An item-level systematic review of the presentation of ADHD in females

Tamara Williams[1,2], Louise Horstmann[1,2], Laiba Kayani,[3] Annabelle Xiao Hui Lim[3], Abigail Russell[4], Tamsin Ford[5], Ann John[6], Kapil Sayal[7,8], Anita Thapar[1,2], Kate Langley[9], Joanna Martin[1,2]

[1]Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK

[2]Wolfson Centre for Young People's Mental Health, Cardiff University, Cardiff, UK

[3]School of Medicine, Cardiff University, Cardiff, UK

[4]University of Exeter Medical School, University of Exeter, Exeter, UK.

[5]Department of Psychiatry, University of Cambridge, Cambridge, UK

[6]Institute of Suicide Prevention and Mental Health, Swansea University, Swansea, UK

[7]Centre for ADHD and Neurodevelopmental Disorders Across the Lifespan, Institute of Mental Health, Nottingham, UK

[9]Unit of Mental Health and Clinical Neurosciences, School of Medicine, University of Nottingham, UK

[9]School of Psychology, Cardiff University, Cardiff, UK

Corresponding author:

Tamara Williams, Centre for Neuropsychiatric Genetics and Genomics, Hadyn Ellis Building, Cardiff University, Maindy Road, Cardiff CF24 4HQ, UK.

Email: williamstl6@cardiff.ac.uk

Contents

Abstract

Previous studies examining sex differences in attention deficit hyperactivity disorder (ADHD) have primarily examined total or subscale scores. This systematic review aimed to examine which symptoms contribute to the female presentation of ADHD at an item-level.

Six research literature databases were searched for studies comparing ADHD symptoms and their impact at an item-level in females with ADHD compared with: 1) males with ADHD and 2) females without ADHD.

Thirteen studies were included. In childhood, females were more likely to display the symptoms ‘fails to sustain attention in tasks’ and ‘often easily distracted’, whereas males were more likely to display the symptoms ‘often fidgets’, ‘difficulty remaining seated when required’, ‘runs/climbs in situations when inappropriate’, ‘always on the go’, ‘often noisy in playing’, ‘difficulty waiting turn’, ‘often blurts out answers’ and ‘often interrupts others’. In adulthood, females were more likely to endorse the symptoms ‘easily distracted’, ‘difficulty organising tasks’, ‘blurts out answers’ and ‘talks excessively’, as well as to report mind wandering and adverse home impacts.

Females with ADHD differ in their symptom profile to males with ADHD, highlighting the need for future research to identify and characterise symptoms typical of female ADHD.

Keywords: ADHD, Attention deficit hyperactivity disorder, Sex Differences, Symptoms, Item-level, Missed Diagnosis, Impact

1. Introduction

Attention deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental condition, characterised by inattention and hyperactivity-impulsivity, that has an estimated global prevalence of 5.3% (Polanczyk et al., 2014), ranging up to 8% in children and adolescents (Ayano et al., 2023). It is a highly impairing condition associated with a range of adverse outcomes (French et al., 2024), including peer rejection, criminality, poor educational and employment outcomes (Dalsgaard et al., 2013; Gershon & Gershon, 2002; Nijmeijer et al., 2008; Young et al., 2020), mental health and physical health conditions and premature mortality (Cortese et al., 2016; Galera et al., 2023; Schiavone et al., 2022; Young et al., 2020). Timely identification and treatment of ADHD is important as treatment can reduce symptoms and potentially improve outcomes (Daley et al., 2019; Dalsgaard et al., 2013; Shaw et al., 2012).

Sex differences in the prevalence of ADHD are well reported in the literature (Martin, 2024; Young et al., 2020), with childhood ADHD diagnosed 7-8 times more frequently in males than females, despite a population sex ratio of 3-4:1 (Faraone et al., 2015; Willcutt, 2012). This sex difference was previously assumed to be due to a genuine prominent male excess in ADHD risk (Arnett et al., 2015). However, recent research suggests that this may not be the only explanation and that at least part of the difference is due to under-recognition of ADHD in females (Martin, 2024; Young et al., 2020). In addition, females often receive an ADHD diagnosis later than males (Grevet et al., 2006; Wimberley et al., 2020), with the mean age at first diagnosis being around 10.9 years in males and 12.6 years in females (Martin et al., 2024).

ADHD is reportedly under-recognised and under-diagnosed, particularly in females (Quinn & Madhoo, 2014; Young et al., 2021) for several possible reasons (Martin, 2024; Young et al., 2020). ADHD symptom profiles may differ by sex, with females reportedly displaying more inattentive symptoms and fewer hyperactive and impulsive symptoms than males (Gershon & Gershon, 2002; Quinn & Madhoo, 2014). Further, the field trials for establishing the Diagnostic and Statistical Manual of Mental disorders (DSM) version IV criteria (American Psychiatric Association, 2000) for ADHD were developed and validated using a majority male sample (79% males) (Lahey et al., 1994). As such, the diagnostic criteria may be biased towards the male manifestation of ADHD, with males more likely than females to meet the diagnostic criteria (Willcutt, 2012). Additionally, co-occurring anxiety and emotional difficulties are more common in females and tend to be less overt or disruptive than associated conduct difficulties that are more common in males (Quinn & Madhoo, 2014). This may also contribute to females being more likely to be overlooked for an ADHD diagnosis (Quinn & Wigal, 2004; Quinn & Nadeau, 2002) and instead receive a primary diagnosis of depression or anxiety (Martin et al., 2024; Powell et al., 2021), delaying diagnosis of ADHD.

Several literature reviews have examined sex differences in ADHD symptom profiles based on total scores, hyperactive-impulsive and inattention sub-scales and impact scores. These reviews (Gershon & Gershon, 2002; Quinn & Madhoo, 2014) and meta-analyses (Loyer Carbonneau et al., 2021), using both clinical and community populations, have suggested that females with ADHD may display a different symptom profile than

males with ADHD. Gershon & Gershon (2002) reported that females with ADHD were rated by parents and teachers as having fewer symptoms of hyperactivity, impulsivity, inattention, and behavioural problems, but more emotional problems than males with ADHD. Quinn & Madhoo (2014), in a selective review of the literature, suggested that females with ADHD predominately display inattentive symptoms, whereas males with ADHD display predominately hyperactive and impulsive symptoms. Additionally, females with ADHD demonstrate more difficulty with peer relationships than males with ADHD, and more difficulty with social behaviours, peer functioning and interpersonal relationships, including having fewer friends and less stable relationships, than females without ADHD (Quinn & Madhoo, 2014). Further, Loyer Carbonneau et al. (2021) conducted a meta-analysis of 54 studies and concluded that in children and adolescents, males with ADHD expressed significantly more hyperactivity symptoms than females with ADHD. There were no differences in the expression of inattentive or impulsive symptoms. Further, when results were analysed separately by rater, teacher-reports identified that hyperactive-impulsive symptoms were higher in males, whereas parent-reported symptoms were similar in males and females with ADHD (Loyer Carbonneau et al., 2021).

Overall, existing research findings imply that females with ADHD may have different symptom profiles to males with ADHD, including being less likely to manifest symptoms that are overt and impactful on others. However, these reviews and meta-analyses only included comparisons of total or subscale scores. Understanding sex differences in ADHD symptoms at an item-level may better help us to understand in more detail the female manifestation of ADHD, which is needed to improve recognition, identification and refinement of the phenotype of ADHD in females.

The overarching aim of this systematic review was to examine if there are specific symptoms that characterise the manifestation of ADHD in females compared to: 1) males with ADHD and 2) females without ADHD. The specific aims were to determine whether there are: (1) sex differences in individual ADHD symptom items as defined by DSM-5 or impact related to ADHD, (2) sex differences in symptoms of co-occurring mental health or neurodevelopmental conditions, and (3) specific co-occurring mental health or neurodevelopmental symptoms in females with ADHD compared to females without ADHD.

2. Methods

The protocol for this systematic review has been registered in the International Prospective Register of Systematic Reviews (PROSPERO) database (www.crd.york.ac.uk/PROSPERO, CRD42023395625). It was developed according to the Preferred Reporting Items for Systematic Review and Meta-Analyses Protocol (PRISMA-P) guidelines (Moher et al., 2009). The full PRISMA-P checklist is included in the Supplementary Materials (**Table S1**).

2.1. Eligibility criteria

Studies were eligible for inclusion if they were a primary study or grey literature (i.e., Dissertations and Theses) written in English. There were no country or sample size restrictions. Only studies published from 1987 onwards

were eligible for inclusion, as that was the publication date of the DSM-III-R where the contemporary conceptualisation of ADHD was introduced (American Psychiatric Association [APA], 1987). Eligible studies included participants with a diagnosis of ADHD or hyperkinetic disorder, including either a clinical or DSM/International Classification of Diseases (ICD) research diagnosis, or scoring above a screening threshold for ADHD on a validated questionnaire, as well as a comparison sample of participants without ADHD. Participants from clinical and community samples were included. There were no restrictions on participant age, ethnicity or any other demographic information. To be eligible for inclusion, studies had to include statistical analyses comparing: 1) males and females with ADHD on item-level ADHD symptoms or co-occurring difficulties or 2) comparing females with ADHD to a group of females without ADHD, on item-level co-occurring mental health or neurodevelopmental difficulties. For studies to be included, these outcomes needed to be reported using statistical comparisons of group differences on item-level results, including percentages and effect sizes. Studies were also included if they contained the data needed (e.g. means) to calculate comparisons.

Studies were excluded if they only showed results for total ADHD scores and not item-level statistical results. Qualitative studies, case reports, reviews, systematic reviews, meta-analyses, non-human animal model studies, letters and editorials were not eligible for inclusion.

2.2. Comparison variables

The primary variables examined in this systematic review were core DSM-5 ADHD symptoms and impact of ADHD symptoms on functioning. Both types of variables were examined at an item-level. Impact included domains such as, but not limited to, education, peer relationships, and conduct problems. Co-occurring mental health or neurodevelopmental difficulties and impact were additional outcomes that were considered, including but not limited to emotional difficulties (e.g. anxiety, depression, irritability, and emotional dysregulation), peer and social relationship problems, learning problems, autistic traits, and behavioural difficulties. For details on the comparison variables (i.e. how they were measured/assessed) please see Table 1.

2.3. Information sources

Six electronic research databases were searched on 10/02/2023; Medline, EMBASE, APA PsychInfo (via Ovid), ProQuest (Dissertations & Theses Global), ERIC and British Education Index (via EBSCO) (see **Table S2** for databases searched and the coverage of dates).

2.4. Search strategy

The search strategy was developed based on a scoping search of the existing literature and consultation with a university librarian. The search strategy consisted of three elements: (1) terms related to ADHD, (2) terms related to sex, and (3) terms related to symptoms. Terms within each element were combined with the Boolean operator OR and then all three terms were combined with the operator AND. The search used subject headings (controlled vocabulary) and free text terms. Due to the large number of potential co-occurring difficulties with ADHD, no additional terms were used to search for co-occurring difficulties other than the terms already used relating to

symptoms (i.e. 'symptom'). Results were filtered to only include studies published from 1987 onwards. As a scoping search indicated that a high number of results would be retrieved from database searches, terms related to ADHD were only searched in the title and terms related to sex and symptoms were searched in the title/abstract. The full search strategy for each database is in **Table S3**.

2.5. Screening process

EndNote 20 was used to manage the search results (The EndNote Team, 2013) which automatically deduplicated the initial results. This was followed by manual deduplication. Any results with animal terms in the title or abstract (e.g. rat, mice) were removed. The remaining citations were then imported into Rayyan (Ouzzani et al., 2016). The study selection process was undertaken in two-stages. In stage one, titles and abstracts were screened according to the eligibility criteria. In stage two, full text articles were obtained and screened for eligibility. All screening, data extraction and quality appraisal was independently completed by two reviewers (TW, LH) with any conflicts being resolved through discussion with a third reviewer (JM). The reference lists for reviews, systematic reviews, and meta-analyses identified during the first stage of the screening process, were reviewed for any relevant studies. During stage two of the screening process, lead authors of papers were contacted to enquire about item-level results if these were mentioned but not included in published materials.

2.6. Data extraction process

Data were extracted from eligible studies by two reviewers (TW, LH) who both extracted 100% of the data, with the extractions then checked by both reviewers. Data extraction was managed using Microsoft Excel. Data extracted included study characteristics (i.e., authors, title, year, country, study type and design, sample size, numbers of males and females), participant characteristics (age range and ADHD definition [i.e. how an ADHD diagnosis was described in each paper]), and item-level statistical results. ADHD items were grouped according to the DSM-5 criteria (APA, 2013), where possible, or considered as 'other ADHD' items if they were from previous DSM criteria (i.e. DSM-III-R).

2.7. Quality assessment

The Joanna Briggs Institute (JBI) Critical Appraisal checklist for Analytical Cross-Sectional studies was used to judge the risk of bias (i.e. quality) of each study. The JBI checklist is used to assess the methodological quality of a study and determine the extent to which a study has addressed the possibility of bias in its design, conduct and analysis. The JBI checklist was adapted to suit the needs of the systematic review, with the eight questions being reduced to six (see **Supplementary Text: Quality Assessment**). Studies were judged as 'high risk' if one of the questions was answered 'no' or if three or more questions were answered 'unclear', as 'some concern' if two questions were answered 'unclear,' and as 'low risk' if all questions were answered as 'yes' or if one question, judged and discussed by the research team (TW, LH and JM) to be especially important, was answered as 'unclear'.

2.8. Data synthesis

Studies were grouped by comparison type (female ADHD vs male ADHD or female ADHD vs female comparison). Where possible, studies were grouped by age of participants: children (<13 years), adolescents/young adults (13-24 years) or adults (25+ years). Where sufficient data were available, and study designs were suitably similar (e.g. within the same age range and items relating to the same behaviour/difficulty), fixed effects meta-analyses were conducted per item to examine group comparisons on the outcomes listed above.

For the meta-analyses, available data (e.g. the percentage/number of participants endorsing item-level results) from all studies were transformed into odds ratios (OR) with 95% confidence intervals (CI). The ORs and CIs were then adjusted in Stata 17 (StataCorp, 2021), with the meta-analyses being conducted with the "metan" command using the inverse variance model. Weighting of the meta-analyses was done based on study sample size. To examine heterogeneity statistics of any meta-analyses, I^2 was used.

Where meta-analysis was not feasible, the data was synthesised narratively, based on broad themes/domains (e.g. social impact).

