Article:
Nicholls-Clow, R., Simmonds-Buckley, M. orcid.org/0000-0003-3808-4134 and Waller, G. orcid.org/0000-0001-7794-9546 (2024) Avoidant/restrictive food intake disorder: Systematic review and meta-analysis demonstrating the impact of study quality on prevalence rates. Clinical Psychology Review, 114. 102502. ISSN 0272-7358

https://doi.org/10.1016/j.cpr.2024.102502

Clinical Psychology Review 114 (2024) 102502

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

Avoidant/restrictive food intake disorder: Systematic review and meta-analysis demonstrating the impact of study quality on prevalence rates

Rebecca Nicholls-Clow [a], Melanie Simmonds-Buckley [b], Glenn Waller [b,*]

[a] Department of Psychology, University of Sheffield, Sheffield, UK
[b] Clinical and Applied Psychology Unit, Department of Psychology, University of Sheffield, Sheffield, UK

ARTICLE INFO

Keywords:
Avoidant/restrictive food intake disorder
ARFID
Prevalence
Meta-analysis
Quality effects analysis

ABSTRACT

Objectives: The prevalence of Avoidant/Restrictive Food Intake Disorder (ARFID) is unclear. This paper is the first to present meta-analysis based estimates of the prevalence of ARFID, and to assess the impact of the quality of the research on these estimates.
Design: A pre-registered (Prospero: CRD42023487621) systematic review and meta-analysis.
Methods: PubMed, PsychInfo, Web of Science and CINAHL were searched (final date of retrieval 30th July 2024) for peer reviewed papers published between 2013 and 2024. Random-effects and quality effects meta-analyses were used to compute and compare prevalence estimates and to evaluate the impact of study quality on prevalence rates. Subgroups were also considered (gender, age group, clinical status). Loney et al.'s (1998) *Critical Appraisal of the Health Research Literature: Prevalence or Incidence of a Health Problem* scale was used to assign each study a quality score across three categories - methodological validity (six points); interpretation of results (one point); and applicability of the results (one point).
Results: Twenty-six studies were identified (n = 122,861). Meta-analysis using random-effects indicated a prevalence of 11.14 % (95 % CI 8.16–14.5 %), whereas quality effects prevalence was 4.51 % (95 % CI 0.7–10.68 %). Similar contrasts were evident among subgroups.
Conclusions: Even taking the more conservative estimate of 4.51 %, this review demonstrates that ARFID is a common disorder, meriting further research and clinical and service developments. Future research needs to be more methodologically robust (larger samples; standardised diagnostic measures; clearer data presentation).

1. Introduction

Avoidant/restrictive food intake disorder (ARFID) was first included in the Diagnostic and Statistical Manual Fifth Edition (DSM-5; American Psychiatric Association, 2013), providing a reformulation and expansion of the earlier Feeding Disorder of Infancy or Early Childhood construct. ARFID encompasses several terms (e.g., picky eating; food phobia; selective eating) previously used to describe restrictive eating behaviors that did not meet the criteria for existing eating disorders (Sanchez-Cerezo et al., 2023). The disorder is defined by an avoidance or restriction of food intake that results in a persistent failure to meet appropriate nutritional needs through oral intake of food. Unlike anorexia nervosa, patients with ARFID are not concerned with a fear of or preoccupation with weight gain or changes to body shape and size. Food restriction in ARFID can be due to any of three issues – an inability to tolerate certain sensory qualities of food; a lack of interest in eating or food; or concern about the potential adverse consequences of eating, such as vomiting or choking (American Psychiatric Association, 2013). To meet the diagnostic criteria, one or more of the following key features must be present: significant weight loss; significant nutritional deficiency; dependence on enteral feeding or oral nutritional supplements; or a marked interference with psychosocial functioning (American Psychiatric Association, 2013; Sanchez-Cerezo et al., 2023).

The onset of ARFID can occur at any point in the lifespan, though symptoms typically become apparent in childhood (Bourne et al., 2020; Schmidt et al., 2016). The most typical age of onset is somewhat

* Corresponding author at: Clinical and Applied Psychology Unit, Department of Psychology, University of Sheffield, Floor F, Cathedral Court, 1 Vicar Lane, Sheffield S1 2LT, United Kingdom.
E-mail address: g.waller@sheffield.ac.uk (G. Waller).

https://doi.org/10.1016/j.cpr.2024.102502
Received 27 March 2024; Received in revised form 6 August 2024; Accepted 10 September 2024
Available online 11 September 2024

contentious, as neophobia (refusal to accept new foods) can be a behavior typical of young children. However, whilst this brief aversion usually resolves as the child ages, ARFID symptoms have been shown to persist, and can continue to occur across the lifespan (Dovey et al., 2008). It also has a relatively high rate of identification among males compared to other eating disorders, particularly among younger cases (Norris et al., 2014). When left untreated, the physical and psychosocial complications of ARFID extend well beyond those of developmentally 'picky' or 'selective' eating. They include anxiety, obsessive compulsive disorder, bradycardia, electrolyte abnormalities similar to those with anorexia nervosa, lower bone mineral density and loss of vision (Alberts et al., 2020; Chiarello et al., 2018; Thomas et al., 2017). There is limited research as to the comorbidities that may occur with a presentation of ARFID. It is thought that when compared with other restrictive eating disorders, ARFID may more frequently co-occur with anxiety and neurodevelopmental disorders (Aulinas et al., 2020).

Understanding the prevalence of ARFID is important in predicting the demand on health services as the new therapies being developed for this disorder are rolled out to patients (Dumont et al., 2019; Lock et al., 2019; Thomas et al., 2020; Thomas et al., 2021). However, while ARFID is identified in multiple countries and is more commonly seen in clinical settings (Micali & Cooper-Vince, 2020), the prevalence of ARFID is currently hard to determine. Despite the condition's clinical significance, there are few large-scale epidemiological studies of ARFID, with research largely focusing on child and adolescent populations. Much of our current knowledge about the scope of the disorder is based on small clinical samples (Bourne et al., 2020; Kennedy et al., 2023; Sanchez-Cerezo et al., 2023).

There is a growing body of non-systematic reviews in this domain, which provide valuable insight into existing research and the current understanding of ARFID. However, there are currently no meta-analyses that summarise the prevalence of ARFID in clinical or non-clinical populations. Accurate prevalence data of this sort are essential to promote greater awareness at a community and clinical level, in order to tackle the current likely under-resourcing of clinical and research work in this field. Such prevalence data will enable accurate planning, health care education, diagnosis and, ultimately, effective treatment of ARFID (Kambanis & Thomas, 2023).

Existing prevalence reports vary substantially across individual studies, probably due to significant methodological heterogeneity. Therefore, the first aim of this study was to conduct a systematic review and meta-analysis to assess the literature available on the prevalence of ARFID and to combine the evidence into data-driven estimates of the prevalence of this disorder. Study-level characteristics (e.g., population sampled, gender, age) will be examined to establish whether they drive differences in prevalence estimates.

However, it is also important to consider how, within epidemiological research, inter-study differences can be due in part to the quality of the research (i.e., design-related heterogeneity), such that the quality can introduce bias (Bailey, 1987). The outcome of a random-effects meta-analytic model can be unrepresentative, due to quality-based differences in the individual studies, which in turn can render the conclusions unrepresentative (Doi & Thalib, 2008). In contrast, a quality effects model adjusts for the between-study variability in quality of the studies. This approach can be particularly useful in a relatively new field, as such variability is likely to be high in preliminary studies, with their very different designs and methodologies. Redistributing the study weights by quality should produce more accurate prevalence statistics for ARFID (Doi & Thalib, 2008). Therefore, the second aim of this research is to determine whether a quality effects model yields comparable outcomes to the more commonly used random-effects model. This outcome will show whether the quality of the available research influences reported prevalence estimates of ARFID, providing a more generalisable and valid summary. The results will be used to inform recommendations to enhance the reliability and validity of future research into the prevalence of ARFID.

2. Method

2.1. Protocol registration

This review was conducted and reported in accordance with the Preferred Reporting Item for Systematic Reviews and Meta-Analysis (PRISMA) standards (Page et al., 2021). The protocol was pre-registered on PROSPERO (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=487621) on 29th November 2023. It was amended on 9th February 2024 to include a quality effects meta analytical approach in the Intended methodology in order to provide a comparison to the random effects results. A final amendment was made on the 11th March 2024 to include the quality effects method in the title. All review stages were conducted by the first author.

2.2. Search strategy and eligibility criteria

A literature search was conducted to identify eligible studies from four electronic databases (PsycInfo, PubMed, Cumulated Index to Nursing and Allied Health Literature (CINAHL), and Web of Science). The final date of retrieval was 30th July 2024. Databases were searched using variations of the following search terms; "prevalence", "ARFID", "avoidant restrictive food intake disorder", "picky eat*", "selective eat*", "food refusal", "emotional restriction", "feeding disorder of early childhood", "ASD", "Autism Spectrum Disorder" and "autism*". Full details of the exact search terms used for each database, including Boolean phrases as appropriate, can be seen in the Supplementary Material, Table 1. Papers were restricted to those published between 2013 and 2024 in order to align with the introduction of ARFID in DSM-5. Only peer-reviewed English language papers were included in the results. Grey literature was not searched due to the infancy of much of the literature in this area. Google Scholar was not used as it is limited in the specificity of the filters and search terms that can be used, making it difficult to organise the large search yields reliably and accurately (Mahood et al., 2014).

