Clinical Psychology Review 83 (2021) 101954

Contents lists available at ScienceDirect

Clinical Psychology Review

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

REV

Cognitive behavior therapy for externalizing disorders in children and adolescents in routine clinical care: A systematic review and meta-analysis

Eili N. Riise [a,*], Gro Janne H. Wergeland [b,c], Urdur Njardvik [d], Lars-Göran Öst [e,f]

[a] Department of Child and Adolescent Psychiatry, District General Hospital of Førde, Førde, Norway
[b] Department of Child and Adolescent Psychiatry, Division of Psychiatry, Haukeland University Hospital, Bergen, Norway
[c] Department of Clinical Medicine, Faculty of Medicine, University of Bergen, Norway
[d] Department of Psychology, University of Iceland, Reykjavik, Iceland
[e] Department of Psychology, Stockholm University, Stockholm, Sweden
[f] Bergen Center for Brain Plasticity, Haukeland University Hospital, Bergen, Norway

ARTICLE INFO

Keywords:
Externalizing disorders
Routine care
Cognitive behavior therapy
Children
Meta-analysis

ABSTRACT

Various Cognitive Behavioral Therapy (CBT) programs for externalizing disorders in children and adolescents are supported by a substantial body of empirical evidence. Most of the research evidence comes from efficacy studies conducted in university settings, but there is less knowledge about the effect of these treatments in routine clinical care. The purpose of this meta-analysis was to investigate the effectiveness of CBT in non-university settings for Attention Deficit Hyperactivity Disorder (ADHD), Conduct Disorder (CD) and Oppositional Defiant Disorder (ODD). Embase OVID, Ovid MEDLINE and PsycINFO were systematically searched for eligible studies published up to May 2020. In total, 51 treatment effectiveness studies involving 5295 patients were included. The average within-group effect size at post-treatment was significant ($g = 0.91$), and there were large effect sizes for both ADHD ($g = 0.80$) and CD/ODD ($g = 0.98$). At post treatment, remission rates were 38% for ADHD and 48% for CD/ODD, and the overall attrition rate was 14%. Benchmarking against efficacy studies showed that CBT in routine clinical care yields remission rates, within-group effect sizes and attrition rates that are very similar to those found in university settings. The findings support the transportability of CBT for externalizing disorders from university settings to routine clinical care.

PROSPERO registration: CRD42020147524.

Externalizing behavior disorders are among the most common reasons for referral to mental health services for children and adolescents and include the diagnoses of Attention Deficit Hyperactivity Disorder (ADHD), Conduct Disorder (CD) and Oppositional Defiant Disorder (ODD). ADHD has an estimated prevalence of 7.2% (Thomas, Sanders, Doust, Beller, & Glasziou, 2015) whereas CD and ODD have estimated prevalence rates of 3.5% and 2.8%, respectively (O'Connell, Boat, & Warner, 2009). Although ADHD is classified as a neurodevelopmental disorder and CD and ODD as disruptive behavior disorders they are all characterized by problems with aggression, impulse-control and rule-breaking behavior. Children with externalizing disorders have poorer academic and social functioning than other children (Clark, Prior, & Kinsella, 2002; DuPaul, Morgan, Farkas, Hillemeier, & Maczuga, 2016) and have a higher risk of negative outcomes in adulthood, including an increased risk of unemployment, criminality, financial problems, and increased mortality rates (Franke et al., 2018; Scott et al., 2017). They are also at a higher risk of developing emotional problems, as externalizing disorders are highly comorbid with both anxiety and depression, which appears to lead to more severe impairment than when children are diagnosed with either disorder alone (Jarrett & Ollendick, 2008; Schatz & Rostain, 2006; Wolff & Ollendick, 2006). Furthermore, this vulnerability for the development of emotional disorders appears to be present in adulthood as well, as childhood ODD has been found to predict depression in young adulthood (Copeland, Shanahan, Costello, & Angold, 2009). The relatively high prevalence of externalizing disorders and the severe consequences for the individual and the society highlights the necessity of providing evidence-based interventions.

Various treatment programs for externalizing disorders have been supported by findings from numerous randomized controlled trials (RCTs). There is evidence for the efficacy of a wide range of CBT-

* Corresponding author at: Department of Child and Adolescent Psychiatry, District General Hospital of Førde, N-6807 Førde, Norway.
E-mail address: eiliriise@gmail.com (E.N. Riise).

https://doi.org/10.1016/j.cpr.2020.101954
Received 25 June 2020; Received in revised form 9 October 2020; Accepted 23 November 2020
Available online 7 December 2020

oriented treatments. Programs with well-established efficacy for externalizing disorders include behavior therapy programs for parents, children and/or teachers and cognitive interventions aimed at increasing children's self-control and problem solving skills as well as programs addressing the organizational problems of children with ADHD (Evans, Owens, & Bunford, 2014; Kaminski & Claussen, 2017; McCart & Sheidow, 2016). Although it is encouraging that a growing body of methodologically sound trials are carried out, it has been argued that more attention should be directed at investigations of clinical effectiveness in routine clinical care (Tolin, McKay, Forman, Klonsky, & Thombs, 2015). Efficacy trials from university settings are often carried out with a methodologically stringent procedure which produces results with high internal validity. However, the circumstances in such trials may differ from "real world" clinical settings. For instance, participants are usually actively recruited through advertisements and willing to risk being randomized to a control condition. Also, it is common for efficacy studies to apply more exclusion criteria in order to obtain a more homogenous sample of participants. ADHD is, for example, highly comorbid with other disorders (e.g. Owens & Hinshaw, 2013; Steinhausen et al., 2006) and a comorbid diagnosis of ADHD and ODD or CD has been found to increase the severity of ADHD symptoms (Efron, Bryson, Lycett, & Sciberras, 2016; Jerrell, McIntyre, & Park, 2015; Takeda, Ambrosini, deBerardinis, & Elia, 2012). Similarly, children and adolescents with CD are frequently diagnosed with comorbid depression and typically show greater impairment than children diagnosed with either disorder alone (Wolff & Ollendick, 2006). This could result in patients in efficacy trials having less severe symptoms and that the sample may not be representative of patients in routine clinical settings. Furthermore, the therapists involved in efficacy trials are often experts in the intervention under investigation, and are dedicated part- or full time to the research project as opposed to clinicians in routine care who carry large caseloads with a variety of disorders. The differences between university settings and routine clinical care suggest that findings from methodologically stringent efficacy trials may not be transferable to clinical practice.

Previous reviews comparing studies from routine clinical care to efficacy studies from university settings have shown different results for different disorders. A meta-analysis of effectiveness studies for children with autism spectrum disorder found smaller effect-sizes for community-based studies than for university-based clinical trials, indicating a gap between research settings and routine clinical care (Nahmias, Pellecchia, Stahmer, & Mandell, 2019). Conversely, in a recent meta-analysis of CBT for internalizing disorders in children and adolescents in routine clinical care, Wergeland, Riise and Öst (2021) reported treatment outcomes in routine clinical care comparable to those in university settings. To better understand how CBT-oriented treatments for externalizing disorders work in clinical practice, it is important to evaluate how evidence-based treatment programs perform in real world settings. During the last 10 years a substantial number of effectiveness studies in non-university settings have been published, and findings from such studies provide important information to clinicians and decision makers as to whether empirically supported treatments can be expected to yield the intended results when implemented in routine clinical care. Despite the importance of effectiveness studies for dissemination of empirically validated treatments, systematic reviews of treatments for externalizing disorders in routine clinical care are scarce. Several systematic reviews of psychological treatment have been performed in the past decade for ADHD and CD/ODD respectively, but have not focused on effectiveness studies specifically (Bakker, Greven, Buitelaar, & Glennon, 2017; Burkey et al., 2018; Fabiano, Schatz, Aloe, Chacko, & Chronis-Tuscano, 2015; Kaminski & Claussen, 2017; Schatz et al., 2020; Sibley, Kuriyan, Evans, Waxmonsky, & Smith, 2014; Sonuga-Barke et al., 2013). Lee, Horvath, and Hunsley (2013) carried out a review of effectiveness studies of child therapy and reported an increased number of effectiveness studies from 2007, high completion rates, and maintenance of treatment gains in follow-up assessments. They also found encouraging evidence for the effectiveness of parent training for disruptive behavior problems in routine clinical practice. Although a variability in treatment outcome data was found, the authors concluded that positive results could be obtained in various treatment settings with children presenting with multiple problems (Lee et al., 2013). This review was, however, limited to 13 studies of parent training which did not report inclusion criteria, making it difficult to ascertain whether the results apply to specific disorders such as ADHD, CD or ODD. Therefore, further assessment of effectiveness studies seems called for and a meta-analysis of the effectiveness of CBT in the treatment of externalizing disorders in routine clinical care is warranted.

The present meta-analysis is the first to review effectiveness studies of CBT and behavior therapy (BT) for ADHD, CD and ODD in children and adolescents. The included studies investigated the effect of empirically supported treatment programs delivered by practising clinicians in non-university settings to patients referred for treatment through usual clinical routes (i.e. not actively recruited to a research project).

The aims of the present meta-analysis are:

1. To evaluate the effectiveness of CBT for children and adolescents with ADHD, CD and ODD in routine clinical care.
2. To investigate moderators of treatment outcome in effectiveness studies.
3. To investigate whether there are differences between effectiveness and efficacy studies in terms of background data or treatment data.
4. To evaluate whether there are differences in the effect of CBT for externalizing disorders between effectiveness and efficacy studies.

1. Method

The protocol for this meta-analysis was pre-registered at PROSPERO with ID CRD42020147524. The meta-analysis was conducted according to the PRISMA guidelines (Liberati et al., 2009), and reported according to AMSTAR 2 (Shea et al., 2017), see online Supplement S7 and S8. Two independent raters were involved during each step of the study, except for the screening of title and abstract conducted by one rater only.

The meta-analysis was designed according to the PICOS acronym in the following way:

- Population: children and adolescents with ADHD, ODD and/or CD
- Intervention: CBT or BT delivered in routine clinical care
- Comparison: within-group change. i.e. pre vs. post-data
- Outcome: primary continuous measure and remission
- Study design: RCTs and open trials

1.1. Literature search

Studies were identified by a systematic and comprehensive literature search of electronic databases and scanning reference lists of articles. The search was applied to Ovid MEDLINE, Embase OVID, and PsycINFO from the start of the data bases to June 11th 2019. An updated search was done May 5th 2020. The list of search terms utilized to identify potential studies were generated by all four authors in collaboration with a university librarian, who conducted the database searches. We used the following search terms to search the databases: (Cognitive therapy OR behav* therapy OR cognitive behav* therapy OR anger management therapy OR exposure therapy) AND (attention deficit disorder OR attention deficit hyperactivity disorder (including the different subtypes) OR conduct disorder OR oppositional defiant disorder OR Hyperkinetic disorder OR Externali* disorder) AND (open study OR clinical study OR community trial OR intervention study OR Pre post study OR randomized controlled trial) AND (youth OR child* OR adolescent OR pediatric). For full search strategy for Ovid MEDLINE, Embase OVID and PsychINFO, see the online Supplement, S1.

The first author read the titles and abstract of all the papers from this

initial search to decide whether a study warranted a more detailed reading. When there was an indication of a group of patients receiving the particular cognitive-behavioral treatment in a non-university setting the full-text was retrieved. The reference lists in the retrieved articles were then checked against the database search and any other articles that might fulfill the inclusion criteria were retrieved. Although research articles were the target of the search, review articles were also examined for additional references. Key authors were searched in databases to identify additional publications but were not contacted in personal communication. In total, 730 full-text articles were considered for inclusion. The final decision for article inclusion was made using a stricter set of inclusion and exclusion criteria detailed below. The full text articles were read by different pairs of the authors and any disagreements were resolved by consensus discussion. It was determined that 51 articles could be included in the present meta-analysis.

