice versa (e.g. Anorexic patients with neurodiversity who displayed sensory sensitivities to food). Nevertheless, the eating disorders team seemed to routinely dismiss clients on the basis that their primary motivator for restricting food was not weight/shape concerns, even if these concerns were still significant. I wanted to challenge this view of ARFID as wholly distinct from other eating disorders through my Theory Driven Research Project (TDRP).

My Systematic Review of Literature (SRL) also had a view to challenge prevailing narratives. My interest in CBT for sleep disorders for psychosis patients started during my undergraduate degree at the University of Oxford. The second-year exams were brutally stressful, and in the run up to my second-year exams, I simply could not sleep. I found myself in a vicious cycle where stress prevented me from sleeping, but lack of sleep made me more stressed. After surviving these exams, I chose to investigate the links between sleep and mental health for my final-year dissertation topic (the main conclusion being that sleep is *very* important!). During this dissertation, I reached out to the Oxford Cognitive Approaches for Psychosis (O-CAP) team to find out more about their work, and they subsequently offered me a summer research internship to review the literature for sleep and psychosis. This allowed me to immerse myself in the field. I was intrigued by how management of psychotic disorders had traditionally been very medicalised involving potent drugs with unpleasant side-effects. O-CAP offered a very different approach by listening closely to patients' concerns about what was affecting their quality of life and targeting key maintenance factors of psychosis such as worry, low mood and sleep. They were developing randomised-controlled trials of CBT for sleep in this population. No known literature review had investigated CBT interventions for sleep in this population, and so I was keen to summarise the evidence and promote awareness of this when I began my doctoral training.

Cultivating the projects: methodology and supervision

Research is something which I find very fulfilling. There is something beautiful about the process of having a question, rigorously gathering evidence, and generating a report to share with the world the answers that have been found. Whilst quantitative methods are my comfort-zone, I was stretched to use reflexive thematic analysis for my SIP. This was a good learning experience but mixed with some frustration as transcribing and coding were arduous and time-consuming. I am not so patient!

A uniting theme across all the projects, is the fantastic research supervision I have received. My supervisors have all been kind, helpful and wise, and they have shared a genuine passion in the areas we have collaborated on. From each one, I have learnt what it means to be a 'good research supervisor' and experienced the difference that it makes to receive good guidance, grow in confidence, and be enabled to develop my research skills. It is something which I will hold onto for my professional development in the future.

Branching out: taking this forwards

I feel a responsibility to disseminate the results in a way that creates meaningful change. I am hoping to publish them and, following the SIP, I plan to deliver training to nurses and doctors to increase parents' support. Conducting these research projects has given me the confidence to continue research alongside my clinical work where I perceive there to be a need. I look forward to developing new ideas and to also perhaps supervising projects in the future.

Acknowledgements

First and foremost, I give thanks to God, who has answered many of my prayers along the way and given me the abilities, opportunities, and support to complete this project. I pray that I put this training into the care and service of others.

Thank you to the participants and experts-by-experience who were willing to give their time and share their perspectives in taking part in these studies.

To my wonderful research supervisors Myra, Alex, Louise, Matt, Jacinta, Kat, and Anne, thank you for always being very kind and patient, welcoming my questions, offering brilliant guidance and making these projects the best they could be. Thank you so much Ciorsdan for being an inspiration through your constant warmth, compassion, encouragement and wisdom, which helped me navigate through some very difficult moments over the past three years.

Lewis, my soon-to-be husband, you have been the best friend that anyone could wish for and unfailingly been there to support me. I am so grateful for all your love and understanding and the many times I have been able to count on you for help. To our dearest Blessing. To my mum and dad, thank you too for your love and guidance over the years that has helped me reach this dream of becoming a clinical psychologist.

Appendix: Systematic Literature Review

Appendix 1 – PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) checklist: recommended items to address in a systematic review protocol

Section and topic	Item No	Checklist item	
ADMINISTRATIVE INFORMATION			
Title:		Cognitive behavioural therapy for sleep problems in psychosis: A systematic review of effectiveness and acceptability	
Identification	1a	Identify the report as a protocol of a systematic review	V
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	V
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	V
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	V
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	V
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	V
Support:			
Sources	5a	Indicate sources of financial or other support for the review	V
Sponsor	5b	Provide name for the review funder and/or sponsor	V
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	V
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	V
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	V
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	V
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	V
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	V
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	V
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	V

Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	V
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	V
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	V
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	V
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	V
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	V
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	N/A
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	V
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	N/A
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	V

Appendix 2 – PICO table

Table 2.1 *Inclusion criteria*

Dimension	Inclusion Criteria
General	The study either (a) has an experimental design (i.e., a trial with a treatment arm delivering a CBT-based intervention and obtaining quantitative outcomes; the trial may be controlled or non-controlled); OR (b) uses qualitative research methods evaluating the subjective experiences of the CBT delivered in (a). Published in a peer-reviewed journal (i.e. not a dissertation, non-peer reviewed conference abstract). Published in English (the language spoken by the researchers).
Population	At least a third of participants have psychotic symptoms as assessed by either (a) formal diagnosis of a psychotic disorder, OR (b) meeting the criteria for 'at risk mental state' with the use of a comprehensive measure, such as the 'Comprehensive Assessment of At-Risk Mental States' (CAARMS)[61]. Of note, DSM-V and ICD-11 psychotic spectrum disorders include the following[114,115]: • schizophreniform disorder • schizoaffective disorder • schizotypal disorders • brief psychotic disorder • acute and transient psychotic disorder • delusional disorder • psychosis not otherwise specified Participants present with a level of sleep difficulties which are *not* in the context of a medical sleep disorder (e.g. narcolepsy or obstructive sleep apnoea) or neurodegenerative disease (e.g. dementia).
Intervention	Trials evaluate a CBT-based intervention that targets an aspect of sleep quality. This includes CBT directed at insomnia (i.e., difficulties with falling asleep, staying asleep, early waking), nightmares, or excessive sleeping. Trials have a minimum of 10 participants per arm to test quantitative outcomes. For studies with a qualitative design, there is no minimum number of participants.
Comparator/control	All trials are required to report comparisons in clinical outcomes pre- and post-treatment. Randomised controlled trials should additionally report comparisons between a treatment and control group. Qualitative studies are not required to have a comparator/control.
Outcome	Trials include quantitative outcomes of the following: sleep quality and/or psychotic symptoms. Overall intervention effects are reported. Qualitative studies describe participant's subjective experiences of engaging in treatment.

Note. 'DSM-V' refers to the Diagnostic and Statistical Manual of Mental Disorders.[114] *'ICD-11' refers to International Classification of Diseases.*[115]

Appendix 3 - Search terms

A.3.1 Search terms for Embase:

Step	Search History
1	("cognitive behavio* therapy" or CBT or CBT-I).ab,ti.
2	cognitive behavioral therapy/
3	1 or 2
4	(sleep or nightmares or insomnia or asleep or hypersomnia).ab,ti.
5	insomnia/
6	nightmare/
7	4 or 5 or 6
8	(psychosis or psychotic or schizophren* or bipolar or hallucinat* or delusion* or persecutory).ab,ti.
9	psychosis/ or affective psychosis/ or schizoaffective psychosis/ or paranoid psychosis/
10	bipolar disorder/
11	8 or 9 or 10
12	3 and 7 and 11

A.3.2 Search terms for Medline:

Step	Search History
1	("cognitive behavio* therapy" or CBT or CBT-I).ab,ti.
2	Cognitive Behavioral Therapy/
3	1 or 2
4	(sleep* or nightmares or insomnia or asleep or hypersomnia).ab,ti.
5	"Sleep Initiation and Maintenance Disorders"/
6	4 or 5
7	(psychosis or psychotic or schizophren* or bipolar or hallucinat* or delusion* or persecutory).ab,ti.
8	Psychotic Disorders/
9	Bipolar Disorder/ or Affective Disorders, Psychotic/
10	Schizophrenia, Paranoid/
11	Schizophrenia/
12	8 or 9 or 10 or 11
13	7 or 12
14	3 or 6 or 13
15	3 and 6 and 13

A.3.3 Search terms for PsycInfo:

Step	Search History
1	("cognitive behavio* therapy" or CBT or CBT-I).ab,ti.
2	Cognitive Behavior Therapy/
3	1 or 2
4	(sleep or nightmares or insomnia or asleep or hypersomnia).ab,ti.
5	Insomnia/
6	Nightmares/
7	4 or 5 or 6
8	(psychosis or psychotic or schizophren* or bipolar or hallucinat* or delusion* or persecutory).ab,ti.
9	psychosis/ or affective psychosis/ or delusional disorder/ or exp paranoid psychosis/ or exp schizophrenia/
10	8 or 9
11	3 and 7 and 10

A.3.4 Search terms for PubMed:

(("cognitive behavior therapy"[Title/Abstract] OR "cognitive behavioral therapy"[Title/Abstract] OR "cognitive behaviour therapy"[Title/Abstract] OR "cognitive behavioural therapy"[Title/Abstract] OR CBT[Title/Abstract] OR CBT-I[Title/Abstract]) OR ("cognitive behavioral therapy"[MeSH Terms])) AND ((sleep[Title/Abstract] OR nightmares[Title/Abstract] OR insomnia[Title/Abstract] OR asleep[Title/Abstract] OR hypersomnia[Title/Abstract]) OR (sleep quality[MeSH Terms]) OR ("sleep initiation and maintenance disorders"[MeSH Terms])) AND ((psychosis[Title/Abstract] OR psychotic[Title/Abstract] OR schizophren*[Title/Abstract] OR bipolar[Title/Abstract] OR hallucinat*[Title/Abstract] OR delusion*[Title/Abstract] OR persecutory[Title/Abstract]) OR ("psychotic disorders"[MeSH Terms]) OR ("affective disorders, psychotic"[MeSH Terms]) OR ("bipolar and related disorders"[MeSH Terms]) OR ("schizophrenia"[MeSH Terms] OR "schizophrenia spectrum and other psychotic disorders"[MeSH Terms]))

Appendix 4 – Qualitative checklist ratings, developed from Braun and Clarke's reviewing guidelines

A.4.1 Quality ratings for Taylor et al. (2020)[35]

	Evaluating the methods and methodology	Yes/ Probably yes	Maybe/ partly	No/ Probably no	Comments
1	Is the use of TA explained (even if only briefly)?	V			Yes: "To explore patterns within the qualitative interview data"
2	Do the authors clearly specify and justify which type of TA they are using?			V	They reference Braun and Clarke's paper but don't actually say what type they are using
3	Is the use and justification of the specific type of TA consistent with the research questions or aims?			V	At a guess, they have used a fairly realist approach which fits with their aim of understanding the intervention acceptability, but it's not clear what they have even done regarding use, and it's not justified
4	Is there a good 'fit' between the theoretical and conceptual underpinnings of the research and the specific type of TA (conceptual coherence)?			V	No information about the kind of TA
5	Is there a good 'fit' between the methods of data collection and the specific type of TA?			V	No information about how data was collected
6	Is the specified type of TA consistently enacted throughout the paper?			V	No information about the type of TA
7	Is there evidence of problematic assumptions about TA? These commonly include (mark 'yes' if evidence of any of these errors, and indicate where they are): -Treating TA as one, homogenous, entity, with one set of – widely agreed on – procedures. -Assuming grounded theory concepts and procedures (e.g. saturation, constant comparative analysis, line-by-line coding) apply to TA without any explanation or justification. -Assuming TA is essentialist or realist, or atheoretical. -Assuming TA is only a data reduction or descriptive approach and thus has to be supplemented with other methods and procedures to achieve other ends.	V			Treat TA as one homogenous entity Seem to assume it is only data reductionist approach Seem to assume TA is atheoretical from the fact that they don't mention a theoretical approach – they simply assume a realist position it seems They don't supplement TA with other methods
8	(Skip if not applicable) Are any supplementary procedures or methods justified and necessary or could the same results have been achieved simply by using TA more effectively?				N/A
9	Are the theoretical underpinnings of the use of TA clearly specified (e.g. ontological, epistemological assumptions, guiding theoretical framework(s)), even when using TA inductively (inductive TA does not equate to analysis in a theoretical vacuum)?			V	
10	Do the researchers strive to 'own their perspectives' (even if only very briefly); their personal and social standpoint and positioning? (This is especially important when the researchers are engaged in social justice-oriented research and when representing the 'voices' of marginal and vulnerable			V	Not really They mention that as a researcher they may have influenced outcomes by being involved in the other stages of research, but this is not their personal/social standpoint

	groups, and groups to which the researcher does not belong.)				
11	Are the analytic procedures used clearly outlined?			V	No description
12	Is there evidence of conceptual and procedural confusion? For example, reflexive TA (Braun & Clarke, 2006) is the claimed approach but different procedures are outlined such as the use of a codebook or coding frame, multiple independent coders and consensus coding, inter-rater reliability measures, and/or themes are conceptualised as analytic inputs rather than outputs and therefore the analysis progresses from theme identification to coding (rather than coding to theme development).		V		Possibly…they cite Braun & Clarke but it's very unclear
13	Have the authors fully understood their claimed approach to TA?			V	No evidence to suggest this
	Evaluating the analysis				
14	Is it clear what and where the themes are in the report? Would the manuscript benefit from some kind of overview of the analysis: listing of themes, narrative overview, table of themes, thematic map?	V			Themes in a clear table
15	Are themes reported domain summaries rather than fully realised themes? Have the data collection questions been used as themes? Are domain summaries appropriate to the purpose of the research? (If so, if the authors are using reflexive TA, is this modification in the conceptualisation of themes explained and justified?) Would the manuscript benefit from further analysis being undertaken and the reporting of fully realised themes? Or, if the authors are claiming to use reflexive TA, would the manuscript benefit from claiming to use a different type of TA (e.g. coding reliability or codebook)?	V V	 V V V		They are domain summaries, and yes, it seems that the data collection questions are the themes
16	Is a non-thematic contextualising information presented as a theme? (e.g. the first theme is a domain summary providing contextualising information, but the rest of the themes reported are fully realised themes) Would the manuscript benefit from this being presented as non-thematic contextualising information?			V	Contextualising information isn't presented as a theme, so this doesn't really apply
17	In applied research, do the reported themes give rise to actionable outcomes?	V			Yes – clear implications for future interventions
18	Are there conceptual clashes and confusion in the paper? (e.g. claiming a social constructionist approach while also expressing concern for positivist notions of coding reliability, or claiming a constructionist approach while treating participants' language as a transparent reflection of their experiences and behaviours)			V	No claims about conceptual standpoint – so no scope for contradiction
19	Is there evidence of weak or unconvincing analysis? Too many or two few themes? Too many theme levels? Confusion between codes and themes? Mismatch between data extracts and analytic claims? Too few or too many data extracts? Overlap between themes?		V		Relatively superficial analysis although the themes, extracts and analytic claims do largely overlap

20	Do authors make problematic statements about the lack of generalisability of their results, and implicitly conceptualise generalisability as statistical-generalisability?		V		They mention this, but it isn't clear that it's a problem given their research aims

A.4.2 Quality ratings for Waite et al. (2018)[53]

	Evaluating the methods and methodology	Yes/ Probably yes	Maybe/ Partly	No/ Probably no	Comments
1	Is the use of TA explained (even if only briefly)?	V			Organising, encoding and identifying patterns
2	Do the authors clearly specify and justify which type of TA they are using?			V	They report using Braun & Clarke's method but that isn't a type of TA
3	Is the use and justification of the specific type of TA consistent with the research questions or aims?			V	No justification given
4	Is there a good 'fit' between the theoretical and conceptual underpinnings of the research and the specific type of TA (conceptual coherence)?		V		Seems to be a realist/essentialist approach – makes sense that they are concerned about accuracy and reliability
5	Is there a good 'fit' between the methods of data collection and the specific type of TA?		V		According with the above, yes, but difficult to say when they haven't specified what type of TA they are using. If it is reflexive TA then no
6	Is the specified type of TA consistently enacted throughout the paper?		V		Not specified but there is potentially consistency
7	Is there evidence of problematic assumptions about TA? These commonly include: -Treating TA as one, homogenous, entity, with one set of – widely agreed on – procedures. -Assuming grounded theory concepts and procedures (e.g. saturation, constant comparative analysis, line-by-line coding) apply to TA without any explanation or justification. -Assuming TA is essentialist or realist, or atheoretical. -Assuming TA is only a data reduction or descriptive approach and thus has to be supplemented with other methods and procedures to achieve other ends.	V V V			
8	(Skip is not applicable) Are any supplementary procedures or methods justified and necessary or could the same results have been achieved simply by using TA more effectively?				
9	Are the theoretical underpinnings of the use of TA clearly specified (e.g. ontological, epistemological assumptions, guiding theoretical framework(s)), even when using TA inductively (inductive TA does not equate to analysis in a theoretical vacuum)?			V	
10	Do the researchers strive to 'own their perspectives' (even if only very briefly); their personal and social standpoint and positioning? (This is especially important when the researchers are engaged in social justice-oriented research and when representing the			V	