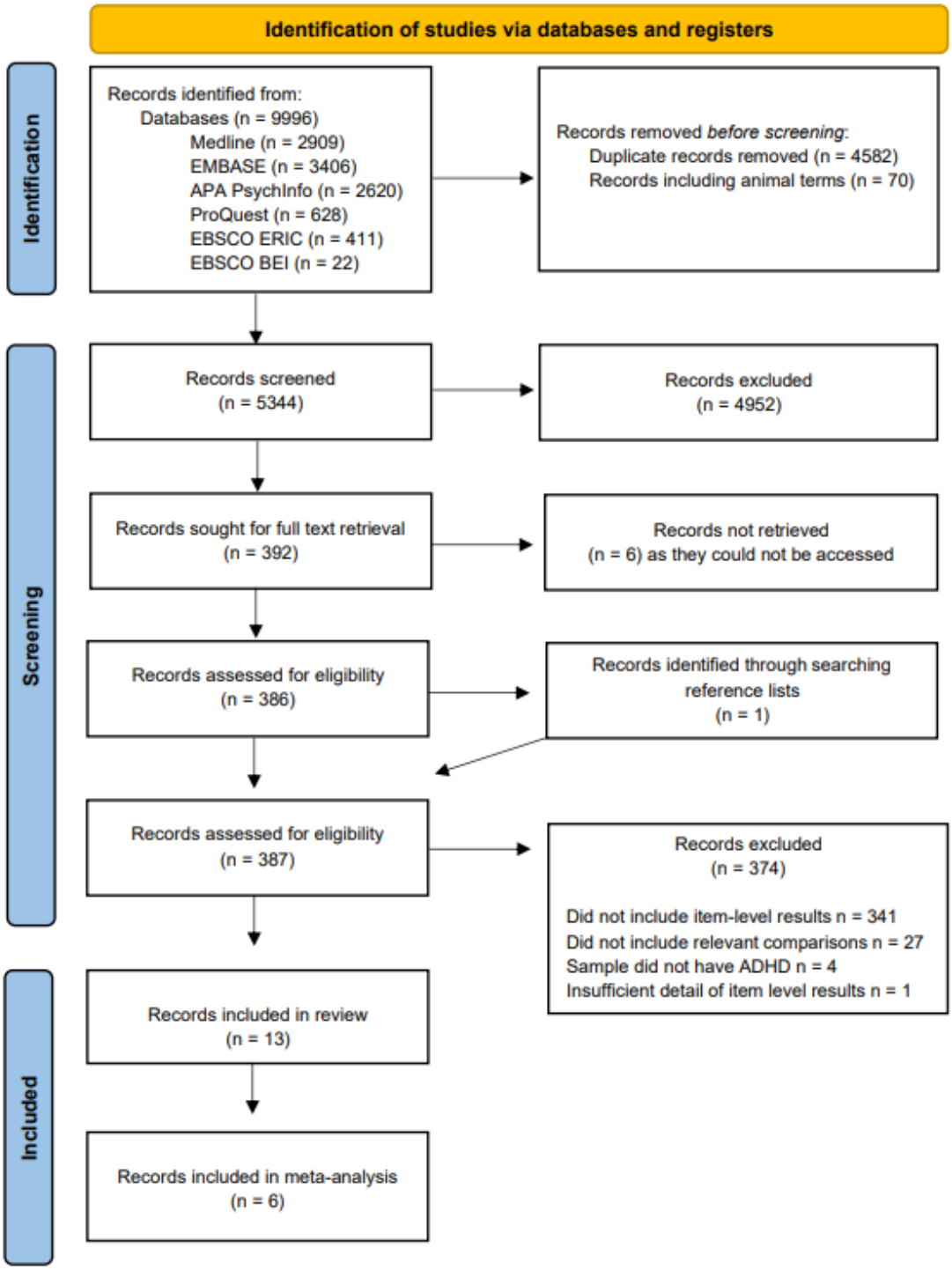
3. Results

3.1. Search selection

The PRISMA flow chart (**Figure 1**) describes the search and selection process. A total of 5344 records were identified through the database search, with one record identified through searching references lists. After the abstract screening, 4952 studies were removed as they did not meet the eligibility criteria. After full-text screening, 13 studies were eligible for inclusion within the review (Biederman et al., 2004; Cortese et al., 2016; Fedele et al., 2012; Ghanizadeh et al., 2019; Graetz et al., 2005; Kamal et al., 2021; Liu et al., 2022; McKay et al., 2023; Meyer et al., 2022; Monuteaux et al., 2010; Moukhtarian et al., 2020; F. Mowlem et al., 2019a; Vildalen et al., 2019). Six of the eligible studies found were included in meta-analyses. Three studies were included in the child/adolescent meta-analyses (Ghanizadeh et al., 2019; Liu et al., 2022; Monuteaux et al., 2010) and included 1098 females and 4399 males. The other three studies were included in the adult meta-analyses (Biederman et al., 2004; Cortese et al., 2016; Vildalen et al., 2019) and included 571 females and 626 males. The remaining seven studies could not be meta-analysed and were instead narratively synthesised. These seven studies included 442 females with ADHD, 613 males with ADHD and 2619 females without ADHD. A list of articles excluded during full-text screening (n=374), including reasons for exclusion, is provided in **Table S4**.

Figure 1

PRISMA flow chart



3.2. Study characteristics

The characteristics of the 13 included studies are described in **Table 1**. Studies were published between 2004 and 2023. The studies were from a variety of countries, including four from the United States, two from the United Kingdom, two from Australia and one each from Iran, Qatar, Norway, Sweden, and China. Five studies included clinical samples, six included general population samples and two studies a mixture of both. Six

studies examined ADHD symptoms at an item-level and seven studies explored impact at an item-level. No studies looked at both symptoms and impact.

Of the included studies, six compared ADHD symptoms in males and females with ADHD, two compared impact in males and females with ADHD and five compared impact in females with and without ADHD. No studies examined co-occurring mental health or neurodevelopmental difficulties at an item-level. Of the 13 eligible studies, one focused on children (<13 years), four focused on adolescents/young adults (13-24 years) and one focused on adults (25+ years). The remaining seven studies reflected samples that crossed these age group boundaries, with four studies including children/adolescents and three including adolescents/adults.

Within the studies, ADHD was confirmed using a variety of methods. 10 studies used research diagnostic interview measures (Biederman et al., 2004; Cortese et al., 2016; Ghanizadeh et al., 2019; Graetz et al., 2005; Liu et al., 2022; McKay et al., 2023; Meyer et al., 2022; Monuteaux et al., 2010; Moukhtarian et al., 2020; F. Mowlem et al., 2019a), two used research diagnostic questionnaires, one completed by expert committees (Kamal et al., 2021) and one used self-report data (Fedele et al., 2012), and another used a questionnaire using teacher-report (Vildalen et al., 2019). Studies also used a range of diagnostic criteria to define ADHD with two using the DSM-III-R, five using the DSM-IV, four using the DSM-5, one using the DSM-IV-TR and another using the ICD-10.

The seven studies examining impact at an item-level all used different measures and examined a variety of impact domains, including home life, friends and school, and another study focused on mind wandering as a symptom of ADHD. Six studies used validated measures of impact (see **Table 1)**, with one study using a questionnaire devised by the lead author and consultant educational psychologist, which was validated by experienced paediatricians and psychologists and another study using the Mind Excessively Wandering Scale (MEWS) (Florence D Mowlem et al., 2019).

Two studies used self-report (Fedele et al., 2012; Moukhtarian et al., 2020), three studies used parent-report (Graetz et al., 2005; McKay et al., 2023; F. Mowlem et al., 2019a), one study used teacher-report (Kamal et al., 2021) and one study used both self- and parent-report (Meyer et al., 2022). As all measures of impact were different, a meta-analysis was not possible, so results were narratively synthesised.

3.3. Risk of bias in studies

The overall risk of bias for all studies was medium to high; a summary of the risk of bias within each study is presented in **Figure S1**. One study was judged as ‘low risk’, five studies judged to have ‘some concerns’, and seven studies judged to be ‘high risk’. The overall risk of bias for question three (‘Was ADHD measured in a valid, objective and reliable way?’) was low, with only one study (Fedele et al., 2012) being judged as ‘high risk’ for this question as they included participants with a self-reported ADHD diagnosis. However, nine of the 13 studies either did not identify any confounding factors or did so but did not deal with them appropriately (e.g.

did not account for age in comparisons), or it was unclear how they did so (e.g. effect of medication status), and therefore four studies were judged as unclear, and five studies were judged as high risk for that question.

Table 1

Characteristics of the studies included in the systematic review.

Study characteristics					Participant characteristics						Variables analysed at an item-level	
					Sample size							
Year	**Author(s)**	**Country**	**Study population**	**Study design / comparison**	**F with ADHD**	**M with ADHD**	**F without ADHD**	**Age range**	**Age group**	**Definition of ADHD**	**ADHD symptoms**	**Non-ADHD items (e.g. impact)**
2004	Biederman et al.*	United States	Clinical sample	M and F with ADHD	82 (69 with item-level results)**	137 (106 with item-level results)**	n/a	37.6 ± 10.5 (mean - ADHD) & 38.7 ± 4.2 (mean - controls)	Adult	DSM-III-R (Structured Clinical Interview)	14 DSM-III-R items (K-SADS-E)	
2005	Graetz et al.	Australia	Population sample	M and F with ADHD	99	225	n/a	6-13 years	Child/adolescent	DSM-IV (DISC-IV)		6 impairment items (DISC-IV)
2010	Monuteaux et al. *	United States	Clinical sample	M and F with ADHD	140	140	n/a	6-17 years	Child/adolescent	K-SADS-E (for those < 18 years) and DSM-III-R (SCID) (for those > 18 years)	14 DSM-III-R symptoms (K-SADS-E and SCID)	
2012	Fedele et al.	United States	Population sample	M and F with ADHD	92	72	n/a	Young adults (college students)	Adolescent/adult	Previously endorsed an ADHD diagnosis or DSM-IV-TR (BCSS-SR)		10 impairment items (BCSS-SR)
2016	Cortese et al.	United States	Population sample	M and F with ADHD	162	178	n/a	18-24 years	Adolescent	DSM-IV (AUDADIS-IV)	18 DSM-IV symptoms (AUDADIS-IV)	

2019	Ghanizadeh et al.	Iran	Clinical sample	M and F with ADHD	280	904	n/a	5.5-19 years	Child/adolescent	DSM-IV diagnostic criteria (psychiatrist)	18 DSM-IV symptoms (Persian version)	
2019	Mowlem et al.	United Kingdom	Population sample	M and F with and without ADHD	32	121	49	7-12 years	Child	DSM-5 (PACS)		5 school impairment items (PACS)
2019	Vildalen et al.	Norway	Clinical sample	M and F with and without ADHD	340	342	522	17-71 years	Adolescent/adult	ICD-10 research criteria (with allowance for the DSM-IV-TR subtypes)	18 DSM-IV-TR symptoms (ASRS)	
2020	Moukhtarian et al.	United Kingdom	Clinical sample and population sample	F with and without ADHD	28	n/a	29	18-65 years	Adolescent/adult	DSM-IV criteria (DIVA)		5 mind wandering items (MEWS)
2021	Kamal et al.	Qatar	Population sample	M and F with and without ADHD	57	93	1001	15 ± 1.5 years (mean age)	Adolescent	DSM-5 (SNAP-IV rating scale)		6 behavioural adaptation questions (academic and social difficulties) devised by lead author
2022	Liu et al.	China	Clinical sample	M and F with ADHD	678	3355	n/a	6-16 years	Child/adolescent	DSM-IV (CDIS)	18 DSM-IV symptoms (ADHD RS-IV)	
2022	Meyer et al.	Sweden	Clinical and population samples	M and F with and without ADHD	105	59	73	15-18 years (ADHD) 14-19 years (controls)	Adolescent	DSM-5 (ADHD module in the MINI-KID)		3 functional impairment items (CSDS)
2023	McKay et al.	Australia	Population sample	M and F with and without ADHD	29	43	18	13-17 years	Adolescent	DSM-5 (DAWBA)		2 friendship items (DAWBA)

Note. *Item-level results provided by authors (not available in published text) **Item-level results only available for a subset of the data

M= male, F = female, n/a = not applicable, DISC-IV = Diagnostic Interview Schedule for Children, AUDADIS-IV = Alcohol Use Disorder and Associated Disabilities Interview Schedule, SNAP-IV = Swanson, Nolan and Pelham Questionnaire, K-SADS-E = Kiddie Schedule for Affective Disorders and Schizophrenia, SCID = Structured Clinical Interview for DSM-III-R, BCSS-SR = Barkley's Current Symptom Scale – Self-report, PACS = Parental Account of Childhood Symptoms, DAWBA = Developmental and Well-being Assessment, MINI-KID = Mini-International

Neuropsychiatric Interview, CDIS = Clinical Diagnostic Interview Scale, MEWS = Mind Excessively Wandering Scale, CSDS = Child Sheehan Disability Scale.

3.4. Meta-analysis results

The child/adolescent meta-analyses comparing females and males with ADHD were conducted on all 18 DSM-IV symptoms (see **Figure 2** note for item list) across three studies. Six DSM-IV items (items 3, 7, 9, 12 and 13) were only available in two studies (Ghanizadeh et al., 2019; Liu et al., 2022) – those items were not measured in Monuteaux et al. (2010) as they used the DSM-III-R. All child/adolescent studies used parent-report.

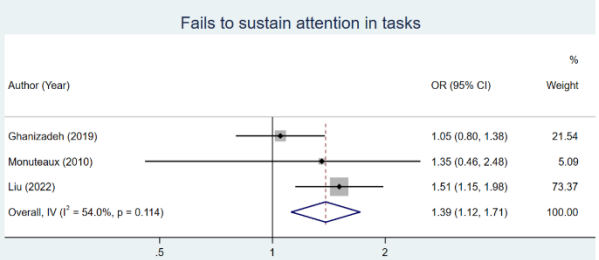
The results suggest that in children with ADHD, parents reported that females were more likely than males to display the symptoms "fails to sustain attention in tasks" (OR= 1.39, 95% CI=1.12, 1.71) and "often easily distracted" (OR=1.54, 95% CI=1.19, 1.96). In contrast, parents were more likely to report males as displaying the symptoms "often fails to follow through on instructions from others", "often fidgets", "difficulty remaining seated when required", "runs/climbs in situations when inappropriate", "always on the go", "often noisy in playing", "difficulty waiting turn", "often blurts out answers" and "often interrupts others". There was no sex difference for the other eight items. The strongest effects were OR=1.54, 95% CI=1.19-1.98 for the item "often easily distracted" and OR=0.86, 95% CI =0.75-0.99 for the item "often blurts out answers".

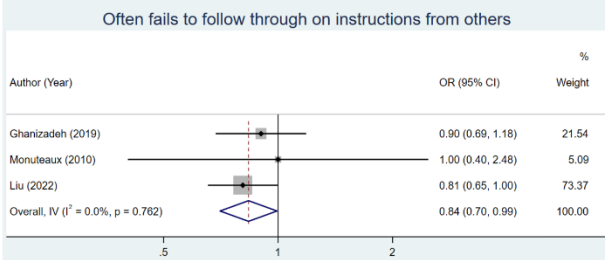
Heterogeneity, as indicated by I^2, ranged from 0% (item 2, 3, 7, 8, 9, 10, 13 and 14) to 84.1% (item 4) (**Figure 2**).

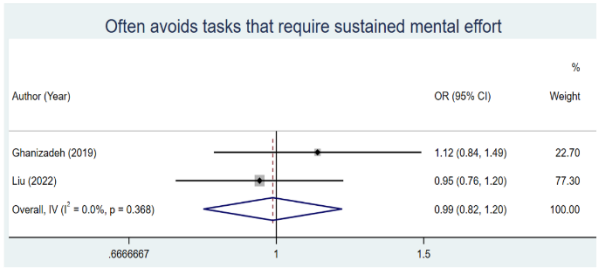
The overall parent endorsement rates for the ADHD symptoms that had statistically significant sex differences were medium-high. The average endorsement rates in females for the ADHD symptoms with significant sex differences ranged from 28.91-76.30%. The item that was, on average, the highest endorsed was "fails to sustain attention in tasks". See **Table S5** for all item endorsement rates.