1.1.1. Inclusion criteria

In order to be included in the review and meta-analysis a study had to:

1. Be published, or in press, in an English language journal.
2. Have participants diagnosed with ADHD, CD, or ODD according to DSM or ICD, or fulfilling a cut-off score on a validated parent or teacher rating scale.
3. Be testing a form of CBT, cognitive therapy (CT) or behavior therapy (BT).
4. Have participants referred for treatment through usual clinical routes.
5. Be an effectiveness study, i.e. carried out in a non-university setting such as clinical routine care or school health care.
6. Have therapists who are practicing clinicians for whom provision of service is a substantial part of the job (Shadish, Matt, Navarro, & Phillips, 2000).
7. Have a treated sample consisting of at least 10 participants.
8. Have a maximum participant age of 18.
9. Provide a measure of the primary disorder treated.

1.1.2. Exclusion criteria

1. The study is a secondary analysis of a previously published study.
2. The study is an evaluation of a service where the results for individual disorders cannot be extracted.
3. The study is not testing a form of CBT, CT, or BT.
4. The study is testing a combination of CBT and pharmacological treatment.

1.2. Cut-off scores for applied parent or teacher rating scales

All ADHD- and 18 of 28 CD/ODD-studies used diagnosis as inclusion criterion, whereas 10 CD/ODD-studies used cut-off scores on the following parent or teacher rating scales: Child Behavior Checklist (CBCL; Achenbach, 1991), Eyberg Child Behavior Inventory (ECBI;

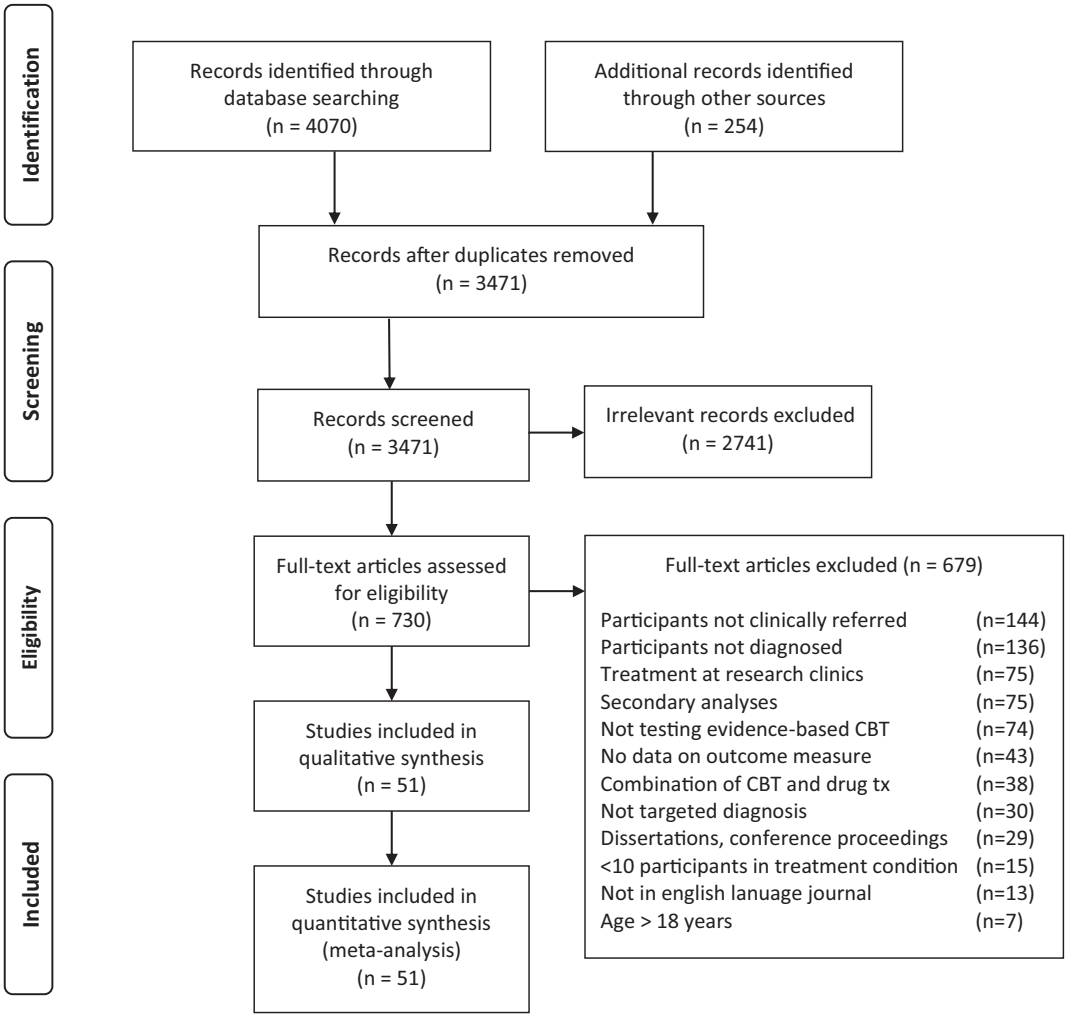

Fig. 1. Flowchart of the inclusion of studies.

Eyberg & Pincus, 1999), and Aversive Behavior Rate (ABR; Patterson, Chamberlain, & Reid, 1982). We transformed the various scores to percentiles and the five studies using CBCL had cut-off scores between 84 and 95, the four studies using ECBI had scores between 90 and 95, and the single study applying the behavior observation instrument ABR used a score of 90.

Fig. 1 shows a flowchart of the inclusion of studies in the present meta-analysis. For references to included studies, see online Supplement S2, and for references to studies excluded in the meta-analyses, see online Supplement S3.

1.3. Potential categorical moderators

In order to include any potential categorical or continuous moderator in the analysis we required that at least 70% of the studies provided information on that variable. With lower proportions it is questionable if the information extracted is representative of the entire body of studies.

1.3.1. Type of study and statistical analysis

Type of study was either RCT (when a CBT-condition was compared with some kind of control condition) or open trial (when only a CBT-condition was used in the study). Statistical analysis was categorized as intent-to-treat (ITT) if all randomized participants were included in the statistical analysis or completers if dropouts were deleted.

1.3.2. Type of treatment and conditions

It was not possible to classify the many treatment methods used based on their names since some methods might have different names but only differ slightly from each other. Instead we categorized the target of treatment as: child, parent, or child and parent, and treatment format as: group, individual, or group and individual.

1.3.3. Parental and teacher involvement

Degree of parental involvement was classified as *low* if parents were not present during sessions but informed about the progress of therapy, *moderate* if parents were present during some therapy sessions full-time or only part-time of all sessions, and *high* if parents were present full-time during all therapy sessions. Teacher involvement was classified as yes or no.

1.3.4. Therapist profession

The profession that the majority of the therapists within a study belonged to was classified as clinical psychologist, social worker, or various professions.

1.3.5. Continent

The country in which the study was carried out was categorized as North America, South America, Europe, Asia, Australia, or Africa.

A coding scheme and manual including the variables of interest were developed. The data extraction and categorizations were done independently by the first and the third author and any disagreements were solved after consensus discussion.

1.4. Potential continuous moderators

The following continuous measures on which at least 70% of the studies provided information were used as potential moderators: number of participants in the study, percent declining participation in the study, percent attrition in the study, percent on drug treatment for the principal disorder, percent boys, mean age, pre-treatment severity (calculated as percentage of the maximum score of the rating scale applied), methodology score (see below), risk-of-bias score (see below), number of therapy sessions, and treatment intensity (hours/week). In addition, we also extracted information on a number of other variables, but these variables did not reach the 70% criterion.

1.5. Methodological quality

1.5.1. The Psychotherapy Outcome Study Methodology Rating Scale (POMRS)

The scale consists of 22 items covering various important aspects of the methodology in psychotherapy outcome research (Öst, 2008). Each item is rated as 0 = poor, 1 = fair, and 2 = good, and each step has a verbal description of one or more sentences. The total score can vary from 0 to 44 points. Since all items were not applicable to all studies the total score was recalculated as a percentage of the maximum score possible for the individual study. The internal consistency of the scale was good with a McDonald's ω of 0.80. The inter-rater reliability of the scale (between the second and the fourth author), based on 20% randomly selected and blindly rated studies was ICC(3,1) = 0.94 (95% CI 0.75–0.99), which according to Cicchetti (1994) is excellent.

1.6. Risk-of-bias

The Cochrane Collaboration tool for assessing risk-of-bias (Higgins, Altman, & Sterne, 2011) was used, and the following domains were rated: random sequence generation, allocation concealment, blinding of outcome assessment, incomplete outcome data, and selective reporting. Blinding of patients and therapists cannot be used in psychotherapy studies. A high risk-of-bias in a domain was given 1 point, an unclear risk 0.5, and a low risk 0 point. Summarizing over the five domains the total score could vary between 0 and 5, with higher scores indicating higher risk-of-bias. Inter-rater reliability was assessed between the second and the fourth author based on 20% randomly selected and blindly rated studies. This yielded an intra-class correlation, ICC(3, 1) = 0.95 (95% CI 0.76–0.99), which is also excellent.

1.7. Effect size measures

We extracted data on both primary and secondary measures in the studies. This was done independently by the third and the fourth authors, and any discrepancies were discussed to reach consensus. Since some studies used proportion of remitted participants as their primary outcome measure, whereas other studies used a continuous rating scale we decided to include both in this meta-analysis.

1.7.1. Remission

Below follows a description of the criteria of remission used in the different studies.

ADHD: 11 out of 23 studies (48%) provided data and in 9 it was that the child's post- or follow-up score was within the normal range on the primary outcome measure, whereas in 2 it was loss of the principal diagnosis.

CD/ODD: 12 out of 28 studies (43%) provided data and in 11 it was that the child's post- or follow-up score was within the normal range on the primary outcome measure (for 2 of which Reliable Change Index must also be fulfilled), and in 1 it was loss of the principal diagnosis.

1.7.2. Continuous rating scales

When a study named its primary outcome measure among rating scales we used that. If none was pinpointed we selected measures according to the following hierarchy: independent assessor or observer rating, teacher report scale, parent report scale, and child self-report scale. All but one of the ADHD-studies and all CD/ODD studies provided data on a continuous rating scale. The various rating scales used for the respective studies are described in the online Supplement S4.

1.7.3. Secondary outcome measures

Besides symptom measures we also aimed to extract data on functioning, broadly defined. However, only 7 out of 23 (30%) ADHD-studies and 10 out of 28 (36%) CD/ODD-studies had used such measures. With these low proportions of studies providing the relevant data

it is questionable if the outcome of a meta-analysis would be representative so we decided not to carry out one.

1.8. Meta-analysis

In order to obtain as large as possible a body of effectiveness studies we included both RCTs and open trials in the meta-analysis since within-group ES can be calculated from both types of studies. Within-group ES was calculated as $(M_{pre} - M_{post})/SD_{pre}$ according to recommendation by Lakens (2013), since there is good reason to assume that the interventions influence not only the means but also the standard deviations. The mean ES was computed by weighting each ES by the inverse of its variance. Rate of remission, with event rate as the effect measure, was analyzed using mixed effect analysis in the subgroup analysis. In this analysis a random effects model is used to combine studies within each subgroup and a fixed effects model is used to compare subgroups and yield the overall effect. When a study presented intent-to-treat data (49%) these were used, if not completer data (51%) were used.

Before pooling the effect sizes we screened for statistical outliers, defined as being outside M ± 2SD. At the post-treatment assessment four (2.2%) of the ESs were outliers, and at follow-up assessment there was one (2.3%). For these ESs *winsorising* (Lipsey & Wilson, 2001) was used by reducing outliers to the exact value of M + 2SD. The software *Comprehensive Meta-Analysis v.3* (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2013) was used for all analyses and to correct for small sample sizes Hedges's *g* was calculated. A random effects model was used since it cannot be assumed that the ESs come from the same population.

Heterogeneity among ES's was assessed with the *Q*- and the *I*-square statistic. The possibility of publication bias was analyzed with the trim-and-fill method of Duval and Tweedie (2000) as well as Egger's regression intercept (Egger, Smith, Schneider, & Minder, 1997). Moderator analyses of continuous variables were carried out with meta-regression and for categorical variables with subgroup analysis using the mixed effect model.