	'voices' of marginal and vulnerable groups, and groups to which the researcher does not belong.)				
11	Are the analytic procedures used clearly outlined?	V			
12	Is there evidence of conceptual and procedural confusion? For example, reflexive TA (Braun & Clarke, 2006) is the claimed approach but different procedures are outlined such as the use of a codebook or coding frame, multiple independent coders and consensus coding, inter-rater reliability measures, and/or themes are conceptualised as analytic inputs rather than outputs and therefore the analysis progresses from theme identification to coding (rather than coding to theme development).			V	
13	Have the authors fully understood their claimed approach to TA?		V		
	Evaluating the analysis				
14	Is it clear what and where the themes are in the report? Would the manuscript benefit from some kind of overview of the analysis: listing of themes, narrative overview, table of themes, thematic map?	V			
15	Are themes reported domain summaries rather than fully realised themes? -Have the data collection questions been used as themes? -Are domain summaries appropriate to the purpose of the research? (If so, if the authors are using reflexive TA, is this modification in the conceptualisation of themes explained and justified?) -Would the manuscript benefit from further analysis being undertaken and the reporting of fully realised themes? -Or, if the authors are claiming to use reflexive TA, would the manuscript benefit from claiming to use a different type of TA (e.g. coding reliability or codebook)?	V	V V V V N A		
16	Is a non-thematic contextualising information presented as a theme? (e.g. the first theme is a domain summary providing contextualising information, but the rest of the themes reported are fully realised themes). Would the manuscript benefit from this being presented as non-thematic contextualising information?			V	Contextualising information isn't one of the theme so it doesn't really apply
17	In applied research, do the reported themes give rise to actionable outcomes?	V			
18	Are there conceptual clashes and confusion in the paper? (e.g. claiming a social constructionist approach while also expressing concern for positivist notions of coding reliability, or claiming a constructionist approach while treating participants' language as a transparent reflection of their experiences and behaviours)		V		Difficult to judge as it doesn't really claim a position

19	Is there evidence of weak or unconvincing analysis? -Too many or two few themes? -Too many theme levels? -Confusion between codes and themes? -Mismatch between data extracts and analytic claims? -Too few or too many data extracts? -Overlap between themes?		V		Perhaps some confusion between what is a code and what is a theme; it's rather superficial and doesn't 'tell a story' – it is broken down into what's helpful/unhelpful etc. which suggests a superficial analysis.
20	Do authors make problematic statements about the lack of generalisability of their results, and implicitly conceptualise generalisability as statistical-generalisability?		V		They do make claims about this, but it isn't fully problematic given their realist positions/aims

Appendix 5 - Means, standard deviations and numbers of participants for outcome measures

A.5.1 Sleep problems

Table S.5.1 *Primary sleep measure: means, standard deviations and participant numbers*

Study	Primary Outcome	Time-point	Results	
			Treatment	Control
Batalla-Martin et al. (2023)[37]	ISI	Baseline Post-treatment Three months	15.0 (4.2); n = 20 8.1 (4.0); n = 20 6.7 (3.5); n = 20	15.6 (4.5); n = 20 15.3 (4.9); n = 20 15.8 (4.3); n = 20
Freeman et al., (2015)[24]	ISI	Baseline Post-treatment Three months	18.6 (3.2); n = 24 9.3 (5.5); n = 22 11.0 (5.6); n = 23	Pre: 18.8 (3.3); n = 26 Post: 15.4 (5.4); n = 25 15.0 (5.7); n = 25
Sheaves et al., (2018)[25]	ISI	Baseline Post-treatment Two weeks Ten weeks	17.1 (6.0); n = 20 8.5 (5.4); n = 20 6.8 (5.2); n = 19 5.8 (4.9); n = 16	16.1 (4.9); n = 20 12.5 (5.5); n = 20 10.1 (5.6); n = 20 8.6 (4.4); n = 18
Sheaves et al., (2019)[26]	DDNSI	Baseline Post-treatment Six months	21.6 (6.9); n = 12 14.2 (8.8); n = 12 12.6 (8.6); n = 11	23.0 (6.4); n = 12 22.6 (7.1); n = 11 22.1 (8.2); n = 9
Waite et al., (2023)[27]	ISI	Baseline Post-treatment Six months	19.2 (2.8); n = 21 6.3 (4.9); n = 21 8.3 (6.3); n = 21	18.7 (3.3); n = 19 14.3 (5.8); n = 18 13.9 (5.8); n = 18
Hwang et al., (2019)[28]*	ISI	Baseline Post-treatment One month	17.7 (0.5); n = 31 12.2 (0.6); n = 31 11.4 (0.6); n = 31	18.5 (0.5); n = 32 18.2 (0.6); n = 32 17.7 (0.6); n = 32
Bradley et al., (2018)[29]	ISI	Baseline Post-treatment One month	17.2 (1.2); n = 11 9.1 (5.1); n = 11 9.1 (4.6); n = 11	None
Haynes et al., (2011)[30]	ISI	Baseline Post-treatment	17.53 (7.33); n = 19 13.36 (7.52); n = 19	None
Holmes et al., (1995)[31]	Subjective sleep rating	Baseline Post-treatment	Means and SDs not reported; simply state that repeated measures ANOVA, looking at change found a significant difference: subjective rating of sleep ($E(1,16) = 6.12$, $p < .05$); n = 18	None
Myers et al., (2011)[32]	ISI	Baseline Post-treatment One month	20.93 (3.45); n = 15 9.13 (5.30); n = 15 10.20 (4.63); n = 15	None
Novak et al., (2020)[33]	ISI	Baseline Post-treatment	18 (no SD reported) 14 (no SD reported)	None
Waters et al., (2020)[34]	PSQI	Baseline Post-treatment	13.4 (3.4); n = 50 10.3 (3.5); n = 40	Not reported

Taylor et al., (2022)[35]	ISI	Baseline Post-treatment	18.45 (4.99); n = 11 12.91 (5.75); n = 11	None

Note: This table reports data assessed by the primary sleep measure in the treatment and control groups at each time-point. Data are presented as mean, (SD), and number of participants completing the assessment measure.
For Hwang et al.'s study[28] study, the asterisk (*) indicates to treat their findings with caution given the unusually small standard deviations.
Primary measures include: ISI = Insomnia Severity Index,[55] DDNSI = Disturbing Dream and Nightmare Severity Index, [56] and PSQI = Pittsburgh Sleep Quality Index.[57] Holmes et al. [31] report a subjective sleep rating that is extracted from a sleep log published elsewhere[58] but provide minimal further details. Where studies did not specify a single primary sleep measure, the most relevant measure to evaluate the intervention (e.g. an insomnia measure for a study of CBT-I) is reported.

A.5.2 Psychotic outcomes

Table S.5.2 *Psychotic symptoms: means, standard deviations and participant numbers*

Study	Primary Outcome		Results	
			Treatment	**Control**
Batalla-Martin et al. (2023)[37]	None			
Freeman et al., (2015)[24]	PSYRATS (delusions)	Baseline	16.1 (3.2); n = 24	15.3 (4.9); n = 26
		Post-treatment	13.9 (4.8); n = 22	13.8 (4.1); n = 25
		Three months	14.0 (4.7); n = 23	12.7 (5.7); n = 25
	PSYRATS (hallucinations)	Baseline	25.1 (12.1); n = 24	26.7 (9.2); n = 26
		Post-treatment	27.5 (9.2); n = 22	25.9 (8.1); n = 25
		Three months	24.6 (11.6); n = 23	22.0 (10.2); n = 25
	GPTS	Baseline	90.8 (28.7); n = 24	90·5 (29.8); n = 26
		Post-treatment	89.6 (36.8); n = 22	96.2 (37.3); n = 24
		Three months	78.3 (34.8); n = 20	88·1 (35.0); n = 25
	PANSS	Baseline	83.6 (16.2); n = 24	79.7 (14.1); n = 26
		Post-treatment	77.5 (12.1); n = 22	79.3 (14.6); n = 24
		Three months	74.8 (14.7); n = 21	75.8 (11.8); n = 24
Sheaves et al., (2018)[25]	PANSS Total	Baseline	68.3 (16.3); n = 20	68.4 (14.9); n = 20
		Post-treatment	56.4 (9.2); n = 20	61.1 (15.5); n = 20
		Two weeks	53.9 (12.6); n = 19	54.9 (14.7); n = 20
		Ten weeks	50.4 (11.7); n = 16	55.8 (19.0); n = 18
	PANSS (positive symptoms)	Baseline	15.3 (6.6); n = 20	15.4 (5.2); n = 20
		Post-treatment	12.2 (4.8); n = 20	12.5 (4.6); n = 20
		Two weeks	11.2 (3.9); n = 19	11.2 (4.3); n = 20
		Ten weeks	9.4 (2.9); n = 16	10.4 (3.5); n = 18
	PANSS (negative symptoms)	Baseline	14.7 (6.0); n = 20	13.9 (4.3); n = 20
		Post-treatment	12.8 (4.1); n = 20	13.8 (5.7); n = 20
		Two weeks	11.9 (4.6); n = 19	13.6 (5.1); n = 20
		Ten weeks	11.9 (3.5); n = 16	14.9 (7.4); n = 18
	PANSS (psychopathology)	Baseline	38.4 (9.2); n = 20	39.2 (8.3); n = 20
		Post-treatment	31.4 (6.6); n = 20	34.7 (8.5); n = 20
		Two weeks	30.8 (8.8); n = 19	30.2 (8.0); n = 20
		Ten weeks	29.1 (8.4); n = 16	30.5 (11.8); n = 18
Sheaves et al., (2019)[26]	CAPS	Baseline	17.7 (7.7); n = 12	18.8 (7.1); n = 12
		Post-treatment	15.5 (7.7); n = 12	16.8 (7.3); n = 10
		Six months	15.8 (7.8); n = 11	16.7 (10.1); n = 9
	GPTS	Baseline	101.2 (35.7); n = 12	109.8 (33.9); n = 12
		Post-treatment	75.3 (37.0); n = 12	109.0 (32.3); n = 10
		Six months	68.5 (39.4); n = 11	100.7 (35.5); n = 9
Waite et al., (2023)[27]	CAARMS	Baseline	50.8 (11.2); n = 21	45.4 (12.1); n = 19
		Post-treatment	42.2 (20.9); n = 19	44.6 (21.6); n = 16
		Six months	38.8 (21.4); n = 20	39.3 (26.6); n = 18
	ČEFSA	Baseline	75.8 (41.2); n = 21	74.0 (31.8); n = 19
		Post-treatment	63.2 (37.0); n = 19	65.6 (38.4); n = 17
		Six months	53.0 (41.2); n = 21	60.7 (41.4); n = 17
	GPTS-A	Baseline	18.7 (8.9); n = 21	17.6 (9.5); n = 19
		Post-treatment	13.2 (8.9); n = 20	15.9 (11.1); n = 17
		Six months	10.4 (9.5); n = 21	15.9 (11.2); n = 18

	GPTS-B	Baseline	18.9 (11.7); n = 21	14.7 (11.9); n = 19
		Post-treatment	14.4 (11.7); n = 20	14.5 (13.3); n = 17
		Six months	10.3 (10.1); n = 21	15.5 (15.5); n = 18
	SPEQ	Baseline	23.8 (9.6); n = 21	23.8 (9.2); n = 19
		Post-treatment	18.8 (12.2); n = 20	18.3 (12.4); n = 18
		Six months	14.6 (10.5); n = 21	16.7 (14.0); n = 18
Hwang et al., (2019)[28]*	PSYRATS Total	Baseline	29.2 (1.3); n = 31	28.8 (1.3); n = 32
		Post-treatment	26.1 (1.3); n = 31	29.2 (1.3); n = 32
		One month	25.8 (1.3); n = 31	29.4 (1.3); n = 32
	PSYRATS (hallucinations)	Baseline	18.3 (1.4); n = 31	19.4 (1.4); n = 32
		Post-treatment	17.0 (1.3); n = 31	19.8 (1.3); n = 32
		One month	16.7 (1.3); n = 31	20.0 (1.3); n = 32
	PSYRATS (delusions)	Baseline	10.9 (0.8); n = 31	9.3 (0.8); n = 32
		Post-treatment	9.2 (0.8); n = 31	9.5 (0.8); n = 32
		One month	9.1 (0.8); n = 31	9.5 (0.7); n = 32
Bradley et al., (2018)[29]	GPTS	Baseline	71.6 (22.7); n = 11	None
		Post-treatment	58.3 (24.1); n = 11	
		One month	53.9 (21.3); n = 11	
	SPEQ	Baseline	13.8 (5.5); n = 11	
		Post-treatment	11.9 (5.8); n = 10	
		One month	9.4 (6.6); n = 11	
Haynes et al., (2011)[30]	None			
Holmes et al., (1995)[31]	None			
Myers et al., (2011)[32]	GPTS-A	Baseline	46.93 (13.27); n = 15	None
		Post-treatment	35.20 (16.89); n = 15	
		One month	34.93 (15.60); n = 15	
	GPTS-B	Baseline	58.27 (15.93); n = 15	
		Post-treatment	39.60 (18.80); n = 15	
		One month	38.93 (20.80); n = 15	
	PSYRATS	Baseline	18.33 (2.72); n = 15	
		Post-treatment	15.07 (3.06); n = 15	
		One month	14.00 (4.11); n = 15	
	CAPS	Baseline	18.07 (10.69); n = 15	
		Post-treatment	13.20 (9.68); n = 15	
		One month	13.60 (9.24); n = 15	
Novak et al., (2020)[33]	None			
Waters et al., (2020)[34]	MINI-p	Baseline	16.1 (7.3); n = 50	Not reported
		Post-treatment	13.5 (6.6); n = 40	
Taylor et al., (2022)[35]	R-GPTS-A	Baseline	11.45 (10.14); n = 11	None
		Post-treatment	7.09 (7.62); n = 11	
	R-GPTS-B	Baseline	15.45 (14.17); n = 11	
		Post-treatment	12.18 (13.90); n = 11	
	SPEQ	Baseline	10.64 (11.05); n = 11	
		Post-treatment	10.27 (11.44); n = 11	

Note: This table reports data assessed by psychotic outcome measures in the treatment and control groups at each time-point. Data are presented as mean, (SD), and number of participants completing the outcome measure.
For Hwang et al.'s study[28], the asterisk (*) indicates to treat their findings with caution given the unusually small standard deviations.
Psychotic measures include: Cardiff Anomalous Perceptions Scale (CAPS) [59]; Černis Felt Sense of Anomaly (ČEFSA) [60]; Comprehensive Assessment of At-Risk Mental States (CAARMs) [61]; Green Paranoid Thoughts Scale (GTPS) [62] which contains subscales of 'Ideas of Reference' (GPTS-A) and 'Ideas of Persecution' (GPTS-B); Adapted Mini International Neuropsychiatric Interview – Psychosis section (MINI-p) [63]; Positive and Negative Symptoms Scale (PANSS) [64] which contains subscales relating to positive symptoms, negative symptoms, and general psychopathology; Psychotic Symptom Rating Scale (PSYRATS) [65] which contains subscales relating to delusions and hallucinations, Revised-Green Paranoid Thoughts Scale (R-GPTS) [66] which contains subscales of 'Ideas of Reference' (GTPS-A) and 'Ideas of Persecution' (GPTS-B) subscales; Specific Psychotic Experiences Questionnaire (SPEQ) [67].

A.5.3 Other clinical outcomes

Table A.5.3 *Other clinical outcomes: means, standard deviations and participant numbers*

Study	Primary Outcome		Results	
			Treatment	**Control**
Batalla-Martin et al. (2023)[37]	Quality of life (ED-5Q VAS)	Baseline	54.8 (19.6); n = 20	56.9 (13.3); n = 20
		Post-treatment	64.1 (20.8); n = 20	49.6 (16.5); n = 20
		Three months	66.9 (14.0); n = 20	47.8 (13.60); n = 20
Freeman et al., (2015)[24]	Psychological recovery (CHOICE)	Baseline	52·2 (21·1); n = 23	55·0 (14·4); n = 26
		Post-treatment	58·0 (22·7); n = 22	49·9 (18·3); n = 24
		Three months	60·0 (22·8); n = 21	57·5 (21·8); n = 23
	Quality of life (ED-5D-5L)	Baseline	0·55 (0·23); n = 24	0·60 (0·22); n = 26
		Post-treatment	0·63 (0·25); n = 22	0·55 (0·22); n = 24
		Three months	0·63 (0·27); n = 22	0·58 (0·20); n = 25
	Fatigue (MFI)	Baseline	43·8 (16·4); n = 23	47·6 (15·3); n = 26
		Post-treatment	29·1 (19·0); n = 22	45·4 (19·6); n = 24
		Three months	25·9 (21·4); n = 21	38·4 (18·1); n = 25
	Mental wellbeing (WEMWBS)	Baseline	35·3 (9·3); n = 24	37·0 (7·8); n = 26
		Post-treatment	36·1 (10·7); n = 22	34·0 (8·9); n = 24
		Three months	39·4 (9·9); n = 21	34·7 (7·9); n = 25
Sheaves et al., (2018)[25]	Suicidal ideation (BSS)	Baseline	4.6 (8.3); n = 20	6.7 (10.1); n = 20
		Post-treatment	0.8 (3.3); n = 20	3.6 (8.7); n = 20
		Two weeks	1.1 (4.6); n = 19	3.0 (6.9); n = 20
		Ten weeks	1.3 (3.5); n = 16	2.0 (5.9); n = 18
	Psychological distress (CORE-10)	Baseline	19.9 (8.4); n = 20	17.2 (9.9); n = 20
		Post-treatment	10.4 (5.6); n = 20	13.3 (7.4); n = 20
		Two weeks	9.9 (6.4); n = 19	10.9 (6.6); n = 20
		Ten weeks	11.7 (8.0); n = 16	11.3 (6.2); n = 18
	Quality of life (ED-5Q)	Baseline	4.4 (2.9); n = 20	3.8 (3.3); n = 20
		Post-treatment	2.3 (2.7); n = 20	2.9 (3.4); n = 20
		Two weeks	2.3 (2.1); n = 19	1.7 (2.1); n = 20
		Ten weeks	2.6 (3.6); n = 16	1.8 (2.1); n = 18
	Mental wellbeing (WEMWBS)	Baseline	39.8 (15.4); n = 20	42.3 (13.1); n = 20
		Post-treatment	47.4 (10.5); n = 20	44.8 (13.4); n = 20
		Two weeks	48.3 (11.7); n = 19	45.6 (10.3); n = 20
		Ten weeks	48.3 (12.3); n = 16	44.4 (12.9); n = 18
	Mania (YMRS)	Baseline	14.6 (9.8); n = 20	13.9 (6.2); n = 20
		Post-treatment	9.4 (6.8); n = 20	11.2 (6.6); n = 20
		Two weeks	8.1 (8.3); n = 19	7.8 (6.4); n = 20
		Ten weeks	5.4 (6.4); n = 16	7.8 (6.7); n = 18
Sheaves et al., (2019)[26]	Suicidal ideation (BSS)	Baseline	17.7 (7.7); n = 12	18.8 (7.1); n = 12
		Post-treatment	15.5 (7.7); n = 12	16.8 (7.3); n = 10
		Six months	15.8 (7.8); n = 11	16.7 (10.1); n = 9
	Anxiety (DASS-21 anxiety)	Baseline	10.8 (5.1); n = 12	14.5 (5.2); n = 12
		Post-treatment	7.4 (5.8); n = 12	13.3 (5.4); n = 10
		Six months	7.6 (5.3); n = 11	11.0 (4.6); n = 9
	Depression (DASS-21 depression)	Baseline	13.2 (5.5); n = 12	14.3 (5.7); n = 12
		Post-treatment	10.8 (7.0); n = 12	13.8 (6.5); n = 10
		Six months	11.5 (5.8); n = 11	10.6 (5.3); n = 9
	Stress (DASS-21 stress)	Baseline	11.8 (4.0); n = 12	15.5 (3.7); n = 12
		Post-treatment	9.0 (5.6); n = 12	15.9 (3.1); n = 10