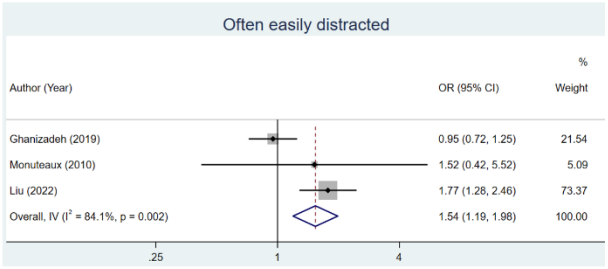
Figure 2.

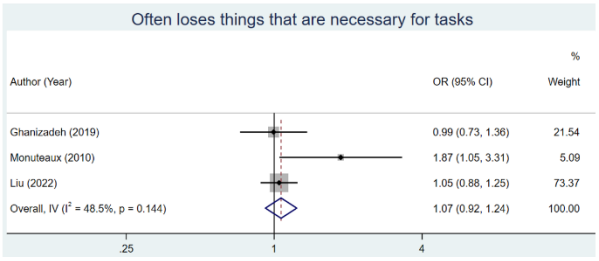
Meta-analysis forest plots for child/adolescent studies comparing males and females for individual ADHD items

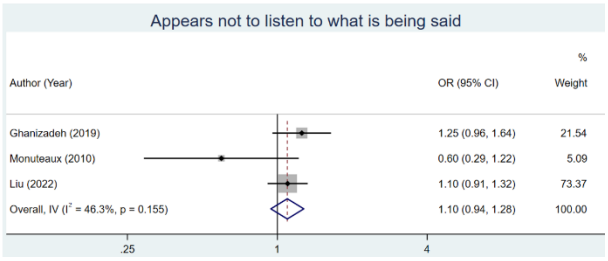

Fails to sustain attention in tasks
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 1.05 (0.80, 1.38) 21.54
Monuteaux (2010) 1.35 (0.46, 2.48) 5.09
Liu (2022) 1.51 (1.15, 1.98) 73.37
Overall, IV (I² = 54.0%, p = 0.114) 1.39 (1.12, 1.71) 100.00
.5
1
2

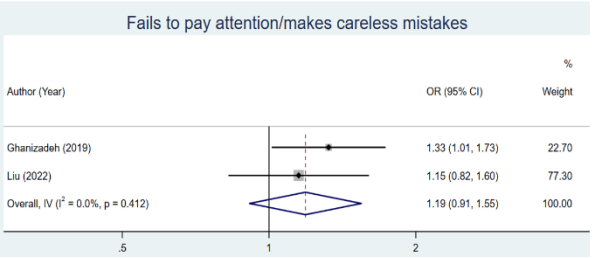

Often fails to follow through on instructions from others
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 0.90 (0.69, 1.18) 21.54
Monuteaux (2010) 1.00 (0.40, 2.48) 5.09
Liu (2022) 0.81 (0.65, 1.00) 73.37
Overall, IV (I² = 0.0%, p = 0.762) 0.84 (0.70, 0.99) 100.00
.5
1
2

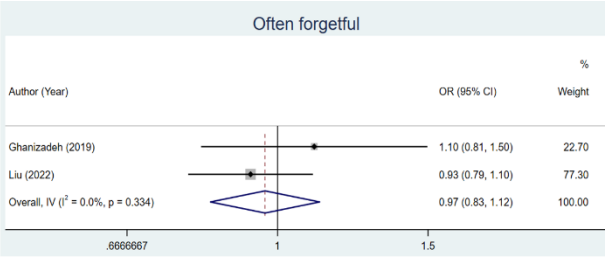

Often avoids tasks that require sustained mental effort
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 1.12 (0.84, 1.49) 22.70
Liu (2022) 0.95 (0.76, 1.20) 77.30
Overall, IV (I² = 0.0%, p = 0.368) 0.99 (0.82, 1.20) 100.00
.6666667
1
1.5

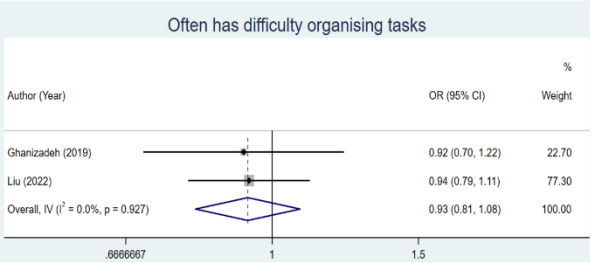

Often easily distracted
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 0.95 (0.72, 1.25) 21.54
Monuteaux (2010) 1.52 (0.42, 5.52) 5.09
Liu (2022) 1.77 (1.28, 2.46) 73.37
Overall, IV (I² = 84.1%, p = 0.002) 1.54 (1.19, 1.98) 100.00
.25
1
4

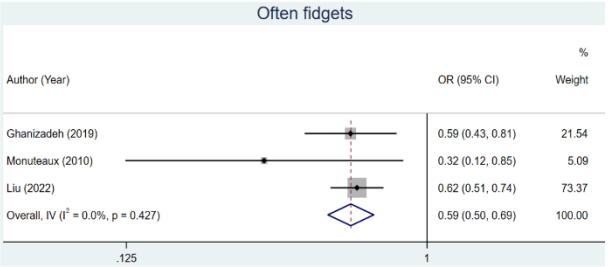

Often loses things that are necessary for tasks
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 0.99 (0.73, 1.36) 21.54
Monuteaux (2010) 1.87 (1.05, 3.31) 5.09
Liu (2022) 1.05 (0.88, 1.25) 73.37
Overall, IV (I² = 48.5%, p = 0.144) 1.07 (0.92, 1.24) 100.00
.25
1
4

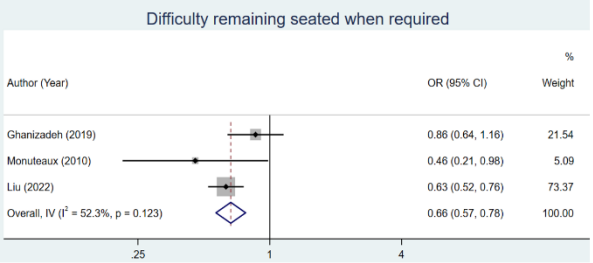

Appears not to listen to what is being said
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 1.25 (0.96, 1.64) 21.54
Monuteaux (2010) 0.60 (0.29, 1.22) 5.09
Liu (2022) 1.10 (0.91, 1.32) 73.37
Overall, IV (I² = 46.3%, p = 0.155) 1.10 (0.94, 1.28) 100.00
.25
1
4

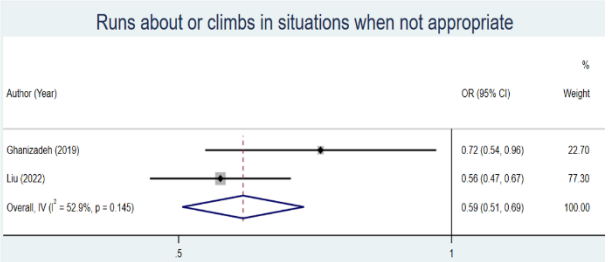

Fails to pay attention/makes careless mistakes
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 1.33 (1.01, 1.73) 22.70
Liu (2022) 1.15 (0.82, 1.60) 77.30
Overall, IV (I² = 0.0%, p = 0.412) 1.19 (0.91, 1.55) 100.00
.5
1
2

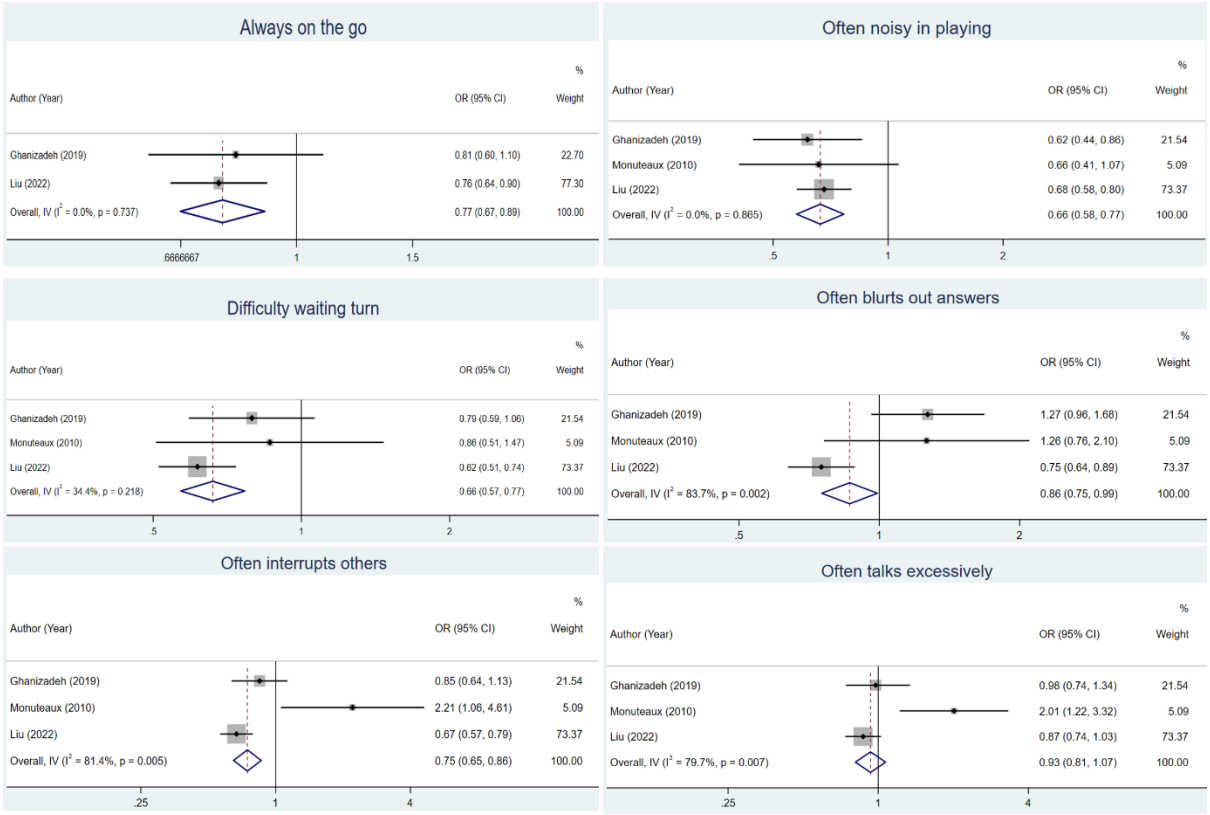

Often forgetful
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 1.10 (0.81, 1.50) 22.70
Liu (2022) 0.93 (0.79, 1.10) 77.30
Overall, IV (I² = 0.0%, p = 0.334) 0.97 (0.83, 1.12) 100.00
.6666667
1
1.5


Often has difficulty organising tasks
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 0.92 (0.70, 1.22) 22.70
Liu (2022) 0.94 (0.79, 1.11) 77.30
Overall, IV (I² = 0.0%, p = 0.927) 0.93 (0.81, 1.08) 100.00
.6666667
1
1.5


Often fidgets
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 0.59 (0.43, 0.81) 21.54
Monuteaux (2010) 0.32 (0.12, 0.85) 5.09
Liu (2022) 0.62 (0.51, 0.74) 73.37
Overall, IV (I² = 0.0%, p = 0.427) 0.59 (0.50, 0.69) 100.00
.125
1


Difficulty remaining seated when required
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 0.86 (0.64, 1.16) 21.54
Monuteaux (2010) 0.46 (0.21, 0.98) 5.09
Liu (2022) 0.63 (0.52, 0.76) 73.37
Overall, IV (I² = 52.3%, p = 0.123) 0.66 (0.57, 0.78) 100.00
.25
1
4


Runs about or climbs in situations when not appropriate
Author (Year)
OR (95% CI)
% Weight
Ghanizadeh (2019) 0.72 (0.54, 0.96) 22.70
Liu (2022) 0.56 (0.47, 0.67) 77.30
Overall, IV (I² = 52.9%, p = 0.145) 0.59 (0.51, 0.69) 100.00
.5
1



Note. Weighted by study's sample size (Ghanizadeh et al., 2019 *n*=1184; Monuteaux et al., 2010 *n*=280; Liu et al., 2022 *n*=4033). An odds ratio of <1 suggests that males were more likely to display an item whereas an odds ratio of >1 suggests that females were more likely to display an item.

Item 1 = "Fails to sustain attention in tasks, Item 2 = "Often fails to follow through on instructions from others", Item 3 = "Often avoid tasks that require sustained mental effort", Item 4 = "Often easily distracted", Item 5 = "Often loses things that are necessary for tasks", Item 6 = "Appears not to listen to what is being said", Item 7 = "Fails to pay attention/makes careless mistakes", Item 8 = "Often forgetful", Item 9 = "Often has difficulty organising tasks", Item 10 = "Often fidgets", Item 11 = "Difficulty remaining seated when required", Item 12 = "Runs about or climbs in situations when not appropriate", Item 13 = "Always on the go", Item 14 = "Often noisy in playing", Item 15 = "Difficulty waiting turn", Item 16 = "Often blurts out answers", Item 17 = "Often interrupts others" and Item 18 = "Often talks excessively".

The adult meta-analyses were conducted on three studies, across 16 DSM-IV items (see **Figure 3** note). Two items (items 2 and 12) could not be meta-analysed as they were only

measured in one study (Cortese et al., 2016). Additionally, five items were only meta-analysed in two studies as four items (3, 7, 8 and 9) were not measured by Biederman et al. (2004) and one item (13) was not measured by Vildalen et al. (2019). All adult studies used self-report.

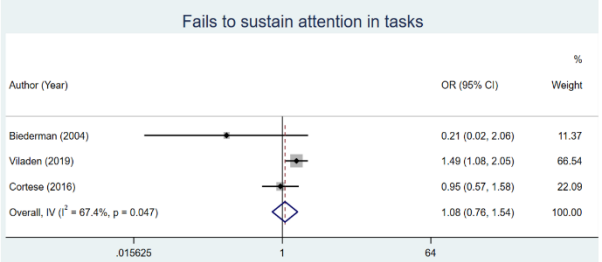
Adult females were more likely than males to endorse “often easily distracted” (OR= 1.49, 95% CI=1.02, 2.16)., “often has difficulty organising tasks” (OR=1.40, 95% CI=1.06, 1.85), “often blurts out answers” (OR=1.32, 95% CI=1.02, .170), and “often talks excessively” (OR= 1.65, 95% CI=1.30, 2.09). The item with the biggest sex difference in those with ADHD was “often easily distracted” and “often talks excessively”, with females more likely than males to endorse these items. There was no sex difference for the other 12 items.

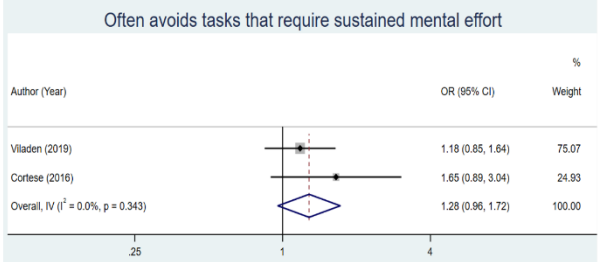
Heterogeneity, as indicated by I^2, ranged from 0% (item 3, 5, 9, 18) to 92.6% (item 13); see **Figure 3**.