1.9. Efficacy studies for comparison

In order to obtain the efficacy studies to be used in comparison of the effect of CBT in effectiveness studies we consulted the most recent evidence base update reviews of psychosocial treatments published in the *Journal of Clinical Child and Adolescent Psychology* for the respective disorders included in the present meta-analysis. For ADHD it was Evans et al. (2014); Evans, Owens, Wymbs, and Ray (2018), and for CD/ODD McCart and Sheidow (2016) and Kaminski and Claussen (2017). In some of these reviews the authors referred to earlier reviews, which we checked in order to get as comprehensive as possible a list of efficacy RCTs. From each of these reviews we listed the RCTs of some kind of cognitive behavioral treatment evaluated as well-established or probably efficacious according to the criteria adopted by the Society of Clinical Child and Adolescent Psychology (Southam-Gerow & Prinstein, 2014). Then we deleted those RCTs we had already included in the body of effectiveness studies. This resulted in the following number of efficacy RCTs for our comparison: ADHD 29 and CD/ODD 33, for a total of 62 trials. These references are listed in the online Supplement S5.

As for the effectiveness studies we extracted data for the primary continuous outcome measure and remission rate, separately at post-treatment and follow-up assessment. In order to compare the two categories of studies on background variables we also extracted data on mean age, proportion of boys, percent with comorbid disorders, proportion on medication for the principal disorders, pre-treatment severity (calculated as percent of maximum score on the continuous measure), treatment time (60 min. hours), and attrition. Other variables were not reported systematically, or not at all in a large enough proportion of studies, which precluded inclusion as a background variable.

1.9.1. Power analysis

In the overall comparison of effectiveness and efficacy studies we have the following number of studies and treatment conditions, which is the unit of analysis: effectiveness studies ADHD 23/28, CD/ODD 28/39 for a total of 51/67; efficacy studies ADHD 29/43, CD/ODD 33/62 for a total of 62/105. The total number is then 113 studies with 172 conditions with an average of 46 participants per condition. According to the formulas for power analysis in meta-analyses by Valentine, Pigott, and Rothstein (2010) we would have 100% power to detect an effect size of 0.20 and 88% for an ES of 0.10, when assuming that the heterogeneity of effect sizes will be high.

2. Results

2.1. Description of the studies

2.1.1. Study characteristics

Background data for the included studies are presented in Table 1. The majority of the 51 studies were done in North America ($n = 25$) and Europe ($n = 23$), whereas only 3 came from Australia, and none from Africa, Asia, or South America. The total number of participants in these studies was 5295 with 3798 in CBT-conditions, 298 in other treatments, e.g. pharmacological, and 1199 in control conditions. For ADHD it was 2223 in 23 studies (1675 in CBT-conditions and 548 in control conditions) and for CD/ODD 2774 in 28 studies (2123 in CBT-conditions and 651 in control conditions). The proportion of eligible participants that declined the offer of treatment was on average 12.7% (ADHD 10% and CD/ODD 15%).

There was an overall majority of boys (77.7%), both in ADHD (78.0%) and CD/ODD (77.3%). The mean age across all studies was 8.2 years (ADHD 8.7 and CD/ODD 7.9). Proportion of participants with comorbidity was reported by only 24 studies (47%), and in an unsystematic fashion. With that in mind, 52.9% of the participants had at least one comorbid disorder (46.7% in ADHD and 63.2% in CD/ODD). Only 33 studies (66.7%) reported what proportion of the participants was on psychotropic medication for their principal disorder at the inclusion to the respective study. The overall mean was 32.5% with a significant difference ($t(32) = 2.38$, $p = 0.024$) between ADHD (41.6%) and CD/ODD (19.5%), which would be expected since central stimulants are evidence based for ADHD.

2.1.2. Treatment data

Treatment data for the included studies are presented in Table 2. The target of treatment was the parent in 25 (49.0%), parent and child in 21 (41.2%), and child only in 5 studies (9.8%). The format of treatment was group in 25 (49.0%), individual in 22 (43.1%), and the combination of group and individual treatment in 4 studies (7.8%). The number of therapists per study was on average 11.1 (range 1–56), which indicates the number of participating therapists working at the routine clinical sites where the studies were done. Treatments were carried out over 15.0 weeks on average (12.0 for ADHD and 17.5 for CD/ODD) and the mean number of sessions was 17.3 (ADHD 17.5, CD/ODD 17.2). The intensity (hours/week) was on average 2.8 (ADHD 3.9, CD/ODD 1.8).

2.2. Methodological data

2.2.1. Methodology ratings

The research methodology score (% of maximum possible score for the individual study) had an overall mean of 47.4 (*SD* 12.3), which corresponds to a raw score of 20.9 points. The mean for ADHD-studies was 46.9 (*SD* 14.3) and for CD/ODD studies 47.7 (*SD* 10.8), a non-significant difference ($t(65) = 0.26$, $p = 0.80$). As could be expected the RCTs (*M* 50.8, *SD* 11.2) across the two disorders had a significantly higher methodology score ($t(65) = 4.44$, $p < 0.001$) than did the open trials (*M* 37.3, *SD* 9.5).

Table 1
Background data of the included studies.

Study	Continent	Type of study	Analysis	Methodology	Severity	% Comorb.	% Decliners	N	% Males	Age range	Mean age	% on meds
ADHD												
Abikoff, 2015	North America	RCT	ITT	0.682	0.752	41.5	0	164	73.8	3–4	3.6	0
Boyer, 2015	Europe	RCT	ITT	0.682	0.474	31.4	0.6	159	73.5	12–17	14.4	78.0
Breider, 2019	Europe	RCT	ITT	0.455	0.323	76.2	29.0	21	71.4	4–12	7.8	
Canu, 2011	North America	Open trial	ITT	0.325	0.719	100	0	16	88	4–12	8.0	
Curtis, 2010; 2013	North America	Open trial	ITT	0.250	0.744		22.0	21	71	7–10	7.8	64.0
Daley, 2013	Europe	RCT	ITT	0.618	0.783		7.1	43	81.4	4–11	7.3	0
Döpfner, 2004	Europe	RCT	Compl.	0.455	0.600	61.3		75	93.3	6–10	8.3	0
Elkins, 2019	North America	Open trial	Compl.	0.350	0.770	52.0		67	76.0	7–10	8.0	72.0
Fehlings, 1991	North America	RCT	ITT	0.523	0.563	0		25	100	7–13	9.5	0
Hannesdottir, 2017	Europe	Open trial	ITT	0.341	0.684			41	70.7	8–10	9.2	78.0
Heath, 2015	North America	Open trial	Compl.	0.375	0.740			47	79.1	7–12	8.0	74.4
Jerrott, 2010	North America	Open trial	Compl.	0.310	0.693	72.5		57	84.2	5–13	10.4	78.0
Loren. 2015	North America	Open trial	Compl.	0.425		27.6		241	75.6	6–12	8.6	
MTA-Group. 1999	North America	RCT	Compl.	0.738	0.467	39.9	4.4	579	80.3	7–9.9	8.5	30.7
Ostberg. 2012	Europe	RCT	Compl.	0.310	0.200	4.3	9.8	92	83.6	7–10	10.9	80.3
Pfiffner. 2013	North America	Open trial	Compl.	0.475	0.600		3.2	57	70.2	7–11	8.1	7.0
Pfiffner. 2016; 2018	North America	RCT	ITT	0.477	0.626	50.4	10.0	135	71	7–11	8.4	8.9
Power. 2012	North America	RCT	ITT	0.500	0.467	27.0	23.0	199	68.3	7–12	9.5	43.0
Power. 2014	North America	Open trial	Compl.	0.500	0.667	36.0	28.0	72	69.4	5–10	7.0	34.7
Sibley. 2018	North America	RCT	Compl.	0.571	0.467		21.8	218	74.2	12–15	13.5	45.9
Tynan. 1999	North America	Open trial	Compl.	0.325	0.607		9.7	65	76.4	5–11	7.5	44.0
Tynan 2004	North America	Open trial	ITT	0.250	0.641		9.5	67	82.4	5–11	7.9	43.1
van den Hoofdakker. 2007	Europe	RCT	Compl.	0.477	0.933	80.9	2.0	96	80.9	4–12	7.4	50.0
CD/ODD												
Axberg. 2012	Europe	RCT	ITT	0.524	0.635		6.1	62	83.9	4–8	6.0	
Axelrad. 2009	North America	Open trial	Compl.	0.425	0.670		0	64	84.0	2–6	4.7	1.6
Bjorseth. 2016	Europe	RCT	ITT	0.591	0.587	74.6	11.0	81	70.4	2–7	5.8	12.3
Burke. 2015	North America	RCT	ITT	0.429	0.753		22.0	252	100		8.9	7.5
Chacko. 2015	North America	RCT	ITT	0.409	0.620			320	67.8	7–11	8.9	33.4
Connolly. 2001	Europe	RCT	Compl.	0.273	0.642			129	86.8	2–9	6.2	2.3
Costin. 2007	Australia	Open trial	Compl.	0.400	0.662	76.4		112	80.9	5–13	9.1	
David. 2014	Europe	RCT	Compl.	0.333	0.601		11.0	130	48.5	4–12	6.2	0
Gardner. 2006	Europe	RCT	ITT	0.575	0.606		28.1	76	73.7	2–9	5.9	
Gavita. 2012	Europe	RCT	ITT	0.425	0.766		14.2	97		5–18	9.5	
Goertz-Dorten. 2019	Europe	RCT	ITT	0.545	0.567		36.4	101	100	6–12	8.7	14.3
Greene. 2004	North America	RCT	Compl.	0.477	0.576	66.0	36.4	50	68.1	4–12	7.2	59.6
Hansson. 2012	Europe	RCT	Compl.	0.357	0.626		0	46	60.9	2–17	15.0	
Harrington. 2000	Europe	RCT	ITT	0.568	0.698		25.0	141	79.0	3–10	6.9	
Helander. 2018	Europe	RCT	ITT	0.568	0.705	67.0	11.7	120	73.0	8–12	9.3	
Hutchings. 2002	Europe	RCT	ITT	0.429	0.696		10.6	41	85.4	2–10	6.0	
Kolko. 2009	North America	RCT	ITT	0.705	0.773	76.3	6.8	139	84.9	6–11	8.8	51.1
Larsson. 2009	Europe	RCT	ITT	0.568	0.623	35.4	8.0	136	79.5	4–8	6.6	
Lees. 2019	Australia	RCT	Compl.	0.476	0.539		10.6	126	68.3	3–7	5.4	
McGilloway. 2012	Europe	RCT	Compl.	0.525	0.621		4.0	149	62.8	2–7	4.8	
McNeil. 1991	North America	Open trial	ITT	0.625	0.717	90.0		12	100	2–7	4.5	0
Nixon. 2003	Australia	RCT	ITT	0.523	0.661			63	70.4	3–5	3.9	0
Patterson. 1982	North America	RCT	Compl.	0.405			34.5	19	68.4	3–11	6.8	
van de Wiel. 2007	Europe	RCT	Compl.	0.455	0.746	62.5		77	88.3	8–13	10.0	40.3
van Manen. 2004	Europe	RCT	ITT	0.523	0.668			97	100	9–13	11.0	
Westermark. 2011	Europe	RCT	ITT	0.405	0.696		7.9	35	51.4	2–18	15.0	
Woltering. 2011	North America	Open trial	ITT	0.342	0.325	21.0		71	72.0	8–12	9.5	25.4
Woltering. 2015	North America	Open trial	Compl.	0.289	0.721			39	79.5	8–12	9.5	25.6

Note: RCT = randomized controlled trial, ITT = intent-to-treat, Compl. = completers, Blank fields = data not provided, Severity = percentage of the maximum score on the primary outcome measure, % Comorb. = proportion having any psychiatric comorbid disorder, % Decliners = proportion of participants who were offered treatment in the study but declined (irrespective of reason), % on meds. = proportion on any psychotropic medication.