Duplicated papers were initially removed automatically by EndNote, with any remaining duplicates removed manually by the lead researcher. Preliminary screening of the titles (to include 'ARFID' or 'prevalence') was carried out, followed by further screening of the abstracts (to include reference to both 'ARFID' and 'prevalence'). Full texts

Table 1
Inclusion and exclusion criteria.

	Inclusion Criteria	Exclusion Criteria
Population	All adult, child, and adolescent samples.	
Diagnosis	A definite or possible (to include wording such as 'borderline' or 'potential') diagnosis of ARFID, in line with the DSM-5 diagnostic criteria.	Studies that do not specify an ARFID diagnosis in line with DSM-5 diagnostic criteria.
Outcomes	Studies that report primary data relating to the prevalence of ARFID.	Studies which do not report primary data relating to the prevalence of ARFID.
Setting	Any clinical or nonclinical setting.	
Study Design	Any quantitative studies reporting primary data.	Qualitative studies. Studies reporting secondary data. Grey literature. Non-English language.

* is a wildcard, used to denote that any variation of the search term, in this case "eat" can be included in the results. For example "eating", "eaters".

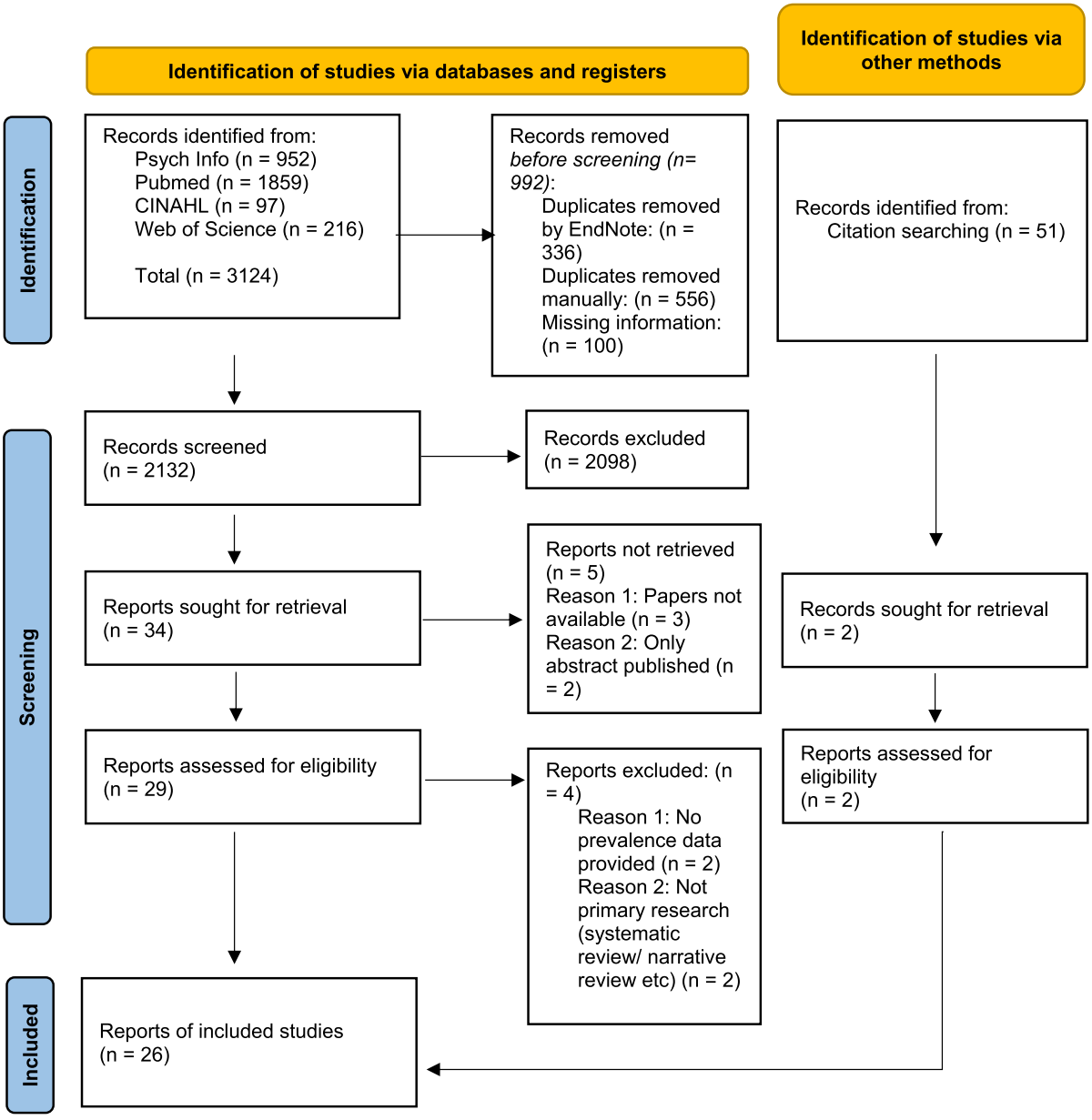
of the remaining papers were then downloaded. For papers that were not openly available, authors were contacted directly. If they did not supply copies, the papers were excluded. Literature was then reviewed based on the pre-established inclusion and exclusion criteria (Table 1). Studies were eligible for inclusion if they reported primary data relating to the prevalence of ARFID in any setting (clinical or nonclinical) and across any sample population (adult and child). Studies reporting either definite or possible cases of ARFID were included if the DSM-5 diagnostic criteria were directly referenced and met. All relevant citations in the excluded literature were also searched manually and screened using the same process. The PRISMA flowchart (Fig. 1) details the number of papers omitted and included at each stage of the process.

2.3. Data extraction

Data extracted included: title; author/s; year of publication; type of publication; doi; sample characteristics (number in the sample, age range, mean age, gender split, recruiting location and setting); study design; diagnostic tools; diagnosis details (potential or definite diagnosis; detail of presentations of ARFID if provided); and prevalence statistics (number of cases included; stated point, weighted or lifetime prevalence; and corresponding confidence intervals). Data extraction on all papers was repeated by a second graduate student, and interrater agreement on the data extraction process was 95 %, with discrepancies resolved by discussion.

2.4. Quality appraisal

Quality appraisal of the studies was completed using the *Critical Appraisal of the Health Research Literature: Prevalence or Incidence of a Health Problem* (Loney et al., 1998). Studies were scored a maximum of eight points over three categories: methodological validity (six points); interpretation of results (one point); and applicability of the results (one point). As the quality scores underpin the quality effects analysis,



*

Fig. 1. PRISMA diagram based on Page et al. (2021).

interrater agreement was measured for all papers, using an uninvolved researcher as the second rater. Percentage agreement was calculated alongside Cohen's kappa to provide a clear indication of reliability across the two measures (McHugh, 2012; Zhao et al., 2022).

To assess the certainty in the synthesised meta-analytic evidence representing the true effect, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system was used (Granholm et al., 2019). This approach is separate to the study quality appraisal as it considers the entire body of evidence rather than critical appraisal of the individual studies. There is no formal guidance for applying GRADE to prevalence reviews. However, use of the baseline risk/overall prognosis GRADE guidelines has been recommended for prevalence reviews in the interim (Migliavaca, Stein, Colpani, Munn, & Falavigna, 2020). Following these suggested GRADE guidelines, initial evidence quality level was set at 'high', due to the observational study designs of the included studies to ascertain prevalence in the population (as opposed to reviews of intervention efficacy where observational designs are set at 'low' quality). To determine the final quality of evidence rating, the meta-analyses were assessed according to five criteria: risk of bias in included studies; degree of imprecision in synthesised estimates; degree of inconsistency represented by unexplained heterogeneity; indirectness of evidence; and extent of publication bias. Quality level was down- or upgraded based on the evidence for each criterion.

2.5. Data synthesis and analysis

Prevalence can be measured in multiple ways. Point prevalence is the proportion of the population suffering from a condition (in this case ARFID) at a given point in time (Migliavaca, Stein, Colpani, Barker, et al., 2020), whereas lifetime prevalence is the proportion of a sample having had at least one episode of illness in their life up to the time of sampling (Streiner et al., 2009). There has been some discussion as to whether the concept of lifetime prevalence should be dropped from the lexicon of epidemiology, as results from such analyses consistently show an unexplained declining trend that is thought to be because of flawed study designs. Taking this into account, along with the data that are available from the included studies, the point prevalence of ARFID was the outcome of interest in this review, with a secondary aim of determining the relevance of study quality to reported prevalence levels. A prevalence estimate for each included study was calculated as the proportion of participants in the sample classified as having ARFID (i.e., number of ARFID cases divided by total number of sample). For those studies that reported multiple samples (for example, both clinical and nonclinical, as can be seen in Schöffel et al., 2021) results from the study were combined into a total sample number and a corresponding total prevalence percentage to be included in the meta analysis. These multiple samples were then used separately in the appropriate subgroup analyses. Estimates of the variability around prevalence rates are known to be at risk of bias due to constraint with a 0–1 proportional scale when proportions are high (close to 1) or low (close to 0). To account for this bias, study proportions were transformed using a Freeman-Tukey double arcsine transformation for synthesis and back-transformed to percentages for interpretation (Barendregt et al., 2013).

