University of Pavia

Department of Brain and Behavioural Sciences

Doctoral Program in "Psychology, Neuroscience and Data Science"

Coordinator: Prof. Gabriella Bottini

XXXII cycle

Interventions aimed at restricted and repetitive behaviours (RRBs) in Autism Spectrum Disorder: systematic review and meta-analysis

Supervisor

Prof. Pierluigi Politi

Doctoral Candidate

Umberto Provenzani

Academic Year 2019/2020

TABLE OF CONTENTS

Summary

Background: Despite being considered a core feature of ASD, restricted and repetitive behaviours (RRBs) have received less attention, if compared to the domain of social interaction and communication, and less frequently targeted by interventions. This is surprising, given their role as major management challenges, obstacles to adaptive functioning and cause of distress for subjects and their families. We conducted a systematic, exhaustive, and up-to-date systematic review and meta-analysis of randomized controlled interventions (RCTs) aimed specifically at RRBs. To avoid methodological limitations found in other reviews, no limitations to the age of the sample and to the type of intervention were set.

Methods*:* Web of Knowledge database (including Web of Science, MEDLINE®, KCI – Korean Journal Database, Russian Science Citation Index and SciELO Citation Index) was searched from inception up to January 1st, 2020. Randomized controlled trials in ASD individuals, specifically aimed at RRBs or both core domains were included, following PRISMA guidelines. A systematic review was analyse the main characteristics of included studies, such as mean age, sample sizes, mean follow up duration, diagnosis of ASD, assessments of IQ and psychiatric comorbidities. Primary outcome of the meta-analysis was the mean reduction of RRBs; effect sizes reported as Hedges' g and 95% CIs, calculated as differences, from baseline to endpoint, between two

compared interventions. Assessments of biases, comprising publication bias, and cumulative analyses were also performed.

***Results**:* Overall, 80 studies (3114 subjects) were included in the systematic review and 46 studies (1339 subjects) in the meta-analytic phase. Included studies were published between 1992 and 2019, mean sample size was 40 patients (in intervention arm), with mean age of 10.5 years and average 19% female participants. Mean follow up was 4 months. IQ assessment was unclear in half of the studies, other psychiatric comorbidities were not disclosed in 61% of the studies. Risk of bias was low in 14 studies (17.5%).

Overall effect size for interventions aimed at RRBs was small, but significantly beneficial (g = 0.37, Cis -0.26 to -0.47), heterogeneity was moderate (I_2 = 43.92%, p < 0.01). Subgroup analyses revealed similar results in the three subtypes of interventions analyzed: pharmacological (g = 0.45, CI -0.26 to -0.64), psychotherapy and education (g = 0.42, CI -0.19 to -0.65) and complementary interventions (g = 0.25, CI -0.12 to -0.38). Differences were not significant among intervention types (p = 0.16). Inspection of forest and funnel plots revealed the presence of five outliers, which presence influenced heterogeneity significantly, but did not affected substantially the magnitude of overall and subgroup effect sizes. Results were also not affected by small-study effects, but publication bias was probably present since grey/unpublished literature was not searched.

Conclusions*:* on the basis of current literature aimed specifically at RRBs, there is no robust evidence to favour any specific intervention for improving RRBs in subjects with ASD, even if small effects were detected for any intervention type analysed.

Introduction

Classification

Since the first descriptions (Kanner, 1943), a broad range of repetitive behaviours have been associated with autism, including motor stereotypy, rituals, compulsions, obsessions, sameness behaviours, echolalia, and self-injury. Together with persistent deficits in social communication and social interests, restricted and repetitive behaviours (RRBs) are now acknowledged as one of the two core symptom domains of Autism Spectrum Disorder (ASD). This domain comprises a broad range of manifestations characterized as rigid, perseverative and limited in scope (Klinger, Dawson, & Renner, 2003; Lawrence Scahill et al., 2015; Turner, 1999).

DSM-5 (APA, 2013) divides RRBs into four subtypes:

1. Stereotyped or repetitive motor movements, use of objects, or speech
2. Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behaviour
3. Highly restricted, fixated interests that are abnormal in intensity or focus
4. Hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment

DSM-5 classification concurs with previous research (Anagnostou et al., 2011; Cuccaro et al., 2003; Lam, Bodfish, & Piven, 2008; Shao et al., 2002; Szatmari et al., 2006) which

analysed the factor structure of RRBs domain, identifying three main factors or dimensions: repetitive sensory and motor behaviours (RSMB), circumscribed interests and insistence on sameness (IS). Furthermore, this subdivision of RRBs in broad categories was already present in earlier literature (Turner, 1999), reflecting a correspondence between the development of both clinical and psychological perspectives on the phenomenon. Repetitions of movement (stereotyped movements, repetitive forms of manipulation, self-injurious behaviour) were referred by Turner as *lower-level* RRBs, while *higher-level* RRBs were more complex and characterized by maintenance of sameness, object attachments, repetitive language, and circumscribed interests.