2.2.2. Risk of bias

The risk of bias categorization is presented in the online Supplement, Table 6. If we delete the 16 open trials from the evaluation of the first two domains we find the following proportions of a low risk-of-bias: Random sequence 83%, allocation concealment 46%, blinding of assessors 29%, incomplete data 59%, and selective reporting 84%. A high risk-of-bias was found in: random sequence 6%, allocation concealment 9%, blinding of assessors 51%, incomplete data 35%, and selective reporting 0%. Thus, it was much more common that these studies had a low than a high risk-of-bias regarding the evaluated domains.

In order to score the risk-of-bias a low risk was given 0, an unclear risk 0.5, and a high risk 1 point, which means that the total score could vary from 0 to 5 points. The total mean score was 1.87 (*SD* 1.52) and the studies within each disorder had the following mean (*SD*): ADHD 2.11 (1.67) and CD/ODD 1.69 (1.41), a non-significant difference ($t(65) = 1.10$). As was found for methodological quality the RCTs (*M* 1.23, *SD* 1.15) had a significantly lower risk-of-bias ($t(65) = 8.38$, $p < 0.001$) than the open trials (*M* 3.74, *SD* 0.75).

2.3. Meta-analysis

2.3.1. Attrition

Forty-seven of the studies (92.2%) provided information on the number of participants who dropped out of the treatments. Using treatment condition (k = 63) as the unit of analysis the overall attrition rate was 14.4% (95% CI 11.5–18.0, $z = 13.36$, $p < 0.0001$). ADHD-studies (k = 27) had a dropout rate of 12.0% and CD/ODD-studies (k = 36) 16.4%, which was not a significant difference ($Q_{between}$ (1 df) =

1.72, $p = 0.19$).

2.3.2. Primary continuous measure

Table 3 displays the mean effect sizes of the primary continuous measure for all studies at post-treatment and follow-up assessment, which was done on average 10.8 months after the end of therapy. At post-treatment the average ES was large ($g = 0.91$) and significantly different from zero. Heterogeneity was significant and large as indicated by the Q- and I^2-values. CD/ODD-studies ($g = 0.98$) had a somewhat higher ES than ADHD-studies ($g = 0.80$) but the difference was not significant.

At follow-up the mean ES ($g = 1.01$) was insignificantly higher than at post-treatment ($Q_{between}$ (1 df) = 0.75, $p = 0.39$), but also significantly heterogeneous. The comparison between disorders was still not significant. Thus, the treatment effects seen at post-assessment were maintained at follow-up.

Publication Bias. The possibility of publication bias was investigated, using Duval and Tweedie's trim-and-fill method and Eggers regression intercept. Regarding the post-treatment data the trim-and-fill method suggested trimming 17 conditions to the left of the mean which would lower the ES from 0.91 to 0.72 (95% CI 0.66–0.75). The regression intercept had a significant $t(60) = 4.66$, $p < 0.01$, indicating that publication bias probably is an issue for this body of studies.

2.3.3. Remission

The remission rates at post-treatment and follow-up are presented in Table 4. At post-treatment 43.8% of the participants had remitted, which was significantly heterogeneous. A subgroup analysis indicated that CD/ODD-studies had about 10 percentage points higher remission rate than ADHD-studies but the difference was not significant.

At follow-up assessment the overall remission rate had increased to 51.4% but the difference to the post-assessment rate was not significant ($Q_{between}$ (1 df) = 1.38, $p = 0.24$). The subgroup analysis also showed that the disorders did not differ significantly from each other. Thus, as was the case for continuous measures the effect sizes seen at post-assessment were maintained at follow-up.

Publication Bias. Eggers regression intercept did not yield a significant t-value (1.15, $p = 0.26$). The trim-and-fill method suggested trimming 7 studies, which would have reduced the remission rate from 43.8% to 37.4% (95% CI 30.6–44.8). Thus, publication bias does not seem to be an important problem regarding the remission rate.

2.3.4. Moderator analyses

As the mean ES was significantly heterogeneous we followed up with moderator analyses. Table 5 shows the results for categorical variables using subgroup analysis with the results for effect size in the left column and remission rate in the right. There was no significant difference between RCTs and open trials. Regarding statistical analysis studies with intent-to-treat analysis yielded non-significantly lower ES than studies using completer analysis. There was no significant differences depending on target or format of therapy, degree of parental involvement, teacher involvement or therapist profession. The continent at which the study was carried out was associated with a significant difference; studies from Australia had the highest ES. However, since the ES for Australia is made up of only 5 conditions this difference should be interpreted with caution.

Continuous variables on which at least 70% of the studies provided information were analyzed with the meta-regression module in the CMA program using the fixed effects analysis (see Table 6). Since 11 variables were included we used the Holm-Bonferroni correction (see Jaccard & Guilamo-Ramos, 2002). For effect size there were four negative moderators: mean age of the participants in the condition, proportion of participants on drug treatment for the principal disorder, percent attrition, and risk-of-bias, i.e. the higher the values on these moderators the lower the ES. There was one positive moderator, pre-treatment severity, i.e. higher severity was associated with higher ES. For remission rate there were three negative moderators: mean age, percent on drug treatment, and attrition rate. In addition, there was one positive moderator: number of sessions, i.e. higher number of sessions was associated with higher remission rate.

2.4. Efficacy-effectiveness comparison

In the following section (Tables 7-10) we compare data for the effectiveness studies reviewed so far with data for the efficacy studies obtained from the evidence base update reviews on ADHD and CD/ODD published in the *Journal of Clinical Child and Adolescent Psychology*.

2.4.1. Background and treatment variables

Table 7 presents a comparison between efficacy and effectiveness studies on some background variables and one treatment variable. Since there are 7 variables and 14 t-tests the Holm-Bonferroni correction was used. The only significant difference emerged on percent comorbidity for the CD/ODD studies where effectiveness studies had a higher proportion than efficacy studies. However, this result must be treated with caution since only 43.8% of the effectiveness and 46.7% of the efficacy studies provided data on this variable. There were no significant difference between the two types of studies regarding mean age, proportion of participants on medication for the principal disorder, and percent attrition. If we use a p-value of 0.05 effectiveness studies had higher proportion of boys in both disorders. a higher pre-treatment severity in CD/ODD, and a lower treatment time in ADHD-studies, but these differences can only be regarded as trends since they don't fulfill the Holm-Bonferroni correction criterion. Thus, judging from the background treatment variables which could be extracted the effectiveness studies do not consist of participants who are easier to treat than do the efficacy studies.

2.4.2. Effect size on primary outcome measure

Table 8 presents the subgroup analyses comparing the within-group effect size (Hedges' g) for effectiveness and efficacy studies within each disorder. Neither at post-treatment assessment (upper part) nor at follow-up (lower part) were there any significant differences between the two types of studies. For both types the ESs were significant and the effects were maintained, or somewhat higher, at follow-up, which was done on average 9.9 months after post-assessment for ADHD- and 11.3 months for CD/ODD-studies ($t(63) = 0.45$, $p = 0.66$).

2.4.3. Remission

Table 9 contains subgroup analyses comparing the remission rates at post-treatment and follow-up assessment. Neither the post-treatment (upper part of Table 9) nor the follow-up (lower part) remission rates differed between study types for ADHD or CD/ODD. When combining the remission rates for the different disorders the post-treatment means were 43.8% and 46.1% for effectiveness and efficacy studies, respectively. The corresponding follow-up means were 51.4% and 53.7%, which means that the increase from post- to follow-up assessment was 7.6 percentage points for both categories of studies.

2.4.4. Comparison of RCTs only

Since the outcomes presented in Tables 8 and 9 might have been unduly influenced by open trials we repeated the analyses using only RCT effectiveness studies. Table 10 summarizes the results across disorders and there were no significant differences between effectiveness and efficacy studies regarding effect size or remission rates at post-treatment or follow-up assessment.

3. Discussion

The primary aim of the present meta-analysis was to investigate the

Table 2
Treatment data of the included studies.

Study	Treatment program	CBT method	Profession	Tx target	Tx form.	PMT	Parent involvem.	Teacher involvem.	No. of sessions	Hrs/ week	% Attrition	F-Up
ADHD												
Abikoff. 2015	New Forrest Parenting Package (NFPP)	PBT (with child) + TI	Psychologist	P	I	Y	High	N	8	1	9.0	6
	Helping the non-compliant child (HNC)	PBT (with child)	Psychologist	P	I	Y	High	N	8	1	6.3	6
Boyer, 2015	Plan my Life (PML)	TI	Psychologist	C	I	N	Moderate	N	10	1.1	4.8	3
	Solution focused therapy (SFT)	TI	Psychologist	C	I	N	Low	N	10	1.1	5.3	3
Breider, 2019	Blended parent training	PBT	Psychologist	P	G	Y	High	N	16	0.9	90.9	0
	Face-to-face parent training	PBT	Psychologist	P	I	Y	High	N	4	1	40.0	0
Canu, 2011	Defiant children	PBT	Psychologist	P	G	Y	High	N	6	1.5	0	0
Curtis, 2010; 2013	Family Stars	PBT + ChBT	Various	C, P	I	Y	Moderate	N	10	1	0	0
Daley, 2013	New Forest Parenting Package Self-help	PBT	NA (Self-help)	P	I	Y	High	N	1	0.3	12.5	0
Döpfner, 2004	Behavior therapy + Psychoeducation	PBT + ChBT + TT		C, P	I	Y	Moderate	Y	18	3	2.9	0
Elkins, 2019	Family Stars	PBT + ChBT		C, P	I	Y	Moderate	N	10	1	17.9	0
Fehlings, 1991	CBT	ChBT		C	I	N	Moderate	N	20	1.8	0	5
Hannesdottir, 2017	OutSMART	ChBT + TI	Psychologist	C	G	N	Low	N	10	4	6.3	3
	Parent training program	PBT	Psychologist	P	G	Y	High	N	7	2	18.2	3
Heath, 2015	Family Stars	PBT + ChBT	Psychologist	C, P	I	Y	Moderate	N	10	1	8.5	0
Jerrott, 2010	CBT	PBT + ChBT	Various	C, P	G	Y	Moderate	Y	32	2		39
Loren, 2015	Parent group program	PBT	Psychologist	P	G	Y	High	N	8	2	31.5	0
MTA-Group, 1999	Behavioral treatment	PBT + ChBT + TT		C, P	G, I	Y	Moderate	Y	91	6.9	2.1	82
Ostberg, 2012	Strategies in Everyday Life (SEL)	PBT + TT		P	G	Y	Moderate	Y	18	2.8	16.7	3
Pfiffner, 2013	Collaborative Life Skills Program (CLS)	PBT + ChBT + TT	Various	C, P	G	Y	Moderate	Y	22	1.5	5.0	0
Pfiffner, 2016; 2018	Collaborative Life Skills Program (CLS)	PBT + ChBT + TT	Various	C, P	G	Y	Moderate	Y	23	1.5	0	7
Power, 2012	Family School Success (FSS)	PBT + ChBT + TT	Psychologist	P	G, I	Y	Moderate	Y	12	1.3	4.3	3
Power, 2014	Partnering to Achieve School Success (PASS)	PBT + TT	Psychologist	P	I	Y	High	Y	9	0.5	24.0	0
Sibley, 2018	High Intensity skills-based summer intervention	PBT + ChBT	Various	C, P	G	Y	Moderate	N	40	45		8
	Low Intensity skills-based summer intervention	PBT + ChBT	Various	C, P	G	Y	Moderate	N	8	1.5		8
Tynan, 1999	Parent training and child social-skills training	PBT + ChBT		C, P	G	Y	Moderate	N	8	2	15.4	0
Tynan 2004	Parent training and child social-skills training	PBT + ChBT		C, P	G	Y	Moderate	N	8	2	23.9	0
van den Hoofdakker, 2007	Behavioral parent training	PBT	Psychologist	P	G	Y	High	N	12	1.2	12.5	6
CD/ODD												
Axelrad, 2009	Brief Behavioral Intervention (BBI)	PBT		P	I	Y	High	N	5	0.83	14.1	12
Bjorseth, 2016	Parent-Child Interaction Therapy (PCIT)	PBT (with child)	Psychologist	P	I	Y	High	N	21.1	1	12.5	12
Burke, 2015	Stop Now And Plan (SNAP)	PBT + ChBT		C, P	G, I	Y	Moderate	N	24	2		12
Chacko, 2015	Multiple Family Group (MFG)	PBT (with child) + FI	Various	C, P	G	Y	High	N	16	1.8	34.2	12
Connolly, 2001	Incredible Years (IY)	PBT		P	G	Y	High	N	12	1.8	50.0	6
	Incredible Years (IY) + Child Intervention	PBT		P	G	Y	Moderate	N	12	1.8	45.1	6
Costin, 2007	Skilled Parenting Program	PBT		P	G	Y	High	N	8	2	5.3	5
David, 2014	Rationale Positive Parenting Program	PBT	Various	P	G	Y	High	N	10	1.5	20.0	1
	CBT Standard	PBT	Various	P	G	Y	High	N	10	1.5	14.9	1
Gardner, 2006	Incredible Years (IY)	PBT	Various	P	G	Y	High	N	14	2	11.6	12
Gavita, 2012	Short cognitive-behavioral group parenting program	PBT		P	G	Y	High	N	4	1.3	21.4	3
Goertz-Dorten, 2019	Treatm. Progr. for Children with Aggr. Beh. (THAV)	ChBT	Psychologist	C	I	N	Low	N	24	0.8	0	0
Greene, 2004	Collaborative Problem Solving (CPS)	PBT	Psychologist	P	I	Y	High	N	11	1	6.7	4
	Parent Training (PT)	PBT	Psychologist	P	I	Y	High	N	10	1	5.0	4
Hansson, 2012	Multidimensional Treatment Foster Care (MTFC)	MCT	Various	C, P	I	Y	Moderate	Y			15.8	12
Harrington, 2000	Parent training – community	PBT		P	G	Y	High	N			48.6	9
	Parent training – hospital	PBT		P	G	Y	High	N			40.6	9
Helander, 2018	KOMET	PBT	Psychologist	P	G	Y	High	N	11	2.5	8.0	0
		PBT + ChBT	Psychologist	P	G	Y	Moderate	N	26	4.3	17.2	0