		Six months	10.4 (6.2); n = 11	13.3 (5.1); n = 9
	Dissociation (DES-B)	Baseline	1.7 (0.7); n = 12	2.1 (0.8; n =12
		Post-treatment	1.5 (0.9); n = 12	2.4 (0.6); n = 10
		Six months	1.2 (1.0); n = 11	1.9 (0.8); n = 9
	Mental wellbeing (WEMWBS)	Baseline	38.6 (7.5); n = 12	34.0 (9.7); n = 12
		Post-treatment	44.5 (12.2); n =12	34.2 (10.4); n = 10
		Six months	44.1 (12.4); n = 11	41.0 (9.9); n = 9
Waite et al., (2023)[27]	Suicidal ideation (CSSRS)	Baseline	2·1 (1·5); n=21	1·5 (1·7); n=19
		Post-treatment	1·2 (1·3); n=18	1·4 (1·9); n=16
		Six months	0·9 (1·3); n=18	1·0 (1·6); n=16
	Negative schemas (BCSS-negative)	Baseline	15·2 (5·4); n=21	12·6 (8·0); n=19
		Post-treatment	9·5 (6·8); n=19	10·4 (7·2); n=17
		Six months	7·9 (6·1); n=21	9·7 (6·6); n=17
	Positive schemas (BCSS-positive)	Baseline	5·2 (4·2); n=21	6·7 (4·2); n=19
		Post-treatment	9·8 (4·5); n=19	7·9 (5·2); n=17
		Six months	10·2 (5·6); n=21	9·1 (6·1); n=17
	Anxiety (DASS-21 anxiety)	Baseline	26.0 (10.1); n = 21	20.3 (12.6) n = 19
		Post-treatment	15.9 (13.3); n = 19	17.9 (14.0); n = 17
		Six months	15.9 (13.3); n = 19	18.6 (11.4); n = 16
	Depression (DASS-21 depression)	Baseline	30.4 (10.5); n = 21	27.5 (12.4); n = 19
		Post-treatment	19.5 (13.4); n = 19	21.3 (14.6); n = 17
		Six months	14.7 (10.2); n = 21	19.8 (14.8); n = 16
	Stress (DASS-21 stress)	Baseline	29.1 (8.9); n = 21	25.1 (9.6); n = 19
		Post-treatment	19.5 (12.7); n = 19	23.4 (12.3); n =17
		Six months	20.3 (12.4); n = 21	22.4 (10.3); n = 16
	Worry (DWQ)	Baseline	34·8 (6·5); n=21	29·8 (9·7); n=19
		Post-treatment	25·9 (12·3); n=19	24·9 (13·5); n=17
		Six months	21·4 (12·3); n=21	26·2 (10·7); n=17
	Quality of life (EQ-5D-5L-Index)	Baseline	0.5 (0.3); n = 21	0.5 (0.3); n = 19
		Post-treatment	0.6 (0.3); n = 19	0.5 (0.3); n = 17
		Six months	0.6 (0.3); n = 21	0.5 (0.3); n = 17
	Quality of life (EQ-5D-5L-VAS)	Baseline	48.1 (24.3); n = 21	55.1 (24.8); n = 19
		Post-treatment	61.4 (24.4); n = 19	57.4 (27.8); n = 17
		Six months	68.0 (22.6); n = 21	60.9 (18.6); n = 17
	Agoraphobic avoidance (O-AS-A)	Baseline	3·0 (2·4); n=21	2·0 (2·2); n=17
		Post-treatment	2·2 (2·4); n=19	2·4 (2·3); n=17
		Six months	1·1 (1·9); n=20	1·3 (1·6); n=17
	Agoraphobic distress (O-AS-D)	Baseline	42·8 (23·2); n=21	35·9 (19·7); n=18
		Post-treatment	31·2 (23·9); n=19	35·6 (20·5); n=17
		Six months	20·4 (19·1); n=21	30·9 (20·5); n=15
	Depression (PHQ-15)	Baseline	15.4 (4.4); n = 21	13.8 (6.1); n = 19
		Post-treatment	11.7 (7.1); n = 19	12.6 (6.0); n = 17
		Six months	10.8 (5.5); n = 20	11.8 (6.4); n = 17
	Recovery (QPR)	Baseline	28·9 (13·5); n=21	28·2 (13·5); n=19
		Post-treatment	34·7 (11·6); n=19	31·2 (13·7); n=17
		Six months	36·6 (10·7); n=21	31·9 (11·1); n=17
	Quality of life (ReQol)	Baseline	20·3 (8·1); n=21	28·5 (16·3); n=19
		Post-treatment	39·8 (18·8); n=19	35·4 (17·4); n=16
		Six months	44·6 (14·3); n=21	32·1 (17·2); n=17
	Functioning (WSAS)	Baseline	27·4 (6·0); n=21	22·4 (9·5); n=19

		Post-treatment	18·3 (12·0); n=19	22·4 (13·0); n=17
		Six months	11·1 (9·2); n=20	19·8 (9·5); n=19
Hwang et al., (2019)[28]*	Anxiety (ASI)	Baseline	18.4 (2.4); n = 31	15.3 (2.4); n = 32
		Post-treatment	15.2 (2.3); n = 31	16.0 (2.3); n = 32
		One month	13.0 (2.1); n = 31	16.1 (2.0); n = 32
	Depression (BDI)	Baseline	16.4 (2.2); n = 31	12.2 (2.1); n = 32
		Post-treatment	10.9 (1.6); n = 31	13.3 (1.6); n = 32
		One month	9.0 (1.6); n = 31	13.4 (1.6); n = 32
Bradley et al., (2018)[29]	Anxiety (DASS-21 anxiety)	Baseline	8.2 (4.7); n = 11	None
		Post-treatment	6.2 (3.8); n = 10	
		One month	7.3 (5.9); n = 11	
	Depression (DASS-21 depression)	Baseline	12.0 (6.0); n = 11	
		Post-treatment	8.3 (6.5); n = 10	
		One month	7.8 (7.8); n = 11	
	Stress (DASS-21 stress)	Baseline	12.2 (4.8); n = 11	
		Post-treatment	8.2 (6.1); n = 11	
		One month	8.2 (5.7); n = 11	
	Mental wellbeing (WEMWBS)	Baseline	35.4 (7.9); n = 11	
		Post-treatment	41.1 (6.2); n = 11	
		One month	42.1 (8.1); n = 11	
	Functioning (WSAS)	Baseline	24.2 (7.5); n = 11	
		Post-treatment	18.6 (9.5); n = 11	
		One month	16.6 (10.4); n = 11	
Haynes et al., (2011)[30]	None			
Holmes et al., (1995)[31]	None			
Myers et al., (2011)[32]	Anxiety (DASS-21 anxiety)	Baseline	21.80 (9.05); n = 15	None
		Post-treatment	11.73 (9.65); n = 15	
		One month	13.47 (10.53); n = 15	
	Depression (DASS-21 depression)	Baseline	23.13 (15.41); n = 15	
		Post-treatment	12.87 (14.37); n = 15	
		One month	16.07 (15.88); n = 15	
Novak et al., (2020)[33]	None			
Waters et al., (2020)[34]	Impulsivity (BIS)	Baseline	18.7 (18.5); n = 50	Not reported
		Post-treatment	17.9 (3.6); n = 40	
	Quality of life (MANSA)	Baseline	37.3 (9.5); n = 50	
		Post-treatment	38.0 (8.7); n = 40	
	Anxiety and depression (PHQ-4)	Baseline	6.5 (6.6); n = 50	
		Post-treatment	4.8 (3.3); n = 40	
Taylor et al., (2022)[35]	Anxiety (DASS-21 anxiety)	Baseline	10.36 (8.85); n = 11	None
		Post-treatment	7.64 (6.31); n = 11	
	Depression (DASS-21 depression)	Baseline	19.82 (13.58); n = 11	
		Post-treatment	14.18 (13.43); n = 11	

Stress (DASS-21 stress)	Baseline	18.55 (10.28); n = 11
	Post-treatment	16.00 (10.81); n = 11
Mental wellbeing (WEMWBS)	Baseline	40.18 (12.05); n = 11
	Post-treatment	43.36 (12.52); n = 11
Functioning (WSAS)	Baseline	23.18 (11.37); n = 11
	Post-treatment	20.18 (11.19); n = 11

Note: This table reports data assessed by other clinical outcome measures in the treatment and control groups at each time-point. Data are presented as mean, (SD), and number of participants completing the outcome measure.
For Hwang et al.'s study [28] study, the asterisk (*) indicates to treat their findings with caution given the unusually small standard deviations. The Measures column in the table states the clinical symptom that was assessed, followed by the assessment measure in brackets. Assessment measures include: Anxiety Sensitivity Index (ASI) [68]; Brief Core Schema Scale (BCSS) [69]; Beck Depression Inventory (BDI) [70]; Beck Suicide Scale (BSS) [71]; CHoice of Outcome In CBT for psychoses (CHOICE) [72]; Clinical Outcomes in Routine Evaluation 10-item scale (CORE-10) [73]; Columbia Suicide Severity Rating Scale (CSSRS) [74]; Depression Anxiety and Stress Scale–21-item (DASS-21) [75] version with subscales for anxiety (DASS-21 anxiety), depression (DASS-21 depression) and stress (DASS-21 stress); Brief Dissociative Experiences Scale (DES-B) [76]; Dunn Worry Questionnaire (DWQ) [77]; Euroqol 5D questionnaire (ED-5Q) [78]; Euroqol 5D questionnaire 5 Level version (ED-5D-5L) [79], Multidimensional Fatigue Inventory (MFI) [80]; Oxford Agoraphobic Avoidance Scale (O-AS) [81]; Patient Health Questionnaire (PHQ-9) [82]; Brief Patient Health Questionnaire (PHQ-4) [82]; Questionnaire about the Process of Recovery (QPR) [83]; Recovering Quality of Life (ReQol) [84]; Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) [85]; Work and Social Adjustment Scale (WSAS) [86]; Young Mania Rating Scale (YMRS) [87].

Appendix 6 – List of included papers

Batalla-Martin D., Martorell-Poveda M.-A., Belzunegui-Eraso A., Marieges Gordo A., Batlle Lleal H., Pasqual Melendez R., et al. A Pilot Nurse-Administered CBT Intervention for Insomnia in Patients with Schizophrenic Disorder: A Randomized Clinical Effectiveness Trial. J Clin Med. 2023;12(19):6147.

Freeman D, Waite F, Startup H, Myers E, Lister R, McInerney J, et al. Efficacy of cognitive behavioural therapy for sleep improvement in patients with persistent delusions and hallucinations (BEST): a prospective, assessor-blind, randomised controlled pilot trial. Lancet Psychiatry. 2015 Nov;2(11):975–83.

Sheaves B, Isham L, Bradley J, Espie C, Barrera A, Waite F, et al. Adapted CBT to Stabilize Sleep on Psychiatric Wards: a Transdiagnostic Treatment Approach. Behav Cogn Psychother. 2018 Nov;46(6):661–75.

Sheaves B, Holmes EA, Rek S, Taylor KM, Nickless A, Waite F, et al. Cognitive behavioural therapy for nightmares for patients with persecutory delusions (nites): An assessor-blind, pilot randomized controlled trial. Can J Psychiatry Rev Can Psychiatr. 2019;64(10):686–96.

Waite F, Černis E, Kabir T, Iredale E, Johns L, Maughan D, et al. A targeted psychological treatment for sleep problems in young people at ultra-high risk of psychosis in England (SleepWell): a parallel group, single-blind, randomised controlled feasibility trial. Lancet Psychiatry. 2023 Sep 1;10(9):706–18.

Hwang DK, Nam M, Lee YJG. The effect of cognitive behavioral therapy for insomnia in schizophrenia patients with sleep disturbance: A non-randomized, assessor-blind trial.

Psychiatry Res. 2019 Apr;274:182-188. Doi: 10.1016/j.psychres.2019.02.002. Epub 2019 Feb 3. PMID: 30807969.

Bradley J, Freeman D, Chadwick E, Harvey AG, Mullins B, Johns L, et al. Treating sleep problems in young people at ultra-high risk of psychosis: A feasibility case series. Behav Cogn Psychother. 2018;46(3):276–91.

Haynes PL, Parthasarathy S, Kersh B, Bootzin RR. Examination of insomnia and insomnia treatment in psychiatric inpatients. Int J Ment Health Nurs. 2011 Apr;20(2):130–6.

Holmes EP, Corrigan PW, Knight S, Flaxman J. Development of a sleep management program for people with severe mental illness. Psychiatr Rehabil J. 1995;19(2):9–15.

Myers E, Startup H, Freeman D. Cognitive behavioural treatment of insomnia in individuals with persistent persecutory delusions: a pilot trial. J Behav Ther Exp Psychiatry. 2011;42(3):330–6.

Novak C., Packer E., Paterson A., Roshi A., Locke R., Keown P., et al. Feasibility and utility of enhanced sleep management on in-patient psychiatry wards. BJPsych Bull. 2020;44(6):255–60.

Waters F, Chiu VW, Dragovic M, Ree M. Different patterns of treatment response to Cognitive-Behavioural Therapy for Insomnia (CBT-I) in psychosis. Schizophr Res. 2020 Jul;221:57-62. doi: 10.1016/j.schres.2020.03.054. Epub 2020 Apr 18. PMID: 32317223.

Waite F, Bradley J, Chadwick E, Reeve S, Bird JC, Freeman D. The Experience of Sleep Problems and Their Treatment in Young People at Ultra-High Risk of Psychosis: A Thematic Analysis. Front Psychiatry. 2018 Aug 24;9:375.

Appendix 7 – Quality Assessment using 'The Measurement Tool to Assess systematic Reviews 2' (AMSTAR)

The scoring indicates a 'high' rating of overall confidence in the review.

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

1. Did the research questions and inclusion criteria for the review include the components of PICO?

For Yes:
- ☑ Population
- ☑ Intervention
- ☑ Comparator group
- ☑ Outcome

Optional (recommended)
- ☐ Timeframe for follow-up

☑ Yes
☐ No

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

For Partial Yes:
The authors state that they had a written protocol or guide that included ALL the following:
- ☑ review question(s)
- ☑ a search strategy
- ☑ inclusion/exclusion criteria
- ☑ a risk of bias assessment

For Yes:
As for partial yes, plus the protocol should be registered and should also have specified:
- ☑ a meta-analysis/synthesis plan, if appropriate, *and*
- N/A a plan for investigating causes of heterogeneity
- ☑ justification for any deviations from the protocol

☑ Yes
☐ Partial Yes
☐ No

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:
- ☐ *Explanation for* including only RCTs
- ☐ OR *Explanation for* including only NRSI
- ☑ OR *Explanation for* including both RCTs and NRSI

☑ Yes
☐ No

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following):
- ☑ searched at least 2 databases (relevant to research question)
- ☑ provided key word and/or search strategy
- ☑ justified publication restrictions (e.g. language)

For Yes, should also have (all the following):
- ☑ searched the reference lists / bibliographies of included studies
- ☐ searched trial/study registries
- ☐ included/consulted content experts in the field
- ☐ where relevant, searched for grey literature
- ☑ conducted search within 24 months of completion of the review

☐ Yes
☑ Partial Yes
☐ No

5. Did the review authors perform study selection in duplicate?

For Yes, either ONE of the following:
- ☐ at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include
- ☑ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

☑ Yes
☐ No

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:

- ☐ at least two reviewers achieved consensus on which data to extract from included studies
- ☑ OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

☑ Yes
☐ No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:
- ☐ provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:
- ☑ Justified the exclusion from the review of each potentially relevant study

☑ Yes
☐ Partial Yes
☐ No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (ALL the following):
- ☑ described populations
- ☑ described interventions
- ☑ described comparators
- ☑ described outcomes
- ☑ described research designs

For Yes, should also have ALL the following:
- ☑ described population in detail
- ☑ described intervention in detail (including doses where relevant)
- ☑ described comparator in detail (including doses where relevant)
- ☑ described study's setting
- ☑ timeframe for follow-up

☑ Yes
☐ Partial Yes
☐ No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs

For Partial Yes, must have assessed RoB from
- ☑ unconcealed allocation, *and*
- ☑ lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)

For Yes, must also have assessed RoB from:
- ☑ allocation sequence that was not truly random, *and*
- ☑ selection of the reported result from among multiple measurements or analyses of a specified outcome

☑ Yes
☐ Partial Yes
☐ No
☐ Includes only NRSI

NRSI

For Partial Yes, must have assessed RoB:
- ☑ from confounding, *and*
- ☑ from selection bias

For Yes, must also have assessed RoB:
- ☑ methods used to ascertain exposures and outcomes, *and*
- ☑ selection of the reported result from among multiple measurements or analyses of a specified outcome

☑ Yes
☐ Partial Yes
☐ No
☐ Includes only RCTs

10. Did the review authors report on the sources of funding for the studies included in the review?

For Yes
- ☐ Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies

☐ Yes
☑ No

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

RCTs

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present.
  - ☐ AND investigated the causes of any heterogeneity

☐ Yes
☐ No
☑ No meta-analysis conducted

For NRSI

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present
  - ☐ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available
  - ☐ AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review

☐ Yes
☐ No
☑ No meta-analysis conducted

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

For Yes:

- ☐ included only low risk of bias RCTs
- ☐ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.