Synthesis of the prevalence estimates was conducted using the Meta XL add-on for Excel. Prevalence data were pooled initially using a random effects (RE) model (employing a Der Simonian and Laird estimator), with studies weighted using the inverse-variance method. Synthesis of prevalence rates were then repeated using a quality effects (QE) meta-analysis. Whilst Meta XL suggests using a safeguard score designed by the programme's developers, a general risk of bias quality assessment score can be used (Barendregt & Doi, 2016). The Loney et al. (1998) scale was used as it is more specific to studies of prevalence.

Quality effects analysis adjusts for the between-study variability using an assessment of the varying quality of the studies. The quality appraisal scores were converted into quality ranks between 0 and 1, allowing for the redistribution of the study weighting by quality (Barendregt et al., 2013). Redistributing the study weights by quality should allow for the more accurate prevalence statistics, due to the relatively greater weight redistribution when weighting by precision (Doi & Thalib, 2008). Random effects models have been shown to result in potentially inflated outcomes (Kelley & Kelley, 2020; Liu et al., 2021) when compared with the more recently developed quality effects models. This difference arises because the quality effects models can weaken the influence of heterogeneity, tackling an issue that is inherent in the random effects model.

Initially, the random and quality effects meta-analyses were conducted using all of the samples from the included studies in order to create the two pooled prevalence estimates. These results were then visualised in corresponding forest plots, which display both the individual study estimates and the overall pooled prevalence with a corresponding 95 % confidence interval. Heterogeneity was assessed through Q and I^2, which is the percentage of total variance across the included studies that stems from actual difference rather than random error. I^2 values were interpreted as >25 %, >50 % and > 75 % indicating low, moderate and high heterogeneity respectively (Higgins et al., 2003).

However, in the case of ARFID, subgroup analysis is particularly important (e.g., gender, age group). In order to assess whether the prevalence of ARFID would differ across setting and sample, the following subgroups were devised: all children; all adults; all males; all females; female children; female adults; male children; male adults; all clinical samples; and all non-clinical samples. The RE and QE meta analyses were conducted on each of these groups in order to create a pooled prevalence estimate with a corresponding 95 % confidence interval. Heterogeneity of the subgroups was also assessed through Q and I^2. However, it is acknowledged that the power of these tests will be very low due to the small size of the subgroups available (Cuijpers et al., 2021).

Whilst the implication of publication bias may be less influential in proportional meta-analyses (Simmonds-Buckley et al., 2022), an assessment of potential reporting bias was made through a visual inspection of the funnel plot. A multiple regression was also performed on Microsoft Excel to test whether there was a relationship between quality scores for each paper and potentially related features of the studies - date of publication, level of reported prevalence, or setting (clinical or non-clinical.

3. Results

The PRISMA diagram (Fig. 1) outlines and illustrates the selection of studies eligible for inclusion in the review. Supplementary Table 2 provides a list of excluded papers from the full text retrieval stage of the review and reasons for exclusion. After removing duplicates and papers that were missing key information, 2132 records were identified for initial screening. Of these, 2098 were excluded as not meeting the eligibility criteria, leaving 34 for full text retrieval, of which 29 were available. Four of the 29 were excluded, leaving 25 papers. A further 51 papers were identified from citation searches, of which one was included in the final review. In total, twenty-six papers were eligible for inclusion. The full titles of the included papers can be seen in the Appendix.

3.1. Characteristics of included studies

Table 2 summarises characteristics of the 26 included studies. All studies included were published between 2014 (the year after the start point of the search) through to 2024. Studies were included from twelve countries: USA *(k = 11);* Australia *(k = 2)*, Sweden *(k = 2),* Japan *(k = 2),* Germany *(k = 2)*; UK *(k = 1)*; Amsterdam *(k = 1)*; Taiwan *(k = 1);* Malaysia *(k = 1);* Canada *(k = 1)*; Portugal *(k = 1)* and Pakistan *(k = 1)*. Out of the 26 studies included in the final analyses, two report definite cases of ARFID, eight report possible cases of ARFID, and 16 did not specify. The measurement/diagnostic tools used included the NIAS ($k = 6$), DSM-5 criteria ($k = 11$), EDY-Q ($k = 3$), and the PARDI-AR-Q ($k = 2$)

Table 2
Characteristics of the included papers.

First author and year	Study design	Region	Included in which Subgroup Analyses	Study Setting	Adult/ Child/ Mixed	Sample Age Range	Diagnostic tool	N	N ARFID	Prevalence %	Quality Score /8
Atkins et al., 2023	Retrospective Chart Review	USA	Clinical; adult; child	Clinical – tertiary care centre adult and child neuro-gastrology clinic	Mixed	6 to 90 years	DSM-5 criteria	574	130	22.65 %	8
Van Buuren et al., 2023	Cross-sectional	Australia	Nonclinical; child; male; Female	None - schools	Child	11 to 19 years	EVERYbody online survey cross referenced with DSM criteria	4896	97	1.98 %	8
Burton-Murray et al., 2024	Cohort Study	USA	Clinical; adult; male; female	Clinical – Gastro unit at a general hospital	Adult	22 to 80 years	NIAS, EDE-Q8	101	11	11 %	6
Chen et al., 2019	Cross-sectional (national survey)	Taiwan	Nonclinical; child	None	Child	7 to 14 years	Mandarin version of the K-SADS-E for DSM5	4816	40	0.5 %	7
Chua et al., 2022	Online questionnaire	Malaysia	Nonclinical; adult; male; female	None	Adult	18 to 73 years	Stanford-Washington Eating Disorder Screen	818	39	4.8 %	6
D'Adamo et al., 2023	Cross-sectional online	USA	Nonclinical; adult	None	Adult	NA	Questions based on DSM5 criteria	50,082	2378	4.7 %	7
Dinkler et al., 2023	Cross-sectional	Sweden	Nonclinical; child; male; female	None	Child	6 to 12 years	DSM5 criteria	33,902	682	2 %	7
Dinkler et al., 2022	Cross-sectional	Japan	Nonclinical; child; male; female	None	Child	4 to 7 years	ARFID screener (questionnaire developed by researchers)	3746	49	1.5 %	5
Eddy et al., 2015a	Retrospective Chart Review	USA	Clinical; child	Clinical – 19 paediatric gastroenterology clinics	Child	8 to 18 years	DSM-5 diagnostic checklist	2231	33	1.5 %	8
Farag et al., 2021	Case Control Study	UK	Clinical; child; male; female	Clinical – tertiary multidisciplinary feeding service	Child	10 months to 19 years	DSM-5 criteria	536	263	49.1 %	8
Goldberg et al., 2020	Cross-sectional	Canada	Clinical; child	Clinical – tertiary care paediatric and adolescent gynecology clinic	Child	8 to 18 years	3 part self-administered questionnaire package (1. demographics and anthropometric info, reason for referral, current and past medical history, medication use, self-reported psych diagnosis of anxiety and depression. 2. menstrual history and function. 3. EDY-Q)	190	7	3.7 %	6
Gonçalves et al., 2018	Questionnaire	Portugal	Nonclinical; child	None – primary schools	Child	5 to 10 years	ARFID questionnaire based on DSM5 criteria	330	51	15.5 %	7
Haqqi & Irfan, 2024	Cross-sectional	Pakistan	Nonclinical; adult	None	Adult	18 to 25 years	PARDI-AR-Q	660	10	2.8 %	6
Hay et al., 2017a	Cross-sectional	Australia	Nonclinical	None	Adult	15 years and above	Questions adapted from the Eating Disorder Examination as	5737	18	0.3 %	7

(*continued on next page*)

Table 2 (*continued*)