(*continued on next page*)

Table 2 (*continued*)

Study	Treatment program	CBT method	Profession	Tx target	Tx form.	PMT	Parent involvem.	Teacher involvem.	No. of sessions	Hrs/ week	% Attrition	F-Up
	KOMET + The Coping Power Program											
Hutchings, 2002	Standard parent training	PBT	Various	P	I	Y	High	N			31.6	0
	Intensive parent training	PBT (with child)	Psychologist	P	I	Y	High	N			4.5	0
Kolko, 2009	Modular treatment – community	MCT (comm)	Social worker	P, C	I	Y	Moderate	Y	21.6	1.4	2.9	36
	Modular treatment – clinic	MCT (clinic)	Social worker	P, C	I	Y	Moderate	Y	15.4	1.1	20.9	36
Larsson, 2009	Incredible Years (IY)	PBT	Various	P	G	Y	High	N	13	2	4.3	12
	IY + Dinosaur School Program	PBT + ChBT	Various	P, C	G	Y	Moderate	N	31	3.4	0	12
Lees, 2019	IY	PBT	Various	P	G	Y	High	N	13	2	34.9	6
	IY + Home Parent Support (HPS)	PBT (+ HPS)	Various	P	G, I	Y	High	N	23	2.8	17.5	6
McGilloway, 2012	IY	PBT	Various	P	G	Y	High	N	14	2	4.0	0
McNeil 1991	Parent-Child Interaction Therapy (PCIT)	PBT (with child)	Psychologist	P	I	Y	High	N	14	1	0	18
Nixon, 2003	PCIT standard	PBT (with child)	Psychologist	P	I	Y	High	N	12	1.3	22.7	6
	PCIT abbreviated	PBT (with child)	Psychologist	P	I	Y	High	N	10	0.9	13.0	6
Patterson, 1982	Parent Management Training Oregon (PMTO)	PBT (with child)	Various	P	I	Y	High	N	17	1	0	0
van de Wiel, 2007	Utrecht Coping Power Program (UCPP)	PBT + ChBT	Psychologist	P, C	G	Y	Moderate	N	38	1	10.5	0
van Manen, 2004	Social Cognitive Intervention Program (SCIP)	ChBT	Psychologist	C	G	N	Low	N	11	1.2	0	12
	Social Skills Training Program (SST)	ChBT	Psychologist	C	G	N	Low	N	11	1.2	0	12
Westermark, 2011	Multidimensional Treatment Foster Care (MTFC)	MCT	Various	P, C	I	Y	Moderate	Y	40	1	10.0	12
Woltering, 2011	Stop Now And Plan (SNAP)	PBT + ChBT	Various	C, P	I	Y	Moderate	N	14	3		0
Woltering, 2015	Stop Now And Plan (SNAP)	PBT + ChBT	Various	C, P	I	Y	Moderate	N	14	3		12

Note: PBT = parent behavior therapy, TI = training intervention, ChBT = child behavior therapy, TT = teacher training, Tx target = target of treatment, C = child, P = parent, Tx form. = treatment format, G = group, I = individual, PMT = parent management training, Y = yes, N = no, Parent involvem. = degree of parental involvement, % Attrition = proportion dropping out of those participating in at least one session, F-Up = follow up in months.

effectiveness of CBT for ADHD, CD and ODD for children and adolescents in routine clinical care. The overall within-group effect size across the disorders was large and statistically significant, thereby demonstrating that the treatments are effective in real-world settings. The results showed that 44% of children and adolescents with externalizing disorders achieved remission after treatment in routine clinical care and that more than half of the participants were remitted at follow-up. Furthermore, the results showed that 86% of the participants with externalizing disorders completed CBT. The completion rate is slightly higher than what was reported in a review of treatment effectiveness studies for psychological problems in children and adolescents (Lee et al., 2013), which found that 75% of participants completed parent training for disruptive disorders in real world settings. Lee et al. (2013) included 13 studies and reported significant effect sizes across studies for disruptive disorders, concluding that positive results from parent training can be obtained in a range of treatment settings. Our results extends the findings of Lee et al. (2013) by including 51 studies and providing meta-analytical evidence of treatment outcome as well as updated remission and attrition rates for CBT for child and adolescent externalizing disorders in routine care. The results are encouraging and suggest that CBT for externalizing disorders are effective in routine clinical care, acceptable to clinically referred participants, and that a majority of the patients achieves remission at follow-up.

Regarding the magnitude of the effect size for ADHD ($g = 0.80$), our finding is similar to that of Fabiano et al. (2015) who conducted a systematic review of published meta-analyses on psychosocial treatments for ADHD. Of note, the meta-analyses were largely based on studies from university settings. Fabiano et al. (2015) concluded that the effect sizes across meta-analyses were generally moderate to large for behavioral outcomes. However, direct comparisons of effect size estimates to specific studies are not applicable since many previous meta-analyses have investigated subtypes of cognitive behavioral interventions including school-based treatments (DuPaul & Eckert, 1997; DuPaul, Eckert, & Vilardo, 2012) and parent training (Charach et al., 2013; Corcoran & Dattalo, 2006; Lee, Niew, Yang, Chen, & Lin, 2012; Zwi, Jones, Thorgaard, York, & Dennis, 2011). There are meta-analyses that have included a wider range of cognitive behavioral treatments for ADHD, but most of them have excluded open trials and reported between-group ES only (Klassen, Miller, Raina, Lee, & Olsen, 1999; Sonuga-Barke et al., 2013; Van der Oord, Prins, Oosterlaan, & Emmelkamp, 2008) making comparisons to the current ES estimates difficult. However, a meta-analysis by Fabiano et al. (2009) was similar in study inclusion criteria and effect size calculation. That meta-analysis investigated the effect of a wide range of behavioral interventions for ADHD and included both RCTs and open trials. Unlike the present meta-analysis they included studies regardless of their status as efficacy or effectiveness studies. Interestingly, they found a within-group effect size of 0.70, which is only slightly lower than the effect size for ADHD in the present study.

Most meta-analyses on treatment of disruptive behavior disorders have also reported between-group ES only, typically demonstrating small to moderate effect sizes (Bakker et al., 2017; Erford, Paul, Oncken, Kress, & Erford, 2014; McCart, Priester, Davies, & Azen, 2006). However, two meta-analyses have investigated the effect of psychosocial treatments for conduct problems and included both RCTs and open trials and reported large within-group effect sizes of 0.95 (Fossum, Handegård, Martinussen, & Mørch, 2008) and 1.05 (Fossum, Handegård, Adolfsen, Vis, & Wynn, 2016). Considering that these two meta-analyses included a large number of studies from research settings it is encouraging that the ES for CD/ODD in the present meta-analysis ($g = 0.98$) was similar to those reported by Fossum et al. (2016, 2008). Thus, the current results are congruent with previous meta analyses, which have

Table 3
Within-group effect size (Hedges' *g*) for all studies (RCTs and open trials) divided by disorder with treatment condition as unit of analysis.

Disorder	k	*g*	95% CI	z-value	Q_{-value}	p_{-value}	I^2
Post-treatment							
Both disorders	65	0.91	0.81–1.00	18.27[a]	208.96[a]		69.4
ADHD	26	0.80	0.67–0.93	11.76[a]	3.50*	0.06	
CD-ODD	39	0.98	0.85–1.12	14.31[a]			
Follow-up							
Both disorders	45	1.01	0.87–1.14	15.00[a]	188.38[a]		76.6
ADHD	14	0.88	0.70–1.07	9.52[a]	1.93*	0.17	
CD-ODD	31	1.06	1.89–1.24	11.87[a]			

Note: k = number of treatment conditions. [a] $p < 0.0001$. * Comparison between the disorders.

Table 4
Rates of remission for all studies divided by disorder.

Disorder	k	%	95% CI	z-value†	Q_{-b}	p_{-value}	I^2
Post-treatment							
Both disorders	31	43.8	36.5–51.4	1.61	214.34[a]		86.0
ADHD	12	37.7	26.7–50.2	1.93	1.51*	0.22	
CD-ODD	19	47.6	38.0–57.3	0.49			
Follow-up							
Both disorders	17	51.4	41.3–61.3	0.26	106.44[a]		85.0
ADHD	5	38.2	21.8–57.9	1.18	2.68*	0.10	
CD-ODD	12	56.8	46.5–66.5	1.29			

Note: k = number of treatment conditions. [a] $p < 0.0001$. † Test if significantly different from 50%. * Comparison between the two disorders.

reported within-group ES, but slightly higher than previous meta analyses on disruptive behavior disorders which have only reported between-group ES.

The fact that the effect size for CD/ODD ($g = 0.98$) was slightly higher than for ADHD ($g = 0.80$) is interesting. It can be argued that this finding is contrary to expectations, considering that the comprehensive review of meta-analyses for ADHD (Fabiano et al., 2015) reported moderate to large effect sizes for behavioral outcomes, whereas meta-analyses on disruptive behavior disorders have reported small to moderate effect sizes. However, the findings from a meta-analysis of CBT for externalizing disorders (Battagliese et al., 2015) found a moderate and statistically significant effect size, and in accordance with our findings, they found better outcomes for CD and ODD than for ADHD, with large effect sizes for CD and ODD and moderate effect sizes for ADHD.