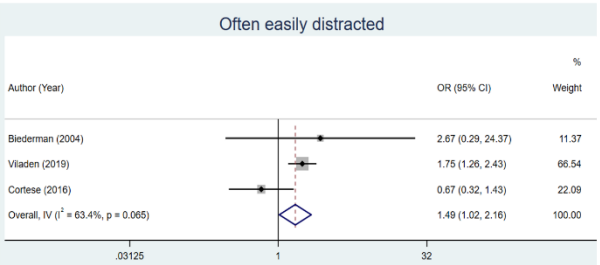
The overall endorsement rates for the ADHD symptoms that had statistically significant sex differences were high. The average endorsement rates for the ADHD symptoms with significant sex differences ranged from 52.66-86.24%. The item that was, on average, the highest endorsed was “often easily distracted”. See **Table S6** for all item endorsement rates.

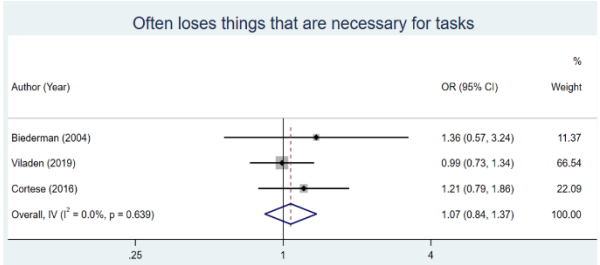
Figure 3

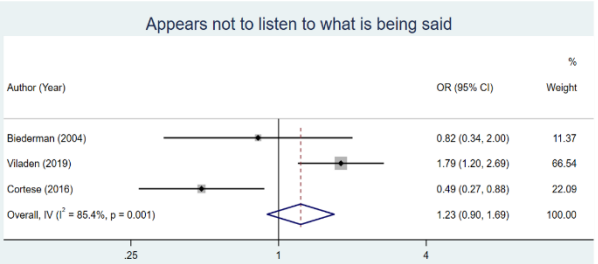
Meta-analysis forest plots for adult studies comparing males and females for individual ADHD items

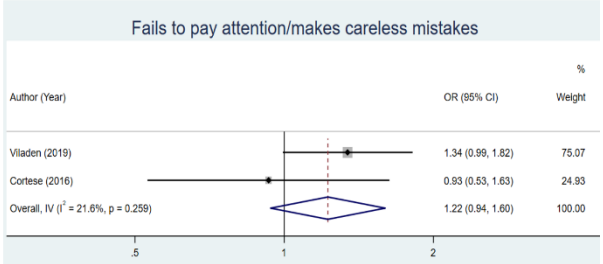

Fails to sustain attention in tasks
Author (Year)
OR (95% CI)
%
Weight
Biederman (2004)
0.21 (0.02, 2.06)
11.37
Viladen (2019)
1.49 (1.08, 2.05)
66.54
Cortese (2016)
0.95 (0.57, 1.58)
22.09
Overall, IV (I² = 67.4%, p = 0.047)
1.08 (0.76, 1.54)
100.00
.015625
1
64

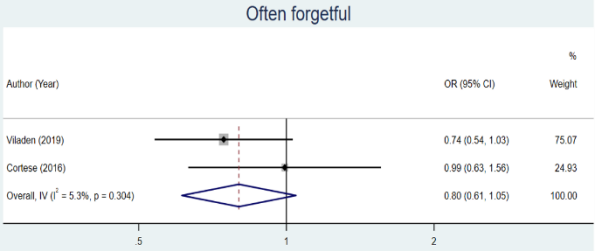

Often avoids tasks that require sustained mental effort
Author (Year)
OR (95% CI)
%
Weight
Viladen (2019)
1.18 (0.85, 1.64)
75.07
Cortese (2016)
1.65 (0.89, 3.04)
24.93
Overall, IV (I² = 0.0%, p = 0.343)
1.28 (0.96, 1.72)
100.00
.25
1
4

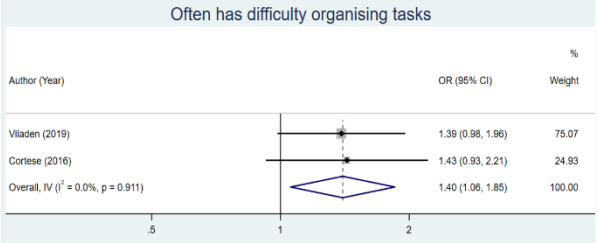

Often easily distracted
Author (Year)
OR (95% CI)
%
Weight
Biederman (2004)
2.67 (0.29, 24.37)
11.37
Viladen (2019)
1.75 (1.26, 2.43)
66.54
Cortese (2016)
0.67 (0.32, 1.43)
22.09
Overall, IV (I² = 63.4%, p = 0.065)
1.49 (1.02, 2.16)
100.00
.03125
1
32

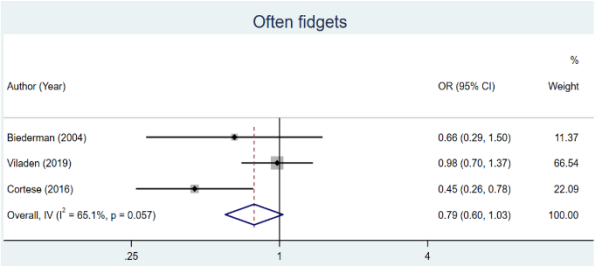

Often loses things that are necessary for tasks
Author (Year)
OR (95% CI)
%
Weight
Biederman (2004)
1.36 (0.57, 3.24)
11.37
Viladen (2019)
0.99 (0.73, 1.34)
66.54
Cortese (2016)
1.21 (0.79, 1.86)
22.09
Overall, IV (I² = 0.0%, p = 0.639)
1.07 (0.84, 1.37)
100.00
.25
1
4

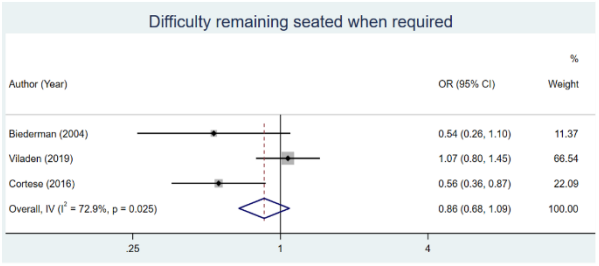

Appears not to listen to what is being said
Author (Year)
OR (95% CI)
%
Weight
Biederman (2004)
0.82 (0.34, 2.00)
11.37
Viladen (2019)
1.79 (1.20, 2.69)
66.54
Cortese (2016)
0.49 (0.27, 0.88)
22.09
Overall, IV (I² = 85.4%, p = 0.001)
1.23 (0.90, 1.69)
100.00
.25
1
4

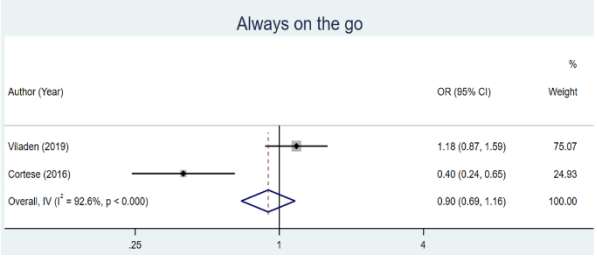

Fails to pay attention/makes careless mistakes
Author (Year)
OR (95% CI)
%
Weight
Viladen (2019)
1.34 (0.99, 1.82)
75.07
Cortese (2016)
0.93 (0.53, 1.63)
24.93
Overall, IV (I² = 21.6%, p = 0.259)
1.22 (0.94, 1.60)
100.00
.5
1
2

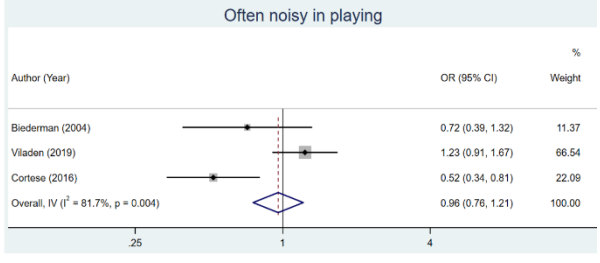

Often forgetful
Author (Year)
OR (95% CI)
%
Weight
Viladen (2019)
0.74 (0.54, 1.03)
75.07
Cortese (2016)
0.99 (0.63, 1.56)
24.93
Overall, IV (I² = 5.3%, p = 0.304)
0.80 (0.61, 1.05)
100.00
.5
1
2

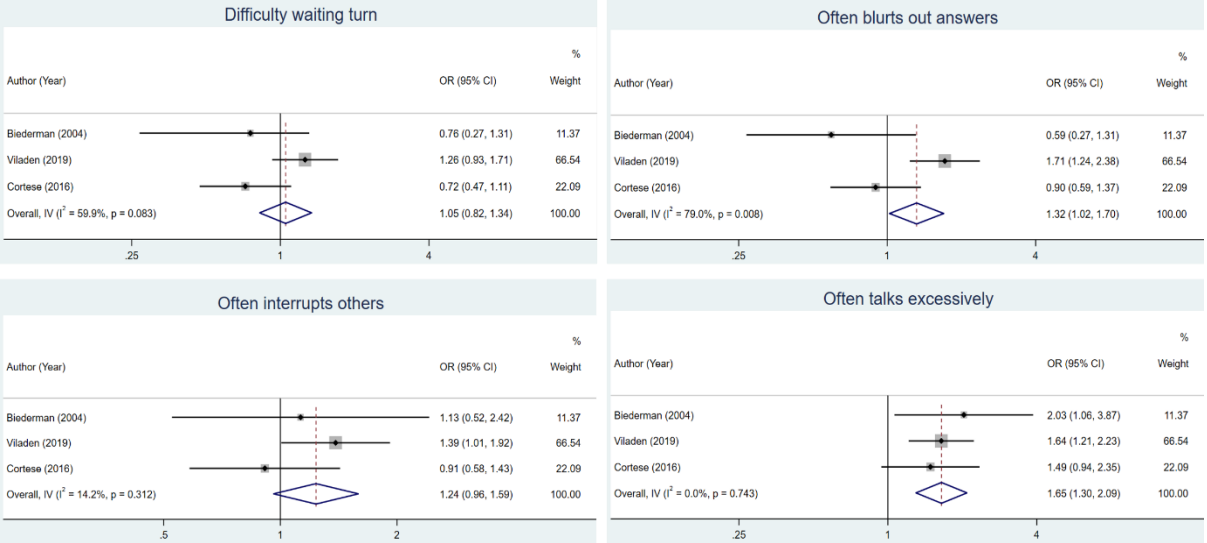

Often has difficulty organising tasks
Author (Year)
OR (95% CI)
%
Weight
Viladen (2019)
1.39 (0.98, 1.96)
75.07
Cortese (2016)
1.43 (0.93, 2.21)
24.93
Overall, IV (I² = 0.0%, p = 0.911)
1.40 (1.06, 1.85)
100.00
.5
1
2


Often fidgets
Author (Year)
OR (95% CI)
%
Weight
Biederman (2004)
0.66 (0.29, 1.50)
11.37
Viladen (2019)
0.98 (0.70, 1.37)
66.54
Cortese (2016)
0.45 (0.26, 0.78)
22.09
Overall, IV (I² = 65.1%, p = 0.057)
0.79 (0.60, 1.03)
100.00
.25
1
4


Difficulty remaining seated when required
Author (Year)
OR (95% CI)
%
Weight
Biederman (2004)
0.54 (0.26, 1.10)
11.37
Viladen (2019)
1.07 (0.80, 1.45)
66.54
Cortese (2016)
0.56 (0.36, 0.87)
22.09
Overall, IV (I² = 72.9%, p = 0.025)
0.86 (0.68, 1.09)
100.00
.25
1
4


Always on the go
Author (Year)
OR (95% CI)
%
Weight
Viladen (2019)
1.18 (0.87, 1.59)
75.07
Cortese (2016)
0.40 (0.24, 0.65)
24.93
Overall, IV (I² = 92.6%, p < 0.000)
0.90 (0.69, 1.16)
100.00
.25
1
4


Often noisy in playing
Author (Year)
OR (95% CI)
%
Weight
Biederman (2004)
0.72 (0.39, 1.32)
11.37
Viladen (2019)
1.23 (0.91, 1.67)
66.54
Cortese (2016)
0.52 (0.34, 0.81)
22.09
Overall, IV (I² = 81.7%, p = 0.004)
0.96 (0.76, 1.21)
100.00
.25
1
4



Note. Weighted by study's sample size (Biederman et al., 2004 *n*=175; Vildalen et al., 2019 *n*=1024; Cortese et al., 2016 *n*=340). An odds ratio of <1 suggests that males were more likely to endorse an item whereas an odds ratio of >1 suggests that females were more likely to endorse an item.

Item 1 = "Fails to sustain attention in tasks, Item 3 = "Often avoid tasks that require sustained mental effort", Item 4 = "Often easily distracted", Item 5 = "Often loses things that are necessary for tasks", Item 6 = "Appears not to listen to what is being said", Item 7 = "Fails to pay attention/makes careless mistakes", Item 8 = "Often forgetful", Item 9 = "Often has difficulty organising tasks", Item 10 = "Often fidgets", Item 11 = "Difficulty remaining seated when required", Item 13 = "Always on the go", Item 14 = "Often noisy in playing/doing leisure activities quietly", Item 15 = "Difficulty waiting turn", Item 16 = "Often blurts out answers", Item 17 = "Often interrupts others" and Item 18 = "Often talks excessively".

3.5. Narrative synthesis

Mind wandering. Moukhtarian et al. (2020) found that spontaneous self-reported mind wandering was greater in intensity in adult females with compared to without ADHD, across all five items measured (see **Supplementary Text: Mind wandering item measurement**).

Home impact. Overall, females with ADHD were found to be more impaired on items measuring their home life than both males with ADHD and females without ADHD. In adults with ADHD, females endorsed significantly higher impact than males (Fedele et al., 2012). Females with ADHD had higher self-reported home impact than males with ADHD, and higher self- and

parent-reported impact than females without ADHD (Meyer et al., 2022). Further, sex by ADHD-subtype interactions generally found that males were more impaired than females in the combined and hyperactive-impulsive subtype and were equally or less impaired than females with the inattentive subtype (Graetz et al., 2005). In contrast, Graetz et al. (2005) found no significant sex differences on ratings of annoyance to parents and interference with family activities.

Social impact. Females with ADHD were more impaired on items measuring social impact than females without ADHD, including being more impaired with friendships (Meyer et al., 2022) and finding it harder than average to make and maintain friends (McKay et al., 2023). Further, females with ADHD were more impaired than females without ADHD on making and maintaining friends and experiencing friend-related distress (Kamal et al., 2021). There were mixed findings when examining sex differences. Some studies reported that females were more impaired than males in their social life (Fedele et al., 2012) and with friendships (Meyer et al., 2022), including finding it harder than average to make and maintain friends (McKay et al., 2023). However, some studies reported no significant sex differences in social difficulties or impact of ADHD symptoms on peer activities (Graetz et al., 2005; Kamal et al., 2021).

School impact in children. Females with ADHD were more impaired on items measuring school impact than females without ADHD (Kamal et al., 2021; Meyer et al., 2022; F. Mowlem et al., 2019). Females with ADHD were found to receive more special education provisions and complaints about hyperactive behaviour (F. Mowlem et al., 2019) and more likely to be perceived as a burden to the teacher or class, impaired in class learning and experienced more difficulty with emotions, concentration, and behaviour at school than females without ADHD (Kamal et al., 2021). Further, compared to females without ADHD, females with the hyperactive-impulsive and combined subtype were more impaired in classroom learning and emotions and behaviour, while only those with the combined subtype were more impaired in burden to the teacher or classroom (Kamal et al., 2021). There were mixed findings when examining sex differences. Fedele et al. (2012) and Meyer et al. (2022) reported that females were more impaired at educational activities and at school than males. In contrast, Graetz et al. (2005) reported that males were more likely than females to have problems with their schoolwork and grades than females. Graetz et al. (2005) also reported that males were more likely than females to be considered annoying by teachers. Some studies reported no sex differences in school related-impact or academic difficulties (Kamal et al., 2021; F. Mowlem et al., 2019). Additionally, males with ADHD were rated as more impaired than females in the combined and

hyperactive-impulsive groups on problems with schoolwork and grades and annoyance to teachers, but equally impaired in the inattentive group (Graetz et al., 2005).