First author and year	Study design	Region	Included in which Subgroup Analyses	Study Setting	Adult/ Child/ Mixed	Sample Age Range	Diagnostic tool	N	N ARFID	Prevalence %	Quality Score /8
							part of a larger health survey				
Hilbert et al., 2021	Cross-sectional	Germany	Nonclinical; adult; male; female	None	Adult	18 to 94 years	EDY-Q, EDE-Q8	2424	64	0.8 %	7
Kaul et al., 2024	Longitudinal	USA	Clinical; child	Clinical – Texas Children's Hospital	Child	10 to 17 years	NIAS, PARDI-AR-Q, ARFID checklist	171	71	41.5 %	6
Koomar et al., 2021	Cross-sectional	USA	Nonclinical; adult; child	None	Adult & Child separate analyses	NA	Adaptation of NIAS	10,142	1930	19 %	8
Krom et al., 2019	Cross-sectional	Amsterdam	Clinical; child	Clinical - Diagnostic Centre for Feeding Problems in the Emma Children's Hospital/ Amsterdam UMC	Child	NA	DSM-5 Criteria	100	64	64 %	6
Nakai et al., 2017	Retrospective Chart Review	Japan	Clinical; female	Clinical – Kyoto University Hospital Eating Disorder Unit	Mixed	15 to 40 years	DSM5 Criteria	245	27	11.2 %	6
Nicely et al., 2014	Retrospective Chart Review	USA	Clinical; child	Clinical – Penn State Hershey Children's hospital, day programme for ED	Child	7 to 17 years	DSM5 Criteria	173	39	22.5 %	6
Nygren et al., 2021	Retrospective Chart Review	Sweden	Nonclinical; child	None	Child	0 to 6 years	DSM5 Criteria	46	13	28.26 %	6
Robelin et al., 2021	Single Centre Cross-Sectional (pilot study)	USA	Clinical; adult	Clinical – Inflammatory Bowel Disease Mayo Clinic	Adult	18 to 40 years	NIAS	98	10	10.2 %	6
Schöffel et al., 2021	Cross-sectional	Germany	Nonclinical; clinical	Clinical – University Hospital Leipzig General and Neuropediatric Clinic	Child	8 to18 years	EDY-Q	910	20	2.2 %	6
Williams et al., 2015	Retrospective Chart Review	USA	Clinical; child	Clinical – Multidisciplinary paediatric feeding programme	Child	4 months to 219 months	DSM5 Criteria	442	133	32 %	5
Yelencich et al., 2022	Cross-Sectional	USA	Clinical; adult	Clinical – Ambulatory Care at UCLA	Adult	NA	NIAS and medical records	161	28	17 %	6
Zickgraf et al., 2023	Single Centre Cross-Sectional	USA	Clinical	Clinical – Midwestern Gender Clinic	Mixed	12 to 23 years	NIAS	164	36	22 %	6

with four papers using more than one tool to screen for ARFID. The remaining papers used other measures, including ones developed by the authors. The number of participants included in each study ranged between 46 and 50,082. Age range was reported in 19 studies. The reported age ranged between 4 months and 94 years old. Mean age was reported in 14 studies, and produced a mean age for the overall sample of 23.21 years. Studies included male ($k = 7$); female ($k = 8$); clinical ($k = 14$); nonclinical ($k = 13$); adult ($k = 10$) and child ($k = 16$) samples. Only one paper included any self-identified non-binary participants, so it was not possible to consider non-binary gender identity further in the analyses.

Prevalence reported in the studies ranged from 0.8 % to 28 % in the non-clinical samples and 0.8 % to 64 % in the clinical samples. The full range of data used in the subgroup analyses can be found in Table 3 in the supplementary material.

3.2. Quality appraisal

Results for each study can be seen in Table 3. All studies met the criteria for using a random sample or a whole population sample and for using an unbiased sampling frame. The largest quality issue identified was that only seven studies provided adequately robust data. In particular, confidence intervals were rarely stated (only given in two papers - Eddy et al., 2015; Van Buuren et al., 2023). There were also issues in the size of the samples used. Using this scale, an adequately sized sample is defined as a minimum of 300 participants, and only 15 studies achieved this. There was also a lack of consistency in using standardised methods of diagnosis, with 20 studies achieving this.

The percent agreement between raters on the quality assessment score was 87.5 %, and consensus was reached after discussion. Cohen's kappa (κ) statistic showed substantial agreement between raters, $\kappa = 0.659$, $p < .005$. According to the agreement thresholds suggested by Landis & Koch (1977), these figures depict sufficient interrater

Table 3
Quality assessment of the included studies using Loney et al. (1998) Critical Appraisal of the Health Research Literature: Prevalence or Incidence of a Health problem.

Paper	Random Sample/ Whole Population	Unbiased sampling frame	Sample size > 300	Standard Measures Used for Diagnosis	Unbiased Assessors	Adequate response rate (70 %) and refusers described	CI stated, subgroup analysis where relevant	Study subjects described	Total
Atkins et al. (2023)	1	1	1	1	1	1	1	1	8
Van Buuren et al. (2023)	1	1	1	1	1	1	1	1	8
Burton-Murray et al. (2024)	1	1	0	1	1	1	0	1	6
Chen et al. (2019)	1	1	1	0	1	0	1	1	6
Chua et al. (2022)	1	1	1	0	1	0	0	1	5
D'Adamo et al. (2023)	1	1	1	1	1	1	0	1	7
Dinkler et al. (2022)	1	1	1	0	1	0	0	1	5
Dinkler et al. (2023)	1	1	1	1	1	1	0	1	7
Eddy et al. (2015)	1	1	1	1	1	1	1	1	8
Farag et al. (2021)	1	1	1	1	1	1	1	1	8
Goldberg et al. (2020)	1	1	0	1	0	1	0	1	5
Gonçalves et al. (2018)	1	1	1	0	1	1	0	1	6
Haqqi and Irfan (2024)	1	1	1	1	1	1	0	0	6
Hay et al. (2017)	1	1	1	0	0	0	1	1	5
Hilbert et al. (2021)	1	1	1	1	1	1	0	1	7
Kaul et al. (2024)	1	1	0	1	1	1	0	1	6
Koomar et al. (2021)	1	1	1	1	1	1	1	1	8
Krom et al. (2019)	1	1	0	1	1	1	0	1	6
Nakai et al. (2017)	1	1	0	1	1	1	0	1	6
Nicely et al. (2014)	1	1	0	1	1	1	0	1	6
Nygren et al. (2021)	1	1	0	1	1	1	0	1	6
Robelin et al. (2021)	1	1	0	1	1	1	0	1	6
Schoffel et al. . (2020)	1	1	0	1	1	1	0	1	6
Williams et al. (2015)	1	1	1	0	1	1	0	0	5
Yelencich et al. (2022)	1	1	0	1	1	1	0	1	6
Zickgraf et al. (2023)	1	1	0	1	1	1	0	1	6

agreement.

Results from the multiple regression, testing whether there was a relationship between quality scores for each paper and potentially related features of the studies - date of publication, level of reported prevalence, or setting (clinical or non-clinical) - showed no significance overall [$F(3,12) = 0.218$, $p = .88$, $R^2 = 0.029$]. The individual predictors were all non-significant - year of publication ($t = 0.433$, $p = .67$), setting ($t = 0.560$, $p = .58$) and reported prevalence ($t = -0.020$, $p = .98$).

For the GRADE assessment, the initial quality of both meta-analytic comparisons was set at 'high quality' as the included prevalence studies were largely based on observational cross-sectional, case-control or retrospective study designs. As both meta-analysis syntheses were based on the same set of studies, assessment of three of the five GRADE criteria were identical for both (study limitations, indirectness, and publication bias). The inconsistency and imprecision criteria were assessed separately for the aggregated effect estimate and heterogeneity in each meta-analysis. There were no concerns regarding indirectness of the evidence, partly due to the restriction of search dates to studies published after ARFID was added to the DSM-5 to ensure the population of interest was captured. Evidence quality for both syntheses were downgraded two levels to 'low' quality due to indicated lack of smaller studies with lower prevalence rates (publication bias), high levels of unexplained heterogeneity and considerable variability in individual study estimates of prevalence (inconsistency), and study limitations relating to use of unstandardised diagnostic measures and lack of adequately robust data in many studies. However, the use of the quality effects approach helped to mitigate impacts of study limitations and imprecision in studies, therefore the GRADE rating was uprated to 'moderate' for the quality effects meta-analysis.

3.3. Prevalence meta-analyses

3.3.1. Random effects model

Meta-analysis of the 26 papers ($n = 122,861$) using a random effects model identified a pooled prevalence of 11.14 % (95 % CI 8.16–14.5 %; GRADE rating: low; Fig. 2). There was significant evidence of heterogeneity (Q = 5950.0792, $p = .0$, $I^2 = 99.6$ %, 95 % CI 99.53–99.62 %), $tau^2 = 0.0645$).

3.3.2. Quality effects model

For the quality effects model, the included papers were weighted dependent upon their quality score. This meta-analysis identified a much lower pooled prevalence of 4.51 % (95 % CI 0.1–10.68 %; GRADE rating: moderate; Fig. 3). There was again significant evidence of heterogeneity (Q = 5950.079, p = .0, $I^2 = 99.6$ %, 95 % CI 99.53–99.62 %), Q index = 13.299).

Only one assessment of publication bias was performed as both meta-analyses used the same included study estimates and standard errors (only differed in the approach to weighting studies). Publication bias for both meta-analyses is presumed to be large after a visual inspection of the funnel plot (Fig. 4). The obvious asymmetry to the right indicates a bias towards publishing studies that report a higher prevalence statistic.

3.4. Subgroup analyses

Subgroup analyses of the random and quality effects meta-analyses indicated higher prevalence in clinical samples (18.61 % and 11.97 % respectively) in comparison to non-clinical samples (4.46 % and 2.84 %) (Tables 4 and 5). Using random effects analysis, prevalence rates were higher in child samples (participants between 4 months and 18 years of age) in comparison to adult (13.55 % and 8.8 %). In contrast, the inverse was found when using quality effects (4.73 % and 5.9 %). The pooled prevalence was higher for male samples than for females using both random effects (6.96 % and 4.93 %) and quality effects (3.18 % and 2.61 %).

Subgroup analyses were also completed for child female/child male samples and adult male/adult female samples. The small number of data sets limits the reliability of these analyses, but they offer some potentially important findings and directions for future research. Table 6 shows the subgroup prevalences using random effects analysis. It shows that male children and adults have a slightly higher prevalence than their female counterparts, and that prevalence is higher among children than adults. (See Table 7.)