☐ Yes
☐ No
☑ No meta-analysis conducted

13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

For Yes:

- ☐ included only low risk of bias RCTs
- ☑ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results

☑ Yes
☐ No

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

For Yes:

- ☐ There was no significant heterogeneity in the results
- ☐ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review

☑ Yes
☐ No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes:

- ☐ performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias

☐ Yes
☐ No
☑ No meta-analysis conducted

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes:

☑	The authors reported no competing interests OR	☑ Yes
☐	The authors described their funding sources and how they managed potential conflicts of interest	☐ No

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.

Appendix 8 – Author guidelines for British Journal of Psychiatry

1. AIMS AND SCOPE

The British Journal of Psychiatry (BJPsych) is a leading international peer-reviewed journal, covering all branches of psychiatry with a particular emphasis on the clinical aspects of each topic.

2. MANUSCRIPT CATEGORIES AND REQUIREMENTS

Article type: Review

Definition: Systematic reviews that are ‘first in field’; narrative review articles in exceptional cases only

Systematic reviews submitted to BJPsych should be 'first in field', explicating important findings in fields where no systematic review has yet been published or where information in the literature is conflicting. They are expected to be a clear building block upon which further research in the field is undertaken.

3. PREPARING THE SUBMISSION

We require authors to register the protocol for systematic reviews before data extraction on an accessible, searchable site such as PROSPERO and include the registration number in the manuscript file. If the review has not been registered, we are unable to consider your submission.

The word count should not exceed 6,000 words (excluding references, tables and figure legends) and may include up to 150 essential references in total. Additional references may be included with permission of the Editor in Chief. Only papers directly referred to should be directly referenced in the main manuscript file; large data sets should be included in the supplementary material.

Please include a structured abstract of up to 250-300 words with the headings: Background, Aims, Method (including data sources, study selections, synthesis approach), Results, Conclusions.

The main text should include the following sections: Introduction, Method, Results and Discussion and conclusion.

It is important that the **Method** section clearly describes the search strategy, study selection criteria and synthesis approach in sufficient detail to ensure the method can be replicated to extract the same data with the same or similar analysis. This should include information about the protocol registration, review software, data sources (bibliographic databases such as PubMed/MEDLINE, Embase, CINAHL, PsycINFO and reference lists from journals or books), MeSH and free text search terms and filters, dates included in the search, screening process, language limitations, inclusion and exclusion criteria, study selection and synthesis approach. To ensure a comprehensive review of the literature, we encourage consideration of publications in non-English languages. Ideally the search should be as current as possible

with the search date noted in the manuscript. Please describe how the quality of evidence was evaluated and explain whether heterogeneity between studies was investigated and explored.

The **results** section should report the results and summarise the quality of the evidence. Where necessary, a statement should be made as to how quality has been used to inform interpretation of results and/or analysis.

The **Discussion** section should summarise key findings and describe the impact of the results on clinical practice. Authors should discuss both the strengths and the limitations of the evidence and the review.

Tables and figures identifying the primary finding(s) may be included in the print version of each paper. Additional tables and figures will be included as online only supplementary material. All large tables (exceeding one journal page) will be published as online only supplementary material. Authors are encouraged to present key data within smaller tables for print publication.

Supplementary tables, figures and data should include (in this order):

1. PRISMA-P (or equivalent) table
2. Search strings used for various platforms such as MEDLINE, Scopus etc.
3. PICOS table (if relevant)
4. List of all included papers
5. Additional sensitivity analyses or additional analyses (if relevant)
6. Publication bias: a funnel plot and additional analyses undertaken where necessary
7. Quality Assessment

Authors should include a GRADE assessment of the findings or equivalent approach to place the importance of the findings in context.

Previously published Reviews for groups such as the Cochrane Collaboration and the National Institute for Health and Care Excellence should be submitted with the latest version of the parent review and its status, so an informed decision can be made about the added value of the submitted paper.

Reviews are subject to external peer review.

References – References should be numbered in the order that they appear in the text and listed at the end of the manuscript using the Vancouver style. Unpublished doctoral theses may be cited but no other citation of unpublished work, including unpublished conference presentations, is permissible. Authors are responsible for checking all references for accuracy and relevance before submission.

Appendix: Service Improvement Project

Appendix 1 – Further detail regarding service context

The neonatal surgical pathway typically follows the following trajectory. Families are be referred by following their 12-weeks or 20-week antenatal scan during which a foetal anomaly is identified. The foetal medicine unit (FMT) conducts several scans to monitor the baby's condition and provide specialist diagnosis. Parents are usually offered an amniocentesis in their first appointment with the FMU and tests for genetic syndromes are conducted. At approximately 28-32 weeks, they are offered the 'antenatal consultation' in which they meet with the surgeon who is planned to conduct the surgery for the baby once they are born (this may change if the surgeon is unavailable) and ward staff from the neonatal care units (NCU) that it is likely the baby will stay at after birth. They meet to discuss the surgical plan for the baby once they are born and parents may sometimes be offered the opportunity to visit neonatal wards.

After birth, babies often stay within the neonatal care unit (NCU) before and after surgery during there recovery. This comprises three wards: the intensive treatment unit (ITU), the higher dependency unit (HDU), and the lower dependency unit (LDU). Babies tend to be transitioned from ITU to HDU to LDU as they recover and thus the required intensity of care reduces. ITU has the highest staff ratio per baby that is 1:1 and staff may therefore respond immediately when a machine or monitor sounds an alarm. LDU has the lowest staff to baby ratio and nurses may not be able to respond immediately if, for example, another baby requires care more urgently. They may also spend time in one of the hospital's children's wards particularly if their period of recovery is longer. This journey is not always entirely linear as, for example, some babies return home after being born before going to the hospital to receive surgery, and some babies require multiple surgeries with periods of recovery in

between. Approximately two weeks after surgery, families usually have a follow-up meeting with the surgeon to discuss how the procedure went.

Psychologists are available for parents to speak to on the neonatal units and there is an ad-hoc provision on the children's wards. Psychologists may approach parents who are on the ward, or nurses and medical clinicians may suggest for psychologists to approach specific parents who appear to be struggling.

When their condition is sufficiently stable, babies are often discharged either to a care team in the community (such as their GP and health visitors) or to a hospital that is more local to where the family lives.

Appendix 2 – Further examples of illustrative quotes

Summary of themes, subthemes and illustrative quotes

Theme	Subtheme	Illustrative Quotes
1. "It's constantly preparing for what comes next"	**1.1. Preparing for diagnosis and next steps**	*P2 "I don't want to hear that there is a genetic scan when it's going to be done. It's, it's, it's that process of reiterating the fact that there are other things we can offer. There are other tests that may be done. It's constantly preparing for what is coming next."* *P4 "It's difficult to know what to expect with all these appointments. I think. And I think I think what was missing was when we first went in the room to have the cardiac scan and he put the scanner on my belly, his first comment was, oh, good, there's a heartbeat, which, we hadn't really been told that there was a possibility they wouldn't be."* *P7 "I was thinking, how can you do an operation on a baby that's like this small (gesturing with hands)? And, of course, when I spoke to her surgeon after the surgery, she was like, 'no, you know, she's a really big baby for NICU! I've done loads of operations on, you know, tinier babies than this.' And actually, if someone had maybe said that to me before, I might not have been so, in such a state."*
	1.2. Preparation for entering and moving through the neonatal space	*P6 "[In NICU] I think it was like quite scary, especially as [partner] kind of stood at the door and peered in, and there's just lots of machines and, and you can't actually see any babies, they're all just completely wired up. So I think that was a bit overwhelming for him. And actually, after the birth, that was one of the things he struggled with most - the alarms and things constantly going off."* *P9 "I found it really emotional when I, when I first went in [during tour of NICU]. And then I was able to sort of process that and, and become calm with it again, so that when we actually went there when he was born, I was, I was already sort of much more prepared for that, and that was one less thing to have to worry about. I think if I'd known that we were going to the, the NICU, but not seen it before, that would have been more stressful."* *P7 "They didn't really prepare me for how different [lower dependency] ward is [compared to NICU]. You've gone from your little bubble of just you, your baby, and your nurse, to like everyone's packed in a room, you know, there's not much space, everything you just felt different, and the nurses are obviously looking after more babies…it was quite kind of quite a sudden drop in care and the whole atmosphere was just different…that was really hard."*
2. Parent-centred communication	**2.1 More information**	*P1 "What we didn't know was how they would propose to manage the condition and what that would mean for the birth; and so, we were looking forward to this meeting because it was going to give us some clarity and answers."*

P4 "The trouble is as well, if you don't get enough information, if you're just handed a leaflet and told 'all this will happen to go and read it, this explains everything', and then you have questions, all you do then is go on Google. And there's lots of unhelpful Google information out there about a lot of this stuff."

P5 "I think it would be really good if you had a point of contact…you need some help navigating it all… very confusing, yeah."

2.2 Compassionate versus clinical language

P3 "I think it's more of empathy and thinking about the person, imagine if you're in their shoes and just showing that compassion."

P4 "One of the doctors said, 'I don't know why you're worrying about that because you're likely have a stillbirth' which was certainly very upsetting. Not useful information to be given just like that…. I just again got the sense that we were just a bit of a nuisance. You know, we were. They'd had this meeting to discuss what was going to happen after she was born, but she, she likely wouldn't be [tearful – breaking up with crying]. Sorry, I'm really I'm, I'm so upset because, because of the way he said it…"

P7 "A lot of the time it's, everything's thrown at you in such a rushed manner - you're like, 'Oh, I don't know! I can't even take in what you've just said!' But actually, the nurses are the ones that stick in my mind for, like, being sympathetic, and certainly once [baby] was out of surgery and then NICU, like, the nurses, are kind of more memorable than the doctors just because the doctors throw jargon at you! And you're like 'oh ok!' The nurses will kind of explain it and like, be a bit, take their time."

P8 "For every really understanding staff member, there's probably another that tends to be a little bit clinical. The [doctor] said 'oh, so he's got a slight regurgitation.' And we said, 'Well, what does that mean?' And they said, 'Well, it means sometimes, it means a little bit of the blood's going backwards through the heart.' And we're like, 'well, what does that mean? Like, is this important or serious?' And they're like, 'Oh, oh, yeah. You should speak to your consultant about that.' And it's only once you've spoken to a few doctors about it, you realise it's probably not that serious: 'We've probably only caught it because we're scanning so, in so, so much detail. And by the way, these scans are really inaccurate anyway, so we'll keep an eye on it.'

2.3 Holding out for hope

P2 "You do want to hear that reassurance 'Yes, this is what's happening', but is there something else that is positive out of this? But every single appointment I could see that they were ticking boxes to rule out 'OK, so it's not this syndrome. It's not that syndrome. So is it, can it be possibly that? Or that?' I could hear it. I could feel it, and it's unsettling…you want to feel, yes, we understand that they're investigating but we want to have hope. A realistic hope… You want to see something positive in the middle of all the negative."

P4 "Everything was just so doom and gloom. They thought it was Edwards (syndrome), and everyone was pushing for termination."

		P5 "[Surgeon] was brilliant, really kind of went through the whole process, showed us sort of on, you know, on pictures, what it meant, what they was going to do. So that, so I felt like I really understood what was happening. And she really put us at ease. She's like, you know, we love doing these operations because it is just bread and butter. The outcomes are so good that, you know, it's, it's nothing really to, to worry about."
	2.3. Involving parents in decisions	*P1 "It felt like decisions were being made and we were being informed, not that we were involved in the decisions."*
		P3 "I mentioned [concerns about infection], but [doctor] kind of dismissed it, but later on, we realised that actually what I said was correct…So just actually listening as well…even if you know more than you know the mother or the father just taking on board what they're saying and just, you know, going through with what they suspect and then coming back and saying, 'OK, we've looked at this because you mentioned this.'"
		P4 "Obviously they should be explaining every option, but to be told that you would be cruel to continue your pregnancy and stuff, I think that should be completely your choice and not have that pushed from a midwife, because you're put under so much pressure when you're in that tiny room. If we made the wrong decision, then I would have regretted it for the rest of my life."
		P4 "[Baby's feeding] was on such a strict protocol and it really just wasn't right for her either… we actually had quite a long discussion with the consultant [at a different hospital] and he actually agreed to let us change… and going forward, she, she absolutely thrived when we were allowed to do that. I think sometimes [surgical hospital] are too focused on the way they've told it should be, and not consider that actually individual babies might not benefit from that kind of treatment."
	3.3. Continuity builds trust	*P3 "Sometimes there will be discrepancies because they're measuring things with the baby. So one person will be like 'ohh is pushing against her heart. And like its overly pushing against her heart.' Another doctor will say, 'it's not quite pushing'…so perhaps also sticking to the same doctor that you're seeing, that would help.*"
		P7 "I felt really happy that we'd already met the surgeon so when she did need the surgery, you know we'd met her, we knew how lovely she was and it just kind of put a face to, to the, to the name. And yeah, that it was really sort of comforting, we felt like we were in safe hands."
		P9 "Where it got very stressful was we found that certain notes were not matching what was being said, and there was certain inconsistencies around, around what was happening. So that, that made it very stressful…left us difficult to trust."
3. Addressing support needs	**3.1 Universal to specialist support**	*P1 "I just feel like if we'd had various touch points…the chance to speak to somebody, they could have helped us to heal with all of that"*
		P4 "When you're going through this, you just need someone to talk to."

	P8 "It would help if everyone was a bit more acknowledging of what it might feel like, or even asking us how, how we feel… The midwives do all the soft feeling-y stuff and the doctors don't touch any of that. Whereas it would be nice if they were kind of sharing that though a little bit because then you get, you get cold and then really warm - and then that's, that's a funny balance to deal with."
	P8 "Sleep, nutrition, exercise. If you could just get them doing, if you could provide support to make sure people are doing those three things, it will increase their capacity to do all the other stuff tremendously."
	P6 "[Psychologist] did an amazing job…I was really not in a good place…we had some really good conversations to help normalise some of what I was feeling, and also to help me and [partner] because we had different coping strategies, just to like have a bit of patience with one another about those. That had a meaningful impact in our ability to support one another through the rest of it."
3.2. Shifting from 'the hospital's baby' to 'my baby'	*P5 "[NICU nurses] were brilliant. As soon as you come out, they're like, 'do you want me to get him out so you can hold him and we could do skin to skin?' I think you could have easily have felt quite detached from your baby in that situation. But they was really good in, in sort of saying 'no, let's get him out. You know, you can have a cuddle.' And so it, it didn't feel like such a big switch when we went down to [lower dependency unit]."*
	P6 "[Difficulty establishing breastfeeding] ended up being the biggest, hardest part of the recovery for me…latching on to that psychologically, it's like the one thing that I could do to help my baby when I'm so helpless with respect to the rest of his care and rely on other people to keep him alive."
	P7 "I really, really want to do something for [baby] rather than to just sit here and feel useless. But I am terrified of hurting her.' And [the nurse] was like, 'you won't hurt her.' And she took the time to really show me exactly how to do [nappies]. And then from then on, I just asked to do everything I could."
3.3. Shared experience	*P6 "Meeting people who've had babies who had been sick and then recovered and gone on to live, really happy, healthy lives was nice to hear… it just really reinforced that like, they just had a completely normal life. You know, they're, they're talking about their kids going travelling and all the normal parent worries that, yeah, have nothing to do with their health."*
	P7 "We didn't really have any friends or family that had been through a similar thing… you still felt really alone in your journey, and actually you kind of just wanted to talk over with someone that could understand...If I could have just spoken to someone, who'd been through it before and come out the other side, that would have just been really helpful"

Appendix 3 – Survey Participant Information and Consent

Improving the wellbeing of parents whose babies require surgery for an antenatally diagnosed condition

This is an invitation to take part in a service evaluation for the John Radcliffe neonatal surgical care pathway. This will involve completing a short online survey.

Please take the time to read this Participant Information sheet, and discuss it with others if you wish. If there is anything that is not clear, or if you would like more information, please contact Hannah Wilkinson (Hannah.Wilkinson@oxfordhealth.nhs.uk), Dr Jacinta Cordell (Jacinta.Cordwell@ouh.nhs.uk) or Dr Matthew Hotton (Matthew.Hotton@ouh.nhs.uk) who will do their best to answer your query. However, please note that they do not work at OUH every day, and so it may take up to 14 days to respond to your enquiry.

Participant Information Sheet

What is the purpose of the service evaluation?