Other impacts. In adults, females with ADHD were significantly more impaired in their money management and daily life activities than males with ADHD (Fedele et al., 2012). There were no statistically significant sex differences found in community, dating or marital relationships, work, driving and leisure impact, although on all measures aside from community, females reported higher impact than males (Fedele et al., 2012). Graetz et al. (2005) found no significant sex differences across or within ADHD subtype on ratings of personal distress caused by symptoms in children.

4. Discussion

The aim of this systematic review was to examine if there were specific ADHD symptoms or types of impact that characterise the manifestation of ADHD in females compared to males with ADHD and females without ADHD. Despite using broad search terms to find relevant studies, only 13 eligible studies were found. Six studies of sex differences in ADHD symptoms could be meta-analysed, whereas the seven studies examining impact could only be synthesised narratively. Overall, our results suggest that there are some sex differences in ADHD symptom profile and that females with ADHD are generally more impaired across a range of domains compared to males with ADHD and females without ADHD.

The main meta-analysis results of sex differences in item-level ADHD symptoms demonstrate that in children with ADHD, parents report that females are more likely to display certain inattentive symptoms (“fails to sustain attention in tasks” and “often easily distracted”) whereas males were more likely to display 8 of the 9 DSM-IV hyperactive-impulsive symptoms (all except “talks excessively”). In adults with ADHD, females were more likely than males to endorse a mix of inattentive (“often easily distracted” and “often has difficulty organising tasks”) and hyperactive-impulsive (“often blurts out answers” and “often talks excessively”) symptoms. Overall, the current results are consistent with previous research of diagnostic subtypes and total scores, which highlighted that females with ADHD are more likely to express inattentive symptoms (Quinn & Madhoo, 2014) and less likely to express hyperactive-impulsive symptoms (Loyer Carbonneau et al., 2021) (Gershon & Gershon, 2002; Quinn & Madhoo, 2014) than males with ADHD.

The results also demonstrate that there are fewer significant sex differences in endorsement of symptoms in adults relative to children. This may be due to developmental changes or how

ADHD symptoms were measured, with parent-report used for children and self-report used for adults. Previous work has found that parents are more likely to rate DSM-IV symptoms of ADHD (excluding “talks excessively”) as male-descriptive (Ohan & Johnston, 2005) and also overrate males’ hyperactive-impulsive symptoms compared to objective interviews (F. Mowlem et al., 2019). As such, differential misclassification may be operating, resulting in parents endorsing more robust sex differences in ADHD symptoms, especially on hyperactive-impulsive symptoms. Although, given the male-biased sex ratio of ADHD in childhood reducing to near 1:1 in adulthood (Williamson & Johnston, 2015), parents may be reporting real sex differences in symptoms, rather than this reflecting a bias in reporting. Additionally, another possibility may be that females are more willing to take part in research studies (Glass et al., 2015) and as such there may be a bias in who takes part in adult studies.

Further, when examining sex differences in adults, “talks excessively” was found to have a large effect size, with females being more likely to endorse the symptom than males. This finding is interesting as it is congruent with previous studies that have attempted to characterise ‘female-sensitive’ ADHD behaviours and have included items such as “talks excessively” and “likes to talk a lot” (Grskovic & Zentall, 2010; Ohan & Johnston, 2005) and indeed this was the only hyperactive-impulsive item not showing a male-bias in our meta-analysis of sex differences in children.

In sum, the results suggest that there are sex differences in the core diagnostic symptoms related to ADHD. This could contribute to the under-recognition of ADHD in females. The overall endorsement rates for individual ADHD symptoms with observed sex differences were high in adults and medium to high in children. This indicates that symptoms commonly differ between sexes across samples with ADHD. These findings have useful clinical implications as they highlight which ADHD symptoms clinicians may want to be more aware of when assessing females with suspected ADHD, such as certain inattentive symptoms in childhood (e.g. “fails to sustain attention in tasks”) and hyperactive-impulsive symptoms in adulthood (e.g. “talks excessively”), which may aid more accurate and timely ADHD diagnoses, allowing for earlier treatment, which would promote improved quality of care.

Further, while there was evidence of sex differences in some ADHD symptoms, there were many symptoms where we did not see any sex differences, in both child- and adulthood, particularly for inattentive symptoms. This may indicate that the diagnostic criteria/symptom checklists used may be valid tools to capture inattentive symptoms overall. Although, given that the development of the diagnostic criteria may be biased towards the male presentation (F. D.

Mowlem et al., 2019), there are likely more female-sensitive ADHD-related difficulties omitted (e.g. previously suggested items such as 'doodles instead of completing classwork', 'impulsively changes conversation topics' & 'changes friends without thinking') (Ohan & Johnston, 2005); if included in the diagnostic criteria, such additional items could theoretically better capture female ADHD and help identify ADHD in females at an earlier age. Also, given the eligible studies in our review included females with recognised ADHD, females with different or atypically presenting ADHD symptoms are likely not to have been included. This could also include difficulties related to mind wandering.

Our findings on mind wandering (Moukhtarian et al., 2020) support previous work and literature reviews, suggesting that spontaneous mind wandering is associated with ADHD (Biederman et al., 2006; Lanier et al., 2021), with females with ADHD displaying more intense mind wandering than females without ADHD (Moukhtarian et al., 2020). Mind wandering can have a negative effect for individuals, including reducing overall wellbeing, even after accounting for the effects of ADHD symptoms (Florence D Mowlem et al., 2019). These findings on mind wandering are interesting as they suggest it is associated with more functional impairment when present, requiring further research given the limited literature available.

Our review also examined sex differences in impact related to ADHD at an item-level. In general, females had more impact from ADHD at home than males (Fedele et al., 2012; Meyer et al., 2022). Although one study reported males were generally more impaired than females when comparing children with the same ADHD subtype (Graetz et al., 2005), this may be because the items used to measure home impact aligned closely with descriptions of hyperactivity-impulsivity ("ratings of annoyance" and "interference with family activities"), which are often more likely to be endorsed by parents as male-descriptive (Quinn & Madhoo, 2014). These findings suggest that females compared to males with ADHD are more likely to be impacted in their home life (Biederman et al., 2006).

Females with ADHD were more impaired on items measuring social impact than females without ADHD (Kamal et al., 2021; McKay et al., 2023; Meyer et al., 2022). This is consistent with previous work indicating that females with ADHD are impaired on peer functioning and have lower levels of friendships participation (Kok et al., 2016; Quinn & Madhoo, 2014). Some of the reviewed studies reported that females with ADHD are more socially impaired than males, including with making and maintaining friends (Fedele et al., 2012; McKay et al., 2023; Meyer et al., 2022). These results are in line with a recent systematic review on sex differences in social functioning (Faheem et al., 2022). However, two studies found no evidence of sex differences in

social impairment (Graetz et al., 2005; Kamal et al., 2021), possibly partly due to reliance on teacher reports to accurately compare students' behaviours to descriptors on a checklist, after only knowing students for six months (Kamal et al., 2021). These findings are important as social skill impairment and limited social activities have been suggested to be associated with long-term mental health difficulties in those with ADHD (Mrug et al., 2012).

The results on school impairment suggested that females with ADHD were more impaired than females without ADHD, consistent with previous literature (Biederman et al., 2006). Sex differences in school impairment suggest that females are more impaired in school and educational activities (Fedele et al., 2012; Meyer et al., 2022), with inattentive females more impaired in classroom learning, emotions, concentration and behaviour (Kamal et al., 2021), while males are more impaired in schoolwork/grades and annoyance to teachers (Graetz et al., 2005). Previous work suggests that females with ADHD are more impaired at school than males with ADHD (Wolraich et al., 1996). The mixed results identifying males with ADHD as more impaired in their schoolwork/grades than females with ADHD may be because Graetz et al. (2005) assessed the extent to which an individual's ADHD influenced their schoolwork/grades, not if the individual was academically impaired. This highlights the importance of reporting item-level results to understand the nuances of this issue. Additionally, males with ADHD may be rated as being more of an annoyance to teachers than females with ADHD as they are rated by teachers as displaying more problem behaviour (e.g. aggression) (Derks et al., 2007). This also highlights that comorbidity is likely to vary by sex. The mixed findings on school impairment by Mowlem et al. (2019a) and Kamal et al. (2021) may have been due to a variety of reasons. These include reduced power to find group differences due to a large mismatch between the number of male and female (121 vs 32) participants (F. Mowlem et al., 2019a) and small sample sizes in the inattentive group (12 males and 11 females) (Kamal et al., 2021).

Some studies within the review analysed impact items across and within individual ADHD subtypes (Graetz et al., 2005; Kamal et al., 2021). The findings highlight sex-specific risks associated with different ADHD subtypes on impact, which are often overlooked when studies only examine sex differences across ADHD regardless of subtypes, highlighting that ADHD subtype should be considered when examining ADHD sex comparisons.

Overall, the results on impact found that females with ADHD were more impaired than females without ADHD in terms of school, social and home impact. Females with ADHD compared to females without ADHD also reported more 'ADHD-related' difficulties such as mind wandering. Females with ADHD were also more impaired than males with ADHD in their home life, with

mixed findings on school and social impact. Impact at school and in an individual's social and home life can have further negative knock-on effects, including increased loneliness due to difficulties with social relationships which may have adverse effects on mental health, including contributing to the development of co-occurring mood and anxiety disorders (Houghton et al., 2020; Jong et al., 2024). Difficulties at school such as receiving complaints about hyperactivity (F. Mowlem et al., 2019) can increase the likelihood of suspension (Loe & Feldman, 2007), and along with impairment in classroom learning (Kamal et al., 2021), affect overall academic performance (Keilow et al., 2018), which can result in lower employability and quality of life (Shifrin et al., 2010). These findings suggest that timely identification and diagnosis of ADHD are vital, especially in females who often receive a delayed diagnosis, as it allows for treatment and support, such as facilitating social support and accommodations/interventions at school (Lovett et al., 2023), that can help mitigate or reduce the impact of symptoms.

Strengths and limitations

This systematic review is novel as it is the first to explore and synthesise findings from studies that report item-level sex differences in ADHD symptoms and impact, using rigorous review methods. The eligibility criteria included participants scoring above a screening threshold for ADHD, allowing results to include females who would be sub-threshold for conventional ADHD diagnostic criteria, increasing the chance of finding sex differences on ADHD symptoms or related impact. Findings highlight important differences in ADHD symptom profiles between males and females. The protocol for this review was pre-registered, and made publicly available via PROSPERO, reducing the risk of reporting bias. Additionally, we gathered unpublished data from eligible studies where possible. However, there were some limitations with the review. There were a limited number of studies found, with only six studies able to be meta-analysed. Additionally, the overall risk of bias of the systematic review is medium-high. All studies of ADHD rely on established ADHD criteria. If these are indeed male biased, then female ADHD behaviours not included in these criteria will have been missed. Females may also need greater symptom levels and impact for their ADHD to be recognised and included in a research study which could account for our sex difference findings. This review also only included studies published in English. Further, there was a lack of adjustment for psychiatric comorbidities and medication status which may have influenced the results and could explain why heterogeneity in the meta-analyses was large for some items. The search terms for the review may have been too broad to identify studies examining sex differences in co-occurring mental health or neurodevelopmental conditions and may have benefited from a narrower scope. Additionally,

there are limitations to consider with the studies included in the review. First, only a limited number of studies meeting inclusion criteria have been published, which limits the interpretation of the results. Second, the included studies used a variety of ADHD definitions and measures, as such, some ADHD symptoms could not be meta-analysed, or the analysis included a subset of studies. Similarly, the impairment measures used were all different, making it hard to draw robust conclusions. Finally, the risk of bias assessment suggested that not all the studies included in the review were of high quality, with only one study being deemed low risk overall. Further, many studies reported confounding variables, but not all dealt with them appropriately, increasing the risk of bias.

Clinical implications and future studies

As previously mentioned, these findings have useful clinical implications as they highlight the individual ADHD symptoms that clinicians may want to be aware of when assessing and diagnosing suspected female ADHD. Increasing awareness of how ADHD manifests in girls, including improving future assessment tools to more readily identify ADHD in girls, is essential as it will potentially facilitate earlier ADHD diagnoses. Further, given our findings indicate that females compared to males with ADHD are more likely to report family and interpersonal difficulties, when clinicians diagnose girls with ADHD, they should ask about these and consider what kinds of relevant support could be offered, such as family-based therapies or counselling tailored to the individual. The findings also have implications for future studies. Given the limited research on item-level sex differences in ADHD symptoms, we strongly recommend that future studies include this level of detail, even if it is not the primary analysis. This will provide more detailed results and allow researchers to unpick which specific symptoms are contributing to different presentations of ADHD. Future studies should also address the limitations of the present review, by adjusting and reporting on confounding factors such as mental health comorbidities, ADHD subtype and medication status.

Further, given that the ADHD diagnostic criteria field studies were based mainly on males (Lahey et al., 1994), future research should examine if other difficulties related to ADHD, not in the diagnostic criteria, characterise the manifestation of ADHD in females compared to males and those without ADHD. Other factors may include emotion dysregulation, which has been identified as a potential characteristic of female ADHD (Quinn & Madhoo, 2014), and symptoms previously suggested as 'female-sensitive' (e.g. emotional impulsivity such as changing friends impulsively) (Grskovic & Zentall, 2010; Ohan & Johnston, 2005). Revisions to diagnostic criteria for ADHD could include additional symptoms or refinements to existing criteria, but the evidence

base for these needs to be robust. Finally, future studies should also examine, and report item-level sex differences in mental health or neurodevelopmental comorbidities in individuals with ADHD as no studies were found to be eligible for the present review

Conclusion

This systematic review and meta-analysis provided insights into the sex differences in individual core symptoms and impact related to ADHD. In childhood, females were more likely to display specific inattentive symptoms, such as “fails to sustain attention in tasks” and “often easily distracted”, than males, who were more likely to display most of the hyperactive-impulsive symptoms, such as “often fidgets” and “difficulty remaining seated when required”. In adulthood, females with ADHD were more likely than males with ADHD to endorse the symptoms “often easily distracted”, “often has difficulty organising tasks”, “often blurts out answers” and “often talks excessively”. Further, the results suggested that females with ADHD are more impaired than males with ADHD and females without ADHD on a range of items, including school impact, and their home and social life. Overall, the review highlights the need for future research to identify and characterise symptoms typical of female ADHD, as it may have important implications for clinical practice and aid future development of a more inclusive ADHD assessment tool to help earlier ADHD recognition and diagnosis in females.

Acknowledgements

This study was funded by the Welsh Government through Health and Care Research Wales via a National Institute for Health and Care Research (NIHR) Advanced Fellowship (Ref: NIHR-FS(A)-2022).

This report is independent research supported by the National Institute for Health Research NIHR Advanced Fellowship - Stage 2, Dr Abigail Russell, NIHR300591. KS and TF are NIHR Senior Investigators. The views expressed are those of the authors and not necessarily those of the NHS, NIHR or the Department of Health and Social Care.