In contrast, quality effects analysis shows that male children had a substantially higher prevalence than female children (50 % higher), but that there was a much smaller gender difference between the adult men and women. Again, overall the males had a higher prevalence rate than the females.

4. Discussion

The prevalence of ARFID is important to understand. Without a clear picture as to the burden of this condition, screening and services cannot accurately forecast what is required. Accurate health intelligence allows for effective treatment strategies and for the appropriate allocation of often scarce resources to deliver timely and targeted inventions for the populations most at risk. In this review, we have conducted a meta-analysis using samples from both clinical and nonclinical populations, using random and quality effects models. We included data from 26 papers involving 122,861 individuals.

Due to the highly heterogeneous nature of the papers that were included in this analysis, a common issue in meta-analysis, the subgroup analyses could be argued to provide potentially the most accurate and clinically relevant data. Our findings can be compared and contrasted with those reported by existing systematic reviews that summarise the prevalence of ARFID (Sanchez-Cerezo et al., 2023), where specialised eating disorder clinics have a prevalence range of 32–64 % (higher than shown here) and non-clinical settings have figures ranging from 0.3 to 15 % (closer to those shown here). Previous evidence shows that ARFID patients are more likely to be male in child samples, but the current meta-analyses confirm that pattern for adults, too.

Meta analysis using random effects found a pooled prevalence of 11.14 % across the 26 papers, with higher prevalence in children than in adults and marginally higher prevalence among males than among females. As expected, clinical settings also provided higher pooled prevalence estimates than nonclinical. The quality effects analysis provides a stark contrast - a pattern that has been seen in previous papers comparing the results yielded by the random and quality effects models (Liu et al., 2021). The pooled prevalence estimate here was 4.5 % - less than half of the prior approach. The difference was particularly noteworthy when considering prevalence by age, with apparent overestimation of prevalence among children when using the random-effects model. GRADE assessments indicated quality of evidence in the random effects meta-analysis was low, compared to moderate evidence

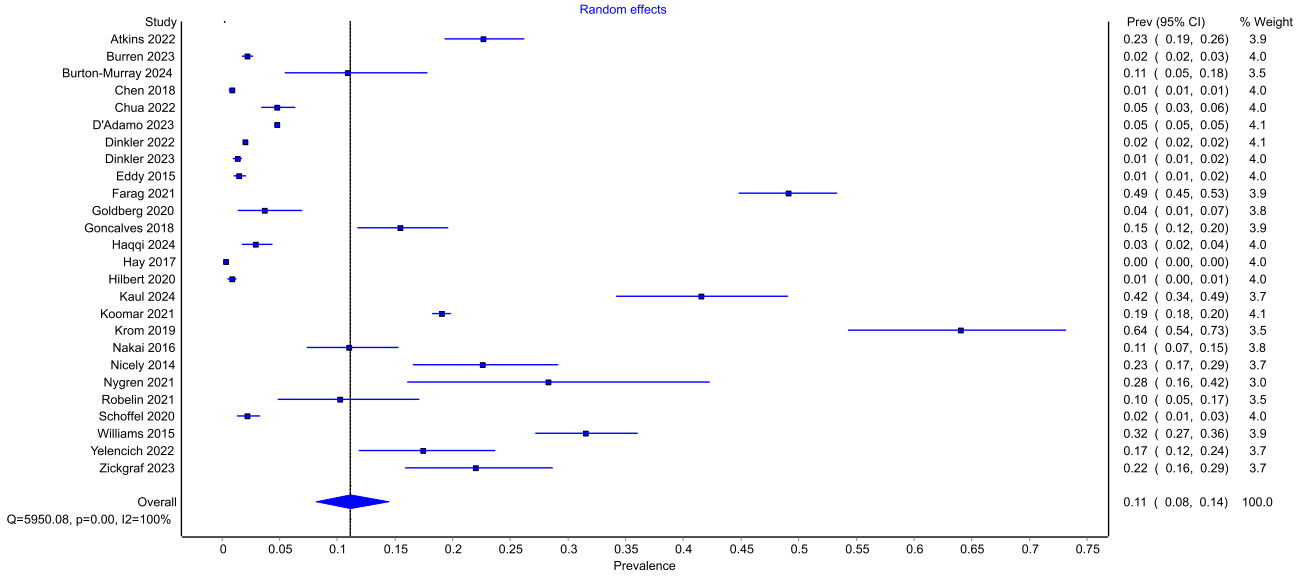


Fig. 2. Forest plot for ARFID prevalence, using a random effects meta-analysis across all papers.

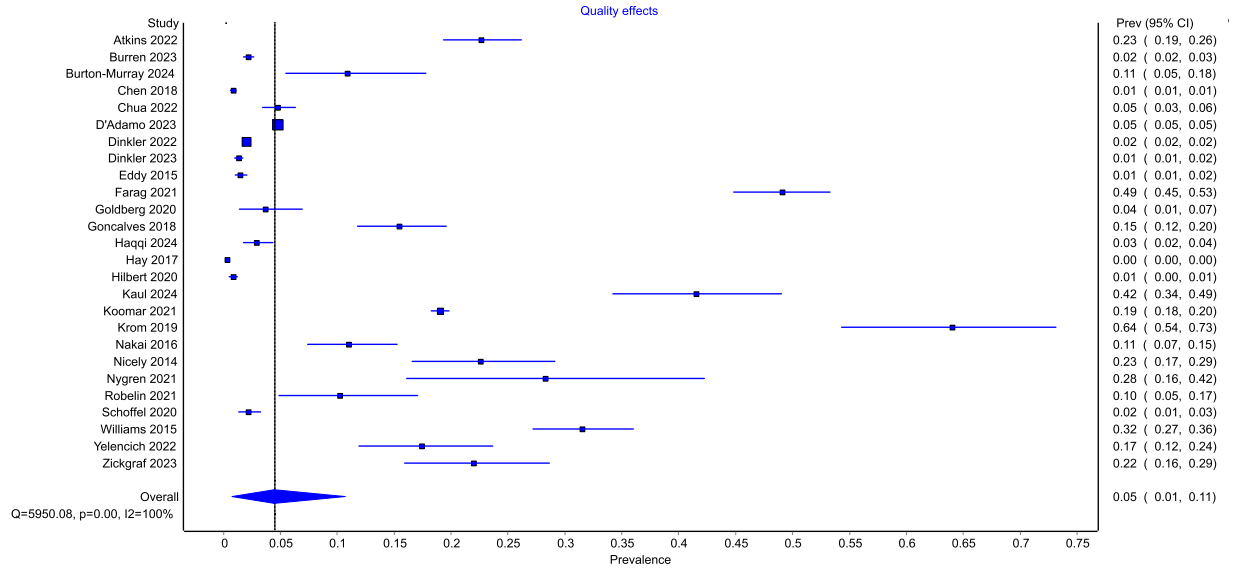


Fig. 3. Forest plot for ARFID prevalence quality effects meta-analysis across all papers.

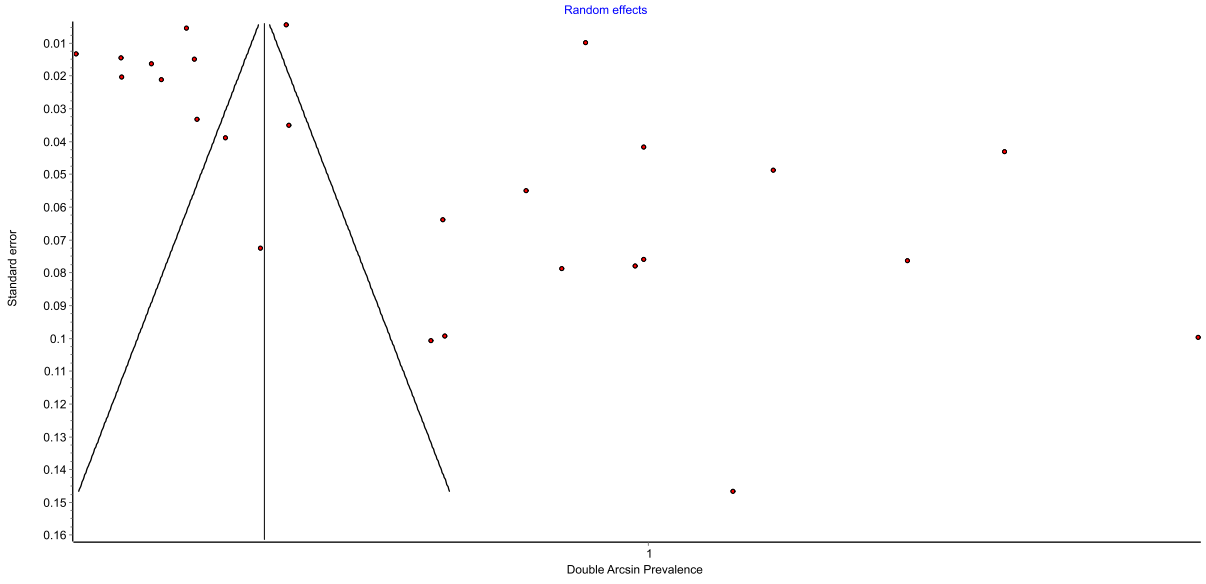


Fig. 4. Funnel plot for ARFID prevalence random effects and quality effects meta-analyses (both meta-analyses used the same included study estimates and standard errors).

quality for the quality effects analysis, indicating that the true effect is likely to be closer to quality effects reported prevalence estimate. Common research quality issues were around sample size, inadequate reporting of statistics, and limited use of objective diagnostic methods.