Subtype categories for RRBs are still useful for description, but they obviously have overlapping boundaries, forming a dimension that runs from the simplest repetitive body movements to highly sophisticated obsessive interests (Leekam, Prior, & Uljarevic, 2011). Lower-level RRBs could be more prominent in younger and more developmentally delayed subjects (while higher-level more often displayed by older and more able cases), but still lower-level RRBs could be seen in high-functioning adult groups (South, Ozonoff, & McMahon, 2005). At the more elaborated end of the dimension, some repetitive behaviours shade into similar features seen in Obsessive Compulsive Disorder (OCD) (Zandt, Prior, & Kyrios, 2007).

Development of RRBs and moderators (age, gender, functioning)

RRBs occur over the entire and normal course of human and nonhuman development, so their presence is not specific to pathological conditions (Eilam, Zor, Szechtman, & Hermesh, 2006) and typically developing children express a wide range of RRBs from infancy through school age. In toddlers and infants, motor stereotypies are quite common and develop in progressive fashion (Wolff, Boyd, & Elison, 2016) gradually giving way to more variable and goal-directed behaviour. Also, rituals and favourite objects are common during development in typical young children (Arnott et al., 2010).

RRBs are thus not specific to ASD, being present in typical individuals and also in a variety of other neurodevelopmental (e.g., Fragile X syndrome, Rett's syndrome), psychiatric (e.g., Obsessive–Compulsive Disorder, Impulse Control Disorders, schizophrenia), neurological (e.g., Tourette syndrome, Parkinson's Disease, dementia) and sensory disorders (blindness, deafness) (M. Lewis & Kim, 2009). Significantly elevated RRBs have been described relative to typically developing children as early as 10 months of age (Werner et al., 2005), but the majority of findings agree that differences begin to emerge more prominently in the second year of life between the RRBs of typical infants and infants with ASD (Elison et al., 2014), and between infants with other developmental delays and those with ASD (Watt, Wetherby, Barber, & Morgan, 2008). Even if children who develop ASD show more repetitive behaviour than their typically developing peers in the first 2 years of life, still the differential diagnosis between RRBs related to ASD or to other developmental delays can be

challenging before age 2 or 3 (Stone et al., 1999). Nevertheless, the occurrence of repetitive behaviours in ASD appears to be characterized by an elevated pattern of occurrence, co-occurrence, and severity relative to other neurodevelopmental disorders (J. W. Bodfish, Symons, Parker, & Lewis, 2000).

The developmental timing of the transition from normative to pathological repetitive behaviour is still not clear, together with biological and environmental mechanisms that mediate this transition and the persistence of repetitive behaviour (M. Lewis & Kim, 2009). Once thought to emerge after core social symptoms, recent findings raise the possibility that RRBs may be among the earliest behavioural manifestations of ASD (Kim & Lord, 2010; Sally Ozonoff et al., 2008). Also, literature findings tend to agree on the notion that discrete types of RRBs may have a unique age-related pattern. Stereotyped movements and restricted interests appear to be more frequent among younger individuals with ASD. Evidence about ritualistic/sameness behaviour is not consistent among studies, appearing more frequent among older individuals (Lam & Aman, 2007) or less frequent and less severe in the same age group, in line with a generalized abatement of RRBs in older subjects (Esbensen, Seltzer, Lam, & Bodfish, 2009).