Parents received treatment in 90% of the effectiveness studies; either as the only recipient (49%) or together with their child (41%). Parent management training is well-established as a stand-alone treatment for ADHD (Evans et al., 2018) and childhood disruptive behavior (Kaminski & Claussen, 2017) and as part of multicomponent treatments for adolescents with disruptive behavior (McCart & Sheidow, 2016). Furthermore, some studies suggest that parenting behavior mediates the effect of parent training (Forehand, Lafko, Parent, & Burt, 2014). Accordingly, it might be expected that higher parental involvement would positively affect outcome. However, whether the treatment targeted parents, children or both, or the degree of parental involvement did not moderate effect size significantly. These results are in congruence with previous studies specifically assessing the potential effect of involving parents or children in treatment, which have typically not found significant effects (Corcoran & Dattalo, 2006; Lundahl, Risser, & Lovejoy, 2006; Thulin, Svirsky, Serlachius, Andersson, & Öst, 2014). It might be that the content of the parent intervention is of greater importance than the quantity. For instance, Kaminski, Valle, Filene, and Boyle (2008) found that treatment programs with components that focused on increasing positive parent–child interactions and emotional communication skills, teaching parents to use time out and the importance of parenting consistency, and requiring parents to practice new skills with their children during parent training sessions were associated with better outcomes. On the other hand, they found that treatments that focused on teaching parents problem solving, teaching parents to promote children's

Table 5
Subgroup analysis of the effect size and remission for all studies at post-treatment.

	Effect size			Remission		
Variable	k	*g*	95% CI	k	%	95% CI
Type of study ($Q_b = 0.02$, $p = 0.90$)*				($Q_b = 0.006$, $p = 0.94$)*		
RCT	50	0.90	0.79–1.02	25	43.9	35.2–53.1
Open Trial	15	0.92	0.77–1.06	6	43.5	35.7–51.5
Statistical analysis ($Q_b = 1.15$, $p = 0.29$)*				($Q_b = 3.07$, $p = 0.08$)*		
Intent-to-treat	33	0.86	0.73–0.99	16	38.4	27.6–50.4
Completers	32	0.97	0.82–1.12	15	50.2	44.6–55.8
Target of treatment ($Q_b = 0.46$, $p = 0.80$)*				($Q_b = 1.51$, $p = 0.47$)*		
Child	7	0.82	0.56–1.09	5	32.4	11.5–63.9
Parent	37	0.91	0.77–1.06	15	48.9	38.9–59.1
Child and Parent	21	0.93	0.77–1.08	11	42.0	32.0–52.8
Treatment format ($Q_b = 3.70$, $p = 0.16$)*				($Q_b = 0.80$, $p = 0.37$)*		
Group	32	0.81	0.68–0.94	19	44.9	36.6–53.6
Individual	29	0.98	0.84–1.11	11	37.5	25.2–51.6
Group + Individual	4	1.11	0.62–1.59	—		
Parental involvement ($Q_b = 1.25$, $p = 0.54$)*				($Q_b = 0.24$, $p = 0.89$)*		
Low	5	0.84	0.46–1.22	4	42.4	17.5–71.9
Moderate	27	0.86	0.72–1.00	12	42.1	33.7–51.1
High	33	0.97	0.82–1.12	15	45.9	34.0–58.2
Teacher involvement ($Q_b = 0.52$, $p = 0.48$)*				($Q_b = 0.88$, $p = 0.35$)*		
Yes	12	0.84	0.62–1.05	7	48.2	42.0–54.5
No	53	0.93	0.82–1.04	24	42.7	33.6–52.4
Therapist profession ($Q_b = 2.51$, $p = 0.29$)*				($Q_b = 1.74$, $p = 0.42$)*		
Psychologist	25	0.90	0.74–1.07	9	37.2	21.2–56.5
Social worker	7	0.79	0.55–1.03	6	50.6	43.9–57.2
Various professions	16	1.06	0.83–1.30	7	52.4	30.0–73.8
Continent ($Q_b = 6.77$, $p = 0.04$)*				($Q_b = 2.36$, $p = 0.31$)*		
Australia	5	1.37	0.99–1.75	4	59.9	37.6–78.8
North America	28	0.91	0.77–1.05	11	42.7	32.0–54.3
Europe	32	0.83	0.69–0.97	16	40.6	30.6–51.4

Note: k = number of treatment conditions, Q_b = Q between subgroups. * The statistic in parenthesis tests if the subgroups within the individual category differ significantly from each other.

Table 6
Meta-regression analysis of the effect size and remission for all studies at post-treatment.

	Effect Size				Remission			
Variable	k	Point	*z*-value	*p*-value	k	Point	*z*-value	*p*-value
Mean age	65	−0.053	−5.53	0.00001*	31	−0.173	−6.10	0.00001*
% drug treatment	40	−0.006	−4.84	0.00001*	19	−0.018	−6.64	0.00001*
Attrition	59	−0.010	−4.91	0.00001*	30	−0.027	−6.72	0.00001*
Pre-treatment severity	64	0.643	3.05	0.0023*	30	−0.337	−0.51	0.61
Risk-of-bias	65	−0.045	−2.81	0.0047*	31	−0.064	−1.65	0.099
Percent boys	64	0.007	2.81	0.0049	29	0.008	1.60	0.11
Intensity of treatment	60	−0.009	−2.76	0.006	29	0.043	2.65	0.008
# of sessions	60	0.004	2.71	0.007	29	0.026	3.10	0.002*
1 outlier deleted	59	0.006	1.85	0.064	–			
Percent declining	47	−0.007	−2.26	0.024	23	0.065	0.64	0.52
Methodology score	65	0.037	1.50	0.13	31	−0.804	−1.72	0.086
# of participants	65	−0.002	2.95	0.003	31	−0.009	−8.35	0.0001
1 outlier deleted	64	−0.000	−0.11	0.91	30	−0.014	−3.34	0.008

Note: k = number of treatment conditions, Point = point estimate. * Significant using the Holm-Bonferroni correction.

Table 7
Some background and treatment data (M and SD) for effectiveness and efficacy studies in the different disorders.

Disorder	k	Age (years)	% Boys	% Comorb.	% on meds.	Severity	% Attrition	Tx time
ADHD		$p = 0.95$	$p = 0.05$	$p = 0.26$	$p = 0.96$	$p = 0.62$	$p = 0.28$	$p = 0.03$
Effectiveness	25	8.9 (2.6)	77.7 (7.7)	41.1 (25.8)	42.4 (31.8)	60.3 (16.2)	14.3 (18.9)	18.2 (8.4)
Efficacy	43	8.9 (3.3)	73.0 (10.1)	49.6 (25.8)	42.0 (27.6)	58.5 (14.2)	10.2 (12.3)	23.5 (17.4)
CD-ODD		$p = 0.82$	$p = 0.03$	$p = 0.001$	$p = 0.10$	$p = 0.04$	$p = 0.15$	$p = 0.87$
Effectiveness	39	7.6 (2.6)	76.9 (14.2)	62.3 (19.9)	21.1 (22.4)	65.0 (8.2)	15.4 (14.5)	24.4 (13.7)
Efficacy	62	7.5 (3.2)	70.2 (14.6)	18.0 (18.3)	11.7 (11.4)	60.6 (10.6)	11.4 (10.0)	23.7 (21.3)

Note: k = number of treatment conditions, % Comorb. = proportion having any psychiatric comorbid disorder, % on meds. = proportion on any psychotropic medication, Severity = percentage of the maximum score on the primary outcome measure. % Attrition = proportion dropping out of those participating in at least one session. Tx time = number of 60 min therapy hours.

Table 8
Effect sizes (Hedges' g) for effectiveness and efficacy studies within the different disorders.

Disorder	Study type	k	*g*	95% CI	*z*-value	Q_b†	*p*-value
	Post-treatment						
ADHD	Effectiveness	26	0.80	0.67–0.93	11.76[a]		
	Efficacy	40	0.74	0.61–0.87	11.05[a]	0.40	0.53
CD-ODD	Effectiveness	38	0.98	0.85–1.12	14.10[a]		
	Efficacy	58	1.07	0.92–1.21	14.77[a]	0.65	0.42
	Follow-up						
ADHD	Effectiveness	14	0.88	0.70–1.07	9.52[a]		
	Efficacy	19	1.06	0.85–1.28	9.69[a]	1.51	0.22
CD-ODD	Effectiveness	30	1.06	0.88–1.24	11.58[a]		
	Efficacy	38	1.10	0.94–1.26	13.79[a]	0.09	0.77

Note: k = number of comparisons. [a] $p < 0.0001$. Q_b = Q between, † Comparison Efficacy vs. Effectiveness within the respective disorders.

Table 9
Remission rates for effectiveness and efficacy studies for the different disorders.

Disorder	Study type	k	%	95% CI	*z*-value*	Q_b†	*p*-value
	Post-treatment						
ADHD	Effectiveness	12	37.7	26.7–50.2	1.93		
	Efficacy	17	33.4	23.2–45.4	2.67[a]	0.26	0.81
CD-ODD	Effectiveness	19	47.6	38.0–57.3	0.49		
	Efficacy	23	54.8	48.1–61.4	1.40	1.43	0.23
	Follow-up						
ADHD	Effectiveness	5	38.2	21.8–57.9	1.18		
	Efficacy	2	44.3	36.5–52.7	1.37	0.33	0.57
CD-ODD	Effectiveness	12	56.8	46.5–66.5	1.29		
	Efficacy	23	54.9	47.3–62.3	1.27	0.08	0.77

Note: k = number of comparisons. [a] $p < 0.01$. * Test if significantly different from 50%. Q_b = Q between, † Comparison Efficacy vs. Effectiveness within the respective disorders.

Table 10
Effect sizes for randomized controlled studies only: ADHD and CD-ODD combined.

Study type	k	*ES*	95% CI	*z*-value	Q_b†	*p*-value
g-value at post-treatment						
Effectiveness	49	0.90	0.78–1.02	14.86[a]		
Efficacy	98	0.93	0.83–1.03	18.16[a]	0.08	0.79
g-value at follow-up						
Effectiveness	38	0.99	0.84–1.14	13.29[a]		
Efficacy	57	1.09	0.96–1.21	17.05[a]	0.97	0.32
Remission rate at post-treatment						
Effectiveness	25	43.9	35.2–53.1	1.30		
Efficacy	40	46.1	39.5–52.9	1.12	0.14	0.71
Remission rate at follow-up						
Effectiveness	16	50.9	40.4–61.4	0.17		
Efficacy	25	53.7	47.0–60.4	1.08	0.19	0.66

Note: k = number of comparisons. [a] $p < 0.0001$. Q_b = Q between, † Comparison Effectiveness vs. Efficacy.

cognitive, academic, or social skills, and providing other, additional services were associated with smaller effect sizes.

The majority of the included studies delivered the treatment in groups, but treatment format (i.e. individual, group or combined individual and group treatment) or involving teachers in treatment did not significantly moderate the effect sizes. Although the involvement of teachers may be beneficial in certain cases, the results suggest that teacher involvement is not a crucial factor for treatment outcome in routine care, and that CBT for externalizing disorders is effective across different formats in routine care. The treatments had a mean number of 17 sessions across a mean of 15 weeks, with a mean intensity of approximately three hours per week. There was no evidence that treatment intensity significantly moderated outcome; however, number of sessions positively affected treatment outcome, i.e. more sessions yielded higher effect sizes. Previous meta-analyses have demonstrated similar findings (Fossum et al., 2016; Shelleby & Kolko, 2015).

Regarding patient characteristics, lower age was associated with a better outcome of treatment. It has been argued that externalizing problems become increasingly resistant to change with age (e.g. Bernazzani, Côté, & Tremblay, 2001; Deković & Stoltz, 2015), and our results fit well with that assumption. Moreover, most of the studies in this meta-analysis involved parent behavior therapy, which is found to have a stronger effect on the youngest children (McCart et al., 2006). In line with these findings, a recent meta-analysis on psychosocial treatment for conduct disorder also found larger reductions in conduct problems in younger than in older children (Fossum et al., 2016). However, the literature is not consistent on this matter; two previous meta-analyses on parent training found no significant association between age and treatment outcome (Cedar & Levant, 1990; Lundahl et al., 2006), whereas another found that older children benefitted more from behavioral parent training (Serketich & Dumas, 1996). Accordingly, no general conclusions can be drawn regarding the effect of age on treatment outcome. Nevertheless, the findings from the present meta-analysis suggest that, in routine care, younger age is associated with greater symptom reductions and higher remission rates after CBT.