What might explain this very large difference in prevalence rates across the two forms of meta-analysis, and particularly the way in which prevalence rates for adults are similar to those for children, but only when using the quality effects analysis? It is particularly noteworthy that the overall strength of the current body of evidence was downgraded due to imprecision and inconsistency (however, some of these concerns were mitigated by the use of the quality effects analysis approach), and there was evidence of a considerable publication bias, with papers with a high prevalence rate being more likely to be published. However no peripheral factors (e.g., year of publication) were found that might account for this difference in prevalence rates between the two approaches to meta-analysis. Therefore, the most plausible explanation is that, until now, relatively weak studies (e.g., underpowered; limited data reporting; overly inclusive case identification) might have dominated our picture of how many cases there are in the population, potentially being over-represented in the literature if they report high prevalence levels that enhance their novelty and publishability. Given the stereotype that ARFID is a disorder that primarily affects children, this pattern might have led to an issue of confirmation bias – where weaker quality studies have been more readily accepted if they relate to younger populations. Consequently, when those quality issues are addressed, studies among children are 'corrected' to show

Table 4
Pooled prevalence for subgroups using random effects analysis.

Subgroup	Number of samples	*N*	Pooled prevalence	95 % CI	I^2
Population					
Clinical	14	5277	18.61 %	8.29–31.64 %	99.0 %
Non-Clinical	13	118,380	4.46 %	1.91–7.93 %	99.79 %
Age group					
Adult	10	59,950	8.8 %	4.71–13.96 %	99.2 %
Child	16	57,727	13.55 %	8.49–19.55 %	99.6 %
Gender					
Male	7	23,069	6.96 %	2.36–13.5 %	99.25 %
Female	8	23,186	4.93 %	2.87–7.49 %	96.77 %

Abbreviations: pp: pooled prevalence.

Table 5
Pooled prevalence for subgroups using quality effects analysis.

Subgroup	Number of samples	*n*	Pooled prevalence	95 % CI	I^2
Population					
Clinical	14	5277	11.97 %	0–31.81 %	99.0 %
Non-Clinical	13	118,380	2.84 %	0.5–8.35 %	99.79 %
Age group					
Adult	10	59,950	5.9 %	0–16.43 %	99.2 %
Child	16	57,727	4.73 %	0.24–12.95 %	99.6 %
Gender					
Male	7	23,069	3.18 %	0–11.68 %	99.25 %
Female	8	23,186	2.61 %	0.67–5.6 %	96.77 %

Abbreviations: pp: pooled prevalence.

Table 6
Subgroup analysis using random effects analysis.

Subgroup	Number of Samples	*n*	Pooled Prevalence	95 % CI	I^2
Female Child	5	21,318	6.34 %	3.25–10.32 %	97.908
Male Child	3	21,493	7.73 %	1.1–18.48 %	99.591
Female Adult	3	1868	2.12 %	0.08–6.03 %	91.368
Male Adult	3	1576	3.35 %	0.14–9.36 %	93.296

Table 7
Subgroup analysis using quality effects analysis.

Subgroup	Number of Samples	*n*	Pooled Prevalence	95 % CI	I^2
Female Child	5	21,318	3 %	0.5–7.10 %	97.908
Male Child	3	21,493	4.47 %	0–14.1 %	99.591
Female Adult	3	1868	1.54 %	0–5.16 %	91.368
Male Adult	3	1576	1.87 %	0–7.17 %	93.296

more accurate, lower prevalence rates than commonly reported. Therefore, there is a greater apparent overestimation of prevalence among children than among adults when using the random effects model, and a greater reduction in prevalence among children than adults when using quality effects analysis. To conclude, ongoing surveillance of the quality of the studies conducted is essential, especially where there is a risk of a potential bias in the way that researchers expect prevalence to manifest.

Of course, even though the quality of research was a key variable explored in this systematic review and meta-analyses, this summary has a baseline limitation due to the limited quality and quantity of the studies that have been published. For instance, even though subthreshold cases were excluded from the review, many of the studies that were included were limited by using the criterion of a 'probable' diagnosis of ARFID. This lack of diagnostic certainty could have led to an over-inclusive set of prevalence calculations. It is also worth noting that some of the subgroup analyses were relatively small, resulting in a lower statistical power, and the results should therefore be treated with caution (Linardon et al., 2016). A further consideration is the lack of data pertaining to individuals who identify as nonbinary, with only one study currently available for this review. These are limitations that should be considered by researchers who seek to add to the prevalence literature in ARFID and beyond.

It is also important to consider any potential impact of the way in which this study was conducted. It is possible that relevant studies were missed because of the focus on English language papers, which can result in overinflated effects (Egger et al., 1997) and limited cross-cultural generalisability and utility of these findings. A further limitation is that the screening of the abstracts and the full texts was completed by one reviewer, so full independent double screening of the literature or data extraction was not conducted. The decision to focus on the time period 2013 onwards reflects the American Psychiatric Association (2013) categorisation of ARFID. Earlier papers could have been considered, based on earlier constructs (e.g., selective eating), but it is likely that this would have introduced unaccountable variability into the overall prevalence scores. That variability would need to be explained by future reviews comparing prevalence between the pre-2013 and post-2013 constructs, if that is possible within the literature. Similarly, the lack of inclusion of the 'grey' literature might mean that some unpublished findings were missed. Given the likely bias shown here towards the publication of studies with higher prevalence rates, that decision not to include the grey literature might mean that some lower prevalence papers were missed, and that the meta-analyses here might represent a slight overestimate of prevalence. Therefore, future meta-analyses in this field should consider the wider global literature and unpublished research.

Notwithstanding these caveats, these meta-analyses provide new insights into the prevalence of ARFID, and the importance of conducting high-quality research in this field. ARFID has severe potential psychological and physical impacts on the individual and impacts on the quality of life of patients and family members alike (Hay et al., 2017; Nicely et al., 2014). Enhancing awareness and identification of ARFID sufferers is needed to influence planning and resource allocation, better training in the identification of the disorder, and the development of evidence-based interventions. Future research should ensure that methodological limitations are addressed to enhance the research quality, to help close the gap in prevalence's shown here between the random-effects model and the quality effects model. The key consideration from the quality analysis is the very limited proportion of these studies (only three out of 26) that provide a definite diagnosis of ARFID, though discussing the prevalence of the condition. This low number of diagnosed cases is a limitation of the value of this research from a clinical perspective. This quality issue should be supported through the use of standardised diagnostic measures with researchers ensuring that the measures used are validated for the diagnosis of the condition in question. For ARFID there are currently five validated tools: Eating Disorders Examination - ARFID module (Schmidt et al., 2019); Pica ARFID and Rumination Disorder Interview (PARDI) (Bryant-Waugh et al., 2019); Nine Item ARFID Screen (NIAS) (Zickgraf & Ellis, 2018); PARDI-AR-Q (PARDI ARFID Questionnaire) (Bryant-Waugh et al., 2022), and the Eating Disorders in Youth Questionnaire (EDY-Q) (Kurz

et al., 2015). Future research should clearly state the statistical procedures and results and consider the role of quality in determining outcomes. Whilst many of the included studies had large numbers of participants, it is important to note that the lack of specificity in the samples means that there is a risk that existing studies limit the specificity and dilute the accuracy of the resulting prevalence estimates. Large scale studies that focus on specific sample populations (e.g., the adult male population) could enhance the accuracy and utility of the resulting prevalence figures.

The Loney et al. (1998) scale lacks a clear cut-off score as to what is to be considered a low, medium or high quality study. In a recent systematic review assessing the quality assessment tools available specifically for use on prevalence studies, the Joanna Briggs Institute Prevalence Critical Appraisal Tool is considered to have high methodological rigor and to address key considerations when making quality assessments on prevalence studies. Future researchers should consider this as the most appropriate tool (Migliavaca, Stein, Colpani, Munn, & Falavigna, 2020; Munn et al., 2014; Munn et al., 2015).

While this is the first meta-analysis that allows for up-to-date prevalence estimates across adult and child populations in both clinical and nonclinical settings, it has also illustrated the importance of considering the influence of the quality of the research used to generate those estimates. The use of quality effects models should be considered beyond ARFID, in order to provide appropriate caution in the presentation of prevalence figures for different disorders.

CRediT authorship contribution statement

Rebecca Nicholls-Clow: Conceptualization, Formal analysis, Data curation, Methodology, Project administration, Writing – original draft. **Melanie Simmonds-Buckley:** Formal analysis, Methodology, Supervision, Writing – review & editing. **Glenn Waller:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no conflict of interest arising from this paper. The work has not been supported by any funding source.

Data availability

No data was used for the research described in the article.

Acknowledgements

The authors would like to acknowledge Heather Duggan for her careful review of the data extracted from and the quality scores assigned to each of the final papers included in the review.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2024.102502.