Receiving a diagnosis for your baby during pregnancy and preparing for surgery can be an incredibly stressful and challenging period for parents. We are looking to enhance our understanding of the most effective way to prepare parents and support their emotional wellbeing during this time, and any barriers there may be to this. We hope that this will enable us to better tailor our service to the needs of parents, and allow for more accessible, parent-led care.

Why have I been invited to take part?

You have been invited to take part because you have been identified as a parent whose baby received surgery following an antenatal diagnosis. We would very much value the opportunity to hear from a range of parents who additionally meet the following requirements:

- Aged over 18 years
- Are able to read English
- Have been discharged from the neonatal care unit at least 1 month ago, and up to 3 years ago

Do I have to take part?

No, taking part is entirely voluntary. The evaluation is in the form of an online survey: you can stop completing the survey at any point. However, as the data collected is anonymous, you will be unable

to withdraw your responses once you have submitted them.

What will happen if I decide to take part?

You will be asked to confirm that you have read this information sheet, and then to complete a short online survey, estimated to take 10-15 minutes. This will involve a combination of 'tick box' answers and written feedback. This will involve some demographic questions, followed by questions asking you about your experiences. We are focusing on a specific period, starting from when the John Radcliffe Foetal Medicine team first contacted you to discuss your baby's diagnosis and treatment plan, through to when your baby received surgery and aftercare at the neonatal care unit.

Are there any disadvantages/risks from taking part?

Taking part will take approximately 10-15 minutes of your time. Asking you to reflect on the time whilst your baby was diagnosed and received surgery may be difficult, and it is possible you might find answering some of the questions upsetting. You are welcome to take a break from the survey and return to it later, or to stop the survey completely. We have consulted with people with lived experience of having their baby cared for on the neonatal unit, and the surveys have been designed with their feedback in mind, in order to try and minimise the potential for distress.

What are the possible benefits of taking part?

Your answers will help us understand how to better support parents' needs and emotional wellbeing whilst receiving care from neonatal surgical care pathway. Your feedback will contribute to the development of more effective, accessible and parent-led psychological care, which may include new initiatives and training within the service.

What happens with my information?

We will adhere to the Oxford University Hospital data management policy. Data will be password-protected, and access will be restricted to the core project team (Hannah Wilkinson, Dr Jacinta Cordwell, Dr Matthew Hotton). Responsible members of the University of Oxford may be given access to the data for monitoring and/or audit of the evaluation. All data is anonymous from the point of collection, however it is possible that you may disclose information in the survey which makes you potentially identifiable. All identifiable information will be removed in the write-up. Demographic information will be destroyed at the end of the study. All other survey data will be stored for at least three years after publication, in accordance with the University of Oxford's Policy on the Management of Research Data & Records.

We hope to use the findings to make recommendations for the neonatal surgical care pathway so that

they can improve parents' support. We may also publish the findings in an academic journal and present them at conferences with relevant healthcare professionals and charities. Any identifiable information will be removed before information is shared with other researchers or results are made public. This evaluation will contribute to Hannah Wilkinson's Doctorate in Clinical Psychology. If you would like to receive information on the findings, please email Hannah.Wilkinson@oxfordhealth.nhs.uk.

Who has reviewed this service evaluation?

This study has been approved by Oxford University Hospital Medical Governance Committee.

What if there is a problem?

If you have a concern about any aspect of this study, please contact Hannah Wilkinson (Hannah.Wilkinson@oxfordhealth.nhs.uk) or her supervisors Dr Jacinta Cordwell (Jacinta.Cordwell@ouh.nhs.uk) and Dr Matt Hotton (Matthew.Hotton@ouh.nhs.uk) who will do their best to answer your query. Please note that it may take up to 14 days to respond to your enquiry. If you remain unhappy or wish to make a formal complaint, please contact complaints@ouh.nhs.uk.

Support services

Advice and support are available from the NHS if you are finding things difficult. There are also charities who can provide advice and support. If you are concerned about your mental health, we encourage you to speak to your General Practitioner (GP)/Family Doctor who can then refer you for further specialist help if required.

Please see below details of where you can access support:

- GP/Family Doctor
- NHS 111: To contact NHS 111 call "111". Available 24 hours a day, 7 days a week.
- The Samaritans: Call "116 123" to contact their helpline which is available 24 hours a day, 7 days a week. A range of support can also be found on their website: www.samaritans.org
- BLISS Charity: A UK charity for babies born premature or sick. It supports families with a baby in neonatal care, website: https://www.bliss.org.uk email: hello@bliss.org.uk

Please note that service evaluation staff email addresses are not always monitored and should not be used in the case of an emergency. If you are experiencing a crisis please contact emergency services.

I'm Interested! What's Next?

Thank you so much for taking the time to read this information sheet! If you are happy to proceed, please select to continue.

Appendix 4 – Interview Participant Information and Consent

Neonatal Surgical Care Pathway – Improving Parents' Support

PARTICIPANT INFORMATION SHEET (Version/Date: V1.0 17/07/2023)

My name is Hannah Wilkinson and I am a Trainee Clinical Psychologist at the University of Oxford, working with the Neonatal team at Oxford University Hospital at Oxford University Hospital. This is an invitation to take part in a service evaluation for the John Radcliffe neonatal surgical care pathway.

Please take the time to read this Participant Information sheet, and discuss it with others if you wish. If there is anything that is not clear, or if you would like more information, please contact Hannah Wilkinson (Hannah.Wilkinson@hmc.ox.ac.uk) or her supervisors Dr Jacinta Cordell (Jacinta.Cordwell@ouh.nhs.uk) or Dr Matthew Hotton (Matthew.Hotton@ouh.nhs.uk) who will do their best to answer your query. However, please note that they do not work at OUH every day, and so it may take up to 14 days to respond to your enquiry.

What is the purpose of the service evaluation?
Receiving a diagnosis for your baby during pregnancy and preparing for surgery can be an incredibly stressful and challenging period for parents. We are looking to enhance our understanding of the most effective way to prepare parents and support their emotional wellbeing during this time, and any barriers there may be to this. We hope that this will enable us to better tailor our service to the needs of parents, and allow for more accessible, parent-led care.

The first part of the service evaluation is a survey to allow us to hear about a wide range of experiences. We are also inviting a smaller number of parents for follow-up interviews to help us gain a deeper understanding of what was challenging, what helped, and any ideas on how to improve the support we can offer.

Why have I been invited to take part?

You have been invited to take part because you have been identified as a parent whose baby received surgery following an antenatal diagnosis. We would very much value the opportunity to hear from a range of parents who additionally meet the following requirements:

- Aged over 18 years
- Are able to read English
- Have been discharged from the neonatal care unit at least 1 month ago, and up to 3 years ago

Do I have to take part?

No, taking part is entirely voluntary. You can stop taking part at any point and are free to answer only what you would like. You can also withdraw your data from the study up to 3 weeks after the interview, as after this time the interviews will be anonymised and pseudonymised.

What will happen to me if I decide to take part?
You will be asked to confirm that you have read this information sheet and consent to taking part. You will then be invited to take part in an online interview about your experiences of the surgical care pathway, estimated to take 30-60 minutes. We are focusing on a specific period, starting from when the John Radcliffe Foetal Medicine team first contacted you to discuss your baby's diagnosis and treatment plan, through to when your baby received surgery and aftercare at the neonatal care unit.

Are there any disadvantages/risks from taking part?
Taking part will take approximately 30-60 minutes of your time. Asking you to reflect on the time whilst your baby was diagnosed and received surgery may be difficult, and it is possible you might find answering some of the questions upsetting. You are welcome to take a break or to stop the interview completely. We have consulted with people with lived experience of having their baby cared for on the neonatal unit, and the interviews have been designed with their feedback in mind, in order to try and minimise the potential for distress.

What are the possible benefits of taking part?
Your answers will help us understand how to better support parents' needs and emotional wellbeing whilst receiving care from neonatal surgical care pathway. Your feedback will contribute to the development of more effective, accessible and parent-led psychological care, which may include new initiatives and training within the service.

What happens with my information?
We will adhere to the Oxford University Hospital data management policy. Data will be password-protected, and access will be restricted to the core project team (Hannah Wilkinson, Dr Jacinta Cordwell, Dr Matthew Hotton). Responsible members of the University of Oxford may be given access to the data for monitoring and/or audit of the evaluation. Interview data will be transcribed (written up) and anonymised (personal identifying details and names/places removed). These transcribed interviews will be stored in an archive for a period of at least 3 years, in accordance with the University of Oxford's Policy on the Management of Research Data & Records.

We hope to use the findings to make recommendations for the neonatal surgical care pathway so that they can improve parents' support. We may also publish the findings in an academic journal and present them at conferences with relevant healthcare professionals and charities. Any identifiable information will be removed before it is shared with other researchers or results are made public. This evaluation will contribute to Hannah Wilkinson's Doctorate in Clinical Psychology. If you would like to receive information on the findings, please email Hannah.Wilkinson@hmc.ox.ac.uk.

Who has reviewed this service evaluation?
This study has been approved by Oxford University Hospital Medical Governance Committee.

What if there is a problem?
If you have a concern about any aspect of this study, please contact Hannah Wilkinson (Hannah.Wilkinson@ouh.nhs.uk) or her supervisors Dr Jacinta Cordwell (Jacinta.Cordwell@ouh.nhs.uk) and Dr Matt Hotton (Matthew.Hotton@ouh.nhs.uk) who will do their best to answer your query. Please note that it may take up to 14 days to respond to your enquiry. If you remain unhappy or wish to make a formal complaint, please contact complaints@ouh.nhs.uk.

Signposting
Advice and support are available from the NHS if you are finding things difficult. There are also charities who can provide advice and support. If you are concerned about your mental health, we encourage you to speak to your General Practitioner (GP)/Family Doctor who can then refer you for further specialist help if required. Please see below details of where you can access support:

- GP/Family Doctor
- NHS 111: To contact NHS 111 call "111". Available 24 hours a day, 7 days a week.
- The Samaritans: Call "116 123" to contact their helpline which is available 24 hours a day, 7 days a week. A range of support can also be found on their website www.samaritans.org
- BLISS Charity: A UK charity for babies born premature or sick. It supports families with a baby in neonatal care, website: https://www.bliss.org.uk email: hello@bliss.org.uk

Please note that service evaluation staff email addresses are not always monitored and should not be used in the case of an emergency. If you are experiencing a crisis please contact emergency services.

Neonatal Surgical Care Pathway – Improving Parents' Support

CONSENT FORM (Version/Date: V1.0 17/07/2023)

Name of Researchers: Hannah Wilkinson (trainee clinical psychologist), Dr Jacinta Cordwell (clinical psychologist), Dr Matthew Hotton (clinical psychologist)

1. I confirm that I have read the information sheet for the above project. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw without giving any reason and without my legal rights being affected.

3. I understand that relevant sections of the data collected during the study may be looked at by members of the research team. I give permission for these individuals to have access to my data.

4. I understand that taking part involves identifiable video recordings recorded on Microsoft Teams which will be kept securely by researchers from the University of Oxford, Department of Clinical Psychology, and will be deleted after being transcribed (written up).

5. I understand that the interview data will be transcribed and anonymised (personal identifying details and names/places removed) and that I will only have 3 weeks to withdraw my data from the study as after this time the interviews will be anonymised and pseudonymised. These transcribed interviews will be stored in an archive for a period of up to 3 years, and may be used for academic publication.

6. I understand that (if given) my contact details will be kept securely and used by researchers from the Clinical Psychology team at the University of Oxford and members of the psychology team at the John Radcliffe Hospital to contact me if I would like to receive updates on the study findings.

7. I understand that (if given) my contact details can be kept securely and used by researchers from the University of Oxford, Department of Clinical Psychology and members of the psychology team at the John Radcliffe Hospital for a period of up to 3 years, only to contact me if I want to know about the outcomes of the research, and that after this my contact details will be destroyed from records.

8. I agree to take part in the above project.

Please tick if you consent or do not consent and write your name and the date.

☐ I consent

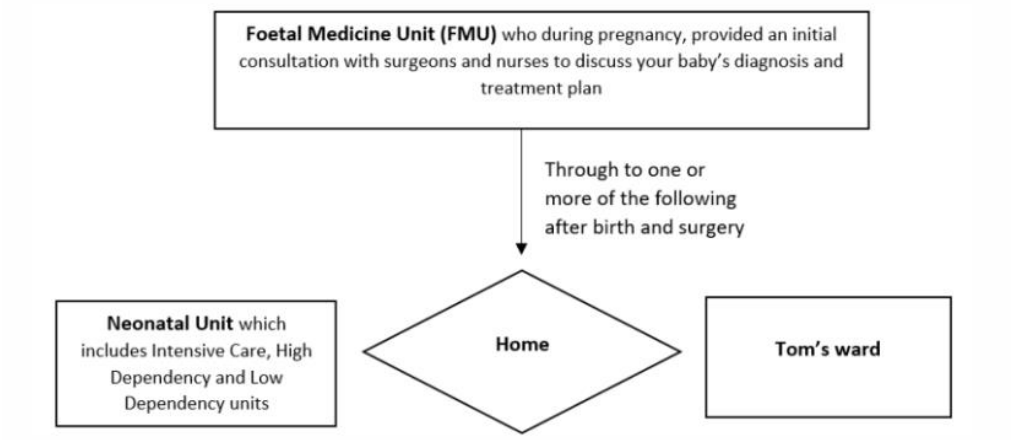
☐ I do not consent

Please write your name.

Please write the date

Appendix 5 – Survey questionnaire

This survey aims to explore the needs and experiences of parents whose babies require surgery for an antenatally diagnosed condition. Our view is to improve the support parents receive from the neonatal surgical care pathway at the John Radcliffe Hospital. We are interested in learning about how the service prepares parents for surgery and the implications of this, through to surgery and immediate after-care. We ask you to focus your responses on your care during a set period. We are specifically interested in when you first met with the surgical team antenatally to discuss your baby's diagnosis and treatment plan, through to when your baby was discharged from the surgical team. Please let us know about your time whilst under the below:



There will be five blocks of questions relating to (1) demographics, (2) the antenatal consultation, (3) your experiences after delivery, and (4) your experiences after surgery, and (5) an overall reflection on what made a difference along this treatment journey.

Block 1: Demographic questions

We are asking these questions to understand what factors affect people's experiences of the surgical pathway.

- What is your gender?
  - Male
  - Female
  - Other/ prefer not to say
- What is your ethnic group? Choose one option that best describes your ethnic group or background.
  - White
  - Mixed/Multiple ethnic groups
  - Asian/Asian British
  - Black/ African/Caribbean/Black British
  - Other ethnic group
- What is your partnership status?
  - Single parent
  - Couple
  - Other/Prefer not to say
- At how many weeks of pregnancy was your baby first identified as having an antenatal condition? [enter]
- Was this your first baby?
  - Y
  - N
- How many children do you have?
  - 1
  - 2
  - 3
  - 4
  - 5+

- What condition(s) was your baby diagnosed with?
    - Abdominal cysts
    - Gastroschisis
    - Exomphalos
    - Congenital cystic adenomatoid malformation
    - Congenital diaphragmatic hernia
    - Duodenal atresia
    - Tracheal oesophageal fistula
    - Meconium ileus
    - If other, please state below: [enter]
- Approximately how many months is it since your baby was discharged from the neonatal unit? [enter]

Block 2: At consultation, *when you went to the John Radcliffe to discuss your baby's condition and treatment plan with the surgeon and nurses, prior to your baby being born…*

- At how many weeks of pregnancy did this consultation take place?
- How easy was it to understand the information?
    - Very hard
    - Somewhat hard
    - Neither easy nor hard
    - Somewhat easy
    - Very easy
- What do you think about the amount of information shared with you?
    - Far too much
    - A bit too much
    - Just the right amount
    - A bit too little
    - Far too little
- How helpful did the information feel at the time?
    - Very helpful
    - Somewhat helpful
    - Neither helpful nor unhelpful
    - Somewhat unhelpful
    - Very unhelpful
- Did you have a central point of contact or know who you could speak to if you had more questions following the consultation?
    - Yes
    - No
- What were your feelings during the consultation (tick the top 3 emotions that fit you best)?

Shock		Anger		Love	
Panic / I can't cope		Pain		Confused	
Disbelief		Calm		Helpless	
Upset / grief		Contentment		Hopeful	
Guilt		Acceptance		Protective	
Other [please specify]					

- Did you feel that your overall psychological wellbeing at the time was acknowledged and considered by surgical team?
    - Yes
    - No
    - Partly
- Were your thoughts and feelings about the diagnosis and the treatment plan acknowledged and considered by the surgical team?
    - Yes
    - No

    - Partly
- Would it have been helpful to speak to a psychological professional at the time of this appointment?
    - Yes
    - No
    - Maybe
- Were you given the opportunity to visit the following units:
    - Maternity ward (Y/N)
    - Neonatal unit (Y/N)
    - Tom's ward (Y/N)
- If so, was this helpful?
    - Yes
    - No
    - Not applicable
- Was there anywhere else that would have been helpful to see? Please specify [ ]
- Were you given the opportunity to meet or discuss with other parents who faced similar challenges to you?
    - Yes
    - No
- How helpful do you think that would have been?
    - Very helpful
    - Somewhat helpful
    - Neither helpful nor unhelpful
    - Somewhat unhelpful
    - Very unhelpful

Block 3: At delivery, *around the time when you had just had your baby…*

- How much had the consultation prepared you for your baby's condition?
    - Very prepared
    - Somewhat prepared
    - Neither prepared nor unprepared
    - Somewhat unprepared
    - Very unprepared
- To what extent were your expectations from consultation met in relation to your baby's condition and care?
    - Far better than expected
    - Better than expected
    - As expected
    - Less than expected
    - Much less than expected
- Did you feel aware of where you could go and what you could do within the hospital (e.g. that you could see your baby whenever you wanted, that you could be taken from the maternity unit in a wheelchair to see your baby, where you could prepare and eat food on the unit)?
    - Yes
    - No
    - Partly
- Do you feel that communication between maternity and the team caring for you baby was good?
    - Very good
    - Somewhat good
    - Neither good nor poor
    - Somewhat poor
    - Very poor

- How did you feel during this period just after delivery (please select the top 3 emotions that fit you best)?