Regarding gender differences, RRBs in girls and boys with ASD can be distinguished on different phenotypic dimensions (Mandy et al., 2012) in line with increasing evidence of a gender-based difference in presentation of ASD, with males showing more

frequent externalizing behaviour problems such as aggression and hyperactivity (Mayes et al., 2012), and females reporting greater internalizing difficulties, but more frequent nonverbal communication and prosocial behaviour (Rynkiewicz et al., 2016). This evidence, summarised in the hypothesis of the "female camouflage effect" (Ratto et al., 2018), may mask RRBs in females with ASD and result in underdiagnosis. Gender differences include a relative lower number of stereotypies in girls with ASD from early childhood to adulthood (May, Cornish, & Rinehart, 2016; J. Rojahn, Barnard-Brak, Medeiros, & Schroeder, 2016) and lower rates of restricted interests for girls in pre-school and school-age samples (Fulceri et al., 2016; Hiller, Young, & Weber, 2014; Schroeder et al., 2014). Conversely, self-injurious behaviour seems more pronounced in girls and women with ASD (Beggiato et al., 2017; Cohen et al., 2010). A recent study on 615 individuals (Antezana et al., 2019) with ASD confirms this hypothesized pattern, showing heightened stereotyped behaviours and restricted interests in boys, while compulsive, sameness, and self-injurious behaviour were more prominent in girls.

Many theories on RRBs share the common element that individuals who exhibit more RRBs are at risk of greater functional impairment (Troyb et al., 2016). However, RRBs have been relatively understudied as prognostic indicators, while early cognitive ability and language functioning have often been favoured predictors of subsequent outcomes in ASD (Charman et al., 2005). Lower-level RRBs (e.g. stereotypies) are often increased in subjects with intellectual disability (ID), while subjects with higher functioning often have more challenges with sameness (Somer L. Bishop, Richler, &

Lord, 2006; Lam et al., 2008; Richler, Huerta, Bishop, & Lord, 2010). Thus, cognitive skills can influence RRBs in ASD, moderating age effects (Hus, Pickles, Cook, Risi, & Lord, 2007) and this connection is intertwined with communication skills, showing how lower functioning children often show the lowest level of pragmatic language skills and the most severe and frequent RRBs (Barrett, Prior, & Manjiviona, 2004).

Hypothesized causes and functions of RRBs

Why do children with ASD keep engaging in RRBs? Although repetitive behaviour is commonly defined as behaviour that seemingly have no purpose (Hutt & Hutt, 1970; Wigham, Rodgers, South, McConachie, & Freeston, 2015), paradoxically much of the literature on RRBs in autism has sought to explore what functions these behaviours serve. The purpose of RRBs in children with ASD is not always clear, nonetheless many children appear to have a strong drive to perform them (Lawrence Scahill et al., 2014).

"Consider a 2-year-old child who is repetitively lining up his crayons in a precise way. When his brother suggests they play with blocks, he does not look over to see what is going on because the crayons are holding his attention. Thus, he misses the opportunity to engage with his brother and play with something different. When it is time to clean up and move to mealtime, he becomes upset because he does not have his line of crayons perfect yet. So, he misses the family meal. Finally, because he is so focused on colour arrangement of his crayons, he resists when his mother tries to teach him to colour with them. These three examples of missed opportunities highlight some of the negative consequences of RRBs on child learning and development. When asked, his parents might describe this behaviour differently, however. They

may see this behaviour as a strength, as their son knows all the colours in a 128-pack of crayons and can count to 128 at a really early age." (Turner-Brown, 2020)

Repetitive and rhythmical behaviour have a systemic effect in the first phases of typical development, playing a role in neuromuscular development, but also communication, language and social interaction (Iverson & Wozniak, 2007). At the end of the first year of life, RRBs become more varied and motor behaviour more goal-directed in parallel with adaptive functioning. Regarding the connection between sensory and motor features, it has been observed that repetitive motor behaviours often have a strong sensory component, with some theories implying a sensory drive of behaviours (Lovaas, Newsom, & Hickman, 1987). There is evidence that sensory and repetitive features in ASD are linked (Boyd, McBee, Holtzclaw, Baranek, & Bodfish, 2009; Chen, Rodgers, & McConachie, 2009) and some types of sensory features could be differentially associated with different repetitive behaviours (Leekam et al., 2011).

Also, repetitive sensory and motor behaviours are a frequent outcome in experiential deprivation or restriction (M. Lewis & Kim, 2009). Thus, even if motor and sensory feedback is not necessary to engage repetition in subjects with ASD (i.e. repetitive actions such as ordering objects), it seems that the role of environment is critical in the formation of RRBs: changes in early experience-dependent development, such as social restriction, could mediate gene expression in the basal ganglia neural circuitry (Marieke Langen et al., 2009). Mouse studies (M. H. Lewis, Tanimura, Lee, & Bodfish, 2007) show selective effects on RRBs and regional brain differences in the motor

cortex and basal ganglia, when mice were confronted with enriched and more complex environmental experiences. It can be hypothesized that, in ASD subjects, early difficulties in the social, communicative, and adaptive domains (arising from extreme social withdrawal) could interfere with brain maturation and experience-dependent development. Nevertheless, so far there is no reliable evidence from post-mortem studies for specific basal ganglia alterations in ASD, while some MRI studies have showed caudate volume associations with stereotypies (M. Langen, Kas, Staal, van Engeland, & Durston, 2011). An interesting hypothesis proposes the presence of three cortico-striatal feedback loops:

1) Sensory-motor loops: linked to motor and premotor cortex, related to lower order RRBs
2) Cognitive-associative loops connected to dorso-lateral prefrontal cortex, implicating rigidity and obsessive routine
3) Limbic-motivation loops connecting to anterior cingulate and orbital-frontal cortex, linked to compulsive and addictive behaviour (M. Langen et al., 2011)

While anxiety (Kamp-Becker, Ghahreman, Smidt, & Remschmidt, 2009) and deficit in executive function (Yerys et al., 2009) have also been indicated as linked to RRBs, evidence indicating arousal as a key trigger, together with unstructured environment, is converging from neurobiological, cognitive and developmental psychology perspectives . RRBs in ASD could function, in the same way as in typically developing

subjects, as self-regulating coping strategies that help to regulate hyperarousal or to increase sensory stimulation in hypo-arousal (Jiujias, Kelley, & Hall, 2017).

Focusing on functional analysis, repetitive behaviours have also been seen as constructive elements of early development which could generate behavioural precision, efficiency, and adaptive outcomes (Ejiri & Masataka, 2001). The functional properties of RRBs were often derived in literature from functional analysis methodologies (J. T. Rapp & T. R. Vollmer, 2005), aimed at identifying the variation of specific behaviours through systematic manipulation of environmental antecedents and consequences (Iwata, Dorsey, Slifer, Bauman, & Richman, 1994). The most commonly environmental contingencies in functional analyses of RRBs include:

- social positive reinforcement (e.g., attention)
- social negative reinforcement (e.g., avoidance)
- non-social positive reinforcement (e.g., self-stimulation, automatic reinforcement)
- non-social negative reinforcement (e.g., escape from an aversive physical stimulus)
- a combination of reinforcements

The current body of literature underlines a substantial heterogeneity of the reinforcement contingencies under which repetitive behaviours may be controlled. RRBs thus comprise a group of behaviours that do not belong to a predetermined response class. Instead, they vary widely not only in form, but in environmental

determination —across individuals, context, setting, and time (Cunningham & Schreibman, 2008). The most cited maintaining reinforcement contingency is self-stimulation or automatic reinforcement (J. T. Rapp & T. R. Vollmer, 2005; Rogers & Ozonoff, 2005), but a growing body of literature specifies that RRBs may be determined by other reinforcement contingencies, for example social and tangible reinforcement contingencies (Ahearn, Clark, Gardenier, Chung, & Dube, 2003).

RRBs as treatment outcomes

Despite being considered a core feature of ASD, RRBs have received less attention, if compared to the domain of social interaction and communication (Richler et al., 2010), and less frequently targeted by interventions. This is surprising, given their role as major management challenges, obstacles to adaptive functioning and cause of distress for subjects and their families (Leekam et al., 2011). Potential factors associated with this divergence could include:

1) Heterogeneity of repetitive behaviours as a construct, ad discussed above. There are a variety of forms which may differ in clinical significance, function, and underlying mechanism

2) Potential phenotypic overlap between repetitive behaviours and other features, such as sensory-motor features or comorbid psychiatric symptoms (e.g., anxiety, OCD and hyperactivity)

3) Moderation exerted by a wide host of factors such as age, cognitive ability, and family factors which can influence the expression of RRBs

4) The assumed “primacy” of social deficits defining features;

(Boyd, McDonough, & Bodfish, 2012).

RRBs are reliable predictors of a stable diagnosis of ASD between age 2 and 9 (Lord et al., 2006), and ‘gold standard’ tools for diagnosis of ASD (ADI-R and ADOS) include valid and specific items to assess this core domain (S. Ozonoff, Goodlin-Jones, & Solomon, 2005). Conversely, when targeting RRBs as treatment outcomes, there is less consensus. Several instruments have been developed for these symptoms in the ASD population, but the rigor of psychometric assessment is variable and few measures have a solid track record as an outcome measure in clinical trials (Lawrence Scahill et al., 2015). Moreover, there is little agreement on terminology (e.g. stereotypic, self-stimulatory, ritualistic, perseverative, gesturing, posturing), while other terms such as abnormal preoccupations, circumscribed interest patterns, abnormal object attachments, and idiosyncratic responses to sensory stimuli often lack specific behavioural referents. (J. W. Bodfish et al., 2000).