References

Alberts, Z., Fewtrell, M., Nicholls, D. E., Biassoni, L., Easty, M., & Hudson, L. D. (2020). Bone mineral density in anorexia nervosa versus avoidant restrictive food intake disorder. *Bone, 134*, Article 115307. https://doi.org/10.1016/j.bone.2020.115307

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.) Arlington, VA: Author.

Aulinas, A., Marengi, D. A., Galbiati, F., Asanza, E., Slattery, M., Mancuso, C. J., & Lawson, E. A. (2020). Medical comorbidities and endocrine dysfunction in low-weight females with avoidant/restrictive food intake disorder compared to anorexia nervosa and healthy controls. *International Journal of Eating Disorders, 53*(4), 631–636. https://doi.org/10.1002/eat.24218

Bailey, K. R. (1987). Inter-study differences: How should they influence the interpretation and analysis of results? *Statistics in Medicine, 6*(3), 351–360. https://doi.org/10.1002/sim.4780060327

Barendregt, J. J., & Doi, S. A. (2016). The Meta XL user guide version 5.3. [online]. Retrieved from http://www.epigear.com/index_files/MetaXL%20User%20Guide.pdf.

Barendregt, J. J., Doi, S. A., Lee, Y. Y., Norman, R. E., & Vos, T. (2013). Meta-analysis of prevalence. *Journal of Epidemiology and Community Health, 67*(11), 974–978. https://doi.org/10.1136/jech-2013-203104

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*, Article 112961. https://doi.org/10.1016/j.psychres.2020.112961

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 10–22. *International Journal of Eating Disorders, 52*(4), 378–387.

Bryant-Waugh, R., Stern, C. M., Dreier, M. J., Micali, N., Cooke, L. J., Kuhnle, M. C., … 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.

Chiarello, F., Marini, E., Ballerini, A., & Ricca, V. (2018). Optic neuropathy due to nutritional deficiency in a male adolescent with avoidant/restrictive food intake disorder: A case report. *Eating and Weight Disorders, 23*, 533–535. https://doi.org/10.1007/s40519-017-0409-6

Cuijpers, P., Griffin, J. W., & Furukawa, T. A. (2021). The lack of statistical power of subgroup analyses in meta-analyses: A cautionary note. *Epidemiology and Psychiatric Sciences, 30*, Article e78.

Doi, S. A., & Thalib, L. (2008). A quality-effects model for meta-analysis. *Epidemiology, 94-100*. https://doi.org/10.1097/EDE.0b013e31815c24e7

Dovey, T. M., Staples, P. A., Gibson, E. L., & Halford, J. C. (2008). Food neophobia and 'picky/fussy'eating in children: A review. *Appetite, 50*(2–3), 181–193. https://doi.org/10.1016/j.appet.2007.09.009

Dumont, E., Jansen, A., Kroes, D., de Haan, E., & Mulkens, S. (2019). A new cognitive behavior therapy for adolescents with avoidant/restrictive food intake disorder in a day treatment setting: A clinical case series. *International Journal of Eating Disorders, 52*(4), 447–458. https://doi.org/10.1002/eat.23053

Eddy, K. T., Thomas, J. J., Hastings, E., Edkins, K., Lamont, E., Nevins, C. M., & Becker, A. E. (2015). Prevalence of DSM-5 avoidant/restrictive food intake disorder in a pediatric gastroenterology healthcare network. *International Journal of Eating Disorders, 48*(5), 464–470. https://doi.org/10.1002/eat.22350

Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C., & Antes, G. (1997). Language bias in randomised controlled trials published in English and German. *The Lancet, 350*(9074), 326–329. https://doi.org/10.1016/S0140-6736(97)02419-7

Granholm, A., Alhazzani, W., & Møller, M. H. (2019). Use of the GRADE approach in systematic reviews and guidelines. *British Journal of Anaesthesia, 123*(5), 554–559. https://doi.org/10.1016/j.bja.2019.08.015

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*, 21. https://doi.org/10.1186/s40337-017-0149-z

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560.

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

Kelley, G. A., & Kelley, K. S. (2020). Yoga, health-related quality of life and mental well-being: A re-analysis of a meta-analysis using the quality effects model. *The Journals of Gerontology: Series A, 75*(9), 1732–1736. https://doi.org/10.1093/gerona/glz284

Kennedy, H. L., Hitchman, L. M., Pettie, M. A., Bulik, C. M., & Jordan, J. (2023). Avoidant/restrictive food intake disorder (ARFID) in New Zealand and Australia: A scoping review. *Journal of Eating Disorders, 11*(1), 196. https://doi.org/10.1186/s40337-023-00922-9

Kurz, S., Van Dyck, Z., Dremmel, D., Munsch, S., & Hilbert, A. (2015). Early-onset restrictive eating disturbances in primary school boys and girls. *European Child & Adolescent Psychiatry, 24*, 779–785.

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics, 33*(2), 363–374.

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

Liu, C., Pan, W., Li, L., Li, B., Ren, Y., & Ma, X. (2021). Prevalence of depression, anxiety, and insomnia symptoms among patients with COVID-19: A meta-analysis of quality effects model. *Journal of Psychosomatic Research, 147*, Article 110516. https://doi.org/10.1016/j.jpsychores.2021.110516

Lock, J., Robinson, A., Sadeh-Sharvit, S., Rosania, K., Osipov, L., Kirz, N., Derenne, J., & Utzinger, L. (2019). Applying family-based treatment (FBT) to three clinical presentations of avoidant/restrictive food intake disorder: Similarities and differences from FBT for anorexia nervosa. *International Journal of Eating Disorders, 52*(4), 439–446. https://doi.org/10.1002/eat.22994

Loney, P. L., Chambers, L. W., Bennett, K. J., Roberts, J. G., & Stratford, P. W. (1998). Critical appraisal of the health research literature: Prevalence or incidence of a health problem. *Chronic Diseases in Canada, 19*(4), 170–176.

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: Challenges and benefits. *Research Synthesis Methods, 5*(3), 221–234. https://doi.org/10.1002/jrsm.1106

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica, 22*(3), 276–282.

Micali, N., & Cooper-Vince, C. E. (2020). Aetiology and epidemiology. In R. Bryant-Waugh, & C. Higgins (Eds.), *Avoidant restrictive food intake disorder in childhood and adolescence: A clinical guide*. Routledge.

Migliavaca, C. B., Stein, C., Colpani, V., Barker, T. H., Munn, Z., Falavigna, M., & Prevalence Estimates Reviews–Systematic Review Methodology Group (PERSyst). (2020). How are systematic reviews of prevalence conducted? A methodological study. *BMC Medical Research Methodology, 20*, 1–9. https://doi.org/10.1186/s12874-020-00975-3

Migliavaca, C. B., Stein, C., Colpani, V., Munn, Z., & Falavigna, M. (2020). Quality assessment of prevalence studies: A systematic review. *Journal of Clinical Epidemiology, 127*, 59–68.

Munn, Z., Moola, S., Lisy, K., Riitano, D., & Tufanaru, C. (2015). Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. *JBI Evidence Implementation, 13*(3), 147–153.

Munn, Z., Moola, S., Riitano, D., & Lisy, K. (2014). The development of a critical appraisal tool for use in systematic reviews addressing questions of prevalence. *International Journal of Health Policy and Management, 3*(3), 123.

Nicely, T. A., Lane-Loney, S., Masciulli, E., Hollenbeak, C. S., & Ornstein, R. M. (2014). Prevalence and characteristics of avoidant/restrictive food intake disorder in a cohort of young patients in day treatment for eating disorders. *Journal of Eating Disorders, 2*(1), 21. https://doi.org/10.1186/s40337-014-0021-3

Norris, M. L., Robinson, A., Obeid, N., Harrison, M., Spettigue, W., & Henderson, K. (2014). Exploring avoidant/restrictive food intake disorder in eating disordered patients: A descriptive study. *International Journal of Eating Disorders, 47*(5), 495–499. https://doi.org/10.1002/eat.22217

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., & Mulrow, C. D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal, 372*, Article n71. https://doi.org/10.1136/bmj.n71

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

Schmidt, R., Kirsten, T., Hiemisch, A., Kiess, W., & Hilbert, A. (2019). Interview-based assessment of avoidant/restrictive food intake disorder (ARFID): A pilot study evaluating an ARFID module for the eating disorder examination. *International Journal of Eating Disorders, 52*(4), 388–397. https://doi.org/10.1002/eat.23063

Schmidt, U., Adan, R., Böhm, I., Campbell, I. C., Dingemans, A., Ehrlich, S., & Zipfel, S. (2016). Eating disorders: The big issue. *The Lancet Psychiatry, 3*(4), 313–315. https://doi.org/10.1016/S2215-0366(16)00081-X

Simmonds-Buckley, M., Osivwemu, E. O., Kellett, S., & Taylor, C. (2022). The acceptability of cognitive analytic therapy (CAT): Meta-analysis and benchmarking of treatment refusal and treatment dropout rates. *Clinical Psychology Review, 96*, Article 102187.