Shock		Anger		Love	
Panic / I can't cope		Pain		Confused	
Disbelief		Calm		Helpless	
Upset / grief		Contentment		Hopeful	
Guilt		Acceptance		Protective	
Other [please specify]					

- Did you feel that your overall psychological wellbeing at the time was acknowledged and considered by the surgical team?
    - Yes
    - No
    - Partly
- Were your thoughts and feelings about your baby's diagnosis and treatment plan were acknowledged and considered by the surgical team?
    - Yes
    - No
    - Partly
- Were you offered to see a psychological professional?
    - Yes
    - No
- [If answered yes] How helpful was it to meet with a psychological professional?/ [If answered no] How helpful do you think it would have been to see a psychological professional?
    - Very helpful
    - Somewhat helpful
    - Neither helpful nor unhelpful
    - Somewhat unhelpful
    - Very unhelpful
- Were you given the opportunity to meet or discuss with other parents who faced similar challenges to you following delivery?
    - Yes
    - No
- [If answered yes] How helpful was that opportunity?/ [If answered no] How helpful do you think this would have been?
    - Very helpful
    - Somewhat helpful
    - Neither helpful nor unhelpful
    - Somewhat unhelpful
    - Very unhelpful
- Did you feel involved in your baby's care?
    - Yes
    - No
    - Partly
- Were you involved in your baby's care as much as you would have liked?
    - I would like to have been involved much more
    - I would like to have been involved a bit more
    - My involvement felt just right
    - I would have liked to have been less involved
    - I would like to have been involved much less

***Block 4: Following surgery**, in the weeks when you received aftercare at the John Radcliffe neonatal care unit...*

- How many weeks old was your baby when they underwent surgery? [enter]
- In total, for how many weeks did your baby stay at the neonatal care unit after they had surgery?

- How did you feel during this period just after surgery? (please select the top 3 emotions that fit you best)?

Shock		Anger		Love	
Panic / I can't cope		Pain		Confused	
Disbelief		Calm		Helpless	
Upset / grief		Contentment		Hopeful	
Guilt		Acceptance		Protective	
Other [please specify]					

- Following surgery, did you feel that your overall psychological wellbeing at the time was acknowledged and considered by the surgical team?
    - Yes
    - No
    - Partly
- Were your thoughts and feelings about your baby's diagnosis and treatment plan were acknowledged and considered by the surgical team?
    - Yes
    - No
    - Partly
- Were you offered to see a psychological professional at this time?
    - Yes
    - No
- [If answered yes] How helpful was it to meet with a psychological professional?/ [If answered no] How helpful do you think it would have been to see a psychological professional?
    - Very helpful
    - Somewhat helpful
    - Neither helpful nor unhelpful
    - Somewhat unhelpful
    - Very unhelpful
- How many weeks after surgery did the surgeon meet with you for a follow-up consultation?
    - Under 1 week
    - Between 1 and 2 weeks
    - Between 2 and 3 weeks
    - More than 3 weeks
- Would you have preferred to meet sooner or later with the surgeon?
    - Much sooner
    - A bit sooner
    - The timing was just right
    - A bit later
    - Much later
- Was the information given by the surgeon at this follow-up meeting easy to understand?
    - Very easy
    - Somewhat easy
    - Neither easy nor difficult
    - Somewhat difficult
    - Very difficult
- Was the information given by the surgeon helpful?
    - Very helpful
    - Somewhat helpful
    - Neither helpful nor unhelpful
    - Somewhat unhelpful
    - Very unhelpful
- What (if any) information would you have liked from this meeting with the surgeon? [enter]
- If you were referred to other clinics or specialists for follow-on care, did you understand what would happen next along this treatment journey?

    - Yes
    - No
    - Partly
- Did you feel there was a good handover to these other clinicians providing follow-on care?
    - Very good
    - Somewhat good
    - Neither good nor poor
    - Somewhat poor
    - Very poor
- Did you feel that this transition in care took place at the right time?
    - Should have happened much sooner
    - Should have happened sooner
    - The timing was about right
    - Should have happened later
    - Should have happened much later
- How smooth and joined up was this transition?
    - Very smooth
    - Somewhat smooth
    - Neither smooth nor disjointed
    - Somewhat disjointed
    - Very disjointed
- How long is it since you were discharged from the John Radcliffe neonatal care team and this transition took place? (__ years and __ months)
- Overall, how accurate were the time-scales presented to you regarding your baby's treatment plan, recovery and discharge?
    - Very accurate
    - Somewhat accurate
    - Neither accurate nor inaccurate
    - Somewhat inaccurate
    - Very inaccurate

Block 5: Overall, *when you consider this journey whilst under the care of the neonatal surgical care pathway...*

- What were the biggest stressors or challenges that you experienced [enter]
- What factors helped your psychological wellbeing [enter]
- What further support would you have liked? [enter]
- We are wanting to hold interviews with parents to explore ideas for how we can improve the surgical pathway for parents. Please leave your email address if you are willing to be contacted about this (otherwise, please leave this box blank).

Appendix 6 – Interview schedule

Introduction/ rationale: Thank parents for the time. Aim for 30-60mins, but flexible if have more or less to say. Going to ask about several stages broadly relating to the support you received from the surgical team at the time from (1) antenatal consultation with doctors, (2) around the time of birth, (3) surgery, and (4) follow-up care. Aim is to improve the support provided to parents by this team so specifically asking about experiences at the JR with them. Do not have to answer any questions that don't want to. Any questions?
Please can you tell me a bit about your experience of receiving a diagnosis from the surgical team?

- How appropriate did the information feel at the time?
    - What do you think about the amount of information?
    - How easy was it to understand?
    - Was it sensitively delivered?
    - Was it at the right time?
    - Did it help you feel prepared?
- How did you feel? Were your thoughts and feelings about your baby's diagnosis and treatment plan acknowledged and considered by the surgical team? And your psychological wellbeing?
- What were the greatest challenges at this time-point?
- What was most helpful?
    - Who - was it family? Friends? Staff? What did they do that made a difference?
- What support would you have liked to receive?
    - How involved were you in any decision-making process?
    - Did you see a psychologist? How might this support have been helpful?

Please can you tell me a bit about your experience around the time of delivery for your baby?

- How did you feel? Were your thoughts and feelings acknowledged and considered by the surgical team? And your psychological wellbeing?
- What were the greatest challenges at this time-point?
- What was most helpful?
    - Who - was it family? Friends? Staff? What did they do that made a difference?
- What support would you have liked to receive?
    - Did you see a psychologist? How might this support have been helpful?

When your baby received surgery…

- How did you feel? Were your thoughts and feelings acknowledged and considered by the surgical team? And your psychological wellbeing?
- How involved were you in any decision-making processes?
- Did you feel that the surgery took place at the right time?
- What were the greatest challenges at this time-point?
- What was most helpful?
    - Who - was it family? Friends? Staff? What did they do that made a difference?
- What support would you have liked to receive?
    - Did you see a psychologist? How might this support have been helpful?

Please can you tell me a bit about your experience of your baby's surgical aftercare and discharge?

- When you had a consultation with the surgeon after surgery, did the information feel appropriate?
    - What do you think about the amount of information?

    - How easy was it to understand?
    - Was it sensitively delivered?
    - Was it at the right time?
    - Did it help you feel prepared?
- How involved were you in any decision-making processes?
- Was your baby discharged to another team? How joined up was this process?
- What were the greatest challenges at this time-point?
- What was most helpful?
    - Who - was it family? Friends? Staff? What did they do that made a difference?
- What support would you have liked to receive?
    - Did you see a psychologist? How might this support have been helpful?

Appendix 7 – Data analysis

Reflective thematic analysis was undertaken following Braun and Clarke's guidance (2006, 2021). Interviews were transcribed and anonymised. Transcripts were coded twice. The first round of coding took place at a dense, fine-grain, surface level where codes were units of meaning that stayed close to the original text; the second round of coding looked for more abstracted concepts that were potentially applicable across multiple interviewees. Supervisors reviewed examples of coding and provided guidance. Similarities and differences were compared for codes to ensure a consistent process. Following this, codes were grouped and organised into potential themes. Themes were considered in light of the original study questions (i.e., focused on identifying parents' needs, the challenges they faced, and suggestions for how the service could have been improved). Maps were made of how the themes might relate to each other – for example, considering whether they overlapped and fell under a greater overarching theme, or were distinctive. Data which conflicted with the theme (or "negative cases") were considered.

The next step involved giving themes labels. These aimed to "tell a story", or to highlight an essential point (Braun & Clarke, 2006, 2021). In line with the critical realist philosophical position, the researchers treated participants' experiences of the neonatal surgical pathway as describing a real-world truth, whilst being actively aware that their interpretations of the interviews was also influenced their own experiences and preconceptions. Themes thus aimed to stay close to what parents had described. Reflection with supervisors helped to maintain reflexivity about personal experiences or preconception which might influence the selection and development of codes and themes.

Appendix 8 – Author guidelines

1. AIMS AND SCOPE

The Journal of Pediatric Psychology is an official publication of the Society of Pediatric Psychology, Division 54 of the American Psychological Association, whose mission is to promote the health and psychological well-being of children, youth and their families through science and an evidence-based approach to practice, education, training, advocacy, and consultation. As such, the journal publishes articles related to theory, research, and professional practice in pediatric psychology.

2. MANUSCRIPT CATEGORIES AND REQUIREMENTS

Article type: Original research

Original Research is the most common type of journal manuscript used to publish full reports of data from original research. Authors are also encouraged to visit the Equator Network for additional information on transparent reporting of all manuscript types.

3. PREPARING THE SUBMISSION

Cover Letter

The cover letter is an essential document and must be included with the submission of all new manuscript submissions and revisions. The cover letter should be addressed to the journal editor and include the following:

- Manuscript title
- Assurance that all authors agree with the content of the manuscript and order of authorship
- Assurance the publication is not currently submitted to another journal
- Information about duplicate and redundant publications
- Notice of any conflicts of interest
- Consideration for a special issue/series or if this is an invited commentary
- Corresponding author contact information

Title Page
In addition to the APA Manual, the academic degrees of authors should be placed on the title page following their names.

Abstract
A structured double-spaced abstract of not more than 250 words should be included. The abstract should include the following parts:
a. Objective (brief statement of the purpose of the study)
b. Methods (summary of the participants, design, measures, procedure)
c. Results (the primary findings of this work)
d. Conclusions (statement of implications of these data)

Body of the Manuscript
a. Introduction
b. Methods - Informed consent and ethical treatment of study participants: Authors should indicate in the Method section of relevant manuscripts how informed consent was obtained and report the approval of the study by the appropriate Institutional Review Board(s).
c. Results
d. Discussion - Clinical relevance of the research should be incorporated into the manuscripts. There is no special section on clinical implications, but authors should integrate implications for practice, as appropriate, into papers.
Word limit: 5000 words.

References – Up to 50 references.

Appendix: Theory-driven Research Project

Appendix 1 – Ethical approval

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

Dr Myra Cooper & Hannah Wilkinson
Oxford Institute for Clinical Psychology Training & Research (OxICPTR)
Isis Education Centre
Warneford Hospital
Oxford

10 November 2023

Dear Dr Cooper and Hannah,

Research Ethics Approval - CUREC 1

Ethics Approval Reference: R88894/RE001

Study title: Do eating disorder cognitions play a role in Avoidant Restrictive Feeding Intake Disorder (ARFID)? A comparison between ARFID and Anorexia Nervosa

Short title: Comparing the thoughts of people with a varied diet and people who have eating difficulties

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 from **10th November 2023** until **31st May 2025.**

Insurance-provided indemnity arrangements are in place for the duration of the approval stated above. It is your responsibility to ensure that you request an extension to the end date for indemnity to remain in place should you continue the research beyond the dates covered.

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

DocuSigned by:
Leah Butts
9F1488902BC540A

Mrs Leah Butts
Research Ethics Administrator

for
Dr Helen Barnby-Porritt
Research Ethics Manager

Appendix 2 – Recruitment posters

Appendix 2.1 Poster advert

Oxford Institute of Clinical Psychology Training and Research
Isis Education Centre
Warneford Hospital,
Headington
Oxford
OX3 7JX

Oxford Institute of Clinical Psychology Training and Research
Research Team Email: Hannah.wilkinson@psy.ox.ac.uk; Myra.cooper@hmc.ox.ac.uk

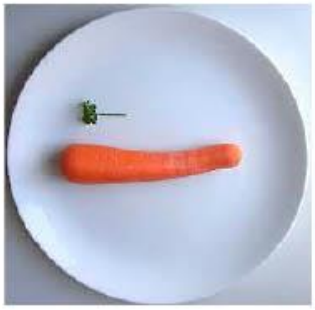
UNIVERSITY OF OXFORD

VOLUNTEERS NEEDED FOR RESEARCH STUDY

Ethics Approval Reference: R88894/RE001. Comparing the thoughts of people with a varied diet and people who have eating difficulties.

Do you have difficulties eating a varied diet, or enough food?

WE NEED YOUR HELP!







We are a group of Clinical and Trainee Psychologists, at the University of Oxford. Our study aims to understand the role of people's thoughts and beliefs in Avoidant Restrictive Feeding Intake Disorder (ARFID). ARFID was first recognised as an eating disorder in 2013. It consists of a range of restricted eating patterns due to a lack of interest in food; or sensory sensitivities to the taste or texture of food, and other food qualities; or a fear of the possible consequences of eating, such as a fear of choking.

Poster advert, version 1.0, September 2023

Oxford Institute of Clinical Psychology Training and Research
Isis Education Centre
Warneford Hospital,
Headington
Oxford
OX3 7JX

Oxford Institute of Clinical Psychology Training and Research
Research Team Email: Hannah.wilkinson@psy.ox.ac.uk; Myra.cooper@hmc.ox.ac.uk

We want to hear from everybody, with or without any eating difficulty. This is so we can understand how people's beliefs about food and themselves varies across three different groups:

(1) people with ARFID
(2) people with anorexia nervosa
(3) people without any eating disorder

We are looking for volunteers aged 18 years and above, who currently live in the UK and are able to read English. Taking part will take about 5-10 minutes and involves answering several questionnaires online.

Please note that you do not need a formal diagnosis to take part. This survey is not a medical assessment and will not provide a diagnosis or medical help. We recommend that you contact your GP if you have any concern about your eating, and physical or mental health. Our hope is that the research findings will be used to improve our ability to identify people who need support and help us improve interventions in future.

If you would like more information, please visit https://psychiatryoxford.qualtrics.com/jfe/form/SV_2laAvFVxbuLpr5c or scan the QR code below.

There is no obligation to participate and you can stop at any point. If you would like to speak to a member of our research team, please contact Hannah.wilkinson@psy.ox.ac.uk or Myra.cooper@hmc.ox.ac.uk. Thank you!

Poster advert, version 1.0, September 2023

Appendix 2.2 Social media advert



Appendix 3 - Measures

Questionnaire items were adapted by omitting repeated items (for example, both the EDDS-V and PARDI-AR-Q asked about weight and height). To minimise the burden to participants, some items were also omitted from the EDDS-V and PARDI-AR-Q where they did not contribute to diagnostic screening and were thus not required for the purposes of this research. On the ED-CBQ-R, an expert-by-experience raised that they did not know the meaning of several words and suggested they were old-fashioned; for these items, definitions of the words were provided to ensure accessibility.

Questionnaires (in their adapted forms, if applicable) are presented below.

Appendix 3.1 Demographic questions

1. What is your age?
    a. [enter in years]

2. What is your sex?
    a. A question about gender identity will follow later on in the questionnaire
    b. [ ] Female
    c. [ ] Male

3. Is the gender you identify with the same as your sex registered at birth?
    a. This question is voluntary
    b. [ ] Yes
    c. [ ] No
    d. (Enter gender identity)

4. What is your ethnic group?
    a. White
    b. Mixed/Multiple Ethnic groups
    c. Asian/ Asian British
    d. Black/African/Caribbean/Black British
    e. Other ethnic group

5. In the last seven days, were you doing any of the following? Tick all that apply. Include casual or temporary work, even if only for one hour
    a. Working as an employee
    b. Self-employed or freelance
    c. Temporarily away from work ill, on maternity or paternity leave, on holiday or temporarily laid off
    d. Doing any other kind of paid work
    e. None of the above

6. Which of the following describes what you were doing in the last seven days? Tick all that apply
    a. Retired (whether receiving a pension or not)
    b. Student

c. Looking after home or family
d. Long-term sick or disabled
e. Other

7. Do you have a brain injury or physical illness that affects diet, smell or taste (e.g. long covid, Crohn's disease, cystic fibrosis, coeliac's disease, diabetic).
   a. Y
   b. N
   c. Please enter what you have: ___

8. Are you pregnant?
   a. Y
   b. N

Appendix 3.2 EDDS – DSM-5 Version

Please carefully complete all questions, choosing NO or 0 for questions that do not apply.