The authors would like to thank Jonathan Jones, the university librarian, for his help developing the search strategy. We also acknowledge the authors of two of the studies included in the systematic review (Biederman et al., 2004; Monuteaux et al., 2010) for providing additional unpublished data.

Disclosures

Tamsin Ford's research group received funds for research methods consultation with Place2Be, a third sector organisation providing mental health interventions and training within schools. Kate Langley has been part of the scientific advisory board for Medice, on topics unrelated to this work.

All other authors report no conflicts of interest.

References

Arnett, A. B., Pennington, B. F., Willcutt, E. G., DeFries, J. C., & Olson, R. K. (2015). Sex differences in ADHD symptom severity. *Journal of Child Psychology and Psychiatry*, *56*(6), 632-639.

Association, A. P. (2000). Diagnostic and statistical manual of mental disorders. *Text revision*.

Ayano, G., Demelash, S., Gizachew, Y., Tsegay, L., & Alati, R. (2023). The global prevalence of attention deficit hyperactivity disorder in children and adolescents: An umbrella review of meta-analyses. *Journal of Affective Disorders*, *339*, 860-866. https://doi.org/https://doi.org/10.1016/j.jad.2023.07.071

Biederman, J., Faraone, S. V., Monuteaux, M. C., Bober, M., & Cadogen, E. (2004). Gender effects on attention-deficit/hyperactivity disorder in adults, revisited. *Biological psychiatry*, *55*(7), 692-700.

Biederman, J., Monuteaux, M. C., Mick, E., Spencer, T., Wilens, T. E., Klein, K. L., Price, J. E., & Faraone, S. V. (2006). Psychopathology in females with attention-deficit/hyperactivity disorder: a controlled, five-year prospective study. *Biological psychiatry*, *60*(10), 1098-1105.

Cortese, S., Faraone, S. V., Bernardi, S., Wang, S., & Blanco, C. (2016). Gender differences in adult attention-deficit/hyperactivity disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *The Journal of clinical psychiatry*, *77*(4), 7626.

Daley, D., Jacobsen, R. H., Lange, A.-M., Sørensen, A., & Walldorf, J. (2019). The economic burden of adult attention deficit hyperactivity disorder: a sibling comparison cost analysis. *European Psychiatry*, *61*, 41-48.

Dalsgaard, S., Mortensen, P. B., Frydenberg, M., & Thomsen, P. H. (2013). Long-term criminal outcome of children with attention deficit hyperactivity disorder. *Criminal Behaviour and Mental Health*, *23*(2), 86-98.

Derks, E. M., Hudziak, J. J., & Boomsma, D. I. (2007). Why more boys than girls with ADHD receive treatment: a study of Dutch twins. *Twin Research and Human Genetics*, *10*(5), 765-770.

Faheem, M., Akram, W., Akram, H., Khan, M. A., Siddiqui, F. A., & Majeed, I. (2022). Gender-based differences in prevalence and effects of ADHD in adults: A systematic review. *Asian journal of psychiatry*, *75*, 103205.

Faraone, S. V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J. K., Ramos-Quiroga, J. A., Rohde, L. A., Sonuga-Barke, E. J., Tannock, R., & Franke, B. (2015). Attention-deficit/hyperactivity disorder. *Nat Rev Dis Primers*, *1*, 15020. https://doi.org/10.1038/nrdp.2015.20

Fedele, D. A., Lefler, E. K., Hartung, C. M., & Canu, W. H. (2012). Sex differences in the manifestation of ADHD in emerging adults. *Journal of attention disorders*, *16*(2), 109-117.

French, B., Nalbant, G., Wright, H., Sayal, K., Daley, D., Groom, M. J., Cassidy, S., & Hall, C. L. (2024). The impacts associated with having ADHD: an umbrella review. *Frontiers in psychiatry*, *15*, 1343314.

Galera, C., Collet, O., Orri, M., Navarro, M., Castel, L., Galesne, C., Reed, C., Brandt, V., Larsson, H., & Boivin, M. (2023). Prospective associations between ADHD symptoms and physical conditions from early childhood to adolescence: a population-based longitudinal study. *The Lancet Child & Adolescent Health*, *7*(12), 863-874.

Gershon, J., & Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of attention disorders*, *5*(3), 143-154.

Ghanizadeh, A., Salehi, A., & Moeini, S. R. (2019). Clinical presentation of attention-deficit hyperactivity disorder symptoms in terms of gender and chronological age. *International Journal of Community Based Nursing and Midwifery*, *7*(3), 241.

Glass, D. C., Kelsall, H. L., Slegers, C., Forbes, A. B., Loff, B., Zion, D., & Fritschi, L. (2015). A telephone survey of factors affecting willingness to participate in health research surveys. *BMC Public Health*, *15*, 1017. https://doi.org/10.1186/s12889-015-2350-9

Graetz, B. W., Sawyer, M. G., & Baghurst, P. (2005). Gender differences among children with DSM-IV ADHD in Australia. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(2), 159-168.

Grevet, E. H., Bau, C. H., Salgado, C. A., Fischer, A. G., Kalil, K., Victor, M. M., Garcia, C. R., Sousa, N. O., Rohde, L. A., & Belmonte-de-Abreu, P. (2006). Lack of gender effects on subtype outcomes in adults with attention–deficit/hyperactivity disorder: Support for the validity of subtypes. *European archives of psychiatry and clinical neuroscience*, *256*, 311-319.

Grskovic, J. A., & Zentall, S. S. (2010). Understanding ADHD in Girls: Identification and Social Characteristics. *International journal of special education*, *25*(1), 171-184.

Houghton, S., Lawrence, D., Hunter, S. C., Zadow, C., Kyron, M., Paterson, R., Carroll, A., Christie, R., & Brandtman, M. (2020). Loneliness accounts for the association between diagnosed attention deficit-hyperactivity disorder and symptoms of depression among adolescents. *Journal of Psychopathology and Behavioral Assessment*, *42*, 237-247.

Jong, A., Odoi, C. M., Lau, J., & J.Hollocks, M. (2024). Loneliness in Young People with ADHD: A Systematic Review and Meta-Analysis. *Journal of attention disorders*, *28*(7), 1063-1081. https://doi.org/10.1177/10870547241229096

Kamal, M., Al-Shibli, S., Shahbal, S., & Yadav, S. K. (2021). Impact of attention deficit hyperactivity disorder and gender differences on academic and social difficulties among adolescents in Qatari Schools. *Qatar medical journal*, *2021*(1), 11.

Keilow, M., Holm, A., & Fallesen, P. (2018). Medical treatment of Attention Deficit/Hyperactivity Disorder (ADHD) and children's academic performance. *PloS one*, *13*(11), e0207905. https://doi.org/10.1371/journal.pone.0207905

Kok, F. M., Groen, Y., Fuermaier, A. B., & Tucha, O. (2016). Problematic peer functioning in girls with ADHD: A systematic literature review. *PloS one*, *11*(11), e0165119.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., Barkley, R. A., Newcorn, J., Jensen, P., & Richters, J. (1994). DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *The American journal of psychiatry*, *151*(11), 1673-1685.

Lanier, J., Noyes, E., & Biederman, J. (2021). Mind Wandering (Internal Distractibility) in ADHD: A Literature Review. *J Atten Disord*, *25*(6), 885-890. https://doi.org/10.1177/1087054719865781

Liu, L., Wang, Y., Chen, W., Gao, Y., Li, H., Wang, Y., Chan, R. C., & Qian, Q. (2022). Network analysis of 18 attention-deficit/hyperactivity disorder symptoms suggests the importance of

"Distracted" and "Fidget" as central symptoms: Invariance across age, gender, and subtype presentations. *Frontiers in psychiatry*, *13*, 974283.

Loe, I. M., & Feldman, H. M. (2007). Academic and Educational Outcomes of Children With ADHD. *Journal of Pediatric Psychology*, *32*(6), 643-654. https://doi.org/10.1093/jpepsy/jsl054

Lovett, B. J., Fitzgerald, K., Schaberg, T. M., & Nelson, J. M. (2023). School-Based Interventions and Accommodations for ADHD. In J. L. Matson (Ed.), *Clinical Handbook of ADHD Assessment and Treatment Across the Lifespan* (pp. 451-467). Springer International Publishing. https://doi.org/10.1007/978-3-031-41709-2_17

Loyer Carbonneau, M., Demers, M., Bigras, M., & Guay, M.-C. (2021). Meta-analysis of sex differences in ADHD symptoms and associated cognitive deficits. *Journal of attention disorders*, *25*(12), 1640-1656.

Martin, J. (2024). Why are females less likely to be diagnosed with ADHD in childhood than males? *The Lancet Psychiatry*, *11*(4), 303-310.

Martin, J., Langley, K., Cooper, M., Rouquette, O. Y., John, A., Sayal, K., Ford, T., & Thapar, A. (2024). Sex differences in attention-deficit hyperactivity disorder diagnosis and clinical care: a national study of population healthcare records in Wales. *Journal of Child Psychology and Psychiatry*.

McKay, E., Kirk, H., Martin, R., & Cornish, K. (2023). Social difficulties in adolescent attention deficit hyperactivity disorder: Social motivation, social anxiety and symptom severity as contributing factors. *Journal of Clinical Psychology*, *79*(4), 1113-1129.

Meyer, J., Alaie, I., Ramklint, M., & Isaksson, J. (2022). Associated predictors of functional impairment among adolescents with ADHD—a cross-sectional study. *Child and adolescent psychiatry and mental health*, *16*(1), 29.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*, t. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269.

Monuteaux, M. C., Mick, E., Faraone, S. V., & Biederman, J. (2010). The influence of sex on the course and psychiatric correlates of ADHD from childhood to adolescence: A longitudinal study. *Journal of Child Psychology and Psychiatry*, *51*(3), 233-241.

Moukhtarian, T. R., Reinhard, I., Morillas-Romero, A., Ryckaert, C., Mowlem, F., Bozhilova, N., Moran, P., Ebner-Priemer, U., & Asherson, P. (2020). Wandering minds in attention-deficit/hyperactivity disorder and borderline personality disorder. *European Neuropsychopharmacology*, *38*, 98-109.

Mowlem, F., Agnew-Blais, J., Taylor, E., & Asherson, P. (2019). Do different factors influence whether girls versus boys meet ADHD diagnostic criteria? Sex differences among children with high ADHD symptoms. *Psychiatry research*, *272*, 765-773.

Mowlem, F. D., Rosenqvist, M. A., Martin, J., Lichtenstein, P., Asherson, P., & Larsson, H. (2019). Sex differences in predicting ADHD clinical diagnosis and pharmacological treatment. *Eur Child Adolesc Psychiatry*, *28*(4), 481-489. https://doi.org/10.1007/s00787-018-1211-3

Mowlem, F. D., Skirrow, C., Reid, P., Maltezos, S., Nijjar, S. K., Merwood, A., Barker, E., Cooper, R., Kuntsi, J., & Asherson, P. (2019). Validation of the mind excessively wandering scale and the relationship of mind wandering to impairment in adult ADHD. *Journal of attention disorders*, *23*(6), 624-634.

Mrug, S., Molina, B. S., Hoza, B., Gerdes, A. C., Hinshaw, S. P., Hechtman, L., & Arnold, L. E. (2012). Peer rejection and friendships in children with attention-deficit/hyperactivity disorder: Contributions to long-term outcomes. *Journal of abnormal child psychology*, *40*, 1013-1026.

Nijmeijer, J. S., Minderaa, R. B., Buitelaar, J. K., Mulligan, A., Hartman, C. A., & Hoekstra, P. J. (2008). Attention-deficit/hyperactivity disorder and social dysfunctioning. *Clinical psychology review*, *28*(4), 692-708.

Ohan, J. L., & Johnston, C. (2005). Gender appropriateness of symptom criteria for attention-deficit/hyperactivity disorder, oppositional-defiant disorder, and conduct disorder. *Child psychiatry and human development*, *35*, 359-381.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*, *5*, 1-10.

Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C., & Rohde, L. A. (2014). ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *International journal of epidemiology*, *43*(2), 434-442.

Powell, V., Agha, S. S., Jones, R. B., Eyre, O., Stephens, A., Weavers, B., Lennon, J., Allardyce, J., Potter, R., Smith, D., Thapar, A., & Rice, F. (2021). ADHD in adults with recurrent depression. *J Affect Disord*, *295*, 1153-1160. https://doi.org/10.1016/j.jad.2021.09.010

Quinn, P., & Wigal, S. (2004). Perceptions of girls and ADHD: results from a national survey. *Medscape general medicine*, *6*(2).

Quinn, P. O., & Madhoo, M. (2014). A review of attention-deficit/hyperactivity disorder in women and girls: uncovering this hidden diagnosis. *The primary care companion for CNS disorders*, *16*(3), 27250.

Quinn, P. O., & Nadeau, K. G. (2002). *Gender issues and AD/HD: Research, diagnosis, and treatment*. Advantage Books.

Schiavone, N., Virta, M., Leppämäki, S., Launes, J., Vanninen, R., Tuulio-Henriksson, A., Järvinen, I., Lehto, E., Michelsson, K., & Hokkanen, L. (2022). Mortality in individuals with childhood ADHD or subthreshold symptoms–a prospective perinatal risk cohort study over 40 years. *BMC psychiatry*, *22*(1), 325.

Shaw, M., Hodgkins, P., Caci, H., Young, S., Kahle, J., Woods, A. G., & Arnold, L. E. (2012). A systematic review and analysis of long-term outcomes in attention deficit hyperactivity disorder: effects of treatment and non-treatment. *BMC medicine*, *10*, 1-15.

Shifrin, J. G., Proctor, B. E., & Prevatt, F. F. (2010). Work performance differences between college students with and without ADHD. *Journal of attention disorders*, *13*(5), 489-496.

StataCorp. 2021. *Stata: Release 17*. Statistical Software. College Stations, TX: StataCorp LLC.

The EndNote Team. (2013). *EndNote*. In (Version EndNote 20) [64-bit]. Clarivate.

Vildalen, V. U., Brevik, E. J., Haavik, J., & Lundervold, A. J. (2019). Females with ADHD report more severe symptoms than males on the adult ADHD self-report scale. *Journal of attention disorders*, *23*(9), 959-967.

Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics*, *9*(3), 490-499.

Williamson, D., & Johnston, C. (2015). Gender differences in adults with attention-deficit/hyperactivity disorder: A narrative review. *Clin Psychol Rev*, *40*, 15-27. https://doi.org/10.1016/j.cpr.2015.05.005

Wimberley, T., Agerbo, E., Horsdal, H. T., Ottosen, C., Brikell, I., Als, T. D., Demontis, D., Børglum, A. D., Nordentoft, M., & Mors, O. (2020). Genetic liability to ADHD and substance use disorders in individuals with ADHD. *Addiction*, *115*(7), 1368-1377.

Wolraich, M. L., Hannah, J. N., Pinnock, T. Y., Baumgaertel, A., & Brown, J. (1996). Comparison of diagnostic criteria for attention-deficit hyperactivity disorder in a county-wide sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(3), 319-324.

Young, S., Adamo, N., Ásgeirsdóttir, B. B., Branney, P., Beckett, M., Colley, W., Cubbin, S., Deeley, Q., Farrag, E., & Gudjonsson, G. (2020). Females with ADHD: An expert consensus

statement taking a lifespan approach providing guidance for the identification and treatment of attention-deficit/hyperactivity disorder in girls and women. *BMC psychiatry*, *20*, 1-27.