Streiner, D. L., Patten, S. B., Anthony, J. C., & Cairney, J. (2009). Has 'lifetime prevalence' reached the end of its life? An examination of the concept. *International Journal of Methods in Psychiatric Research, 18*(4), 221–228. https://doi.org/10.1002/mpr.296

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. *International Journal of Eating Disorders, 53*(10), 1636–1646. https://doi.org/10.1002/eat.23355

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 aetiology and treatment. *Current Psychiatry Reports, 19*, 1–9. https://doi.org/10.1007/s11920-017-0795-5

Van Buuren, L., Fleming, C. A. K., Hay, P., Bussey, K., Trompeter, N., Lonergan, A., & Mitchison, D. (2023). The prevalence and burden of avoidant/restrictive food intake disorder (ARFID) in a general adolescent population. *Journal of Eating Disorders, 11*(1), 104. https://doi.org/10.1186/s40337-023-00831-x

Zhao, X., Feng, G. C., Ao, S. H., & Liu, P. L. (2022). Interrater reliability estimators tested against true interrater reliabilities. *BMC Medical Research Methodology, 22*(1), 232. https://doi.org/10.1186/s12874-022-01707-5

Zickgraf, H. F., & Ellis, J. M. (2018). Initial validation of the nine item avoidant/restrictive food intake disorder screen (NIAS): A measure of three restrictive eating patterns. *Appetite, 123*, 32–42. https://doi.org/10.1016/j.appet.2017.11.111

Final Set of Papers used in the Systematic Review

Atkins, M., Zar-Kessler, C., Madva, E. N., Staller, K., Eddy, K. T., Thomas, J. J., Kuo, B., & Burton Murray, H. (2023). History of trying exclusion diets and association with avoidant/restrictive food intake disorder in neurogastroenterology patients: A retrospective chart review. *Neurogastroenterology and Motility, 35*(3), Article e14513. https://doi.org/10.1111/nmo.14513

Burton-Murray, H., Kiser, K., Gurung, J., Williams, K., Thomas, J. J., & Khalili, H. (2024). Avoidant/restrictive food intake disorder symptoms are not as frequent as other eating disorder symptoms when ulcerative colitis is in remission. *Journal of Crohn's and Colitis, jjae052*. https://doi.org/10.1093/ecco-jcc/jjae052

Chen, Y. L., Chen, W. J., Lin, K. C., Shen, L. J., & Gau, S. S. (2019). Prevalence of DSM-5 mental disorders in a nationally representative sample of children in Taiwan: methodology and main findings. *Epidemiology and Psychiatric Sciences, 29*, Article e15. https://doi.org/10.1017/S2045796018000793

Chua, S. N., Fitzsimmons-Craft, E. E., Austin, S. B., Wilfley, D. E., & Taylor, C. B. (2022). Estimated prevalence of eating disorders in Malaysia based on a diagnostic screen. *International Journal of Eating Disorders, 55*(6), 763–775. https://doi.org/10.1002/eat.23711

D'Adamo, L., Smolar, L., Balantekin, K. N., Taylor, C. B., Wilfley, D. E., & Fitzsimmons-Craft, E. E. (2023). Prevalence, characteristics, and correlates of probable avoidant/restrictive food intake disorder among adult respondents to the National Eating Disorders Association online screen: a cross-sectional study. *Journal of Eating Disorders, 11*(1), 214. https://doi.org/10.21203/rs.3.rs-3,007,049/v1

Dinkler, L., Wronski, M. L., Lichtenstein, P., Lundström, S., Larsson, H., Micali, N., & Bulik, C. M. (2023). Etiology of the broad avoidant restrictive food intake disorder phenotype in Swedish twins aged 6 to 12 years. *JAMA Psychiatry, 80*(3), 260–269. https://doi.org/10.1001/jamapsychiatry.2022.4612

Dinkler, L., Yasumitsu-Lovell, K., Eitoku, M., Fujieda, M., Suganuma, N., Hatakenaka, Y., & Gillberg, C. (2022). Development of a parent-reported screening tool for avoidant/restrictive food intake disorder (ARFID): Initial validation and prevalence in 4–7-year-old Japanese children. *Appetite, 168*(105), 735. https://doi.org/10.1016/j.appet.2021.105735

Farag, F., Sims, A., Strudwick, K., Carrasco, J., Waters, A., Ford, V., & Kelly, V. B. (2021). Avoidant/restrictive food intake disorder and autism spectrum disorder: clinical implications for assessment and management. *Developmental Medicine and Child Neurology, 64*(2), 176–182. https://doi.org/10.1111/dmcn.14977

Goldberg, H. R., Katzman, D. K., Allen, L., Martin, S., Sheehan, C., Kaiserman, J., & Kives, S (2020). The prevalence of children and adolescents at risk for avoidant restrictive food intake disorder in a paediatric and adolescent gynecology clinic. Journal of Paediatric and Adolescent Gynecology, 33(5), 466–469. I:doi:https://doi.org/10.1016/j.jpag.2020.06.004.

Gonçalves, S., Vieira, A. I., Machado, B. C., Costa, R., Pinheiro, J., & Conceiçao, E. (2018). Avoidant/restrictive food intake disorder symptoms in children: Associations with child and family variables. *Children's Health Care, 48*(3), 301–313. https://doi.org/10.1080/02739615.2018.1532796

Haqqi, S. A., & Irfan, S. (2024). Relationship of self-reported pica and avoidant restrictive food intake disorder symptomology with dimensions of impulsivity, perceived stress among Pakistani University students. *Journal of Eating Disorders, 12*(1), 10. https://doi.org/10.1186/s40337-023-00956-z

Hilbert, A., Zenger, M., Eichler, J., & Brähler, E. (2021). Psychometric evaluation of the Eating Disorders in Youth-Questionnaire when used in adults: Prevalence estimates for symptoms of avoidant/restrictive food intake disorder and population norms. *International Journal of Eating Disorders, 54*(3), 399–408. https://doi.org/10.1002/eat.23424

Kaul, I., Burton-Murray, H., Musaad, S., Mirabile, Y., Czyzewski, D., van Tilburg, M. A., … Shulman, R. J. (2024). Avoidant/restrictive food intake disorder prevalence is high in children with gastroparesis and functional dyspepsia. *Neurogastroenterology and Motility, 36*(5), Article e14777. https://doi.org/10.1111/nmo.14777

Koomar, T., Thomas, T. R., Pottschmidt, N. R., Lutter, M., & Michaelson, J. J. (2021). Estimating the prevalence and genetic risk mechanisms of ARFID in a large autism cohort. *Frontiers in Psychiatry, 12*(668), 297. https://doi.org/10.3389/fpsyt.2021.668297

Krom, H., van der Sluijs Veer, L., van Zundert, S., Otten, M. A., Benninga, M., Haverman, L., & Kindermann, A. (2019). Health related quality of life of infants and children with avoidant restrictive food intake disorder. *International Journal of Eating Disorders, 52*(4), 410–418. https://doi.org/10.1002/eat.23037

Nakai, Y., Nin, K., Noma, S. I., Hamagaki, S., Takagi, R., Teramukai, S., & Wonderlich, S. A. (2017). Clinical presentation and outcome of avoidant/restrictive food intake disorder in a Japanese sample. *Eating Behaviors, 24*, 49–53. https://doi.org/10.1016/j.eatbeh.2016.12.004

Nygren, G., Linnsand, P., Hermansson, J., Dinkler, L., Johansson, M., & Gillberg, C. (2021). Feeding problems including avoidant restrictive food intake disorder in young children with autism spectrum disorder in a multiethnic population. *Frontiers in Pediatrics, 9*, 1466. https://doi.org/10.3389/fped.2021.780680

Robelin, K., Senada, P., Ghoz, H., Sim, L., Lebow, J., Picco, M., & Werlang, M. (2021). Prevalence and clinician recognition of avoidant/restrictive food intake disorder in patients with inflammatory bowel disease. *Gastroenterology & Hepatology, 17*(11), 510–514.

Schöffel, H., Hiemisch, A., Kiess, W., Hilbert, A., & Schmidt, R. (2021). Characteristics of avoidant/restrictive food intake disorder in a general paediatric inpatient sample. *European Eating Disorders Review, 29*(1), 60–73. https://doi.org/10.1002/erv.2799

Williams, K. E., Hendy, H. M., Field, D. G., Belousov, Y., Riegel, K., & Harclerode, W. (2015). Implications of avoidant/restrictive food intake disorder (ARFID) on children with feeding problems. *Children's Health Care, 44*(4), 307–321. https://doi.org/10.1080/02739615.2014.921789

Yelencich, E., Truong, E., Widaman, A. M., Pignotti, G., Yang, L., Jeon, Y., … Limketkai, B. N. (2022). Avoidant restrictive food intake disorder prevalent among patients with inflammatory bowel disease. *Clinical Gastroenterology and Hepatology, 20*(6), 1282–1289. https://doi.org/10.1016/j.cgh.2021.08.009

Zickgraf, H. F., Garwood, S. K., Lewis, C. B., Giedinghagen, A. M., Reed, J. L., & Linsenmeyer, W. R. (2023). Validation of the nine-item avoidant/restrictive food intake disorder screen among transgender and nonbinary youth and young adults. *Transgender Health, 8*(2), 159–167. https://doi.org/10.1089/trgh.2021.0021