Over the <u>past 3 months</u>…	Not at all			Slightly			Moderately	Extremely
1. Have you felt fat?	0	1	2	3	4	5	6	
2. Have you had a definite fear that you might gain weight or become fat?	0	1	2	3	4	5	6	
3. Has your weight or shape influenced how you judge yourself as a person?	0	1	2	3	4	5	6	

4. During the past <u>3 months</u> have there been times when you have eaten what other people would regard as an unusually large amount of food (e.g., a pint of ice cream) given the circumstances? . . . …… . YES NO

5. During the times when you ate an unusually large amount of food, did you experience a loss of control (e.g., felt you couldn't stop eating or control what or how much you were eating? . . . . . YES NO

6. How many <u>times per month</u> on average over the <u>past 3 months</u> have you eaten an unusually large amount of food <u>and</u> experienced a loss of control? 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16+

During episodes of overeating with a loss of control, did you…

7. Eat much more rapidly than normal?	YES	NO
8. Eat until you felt uncomfortably full?	YES	NO
9. Eat large amounts of food when you didn't feel physically hungry?	YES	NO
10. Eat alone because you were embarrassed by how much you were eating?	YES	NO
11. Feel disgusted with yourself, depressed, or very guilty after overeating?	YES	NO
12. If you have episodes of uncontrollable overeating, does it make you very upset?	YES	NO

In order to prevent weight gain or counteract the effects of eating, how many <u>times per month</u> on average over the <u>past 3 months</u> have you:

13. Made yourself vomit?	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16+
14. Used laxatives or diuretics?	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16+
15. Fasted (skipped at least 2 meals in a row)?	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16+

16. Engaged in more intense exercise specifically
to counteract the effects of overeating . . . . . 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16+

---

17. How many times per month on avearge over the past 3 months have you eaten after awakening from sleep or eaten an unusually large amount of food after your evening meal and felt distressed by the night eating? . . . . . .
. . . . . 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
15
16+

18. How much do eating or body image problems impact your relationships Not at all Slightly Moderately Extremely
with friends and family, work performance, and school performance? 0 1 2 3 4 5 6

19. How much do you weigh? If uncertain, please give your best estimate. _____
    - lbs. -or- _____kg.

20. How tall are you?
    - ___ft. ____in. -or- _____cm.

Appendix 3.3 PARDI-AR-Q Questionnaire

1. Do you think you have a problem with eating, involving avoidance or restriction of foods or your eating overall?
    a. Yes
    b. No
2. Have other people (for example, doctors, family members, significant others) said that you have a problem with eating, involving avoidance or restriction of foods or your eating overall?
    a. Yes
    b. No
3. Have your eating habits led to difficulty maintaining a sufficient weight or, if you are still growing, difficulty gaining enough weight to keep pace with your growth?
    a. Yes
    b. No
4. Have your eating habits led to you losing weight (in other words, if you have lost weight, this is because of avoidance or restriction and not because of a medical illness, or other reason)?
    a. Yes
    b. No
5. Have others (for example, doctors, family members) been concerned about your weight loss, or been concerned that you are having difficulty gaining enough weight to grow, or having difficulty maintaining your weight due to your eating habits?
    a. Yes
    b. No
6. Have others (for example, doctors, family members) been concerned that you are not growing taller as you should due to your eating habits?
    a. Yes
    b. No
    c. I have finished growing
7. Have you ever been told by any health professional that due to your eating habits you were not growing as expected, or that your height was less than it should be?
    a. Yes
    b. No
8. Over the past month, has any health professional said that you have a nutritional deficiency due to your eating habits (for example, low iron, low vitamin B12, low vitamin C)?
    a. Yes

b. No

9. Over the past month, has a healthcare professional prescribed special supplements (for example, pills, capsules, powders, or drinks containing vitamins and/or minerals and other micronutrients) specifically to help with your nutrition?
    a. Yes
    b. No
10. Over the past month, has a healthcare professional prescribed special supplements (for example, high-calorie drinks or 'shots', or dessert-style high-calorie supplements) specifically to help you maintain or gain weight?
    a. Yes
    b. No
11. Are you currently receiving any tube feeding (receiving food or fluid via a tube in your nose or into your stomach)?
    a. Yes
    b. No
12. Does your eating cause you difficulties in daily functioning - that is, in how you are able to go about things each day? This might be at school/college/work or when you are at home.
    a. Yes
    b. No
13. Does your eating cause you difficulties in interactions with other people (for example, disagreements or arguments with parents, siblings, significant others, co-workers), or difficulty making or sustaining friendships or other close relationships? Please rate below how difficult **interactions with other people** are for you because of your eating, ranging from 0 (= no difficulty) to 6 (= extremely difficult).

	0	1	2	3	4	5	6	
No difficulty	○	○	○	○	○	○	○	Extremely difficult

14. Does your eating cause you difficulties in social situations, for example does it make it difficult for you to go out with friends, eat at school/college/work, or stay away from home? Please rate how difficult **social situations** are for you because of your eating, ranging from 0 (= no difficulty) to 6 (= extreme difficulties/try to avoid all social situations).

	0	1	2	3	4	5	6	
No difficulty	○	○	○	○	○	○	○	Extreme difficulties

15. Over the past month, have you been particularly sensitive to variation in taste (for example, noticing slight differences in the taste of foods), which has put you off eating any foods or trying any new foods? Please rate below how much **sensitivity to taste** has affected your eating, ranging from 0 (= no negative effect/no particular sensitivity to taste) to 6 (= extremely negative effect, for example, leading

to refusal to eat many foods, sticking only to a limited number of preferred foods, or extreme caution when eating).

	0	1	2	3	4	5	6	
No negative effect	○	○	○	○	○	○	○	Extremely negative effect

16. Over the past month have you been particularly sensitive to the texture or consistency of food, which has put you off eating any foods or trying any new foods (for example, do you stick to foods of a certain texture only or have you had difficulty eating foods that have different textures mixed together such as pasta with sauce or sandwiches)? Please rate below how much **sensitivity to texture or consistency** has affected your eating, ranging from 0 (= no negative effect/no particular sensitivity) to 6 (= extremely negative effect, for example, leading to refusal to eat many foods, sticking only to a limited number of preferred foods, or extreme caution when eating).

	0	1	2	3	4	5	6	
No negative effect	○	○	○	○	○	○	○	Extremely negative effect

17. Over the past month, have you been particularly sensitive to the appearance of food, which has put you off eating any foods or trying any new foods (for example, if food does not look "right", such as burnt ends of chips/fries, broken biscuits/cookies, or being the "wrong" colour)? Please rate below how much **sensitivity to the appearance of food** has affected your eating, ranging from 0 (= no negative effect/no particular sensitivity) to 6 (= extremely negative effect, for example, leading to refusal to eat many foods, sticking only to a limited number of preferred foods, or extreme caution when eating).

	0	1	2	3	4	5	6	
No negative effect	○	○	○	○	○	○	○	Extremely negative effect

18. Over the past month, how often have you forgotten to eat or found it difficult to make time to eat? Please rate below how often you have **forgotten to eat or found it difficult to make time to eat**, ranging from 0 (= never) to 6 (= always).

	0	1	2	3	4	5	6	

Never	○	○	○	○	○	○	○	Always

19. Over the past month, how often have you lacked enjoyment in food or eating (even if only certain foods)? Please rate below how often you have **lacked enjoyment in food or eating**, ranging from 0 (= never) to 6 (= always).

	0	1	2	3	4	5	6	
Never	○	○	○	○	○	○	○	Always

20. Over the past month, how often have you felt full before your meal is finished, or stopped eating sooner than others because you had had enough? Please rate below how often you have **felt full or stopped eating early**, ranging from 0 (= never) to 6 (= always).

	0	1	2	3	4	5	6	
Never	○	○	○	○	○	○	○	Always

21. Over the past month have you been avoiding or restricting the amount or type of food you eat, because you were afraid that something bad might happen, like being sick, choking, having an allergic reaction, or being in pain? Please rate below how often being **afraid something bad might happen** has affected your eating, ranging from 0 (= never) to 6 (= always).

	0	1	2	3	4	5	6	
Never	○	○	○	○	○	○	○	Always

22. Over the past month have you avoided eating situations because you were worried something bad might happen, like being sick, choking, having an allergic reaction, or being in pain while eating (for example, because you might be served something you usually avoid for these reasons, or because you

have had a bad experience in the past)? Please rate below how often you have **avoided eating situations** due to such worries, ranging from 0 (= never) to 6 (= always).

	0	1	2	3	4	5	6	
Never	○	○	○	○	○	○	○	Always

23. Over the past month have you had any physical feelings of panic or anxiety (examples might include a racing heart, sweaty palms, feeling sick) when you have seen something that has made you think something bad might happen, like being sick, choking, having an allergic reaction, or being in pain while eating? Please rate below how often you have had **physical feelings of panic or anxiety** due to such thoughts, ranging from 0 (= never) to 6 (= always).

	0	1	2	3	4	5	6	
Never	○	○	○	○	○	○	○	Always

Appendix 3.4 GAD-7

Over the last 2 weeks, how often have you been bothered by the following problems?	Not at all sure	Several days	Over half the days	Nearly every day
1. Feeling nervous, anxious, or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it's hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid as if something awful might happen	0	1	2	3

Appendix 3.5 PHQ-8

Over the last 2 weeks, how often have you been bothered by any of the following problems?	Not at all	Several days	More than half the days	Nearly every day

1	Little interest or pleasure in doing things	0	1	2	3
2	Feeling down, depressed, or hopeless	0	1	2	3
3	Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4	Feeling tired or having little energy	0	1	2	3
5	Poor appetite or overeating	0	1	2	3
6	Feeling bad about yourself — or that you are a failure or have let yourself or your family down	0	1	2	3
7	Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8	Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3

Appendix 3.6 Eating Disorder Core Beliefs Questionnaire - Revised (ED-CBQ-R)

Listed below are a number of different words. People sometimes think these words describe how they feel about themselves as a person. Please read each word carefully and decide how much you feel each word describes how you feel about yourself. Base your answer on what you emotionally believe or feel to be true, not on what you rationally believe to be true. Choose the rating which best describes what you usually believe/feel or what you believe/feel most of the time, rather than how you feel right now. If you are unsure of the meaning of a word you may skip it. Work as quickly as you can. Don't spend too long on each word—your first impression is most important. Place a cross or tick in the box that best describes your response.

		Feels very much untrue	Feels moderately untrue	Feels slightly untrue	Feels neither true nor untrue	Feels slightly true	Feels moderately true	Feels very much true
1	Abandoned							
2	Betrayed							
3	Complaining							
4	Deprived							
5	Inhibited (reserved, shy, repressed)							
6	Meek							
7	Needy							
8	Possessive							
9	Putrid (contanimated, foul, sickening)							
10	Repugnant (loathsome, repellent, nasty)							

11 Repulsive
12 Selfish
13 Submissive (subdued, timid, passive)
14 Unassertive
15 Vile (horrible, dreadful, wicked)

Appendix 3.7 Eating Disorder Beliefs Questionnaire (EDBQ)

Listed below are some thoughts which people sometimes have when eating. Please read each thought carefully and decide how much you believe each thought to be true. Choose the rating which best describes how you usually feel rather than how you feel right now.
[Belief in each item is rated on a Likert (0- 100) scale. End points are anchored at "I do not usually believe this at all" and "I am usually completely convinced that this is true."]

Scale/Item
Factor 1: Negative self-beliefs
1. I'm stupid
2. I'm no good
3. I'm a failure
4. I'm useless
5. I'm dull
6. I'm not a likeable person
7. I'm all alone
8. I don't like myself very much
9. I'm unloveable
10. I'm ugly
Factor 2: Acceptance by others
11. If my thighs are firm it means I'm a better person
12. If my hips are narrow it means I'm successful
13. If my bottom is small people will take me seriously
14. If I gain weight it means I'm a bad person
15. If I gain weight I'm nothing
16. If my body shape is in proportion people will love me
17. If my hips are thin people will approve of me
18. If I lose weight people will be friendly and want to get to know me
19. If I lose weight people will care about me
20. If I lose weight I'll count more in the world
Factor 3: Self-acceptance
21. If my stomach is flat I'll be more desirable
22. If my flesh is firm I'm more attractive
23. If my body is lean I can feel good about myself
24. If I eat desserts or puddings I'll get fat
25. Body fat/flabbiness is disgusting
26. If I eat bad foods such as fats, sweets, bread and cereals they will turn into fat
Factor 4: Control over eating
27. If I've eaten something I have to get rid of it as soon as possible
28. If I binge and vomit I can stay in control
29. If I eat three meals a day like other people I'll gain weight
30. If I eat normally I'll gain weight
31. If I stay hungry I can guard against losing control and getting fat
32. If I eat a forbidden food I won't be able to stop

Appendix 3.8 Eating Disorder Thoughts Questionnaire (TQ)

Listed below are several different attitudes and beliefs which people sometimes hold. Please read each statement carefully and decide how much you agree or disagree with the statement. Base your answer on what you emotionally believe or feel, not on what you rationally believe to be true. Choose the rating which best describes what you usually believe, or what you believe most of the time, rather than how you feel right now.

[Belief in each item is rated on a Likert (0–100) scale, with statements "I do not usually believe this at all" and "I am usually completely convinced that this is true" comprising either end of the continuum.]

Factor 1: Negative thoughts about eating
I'll get fat
I'll have to go on a strict diet
I'll gain weight
My clothes won't fit anymore
I'll look a mess after eating so much—fat and disgusting
I'll hate myself for eating so much
I'm going to go on getting heavier and heavier
I've no self-control
The urge to binge is stronger than my willpower
I can't control my eating
Factor 2: Positive thoughts about eating
If I eat it will stop the pain
If I eat it will stop me feeling frightened
If I eat it will all hurt less inside
If I eat it means I don't have to think about unpleasant things
If I don't eat then I'll be overwhelmed with distressing thoughts and feelings
If I don't eat I'll lose control
Go on, eat more to punish yourself
It's not me doing this
I've nothing apart from eating/bingeing in my life
If I eat it will make me feel better, even if only for a little while
Factor 3: Permissive thoughts
One more bite won't hurt
I deserve something nice
It doesn't matter if I keep eating
If I eat it will stop me feeling bored
I'll just have a little bit more
If I eat it will comfort me, it's a way of being nice to myself

Appendix 4 – Scoring details

Appendix 4.1 Anorexia and bulimia diagnostic screening

The EDDS-V was scored using the SPSS code below to yield classical eating disorder diagnoses, which included anorexia and bulimia nervosa (created by Cara Bohon, provided in an email from Eric Stice, estice@stanford.edu, August 2023). Of note, participants who were identified as having a physical health condition that could also result in restrictive eating and difficulty maintaining a healthy weight (such as Crohn's disease) in the demographic questionnaire were screened out.

```
*Coding Guide for EDDS DSM-5
*Each questions is labelled with the number preceded by EDDS.
*EDDS19 should be coded in total kg for BMI calculation to be accurate.
*EDDS20 should be coded in total cm for the BMI calculation to be accurate.
*EDDS21 should be coded in total kg for weight loss calculation to be accurate.
*Yes is coded as 1, and No is coded as 0.
*Male is coded as 1, and Female is coded as 2.
*To compute BMI.
COMPUTE EDDSBMI=999.
COMPUTE EDDSBMI=(EDDS19)/((EDDS20/100)**2).
*To adjust binge eating frequency in case of errors in recording.
If (EDDS4=0 and EDDS5=0) EDDS6=0.
If (EDDS4=0 and EDDS5=0) EDDS7=0.
If (EDDS4=0 and EDDS5=0) EDDS8=0.
If (EDDS4=0 and EDDS5=0) EDDS9=0.
If (EDDS4=0 and EDDS5=0) EDDS10=0.
If (EDDS4=0 and EDDS5=0) EDDS11=0.
If (EDDS4=0 and EDDS5=0) EDDS12=0.
*For Diagnosis.
COMPUTE FEATURE=SUM.1(EDDS7, EDDS8, EDDS9, EDDS10, EDDS11).
COMPUTE BINGE=999.
If EDDS4=0 BINGE=0.
If EDDS5=0 BINGE=0.
If (EDDS4=1 and EDDS5=1 and EDDS6>=2) BINGE=1.
If (EDDS4=1 and EDDS5=1 and EDDS6>=4) BINGE=2.
RECODE BINGE (999=SYSMIS).
COMPUTE COMPSUM=999.
COMPUTE COMPSUM=SUM.2(EDDS13, EDDS14, EDDS15, EDDS16).
RECODE COMPSUM (999=SYSMIS).
COMPUTE COMPEN=999.
If COMPSUM<2 COMPEN=0.
If COMPSUM>=2 COMPEN=1.
If COMPSUM>=4 COMPEN=2.
RECODE COMPEN (999=SYSMIS).
COMPUTE PURGE=999.
COMPUTE PURGE=SUM(EDDS13, EDDS14).
COMPUTE WTSHAP=999.
If EDDS3<6 WTSHAP=0.
If EDDS3>=2 WTSHAP=1.
If EDDS3>=4 WTSHAP=2.
COMPUTE LOWBMI=999.
If EDDSBMI>10 LOWBMI=0.
If (EDDS23>=18 and EDDSBMI<=18.5) LOWBMI=1.
If (EDDS22=2 and EDDS23=17 and EDDSBMI<=18) LOWBMI=1.
```

```
If (EDDS22=1 and EDDS23=17 and EDDSBMI<=18.5) LOWBMI=1.
If (EDDS22=2 and EDDS23=16 and EDDSBMI<=17.75) LOWBMI=1.
If (EDDS22=1 and EDDS23=16 and EDDSBMI<=18) LOWBMI=1.
If (EDDS22=2 and EDDS23=15 and EDDSBMI<=17.25) LOWBMI=1.
If (EDDS22=1 and EDDS23=15 and EDDSBMI<=17.5) LOWBMI=1.
If (EDDS23=14 and EDDSBMI<=16.75) LOWBMI=1.
If (EDDS23=13 and EDDSBMI<=16.25) LOWBMI=1.
If (EDDS23=12 and EDDSBMI<=15.75) LOWBMI=1.
If (EDDS23=11 and EDDSBMI<=15.25) LOWBMI=1.
If (EDDS23=10 and EDDSBMI<=14.75) LOWBMI=1.
COMPUTE FEARWT=999.
If EDDS2<6 FEARWT=0.
If EDDS2>=2 FEARWT=1.
If EDDS2>=4 FEARWT=2.
RECODE PURGE WTSHAP LOWBMI FEARWT (999=SYSMIS).
COMPUTE WTLOSS=999.
If (EDDS19>(0.9*(EDDS21))) WTLOSS=0.
If (EDDS19<=(0.9*(EDDS21))) WTLOSS=1.
RECODE WTLOSS (999=SYSMIS).
COMPUTE EDDSDX=999.
If EDDS1<=6 EDDSDX=0.
If EDDS17>=4 EDDSDX=8.
If (PURGE>=4 and BINGE=0 and WTSHAP=2) EDDSDX=7.
If (BINGE=1 and FEATURE>=3 and EDDS12=1 and COMPEN=0) EDDSDX=6.
If (BINGE=1 and COMPEN=1 and WTSHAP=2) EDDSDX=5.
If (WTLOSS=1 and FEARWT=2 and WTSHAP=2) EDDSDX=4.
If (BINGE=2 and FEATURE>=3 and EDDS12=1 and COMPEN=0) EDDSDX=3.
If (BINGE=2 and COMPEN=2 and WTSHAP=2) EDDSDX=2.
If (LOWBMI=1 and FEARWT=2 and WTSHAP=2) EDDSDX=1.
If (LOWBMI=1 and COMPEN=2 and WTSHAP=2) EDDSDX=1.
If EDDS18<3 EDDSDX=0.
*1=AN, 2=BN, 3=BED, 4=Atypical AN, 5=low frequency BN, 6=low frequency BED, 7=purging
disorder, 8=night eating syndrome
*For Symptom Composite Scores
*For symptom composite scores, you can use the raw scores unless items are so positively skewed
that the alpha is not acceptable. In those instances, we recommend taking the z-score of items before
averaging.
*DSM-5 symptom composite via average of z-scores.
COMPUTE EDSYMZ=999.
COMPUTE EDSYMZ=Mean(zEDDS1 to zEDDS17).
*DSM-5 symptom composite via sum of raw.
COMPUTE EDSYMRAW=999.
COMPUTE EDSYMRAW=SUM(EDDS1 to EDDS17).
```