More severe symptoms at pre-treatment were associated with a larger effect size. This seems to be a common finding (e.g. Lundahl et al., 2006) and a plausible reason is that it reflects that more severe cases have a larger room for improvement. This explanation is supported by the fact that remission rates were not associated with pre-treatment severity. The proportion of patients on psychotropic drugs at the beginning of treatment was a negative moderator of ES and remission rates. These two findings might seem somewhat contradictory, since it may be assumed that the samples with high pre-treatment severity would also have a high proportion of participants on psychotropic drugs. However, we did not include studies that investigated the effect of combination treatment. Hence, the participants were, as part of routine care, medicated if indicated and on a stable dose before starting CBT-treatment. Therefore, the participants who were already on psychotropic drugs, may in fact have had less severe symptoms at pre-treatment, leaving a smaller room for improvement. It does not, however, explain why a higher proportion of patients on psychotropic drugs was associated with lower remission rates. The finding probably reflects that the ADHD studies had a significantly higher proportion of patients receiving medication while also yielding lower effect size and remission rates than CD/ODD studies (although not significantly lower).

The most important and encouraging finding from the present meta-analysis was that the effectiveness-efficacy comparisons demonstrated no significant differences in ES between effectiveness and efficacy studies. The results are in line with a meta-analysis by Michelson, Davenport, Dretzke, Barlow, and Day (2013) investigating the outcome of parent management training (PMT) across different real-world practice contexts. The authors found that whether PMT was delivered to clinically referred or study recruited samples, in service-oriented or research settings, or by non-specialist or specialist therapists, there was no difference in the overall outcome of PMT. The findings from the present meta-analysis lend further support to the transportability of empirically supported CBT-treatments for externalizing disorders from university settings to routine clinical practice. The fact that the effectiveness studies in the present meta-analysis seemed to have somewhat more severe CD/ODD samples with higher proportion of comorbid disorders provide additional support for the robustness of CBT in routine care.

3.1. Strengths and limitations

A methodological strength is that a power analysis indicated that this meta-analysis has a high power to detect a small effect size (based on the number of effect sizes summarized, the mean number of participants in the treatment conditions, and the observed degree of heterogeneity). Statistical power is rarely calculated in meta-analyses, but the importance of power calculations in the interpretation of results has been highlighted (Hedges & Pigott, 2001, 2004). Another methodological strength is that all extractions of information from the included studies were done in pairs and any disparities were solved in consensus with all authors. A third strength is that the ratings of methodological quality and risk-of-bias was done by one of the authors and independently and blindly by another, yielding excellent inter rater reliability (ICC 0.94; 0.95). Still, the most important strength of the current meta-analysis is that it is the first to investigate the effects of CBT for children and adolescents in routine care with a diagnosis of ADHD, CD or ODD. The results therefore give valuable information for clinicians and decision makers. The findings are encouraging and indicate that externalizing disorders in childhood and adolescence respond well to CBT-oriented treatments in routine care.

The present meta-analysis has limitations that should be considered. First, the quality of a meta-analysis is limited by the quality of the original studies and can never be better than the data it summarizes. Hence, an important limitation is that few studies reported measures of functioning or quality of life or comorbidity. Second, the judgment of inclusion was in some cases challenging due to poor descriptions of the treatment setting, the clinicians and/or referral procedures. This might have caused exclusion of some studies that should have been included. Third, titles and/or abstracts were screened by one rater only. The PRISMA guideline recommends the use of two independent raters when screening title and abstract as this procedure may reduce the possibility of rejecting relevant reports. However, by having two independent raters reading the 730 full-text articles and including 7.0% of them in the current meta-analyses the risk of missing out on relevant reports may be low. The inclusion of both RCTs and open trials is a further limitation. However, in the present meta-analysis we wanted to calculate the amount of improvement that can be achieved in routine clinical care and

compare it with efficacy studies. Both RCTs and open trials could contribute with data for calculating within-group effect size and were therefore included. Finally, as only English-language peer-reviewed journal articles were included we may have missed relevant studies not published in peer-reviewed journals or published in languages other than English.

4. Conclusion

The findings from the present meta-analysis demonstrate the effectiveness of CBT in routine clinical care for children and adolescents with externalizing disorders. The results showed that substantial effects can be obtained in routine care across different formats (i.e. individual or group), with different participants (i.e. parents, child or both), and with varying degrees of parental involvement. Finally, the results suggest that treatment delivered in routine care by practicing clinicians to clinically referred participants is as effective as treatments delivered in university settings.

Funding

No external funding was obtained for this meta-analysis.

Contributors

ENR elaborated the protocol registration in PROSPERO and wrote the first draft of the introduction and discussion. GJW rated the methodology and risk of bias of the included studies. ENR and GJW conducted literature searches in collaboration with an academic librarian. UN co-wrote the first draft of introduction and discussion with ENR. L-GÖ designed the meta-analysis, wrote the coding scheme, co-rated the methodology and risk of bias of the studies, meta-analyzed the included studies, and wrote the first draft of methods and results. All authors contributed in the screening process, the extraction of data, and writing of the paper. All authors have approved the final manuscript.

Declaration of Competing Interest

None of the authors have any conflict of interest to report.

Acknowledgements

The authors wish to thank academic librarian Regina Küfner Lein for assistance with the literature search.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2020.101954.

References

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile.* Burlington, Vt: University of Vermont, Department of Psychiatry.

Bakker, M. J., Greven, C. U., Buitelaar, J. K., & Glennon, J. C. (2017). Practitioner review: Psychological treatments for children and adolescents with conduct disorder problems - a systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry, 58*(1), 4–18. https://doi.org/10.1111/jcpp.12590.

Battagliese, G., Caccetta, M., Luppino, O. I., Baglioni, C., Cardi, V., Mancini, F., & Buonanno, C. (2015). Cognitive-behavioral therapy for externalizing disorders: A meta-analysis of treatment effectiveness. *Behaviour Research and Therapy, 75*, 60–71. https://doi.org/10.1016/j.brat.2015.10.008.

Bernazzani, O., Côté, C., & Tremblay, R. E. (2001). Early parent training to prevent disruptive behavior problems and delinquency in children. *The Annals of the American Academy of Political and Social Science, 578*(1), 90–103. https://doi.org/10.1177/000271620157800106.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2013). *Comprehensive meta-analysis (version 3.3.070).* Englewood, NJ: Biostat.

Burkey, M. D., Hosein, M., Morton, I., Purgato, M., Adi, A., Kurzrok, M., … Tol, W. A. (2018). Psychosocial interventions for disruptive behaviour problems in children in low-and middle-income countries: A systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry, 59*(9), 982–993. https://doi.org/10.1111/jcpp.12894.

Cedar, B., & Levant, R. F. (1990). A meta-analysis of the effects of parent effectiveness training. *The American Journal of Family Therapy, 18*(4), 373–384. https://doi.org/10.1080/01926189008250986.

Charach, A., Carson, P., Fox, S., Ali, M. U., Beckett, J., & Lim, C. G. (2013). Interventions for preschool children at high risk for ADHD: A comparative effectiveness review. *Pediatrics, 131*(5), e1584–e1604. https://doi.org/10.1542/peds.2012-0974.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284.

Clark, C., Prior, M., & Kinsella, G. (2002). The relationship between executive function abilities, adaptive behaviour, and academic achievement in children with externalising behaviour problems. *Journal of Child Psychology and Psychiatry, 43*(6), 785–796. https://doi.org/10.1111/1469-7610.00084.

Copeland, W. E., Shanahan, L., Costello, E. J., & Angold, A. (2009). Childhood and adolescent psychiatric disorders as predictors of young adult disorders. *Archives of General Psychiatry, 66*(7), 764–772. https://doi.org/10.1001/archgenpsychiatry.2009.85.

Corcoran, J., & Dattalo, P. (2006). Parent involvement in treatment for ADHD: A meta-analysis of the published studies. *Research on Social Work Practice, 16*(6), 561–570. https://doi.org/10.1177/1049731506289127.

Deković, M., & Stoltz, S. (2015). Moderators and mediators of treatments for youth who show externalizing problem behavior. In M. Maric, P. J. M. Prins, & T. H. Ollendick (Eds.), *Moderators and mediators of youth treatment outcomes* (pp. 97–122). New York, NY: Oxford University Press.

DuPaul, G. J., & Eckert, T. L. (1997). The effects of school-based interventions for attention deficit hyperactivity disorder: A meta-analysis. *School Psychology Review, 26*(1), 5–27. https://doi.org/10.1080/02796015.1997.12085845.

DuPaul, G. J., Eckert, T. L., & Vilardo, B. (2012). The effects of school-based interventions for attention deficit hyperactivity disorder: A meta-analysis 1996–2010. *School Psychology Review, 41*(4), 387–412. https://doi.org/10.1080/02796015.2012.12087496.

DuPaul, G. J., Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Academic and social functioning associated with attention-deficit/hyperactivity disorder: Latent class analyses of trajectories from kindergarten to fifth grade. *Journal of Abnormal Child Psychology, 44*(7), 1425–1438. https://doi.org/10.1007/s10802-016-0126-z.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.

Efron, D., Bryson, H., Lycett, K., & Sciberras, E. (2016). Children referred for evaluation for ADHD: Comorbidity profiles, and characteristics associated with a positive diagnosis. *Child: Care, Health and Development, 42*, 718–724. https://doi.org/10.1111/cch.12364.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629.

Erford, B. T., Paul, L. E., Oncken, C., Kress, V. E., & Erford, M. R. (2014). Counseling outcomes for youth with oppositional behavior: A meta-analysis. *Journal of Counseling & Development, 92*(1), 13–24. https://doi.org/10.1002/j.1556-6676.2014.00125.x.

Evans, S. W., Owens, J. S., & Bunford, N. (2014). Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. *Journal of Clinical Child & Adolescent Psychology, 43*(4), 527–551. https://doi.org/10.1080/15374416.2013.850700.

Evans, S. W., Owens, J. S., Wymbs, B. T., & Ray, A. R. (2018). Evidence-based psychosocial treatments for children and adolescents with attention deficit/hyperactivity disorder. *Journal of Clinical Child & Adolescent Psychology, 47*(2), 157–198. https://doi.org/10.1080/15374416.2017.1390757.

Eyberg, S. M., & Pincus, D. (1999). *Eyberg child behavior inventory and Sutter-Eyberg student behavior inventory: Professional manual.* Odessa, FL: Psychological Assessment Resouces.

Fabiano, G. A., Pelham, W. E., Coles, E. K., Gnagy, E. M., Chronis-Tuscano, A., & O'Connor, B. C. (2009). A meta-analysis of behavioral treatments for attention-deficit/hyperactivity disorder. *Clinical Psychology Review, 29*(2), 129–140. https://doi.org/10.1016/j.cpr.2008.11.001.

Fabiano, G. A., Schatz, N. K., Aloe, A. M., Chacko, A., & Chronis-Tuscano, A. (2015). A systematic review of meta-analyses of psychosocial treatment for attention-deficit/hyperactivity disorder. *Clinical Child and Family Psychology Review, 18*(1), 77–97. https://doi.org/10.1007/s10567-015-0178-6.

Forehand, R., Lafko, N., Parent, J., & Burt, K. B. (2014). Is parenting the mediator of change in behavioral parent training for externalizing problems of youth? *Clinical Psychology Review, 34*(8), 608–619. https://doi.org/10.1016/j.cpr.2014.10.001.

Fossum, S., Handegård, B. H., Martinussen, M., & Mørch, W. T. (2008). Psychosocial interventions for disruptive and aggressive behaviour in children and adolescents. *European Child & Adolescent Psychiatry, 17*(7), 438–451. https://doi.org/10.1007/s00787-008-0686-8.