Young, S., Asherson, P., Lloyd, T., Absoud, M., Arif, M., Colley, W. A., Cortese, S., Cubbin, S., Doyle, N., & Morua, S. D. (2021). Failure of healthcare provision for attention-deficit/hyperactivity disorder in the United Kingdom: a consensus statement. *Frontiers in psychiatry*, *12*, 649399.

Author contributions & roles:

- Tamara Williams (first author): study design, protocol writing & revision, screening & data extraction, analysis/synthesis, manuscript writing & revision
- Louise Horstmann: protocol/manuscript editing, screening & data extraction
- Annabelle Lim: screening
- Laiba Kayani: screening
- Abby Russel, Kate Langley, Tamsin Ford, Ann John, Kapil Sayal, Anita Thapar: protocol/manuscript editing
- Joanna Martin (senior author): study design, protocol writing, resolving data screening discrepancies, manuscript writing & editing, supervision

Contents

Supplementary text

Quality Assessment

JBI Critical Appraisal Checklist for Analytical Cross Sectional Studies

1. Were the criteria for inclusion in the sample clearly defined?
2. Were the study subjects and the setting described in detail?
3. Was the exposure measured in a valid and reliable way?
4. Were objective, standard criteria used for measurement of the condition?
5. Were confounding factors identified?
6. Were strategies to deal with confounding factors stated?
7. Were the outcomes measured in a valid and reliable way?
8. Was appropriate statistical analysis used?

Question four (“Were objective, standard criteria used for measurement of the condition?”) was removed as it was unnecessary, repeating question three (“was the exposure measured in a valid and reliable way?”) for measurement of ADHD.

Questions five (“Were confounding factors identified?”) and six (“Were strategies to deal with confounding factors stated?”) were collapsed into one question (“Were confounding factors identified and dealt with?”).

Mind wandering item measurement

Moukhtarian et al. (2020)

1. How much is your mind on what you are doing or elsewhere now?
2. Were you thinking about many different things at once now?
3. How often do new thoughts keep popping into your head now?
4. How hard is it to stick your thoughts to one thing at a time?
5. My mind just goes – I cannot switch off.

Tables

Table S1. PRISMA-P checklist.

Table S1. Preferred Reporting Items for Systematic Review and Meta-Analyses Protocol (PRISMA-P) checklist.

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Page 3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page 4-5
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page 5
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 5-6
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page 6 & Supplementary material Table S2
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page 6-7 & Supplementary material Table S3
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Page 6-7
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Page 7
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Page 6
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Page 7
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of	Page 7

Section and Topic	Item #	Checklist item	Location where item is reported
		automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Page 8
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Page 7-8
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Page 8
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page 7-8
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Page 7-8
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Page 8
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Page 7-8
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N.A.
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Page 7-8
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page 8 & Figure 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Supplementary material Table S4
Study characteristics	17	Cite each included study and present its characteristics.	Page 8-10 & Table 2
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Page 10-11 & Supplementary material Figure S1
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Figures 3 & 4
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Page 20-22
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing	Page 15-20 & Figure 3 and 4

Section and Topic	Item #	Checklist item	Location where item is reported
		groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Page 15 & 18
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Page. 15 & 17-18
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N.A.
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Figures 3 & 4
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Page 22-26
	23b	Discuss any limitations of the evidence included in the review.	Page 26
	23c	Discuss any limitations of the review processes used.	Page 26
	23d	Discuss implications of the results for practice, policy, and future research.	Page 26-27
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Page 5
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page 5
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N.A.
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 28
Competing interests	26	Declare any competing interests of review authors.	Page 28
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	N.A.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Table S2 - Databases searched

Table S2. *Databases searched and the coverage of date*

Database	Coverage
Ovid	
Medline	1978 - present
Embase	1978 - present
APA PsychInfo	1978 - present
ProQuest	1978 - present
EBSCO	
Education Resources Information Centre	1978 - present
British Education Index	1978 - present

Table S3 - Search strategies

Table S3. *Search strategies*

APA Psychinfo	
1	adhd.ti.
2	Attention deficit disorder with hyperactivity/
3	Attention deficit hyperactivity disorder.ti.
4	Attention deficit disorder.ti.
5	Hyperkinetic disorder.ti.
6	1 or 2 or 3 or 4 or 5
7	Woman.tw.
8	Women.tw.
9	Female*.tw.
10	Girl*.tw.
11	Gender*.tw.
12	Sex.tw.
13	Sexes.tw.
14	7 or 8 or 9 or 10 or 11 or 12 or 13
15	Symptom*.tw.
16	Trait*.tw.
17	Item*.tw.
18	15 or 16 or 17
19	6 and 14 and 18
20	19
21	Limit 19 to yr=”1987-Current”
EMBASE	
1	adhd.ti.
2	Attention deficit disorder with hyperactivity/
3	Attention deficit hyperactivity disorder.ti.
4	Attention deficit disorder.ti.
5	Hyperkinetic disorder.ti.
6	1 or 2 or 3 or 4 or 5
7	Woman.tw.
8	Women.tw.
9	Female*.tw.
10	Girl*.tw.
11	Gender*.tw.
12	Sex.tw.
13	Sexes.tw.
14	7 or 8 or 9 or 10 or 11 or 12 or 13

15	Symptom*.tw.
16	Trait*.tw.
17	Item*.tw.
18	15 or 16 or 17
19	6 and 14 and 18
20	19
21	Limit 19 to yr=”1987-Current”
MEDLINE	
1	adhd.ti.
2	Attention deficit disorder with hyperactivity/
3	Attention deficit hyperactivity disorder.ti.
4	Attention deficit disorder.ti.
5	Hyperkinetic disorder.ti.
6	1 or 2 or 3 or 4 or 5
7	Woman.tw.
8	Women.tw.
9	Female*.tw.
10	Girl*.tw.
11	Gender*.tw.
12	Sex.tw.
13	Sexes.tw.
14	7 or 8 or 9 or 10 or 11 or 12 or 13
15	Symptom*.tw.
16	Trait*.tw.
17	Item*.tw.
18	15 or 16 or 17
19	6 and 14 and 18
20	19
21	Limit 19 to yr=”1987-Current”
ProQuest	
noft(adhd OR "attention deficit hyperactivity disorder" OR “attention deficit disorder” OR “hyperkinetic disorder”) AND noft(woman OR women OR female* OR girl* OR gender* OR sex OR sexes) AND noft(symptom* OR trait* OR item*)	
EBSCO (Education Resource Information Centre (ERIC) & British Education Index (BEI))	
(adhd OR “attention deficit hyperactivity disorder”) AND (woman OR women OR female* OR girl* OR gender* OR sex OR sexes) AND (symptom* OR trait* OR item*)	

Table S4 - Studies excluded

Table S4. *Studies excluded from the systematic review, including reasons for exclusion*

Year	Author	Reason for exclusion
1987	Gada	Did not include item-level results
1987	Levy et al.	Did not include item-level results
1987	Roth	Did not include item-level results
1989	Horn et al.	Did not include item-level results
1989	Shealy	Did not include item-level results
1990	Shekim et al.	Did not include item-level results
1991	Brown et al.	Did not include item-level results
1991	DuPaul	Did not include item-level results
1992	Zohar	Did not include item-level results
1994	Wilson & Berman	Did not include item-level results

1996	Murphy et al.	Did not include relevant comparisons
1996	Thomeer	Did not include item-level results
1997	Carlson et al.	Did not include item-level results
1997	Kern	Did not include item-level results
1997	Martin	Did not include relevant comparisons
1997	McCoy et al.	Did not include item-level results
1997	March et al.	Did not include item-level results
1997	Millstein	Did not include relevant comparisons
1997	Roy-Byrne et al.	Did not include item-level results
1997	Rucklidge & Kaplan	Did not include item-level results
1998	Dominquez & Shapiro	Did not include item-level results
1998	Holl et al.	Did not include item-level results
1998	Ambler	Did not include item-level results
1998	Katz et al.	Did not include item-level results
1999	Austin	Did not include item-level results
1999	Bu-Haroon et al.	Did not include item-level results
1999	Gomez et al.	Did not include item-level results
1999	Biederman et al.	Did not include relevant comparisons
1999	Nolan et al.	Did not include item-level results
1999	Pineda et al.	Did not include item-level results
1999	Scahill et al.	Did not include item-level results
2000	Faraone et al.	Did not include item-level results
2000	Chang & Chuang	Did not include relevant comparisons
2000	Guardiola et al.	Did not include item-level results
2000	Shulman	Did not include item-level results
2000	Rucklidge & Kaplan	Did not include item-level results
2001	Cuffe et al.	Did not include item-level results
2001	DuPaul et al.	Did not include item-level results
2001	Kato et al.	Did not include item-level results
2001	Deshazo	Did not include item-level results
2001	Gadow et al.	Did not include item-level results
2001	Newcorn et al.	Did not include item-level results
2001	Nolan et al.	Did not include item-level results
2001	O'Donnell et al.	Did not include relevant comparisons
2001	Owens & Hoza	Did not include item-level results
2001	Rucklidge & Tannock	Did not include item-level results
2002	Hartung et al.	Did not include item-level results
2002	Gadow & Nolan	Did not include item-level results
2002	Thunstrom	Did not include item-level results
2002	Shaikh	Did not include item-level results
2002	Young	Did not include item-level results
2003	Hoksbergen et al.	Did not include item-level results
2003	Kumar & Steer	Did not include item-level results
2003	McCann	Did not include item-level results
2003	Lewczyk et al.	Did not include item-level results

2003	Mikami & Hinshaw.	Did not include item-level results
2003	Morgan et al.	Did not include item-level results
2004	Klassen et al.	Did not include item-level results
2004	Biederman & Faraone	Did not include item-level results
2004	Airaksinen et al.	Did not include item-level results
2004	Ersan et al.	Did not include item-level results
2004	Marks	Did not include relevant comparisons
2004	Oncu et al.	Did not include item-level results
2005	Cuffe et al.	Did not include item-level results
2005	Fleming et al.	Did not include item-level results
2005	Kooij et al.	Did not include item-level results
2005	Diamantopoulou et al.	Did not include item-level results
2005	Lee	Did not include item-level results
2005	Levy et al.	Did not include item-level results
2005	Shaw-Zirt et al.	Did not include item-level results
2005	Neuman et al.	Did not include item-level results
2005	Parker et al.	Did not include relevant comparisons
2006	Gross-Tsur et al.	Did not include item-level results
2006	Hinshaw et al.	Did not include item-level results
2006	Al-Haggar et al.	Did not include item-level results
2006	Bener et al.	Did not include item-level results
2006	Lee	Did not include item-level results
2006	Lee et al.	Did not include item-level results
2006	Mugnaini et al.	Did not include item-level results
2006	Novik et al.	Did not include item-level results
2006	Peterson	Did not include item-level results
2006	Waschbusch & King	Did not include item-level results
2007	Diamantopoulou et al.	Did not include item-level results
2007	Lahey et al.	Did not include item-level results
2007	Gadow et al.	Did not include item-level results
2007	Hebrani et al.	Did not include item-level results
2007	Bauermeister et al.	Did not include item-level results
2007	Michanie	Did not include item-level results
2007	Lee et al.	Did not include item-level results
2007	van Lier et al.	Did not include item-level results
2007	Mahone & Hoffman	Did not include item-level results
2007	Ponde & Freire	Did not include item-level results
2007	Posner et al.	Did not include relevant comparisons
2007	Rucklidge et al.	Did not include item-level results
2008	Ek et al.	Did not include item-level results
2008	Bener et al.	Did not include item-level results
2008	Fliers et al.	Did not include item-level results
2008	Gau et al.	Did not include item-level results
2008	Ghanizadeh	Did not include item-level results
2008	Huss et al.	Did not include item-level results

2008	Mikami	Did not include item-level results
2008	Sonuga-Barke et al.	Did not include item-level results
2008	Martel et al.	Did not include item-level results
2008	Lee et al.	Did not include item-level results
2008	Langberg et al.	Did not include item-level results
2008	Lee et al.	Did not include item-level results
2008	Reiersen et al.	Did not include item-level results
2008	Robinson et al.	Did not include item-level results
2008	Thorell & Rydell	Sample did not have ADHD
2008	Ruchkin et al.	Sample did not have ADHD
2009	Coutinho et al.	Did not include item-level results
2009	Billingsley-Jackson	Did not include item-level results
2009	Chang	Did not include item-level results
2009	DeGrass	Did not include relevant comparisons
2009	Serra-Pinheiro et al.	Did not include item-level results
2009	Lavigne et al.	Did not include item-level results
2009	Owens et al.	Did not include item-level results
2009	Lara et al.	Did not include item-level results
2009	McGillivray & Baker	Did not include item-level results
2009	Mahone et al.	Did not include item-level results
2009	Mikami et al.	Did not include item-level results
2009	Wilens et al.	Insufficient detail of item-level results
2009	Leren	Sample did not have ADHD
2009	Martel	Did not include item-level results
2009	Marton	Did not include item-level results
2009	Monahan	Did not include item-level results
2009	Philipsen	Did not include item-level results
2009	Soma et al.	Did not include item-level results
2009	Shuhr et al.	Did not include relevant comparisons
2009	Westerlund et al.	Did not include item-level results
2009	Weiner & Mak	Did not include item-level results
2010	Carducci & Lukomski	Did not include item-level results
2010	Chronis-Tuscano et al.	Did not include item-level results
2010	Joseph	Did not include item-level results
2010	Bitter et al	Did not include item-level results
2010	Groskovic & Zentall	Did not include relevant comparisons
2010	Gilmore	Did not include item-level results
2010	Bathiche	Did not include item-level results
2010	Ghanizadeh	Did not include item-level results
2010	Fergusson et al.	Did not include item-level results
2010	Kopp	Did not include relevant comparisons
2010	Katz	Did not include relevant comparisons
2010	Lipowska et al.	Did not include item-level results
2010	Martel et al.	Did not include item-level results
2010	Rivero	Did not include item-level results

2010	Langberg et al.	Did not include relevant comparisons
2010	Mikami et al.	Did not include item-level results
2010	Retz-Junginger et al.	Insufficient detail of item-level results
2010	Ramtekkar et al.	Did not include item-level results
2010	Sobanski et al.	Did not include item-level results
2011	Gomez & Hafetz	Did not include item-level results
2011	Atwoli et al.	Did not include item-level results
2011	Duric & Elgen	Did not include item-level results
2011	Hassan et al.	Did not include item-level results
2011	Barkely et al.	Did not include item-level results
2011	Ambuabunos et al.	Did not include item-level results
2011	Khamis	Did not include item-level results
2011	Elkins et al.	Did not include item-level results
2011	Kumar et al.	Did not include item-level results
2011	Babinski et al.	Did not include item-level results
2011	Lecendreux et al.	Did not include item-level results
2011	Mikami et al.	Did not include item-level results
2011	Russell et al.	Did not include item-level results
2011	Mendez et al.	Did not include item-level results
2011	McKelvy et al.	Did not include item-level results
2011	Langberg et al.	Did not include item-level results
2011	Lundervold et al.	Did not include item-level results
2011	McClernon et al.	Did not include item-level results
2011	Lehtinen	Did not include item-level results
2011	Leung & Pei	Did not include item-level results
2011	Mick et al.	Did not include item-level results
2011	Muller et al.	Did not include item-level results
2011	Park et al.	Did not include item-level results
2011	Simon	Did not include item-level results
2011	Taylor et al.	Did not include item-level results
2012	Daigle & Vingilis	Did not include item-level results
2012	Jenahi et al.	Did not include item-level results
2012	Connor & Ford	Did not include item-level results
2012	Cahill et al.	Did not include item-level results
2012	Holbrook	Did not include relevant comparisons
2012	Abrines et al.	Did not include item-level results
2012	Das et al.	Did not include relevant comparisons
2012	Jaconis & Hartung	Did not include item-level results
2012	Hinshaw et al.	Did not include item-level results
2012	Adamou et al.	Did not include item-level results
2012	Ebejer et al.	Did not include item-level results
2012	Abrines et al.	Did not include item-level results
2012	Sciberras et al.	Did not include item-level results
2012	Miller et al.	Did not include item-level results
2012	Zwann et al.	Did not include item-level results