Appendix 4.2 ARFID diagnostic screening

Participants who did not meet the diagnostic criteria for classical eating disorder diagnoses on the EDDS-V, and who did not report having a physical health condition that might affect eating on demographic questionnaire items, were screened using the PARDI-AR-Q. A diagnosis of ARFID was determined accordingly:

Item 1 **OR** Item 2: Yes (1)

AND

Significant weight loss or failure to grow/ gain weight (Item 3 (Yes) OR Item 4 (Yes) OR Item 5 (Yes)) OR Item 6 (Yes) OR Item 7 (Yes))

OR

Significant nutritional deficiency (Item 8 (Yes))

OR
Dependence on enteral feeding or nutritional supplements (Item 9 (Yes) OR Item 10 (Yes) OR Item 11 (Yes))
OR
Psychosocial impairment (Item 12 (Yes))

Appendix 4.3 Anxiety symptoms
Individual items from the GAD-7 are scored from 0 to 3 with increasing severity. Items were summed to yield a total score on these questionnaires.

Appendix 4.4 Depressive symptoms
Individual items from the PHQ-8 are scored from 0 to 3 with increasing severity. Items were summed to yield a total score on these questionnaires.

Appendix 4.4 Eating Disorder Core Beliefs Questionnaire - Revised (ED-CBQ-R)
Individual items on the ED-CBQ-R yield scores between 0 and 6. The EDBQ yields a total score (mean of all items); and four subscale scores relating to, 'self-loathing' (mean of items 9, 10, 11, 15), 'unassertive' (mean of items 5, 6, 13, 14), 'demanding' (mean of items 3, 7, 8, 12), and 'abandoned' (mean of items 1, 2, 4).

Appendix 4.5 Eating Disorder Beliefs Questionnaire (EDBQ)
Individual items on the EDBQ yield scores between 0 and 100. The EDBQ yields a total score (mean of all items); and four subscale scores relating to, 'negative beliefs' (mean of items 1, 2, 3, 4, 5, 6, 7, 8, 9, 10), 'weight and shape as a means of acceptance by others' (mean of items 11, 12, 13, 14, 15, 16, 17, 18, 19, 20), 'weight or shape as a means of self-acceptance' (mean of items 21, 22, 23, 24, 25, 26), and 'control over eating', 'negative thoughts about eating' and 'permissive thoughts about eating' (mean of items 27, 28, 29, 30, 31, 32).

Appendix 4.6 Eating Disorder Thoughts Questionnaire (TQ)
Individual items on the TQ yield scores between 0 and 100. The TQ yields a total score (mean all items); and scores on three subscale relating to, 'positive thoughts about eating' (mean of items 1, 2, 3, 4, 5, 6, 7, 8, 9, 10), 'negative thoughts about eating' (mean of items 11, 12, 13, 14, 15, 16, 17, 18, 19, 20) and 'permissive thoughts about eating' (mean of items 21, 22, 23, 24, 25, 26).

Appendix 4.7 Measure of 'Restraint'
Items 13, 14, 15 and 16 of the EDDS-V were summed to yield the total number of times that people engaged in various disordered eating behaviours over the past three months, as an index of 'restraint.'

Appendix 4.8 Measure of 'Weight/shape concerns'
This was assessed using the two subscales of the EDBQ ('weight and shape as a means to acceptance by others' and 'weight and shape as a means to self-acceptance'). Details of their scoring is provided in Appendix B.6.

Appendix 5 – Results of non-parametric (Kruskal-Wallis) test of fourth hypothesis relating to prevalence of weight/shape concerns across diagnostic groups

Prevalence of weight/shape concerns

A Kruskal-Wallis test indicated that there was a significant difference in concerns about weight/shape as a means to acceptance by others across the three diagnostic groups, H (2, 184) = 55.27, p <.001. The mean rank was 122.80 for the Anorexia group, followed by 95.80 for the ARFID group, and 51.38 for the healthy controls. Post-hoc comparisons using Bonferroni corrected pairwise comparisons showed that the mean rank of the Anorexia group was significantly higher than both the ARFID (p =.012) and HC (p <.001) group. The mean rank of the ARFID group was significantly higher than HC (p <.001). Similarly, a Kruskal-Wallis test indicated that there was a significant difference in concerns about weight/shape as a means to self-acceptance across the three diagnostic groups, H (2, 184) = 52.80, p <.001. The mean rank was 125.34 for the Anorexia group, followed by 89.36 for the ARFID group, and 55.38 for the healthy controls. Post-hoc comparisons using Bonferroni corrected pairwise comparisons similarly demonstrated that the mean rank of the Anorexia group was significantly higher than both the ARFID (p =.002) and HC (p <.001) group. The mean rank of the ARFID group was significantly higher than HC (p <.001)

Appendix 6 – Cronbach's reliability results

Table 6.1 Reliability of dependent variable measures across diagnostic groups

Measure	Cronbach α		
	Anorexia	ARFID	Healthy
ED-CBQ-R	.88	.88	.96
EDBQ	.94	.98	.96
TQ	.89	.87	.92
Weight/shape concerns			
Weight/shape for acceptance by others	.93	.97	.94
Weight/shape for self-acceptance	.82	.93	.90

Appendix 7 – Results of stepwise regression

Table 7.1 shows the results of a stepwise linear regression in which ED-related core beliefs, assumptions, and automatic thoughts were entered as predictors of restraint behaviours in the ARFID group. The criteria for variable entry and removal were set at a probability of F-to-enter < .050 and probability F-to-remove >.100. ED-related assumptions were the best predictor of restraint behaviours and therefore entered at the first step. Core beliefs was the next best predictor of restraint after controlling for the effects of assumptions, and was therefore entered at the second step, however it became negatively related to food restriction. This suggests a reversal paradox, since all items were positively correlated to restraint (see correlation matrix in main article, Table 5). Suppression effects were also suggested by the observation that automatic thoughts no longer explained a significant amount of variance after controlling for assumptions or core beliefs, and was therefore excluded from the final model.

Strange behaviour in coefficients including sign-switching is indicative of problems of multicollinearity (Nickerson & Brown, 2019).

Table 7.1. Results of stepwise regression

Model	*β*	*SE*	*t*	*p*	Overall model summary
1 (Constant)	-	-	-	-	$F(1,59) = 46.91$, $p < .001$,
Assumptions	.66	4.01	6.85	<.001	adjusted $R^2 = 43.3\%$
2 (Constant	-	-)	-	-	$F(2,58) = 30.19$, $p < .001$,
Assumptions	.84	.05	7.59	<.001	adjusted $R^2 = 49.3\%$
Core beliefs	-.31	1.17	-2.82	.007	

Appendix 8 – Reproduction of models of ARFID profiles from Thomas & Eddy, 2018 (p. 21-23)

Figure 8.1 General cognitive-behavioural model of ARFID

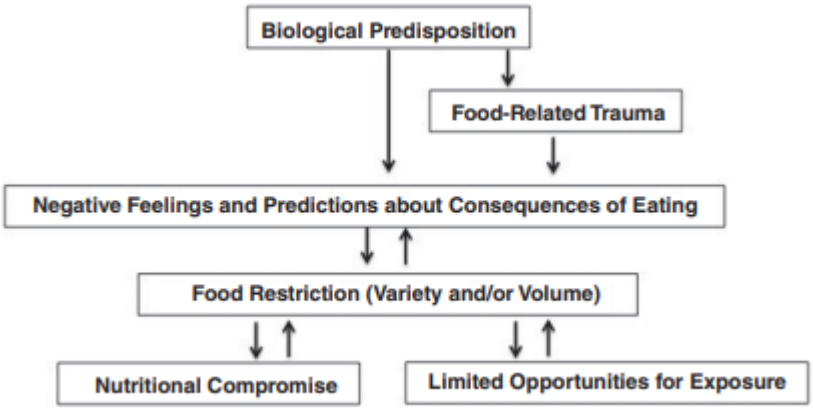


Figure 8.2 Cognitive behavioural model of ARFID with sensory sensitivity

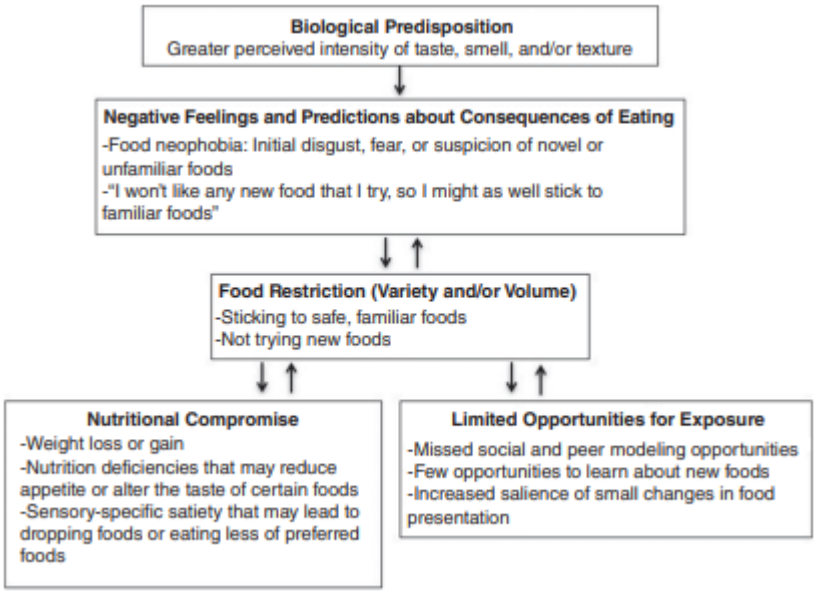


Figure 8.3 Cognitive-behavioural model of ARFID with fear of aversive consequences

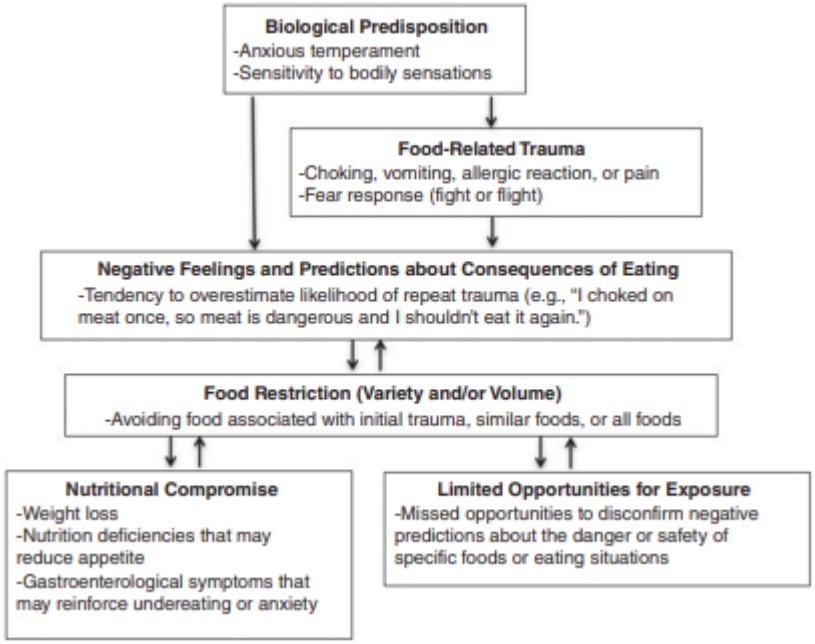
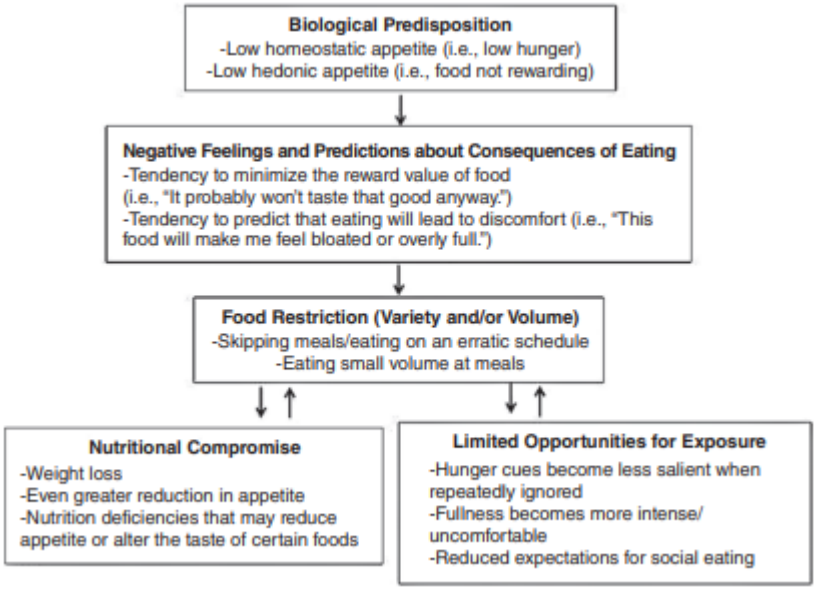


Figure 8.4 Cognitive-behavioural model of ARFID with apparent lack of interest in eating food



Appendix 9 – Author Guidelines for International Journal of Eating Disorders

1. AIMS AND SCOPE

With a mission to advance the scientific knowledge needed for understanding, treating, and preventing eating disorders, the International Journal of Eating Disorders publishes to an international readership of health professionals, clinicians and scientists. We also draw the interest of patient groups, advocates and mainstream media outlets. We welcome state-of-the-art scientific research on theory, methodology, etiology, clinical practice, and policy as well as contributions that facilitate scholarly critique and discussion of science and practice in the field.
There is a 'Special Issue' call for Papers on 'Avoidant Restrictive Food Intake Disorder.'

2. MANUSCRIPT CATEGORIES AND REQUIREMENTS

Original Article: Reports of new research findings that make a significant contribution to knowledge. These contributions report substantive research that is novel, definitive, or complex enough to require a longer communication.

- Word limit excluding abstract, references, tables or figures, is 4500 words
- Recommended abstract structure/word limit is 250 words
- Other requirements/recommendations are:
    - Data Availability Statement
    - IRB Statement
    - Public Significant Statement
    - ≤8 figures/tables

Authors are advised to see the Original Article Submission Checklist (https://onlinelibrary.wiley.com/page/journal/1098108x/homepage/forauthors.html/checklist/original-article?) and IJED Statistical Reporting Guidelines Checklist (https://onlinelibrary.wiley.com/pb-assets/IJED%20Statistical%20Reporting%20Guidelines%20Checklist-1589468996657.pdf) prior to submission.

3. PREPARING THE SUBMISSION

The International Journal of Eating Disorders (IJED) offers Free Format submission for a simplified and streamlined submission process.

Title Page

The Title Page of the manuscript should comprise:

- A brief informative title containing the major keywords. The title should not contain abbreviations.
- All co-author details, including affiliation and email address, and ORCID identifier where possible.
- 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.
- 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.

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). 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

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.

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.

Reproduction of Copyright Material

If excerpts from copyrighted works owned by third parties are included, credit must be shown in the contribution. The corresponding author is responsible for obtaining written permission to reproduce the material "in print and other media" from the copyright owners of the original source, and for supplying Wiley with that permission upon submission. For more information visit Wiley’s Copyright Terms & Conditions FAQ.