Fossum, S., Handegård, B. H., Adolfsen, F., Vis, S. A., & Wynn, R. (2016). A meta-analysis of long-term outpatient treatment effects for children and adolescents with conduct problems. *Journal of Child and Family Studies, 25*(1), 15–29. https://doi.org/10.1007/s10826-015-0221-8.

Franke, B., Michelini, G., Asherson, P., Banaschewski, T., Bilbow, A., Buitelaar, J. K., … Reif, A. (2018). Live fast, die young? A review on the developmental trajectories of

ADHD across the lifespan. *European Neuropsychopharmacology, 28*(10), 1059–1088. https://doi.org/10.1016/j.euroneuro.2018.08.001.
Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods, 6*(3), 203. https://doi.org/10.1037/1082-989X.6.3.203.
Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426. https://doi.org/10.1037/1082-989X.9.4.426.
Higgins, J. P. T., Altman, D. G., & Sterne, J. A. (2011). Assessing risk of bias in included studies. In J. P. T. Higgins, & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions v. 5.1.0. S.I.* The Cochrane collaboration.
Jaccard, J., & Guilamo-Ramos, V. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Issues and recommendations. *Journal of Clinical Child & Adolescent Psychology, 31*(1), 130–146. https://doi.org/10.1207/S15374424JCCP3101_15.
Jarrett, M. A., & Ollendick, T. H. (2008). A conceptual review of the comorbidity of attention-deficit/hyperactivity disorder and anxiety: Implications for future research and practice. *Clinical Psychology Review, 28*(7), 1266–1280. https://doi.org/10.1016/j.cpr.2008.05.004.
Jerrell, J. M., McIntyre, R. S., & Park, Y. M. (2015). Risk factors for incident major depressive disorder in children and adolescents with attention-deficit/hyperactivity disorder. *European Child and Adolescent Psychiatry, 24*, 65–73.
Kaminski, J. W., & Claussen, A. H. (2017). Evidence base update for psychosocial treatments for disruptive behaviors in children. *Journal of Clinical Child & Adolescent Psychology, 46*(4), 477–499. https://doi.org/10.1080/15374416.2017.1310044.
Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36*(4), 567–589. https://doi.org/10.1007/s10802-007-9201-9.
Klassen, A., Miller, A., Raina, P., Lee, S. K., & Olsen, L. (1999). Attention-deficit hyperactivity disorder in children and youth: A quantitative systematic review of the efficacy of different management strategies. *The Canadian Journal of Psychiatry, 44*(10), 1007–1016. https://doi.org/10.1177/070674379904401007.
Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*(863). https://doi.org/10.3389/fpsyg.2013.00863.
Lee, P.-C., Niew, W.-I., Yang, H.-J., Chen, V. C.-H., & Lin, K.-C. (2012). A meta-analysis of behavioral parent training for children with attention deficit hyperactivity disorder. *Research in Developmental Disabilities, 33*(6), 2040–2049. https://doi.org/10.1016/j.ridd.2012.05.011.
Lee, C. M., Horvath, C., & Hunsley, J. (2013). Does it work in the real world? The effectiveness of treatments for psychological problems in children and adolescents. *Professional Psychology: Research and Practice, 44*(2), 81–88. https://doi.org/10.1037/a0031133.
Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., … Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine, 6*(7), Article e1000100. https://doi.org/10.1371/journal.pmed.1000100.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA, US: Sage Publications, Inc.
Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review, 26*(1), 86–104. https://doi.org/10.1016/j.cpr.2005.07.004.
McCart, M. R., & Sheidow, A. J. (2016). Evidence-based psychosocial treatments for adolescents with disruptive behavior. *Journal of Clinical Child & Adolescent Psychology, 45*(5), 529–563. https://doi.org/10.1080/15374416.2016.1146990.
McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology, 34*(4), 525–541. https://doi.org/10.1007/s10802-006-9031-1.
Michelson, D., Davenport, C., Dretzke, J., Barlow, J., & Day, C. (2013). Do evidence-based interventions work when tested in the "real world?" A systematic review and meta-analysis of parent management training for the treatment of child disruptive behavior. *Clinical Child and Family Psychology Review, 16*(1), 18–34. https://doi.org/10.1007/s10567-013-0128-0.
Nahmias, A. S., Pellecchia, M., Stahmer, A. C., & Mandell, D. S. (2019). Effectiveness of community-based early intervention for children with autism spectrum disorder: A meta-analysis. *Journal of Child Psychology and Psychiatry, 60*(11), 1200–1209. https://doi.org/10.1111/jcpp.13073.
O'Connell, M. E., Boat, T., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academies Press.
Öst, L.-G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy, 46*(3), 296–321. https://doi.org/10.1016/j.brat.2007.12.005.
Owens, E. B., & Hinshaw, S. P. (2013). Perinatal problems and psychiatric comorbidity among children with ADHD. *Journal of Clinical Child and Adolescent Psychology, 42*, 762–768. https://doi.org/10.1080/15374416.2013.785359.
Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of a parent-training program. *Behavior Therapy, 13*, 638–650. https://doi.org/10.1016/S0005-7894(82)80021-X.
Schatz, D. B., & Rostain, A. L. (2006). ADHD with comorbid anxiety: A review of the current literature. *Journal of Attention Disorders, 10*(2), 141–149. https://doi.org/10.1177/1087054706286698.
Schatz, N. K., Aloe, A. M., Fabiano, G. A., Pelham, W. E. J., Smyth, A., Zhao, X., … Altszuler, A. R. (2020). Psychosocial interventions for attention-deficit/hyperactivity disorder: Systematic review with evidence and gap maps. *Journal of Developmental & Behavioral Pediatrics, 41*, S77–S87. https://doi.org/10.1097/dbp.0000000000000778.
Scott, J. G., Gjørtz Pedersen, M., Erskine, H. E., Bikic, A., Demontis, D., McGrath, J. J., & Dalsgaard, S. (2017). Mortality in individuals with disruptive behavior disorders diagnosed by specialist services – A nationwide cohort study. *Psychiatry Research, 251*, 255–260. https://doi.org/10.1016/j.psychres.2017.02.029.
Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27*(2), 171–186. https://doi.org/10.1016/S0005-7894(96)80013-X.
Shadish, W. R., Matt, G. E., Navarro, A. M., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions. *Journal of Consulting and Clinical Psychology, 126*, 512–529. https://doi.org/10.1037//0033-2909.126.4.512.
Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., … Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ, 358*, j4008. https://doi.org/10.1136/bmj.j4008.
Shelleby, E. C., & Kolko, D. J. (2015). Predictors, moderators, and treatment parameters of community and clinic-based treatment for child disruptive behavior disorders. *Journal of Child and Family Studies, 24*(3), 734–748. https://doi.org/10.1007/s10826-013-9884-1.
Sibley, M. H., Kuriyan, A. B., Evans, S. W., Waxmonsky, J. G., & Smith, B. H. (2014). Pharmacological and psychosocial treatments for adolescents with ADHD: An updated systematic review of the literature. *Clinical Psychology Review, 34*(3), 218–232. https://doi.org/10.1016/j.cpr.2014.02.001.
Sonuga-Barke, E. J., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., … Sergeant, J. (2013). Nonpharmacological interventions for ADHD: Systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *American Journal of Psychiatry, 170*(3), 275–289. https://doi.org/10.1176/appi.ajp.2012.12070991.
Southam-Gerow, M. A., & Prinstein, M. J. (2014). Evidence base updates: The evolution of the evaluation of psychological treatments for children and adolescents. *Journal of Clinical Child & Adolescent Psychology, 43*(1), 1–6. https://doi.org/10.1080/15374416.2013.855128.
Steinhausen, H. C., Novik, T. S., Baldursson, G., Curatolo, P., Lorenzo, M. J., Pereira, R. R., & Ralston, S.J. og Rothenberger, A.. (2006). Co-existing psychiatric problems in ADHD in the ADORE cohort. *European Child and Adolescent Psychiatry, 15*(s1), 25–29. https://doi.org/10.1007/s00787-006-1004-y.
Takeda, T., Ambrosini, P. J., deBerardinis, R., & Elia, J. (2012). What can ADHD without comorbidity teach us about comorbidity? *Research in Developmental Disabilities, 33*, 419–425. https://doi.org/10.1016/j.ridd.2011.09.024.
Thomas, R., Sanders, S., Doust, J., Beller, E., & Glasziou, P. (2015). Prevalence of attention-deficit/hyperactivity disorder: A systematic review and meta-analysis. *Pediatrics, 135*(4), e994–e1001. https://doi.org/10.1542/peds.2014-3482.
Thulin, U., Svirsky, L., Serlachius, E., Andersson, G., & Öst, L.-G. (2014). The effect of parent involvement in the treatment of anxiety disorders in children: A meta-analysis. *Cognitive Behaviour Therapy, 43*(3), 185–200. https://doi.org/10.1080/16506073.2014.923928.
Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D., & Thombs, B. D. (2015). Empirically supported treatment: Recommendations for a new model. *Clinical Psychology: Science and Practice, 22*(4), 317–338. https://doi.org/10.1111/cpsp.12122.
Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics, 35*(2), 215–247. https://doi.org/10.3102/1076998609346961.
Van der Oord, S., Prins, P. J. M., Oosterlaan, J., & Emmelkamp, P. M. G. (2008). Efficacy of methylphenidate, psychosocial treatments and their combination in school-aged children with ADHD: A meta-analysis. *Clinical Psychology Review, 28*(5), 783–800. https://doi.org/10.1016/j.cpr.2007.10.007.
Wergeland, G. J. H., Riise, E. N., & Öst, L. G. (2021). Cognitive behavior therapy for internalizing disorders in children and adolescents in routine clinical care: A systematic review and meta-analysis. *Clinical Psychology Review, 101918*. https://doi.org/10.1016/j.janxdis.2016.08.003.
Wolff, J. C., & Ollendick, T. H. (2006). The comorbidity of conduct problems and depression in childhood and adolescence. *Clinical Child and Family Psychology Review, 9*(3), 201–220. https://doi.org/10.1007/s10567-006-0011-3.
Zwi, M., Jones, H., Thorgaard, C., York, A., & Dennis, J. A. (2011). Parent training interventions for attention deficit hyperactivity disorder (ADHD) in children aged 5 to 18 years. *Cochrane Database of Systematic Reviews, 12*. https://doi.org/10.1002/14651858.CD003018.pub3.

Eili Nygard Riise earned her Ph.D. degree in psychology from the Faculty of Psychology, University of Bergen, Norway in 2019. She is currently working as clinical psychologist at Sogndal Outpatient Clinic for Child and Adolescent Psychiatry, District General Hospital of Førde. Her research interest areas include treatment for child and adolescent mental disorders and treatment effectiveness in real world settings.

Gro Janne Wergeland earned her Ph.D. from the Faculty of Medicine, University of Bergen, Norway in 2015. She is currently working as Consultant Child and Adolescent Psychiatrist at an outpatient clinic for Child and Adolescent Psychiatry, Haukeland University Hospital, Bergen, and as Associate Professor at Faculty of Medicine, University of Bergen. Her research interest has primarily focused on anxiety disorders in children and treatment effectiveness in real world settings.

Urdur Njardvik earned her Ph.D. degree in clinical child psychology from Louisiana State University in 2000 and completed a clinical internship in pediatric psychology at Johns Hopkins University Medical School the same year. She is a professor of clinical child psychology in the Department of Psychology at the University of Iceland. Her research interests include disruptive behavior problems, ADHD, comorbidity, and pediatric obesity.

Lars-Göran Öst earned his Ph.D. in psychology from Stockholm University, Sweden, in 1976 and Dr. Med. Sci. in psychiatry from Uppsala University, Sweden, in 1985. He was Professor of clinical psychology, at the Dept. of Psychology, Stockholm University for 20 years, and then became Professor emeritus. He has carried out research in CBT for almost 50 years and published 300 journal articles, book chapters, and books. The research has primarily focused on anxiety disorders in both adults and children. He has developed and tested three evidence-based treatment methods: One-session treatment for specific phobias, Applied relaxation for generalized anxiety disorder, and Applied tension for blood phobia.