2012	Lidzba et al.	Did not include item-level results
2012	Miller et al.	Did not include item-level results
2012	Macek et al.	Did not include item-level results
2012	Nelson & Gregg	Did not include item-level results
2012	Sonnby et al.	Did not include item-level results
2012	Takeda et al.	Did not include relevant comparisons
2013	Evans et al.	Did not include item-level results
2013	Donfraesco et al.	Did not include item-level results
2013	Becker et al.	Did not include item-level results
2013	Gomez	Did not include item-level results
2013	Ajinkya et al.	Did not include item-level results
2013	Jahangard et al.	Did not include item-level results
2013	Loya	Did not include item-level results
2013	Major et al.	Did not include item-level results
2013	Skogli et al.	Did not include item-level results
2013	Yucwe et al.	Did not include item-level results
2013	Martel	Did not include item-level results
2013	Martel	Did not include item-level results
2013	Tseng	Did not include item-level results
2013	Usami et al.	Did not include item-level results
2014	Jin et al.	Did not include item-level results
2014	DuPaul et al.	Did not include item-level results
2014	Caci et al.	Did not include item-level results
2014	Burcu Ayaz et al.	Did not include item-level results
2014	Elumour & Thabet	Did not include item-level results
2014	Makransky et al.	Did not include item-level results
2014	Langberg et al.	Did not include item-level results
2014	Liu et al.	Did not include item-level results
2014	Lui	Did not include relevant comparisons
2014	Nazar et al.	Did not include item-level results
2014	O'Callaghan & Sharma	Did not include item-level results
2014	Panevska et al.	Did not include item-level results
2015	Flagg	Did not include item-level results
2015	Ghanizadeh	Did not include item-level results
2015	Arnett et al.	Did not include item-level results
2015	Al-Mamari et al.	Did not include item-level results
2015	Green et al.	Did not include item-level results
2015	Gumus et al.	Did not include item-level results
2015	Kercood et al.	Did not include item-level results
2015	Gao et al.	Did not include item-level results
2015	Lopez et al.	Did not include item-level results
2015	Sasaki et al.	Did not include item-level results
2015	Lefler et al.	Did not include item-level results
2015	Nicolau et al.	Did not include item-level results
2015	Lin et al.	Did not include item-level results

2015	Morstedt et al.	Did not include item-level results
2015	Panevska et al.	Did not include item-level results
2015	Prevatt et al.	Did not include item-level results
2015	Sanchez et al.	Did not include item-level results
2015	Sonnby et al.	Did not include item-level results
2015	Wang et al.	Did not include item-level results
2015	Yell & Sherry	Did not include item-level results
2016	Jarrett	Did not include item-level results
2016	Gomez	Did not include item-level results
2016	Kitsune et al.	Did not include item-level results
2016	Farooq et al.	Did not include item-level results
2016	Lahey et al.	Did not include item-level results
2016	Adler et al.	Did not include item-level results
2016	Simsek et al.	Did not include item-level results
2016	Meinzer et al.	Did not include item-level results
2016	Newark et al.	Did not include item-level results
2016	Noren et al.	Did not include item-level results
2016	Ozten et al.	Did not include item-level results
2016	Rimal & Pokharel	Did not include item-level results
2016	Rinsky	Did not include item-level results
2016	Soendergaard et al.	Did not include item-level results
2017	Ahmad et al.	Did not include item-level results
2017	Huang et al.	Did not include relevant comparisons
2017	Dallos et al.	Did not include item-level results
2017	Gokce et al.	Did not include item-level results
2017	Bendiksen et al.	Did not include item-level results
2017	Bakshi	Did not include item-level results
2017	Davidsson et al.	Did not include item-level results
2017	Corbisiero et al.	Did not include item-level results
2017	Guelzow et al.	Did not include item-level results
2017	Bijlenga et al.	Did not include item-level results
2017	Becker	Did not include item-level results
2017	Leung & Chan	Did not include item-level results
2017	Leno et al.	Did not include item-level results
2017	Lundervold	Did not include item-level results
2017	Mokobane et al.	Did not include item-level results
2017	Owens et al.	Did not include item-level results
2018	Balaz et al.	Did not include item-level results
2018	Cerrillo-Urbina et al.	Did not include item-level results
2018	Gordon	Did not include item-level results
2018	Adamis et al.	Did not include item-level results
2018	Amiri et al.	Did not include item-level results
2018	Barbaresi et al.	Did not include item-level results
2018	Canals et al.	Did not include item-level results
2018	Alex et al.	Did not include relevant comparisons

2018	Loggans	Did not include item-level results
2018	Palacious-Cruz et al.	Did not include relevant comparisons
2018	Lapalme et al.	Did not include item-level results
2018	Pi Davanzo et al.	Did not include item-level results
2018	Madsen et al.	Did not include item-level results
2018	Major	Did not include item-level results
2018	Millenet et al.	Did not include item-level results
2018	Nelson & Liebel	Did not include item-level results
2018	Oh et al.	Did not include item-level results
2018	Oie et al.	Did not include item-level results
2018	Ramy et al.	Did not include item-level results
2018	Weissenberger et al.	Did not include item-level results
2019	Gomez-Benito et al.	Did not include item-level results
2019	Anker et al.	Did not include item-level results
2019	Hayashi et al.	Did not include item-level results
2019	Becker et al.	Did not include item-level results
2019	Ahmad et al.	Did not include item-level results
2019	Babinski et al.	Did not include item-level results
2019	Choi et al.	Did not include item-level results
2019	Kivumbi et al.	Did not include item-level results
2019	Biederman et al.	Did not include relevant comparisons
2019	Ben-Sheetrit et al.	Did not include item-level results
2019	Adler et al.	Did not include relevant comparisons
2019	Leopold et al.	Did not include item-level results
2019	Mowlem et al.	Did not include item-level results
2019	Sevincok et al.	Did not include item-level results
2019	Li et al.	Did not include item-level results
2019	Mahendiran et al.	Did not include item-level results
2019	Martin et al.	Did not include item-level results
2019	Mitchison & Njardvik	Did not include item-level results
2019	Mowlem et al.	Did not include relevant comparisons
2019	Salvi et al.	Did not include item-level results
2019	Slobodin & Davidovitch	Did not include item-level results
2019	Taylor et al.	Did not include item-level results
2020	DuPaul et al.	Sample Did not have ADHD
2020	Al-Yagon et al.	Did not include item-level results
2020	Isaksson et al.	Did not include item-level results
2020	Anker et al.	Did not include item-level results
2020	Figueiredo et al.	Did not include item-level results
2020	Mil et al.	Did not include item-level results
2020	Lan et al.	Did not include item-level results
2020	Mphahlele et al.	Did not include item-level results
2020	Lee et al.	Did not include item-level results
2020	Levy et al.	Did not include item-level results
2020	Mahajnah et al.	Did not include item-level results

2020	Margherio	Did not include item-level results
2020	Mochrie et al.	Did not include item-level results
2020	Molavi et al.	Did not include item-level results
2020	Ntiakoh-Ayipah et al.	Did not include item-level results
2020	Regan & Tubman	Did not include item-level results
2020	Rokeach	Did not include item-level results
2020	Stibbe et al.	Did not include item-level results
2021	Ahmad et al.	Did not include item-level results
2021	Dobrean et al.	Did not include item-level results
2021	Burns et al.	Did not include item-level results
2021	Brancati et al.	Did not include item-level results
2021	Kiraz et al.	Did not include item-level results
2021	Hoang et al.	Did not include item-level results
2021	Babinski et al.	Did not include item-level results
2021	Akca et al.	Did not include item-level results
2021	Anker et al.	Did not include item-level results
2021	DuPaul et al.	Did not include item-level results
2021	Assari	Did not include item-level results
2021	Al-Ani et al.	Did not include item-level results
2021	Lim et al.	Did not include item-level results
2021	Shoval et al.	Did not include item-level results
2021	Looby et al.	Did not include item-level results
2021	Lau et al.	Did not include item-level results
2021	Shim et al.	Did not include item-level results
2021	Lugo-C & Elas	Did not include item-level results
2021	Martin et al.	Did not include item-level results
2021	Masi et al.	Did not include item-level results
2021	Suh	Did not include item-level results
2021	Sultan et al.	Did not include relevant comparisons
2022	Frick et al.	Did not include item-level results
2022	Flores et al.	Did not include item-level results
2022	Bodalski et al.	Did not include item-level results
2022	Kniola & Talpade	Did not include item-level results
2022	Gomes et al.	Did not include item-level results
2022	Fioravante et al.	Did not include item-level results
2022	Al-Yagon & Borenstein	Did not include item-level results
2022	De Rossi et al.	Did not include item-level results
2022	Barney	Did not include item-level results
2022	Krauss & Schellengerg	Did not include item-level results
2022	Leffa et al.	Did not include item-level results
2022	McQuade	Did not include item-level results
2022	Li & Guo	Did not include item-level results
2022	Selinus	Did not include item-level results
2022	Levy et al.	Did not include item-level results
2022	Li et al.	Did not include item-level results

2022	Morgan et al.	Did not include item-level results
2022	Trognon & Richard	Did not include item-level results
2023	Shakeshaft et al.	Did not include item-level results
2023	Riglin et al.	Did not include item-level results

Table S5 – Child ADHD symptom endorsement rates

Table S5. *Average child ADHD symptom endorsement rates for items with significant sex differences*

Item	Gender	Percentage (%)
Fails to sustain attention in tasks	Females	75.27
	Males	73.45
Often fails to follow through on instructions from others	Females	71.01
	Males	72.94
Often easily distracted	Females	76.30
	Males	75.91
Often fidgets	Females	59.63
	Males	68.72
Difficulty remaining seated when required	Females	46
	Males	53.01
Runs about or climbs in situations when not appropriate	Females	28.91
	Males	36.35
Always on the go	Females	42.60
	Males	62.45
Often noisy in playing	Females	40.95
	Males	50.41
Difficulty waiting turn	Females	43
	Males	42.91
Often blurts out answers	Females	52.01
	Males	50.90
Often interrupts others	Females	54.95
	Males	56.64

Table S6 – Adult ADHD symptom endorsement rates

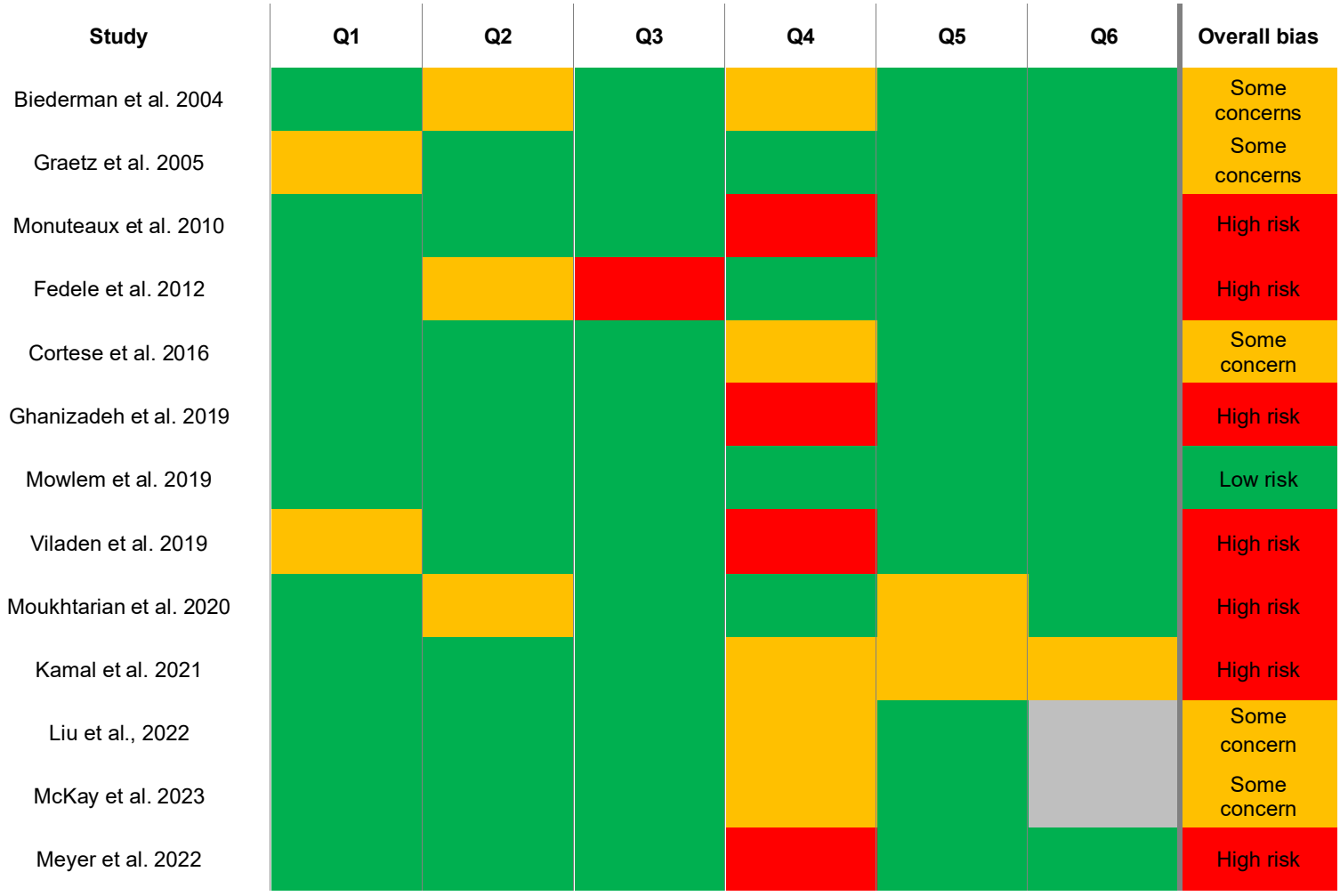
Table S6. *Average adult ADHD symptom endorsement rates for items with significant sex differences*

Item	Gender	Percentage (%)
Often easily distracted	Females	86.24
	Males	80.26
Often has difficulty organising tasks	Females	70.65
	Males	67.71
Often fidgets	Females	59.63
	Males	68.72
Often blurts out answers	Females	69.20
	Males	68.76
Often talks excessively	Females	65.08
	Males	52.66

Figures

Figure S1 – Risk of bias

Figure S1. *Risk of bias of studies included in the systematic review*



Green = yes, orange = unclear, red = no, grey = not applicable

Note. Q1 = Were the criteria for inclusion in the sample clearly defined?, Q2 = Were the study subjects and the setting described in detail?, Q3 = Was ADHD measured in a valid, objective and reliable way?, Q4 = Were confounding factors identified and dealt with?, Q5 = Were the outcomes measured in a valid and reliable way?, Q6 = Was appropriate statistical analysis used?