In 2011, The Autism Speaks Foundation assembled a panel of experts to address this issue and identify which available tools were valid and reliable to assess RRBs as outcomes in youth with ASD (Lawrence Scahill et al., 2015). Twenty-four instruments were evaluated, and five were considered appropriate with conditions for use in clinical trials:

1) *Children’s Yale-Brown Obsessive-Compulsive Scales for Pervasive Developmental Disorder (CY-BOCS-PDD*) (L. Scahill et al., 1997): modified version of the

original CY-BOCS designed to measure symptom severity of OCD in children with PDD. CY-BOCS is a semi-structured clinician interview, with separate symptom checklists for obsessions and compulsions. Checklists are reviewed by both parent and child to identify primary symptoms, then the clinician considers five severity dimensions (time, interference, distress, resistance, degree of control). The modified version excludes the obsessions checklist and adds several RRBs relevant to ASD.

2) *Repetitive Behavior Scale – Revised (RBS-R)* (J. Bodfish, Symons, & Lewis, 1999): it is a parent-rated scale which measures the severity of RRBs in children and adults with ASD. It has 43 items organized into six subscales (stereotyped, self-injurious, compulsive, ritualistic and restricted behaviours).

3) *Aberrant Behavior Checklist: Stereotypic behavior subscale (ABC-S)* (Aman, Singh, Stewart, & Field, 1985): consists in 7 items, among 58 in the complete tool, mostly describing motor stereotypy. It is designed to be completed by any adult in close contact with the subject.

4) *Stereotyped Behavior Scale (SBS)* (Johannes Rojahn, Matlock, & Tassé, 2000): it is a 24-item tool for measuring RRBs, rated by frequency and severity. It has good coverage of stereotypic motor behaviour, but it does not include more complex behaviours or circumscribed interests.

5) *Repetitive Behavior Questionnaire (RBQ)* (Honey, McConachie, Turner, & Rodgers, 2012): it is a parent-rated instrument, consisting in 26 items derived from the Repetitive Behaviour Interview. It includes four domains (repetitive motor behaviour, insistence on sameness, repetitive speech and circumscribed interests).

It must be noted that CY-BOCS, RBS and ABC-S scales are developed by the same authors involved in the panel.

Evidence for interventions aimed at RRBs

RRBs are one of the two core symptom domains of ASD and, to this date, there is no medication approved to treat them (Pandina, Ring, Bangerter, & Ness, 2020). A recently published meta-analysis (Yu et al., 2020) of pharmacotherapy of RRBs summarised the results from 14 randomised controlled trials (RCTs) published between 1994 and 2018. Nine studies were included in the final meta-analysis, addressing changes in RRBs or compulsive behaviour and stereotypies measured using 3 scales: Repetitive Behaviour Scale Revised (RBS); Aberrant Behaviour Checklist (ABC), Yale Brown Obsessive Compulsive Scale (Y-BOCs) and its child version (CY-BOCS). The results underlined that even if five medications (risperidone, fluvoxamine, fluoxetine, buspirone, divalproex sodium) showed an effect in individual studies, the meta-analysis did not show a significant difference with placebo. Also, the results showed further evidence that SSRIs have little effect on treatment of RRBs in children and adult subjects with ASD. Another meta-analysis in press (Zhou et al., 2020) summarises the results from 64 pharmacological RCTs published until November 2019. The study reports a small statistically significant improvement in RRB outcomes after treatment with antipsychotics, with greatest benefit seen using risperidone and aripiprazole. The authors underline that, given the small effect size of the benefit, clinicians need to balance it with the considerable side-effects of these type of medications. Other

classes of medications (antidepressants, oxytocin, methylphenidate, naltrexone, atomoxetine, secretin, divalproex bumetanide) and active ingredients analysed (omega-3 fatty acids, N-acetylcysteine, folinic acid) did not prove a substantial clinical benefit, even if larger significant positive effects were seen in preliminary individual studies. The results of these two recent meta-analyses are in line with previous research, showing mixed findings for risperidone, oxytocin (Soorya, Kiarashi, & Hollander, 2008) and SSRI (Wink, Erickson, & McDougle, 2010), corroborating the evidence that interventions directed at core symptoms of ASD have not been encouraging so far.

Regarding the effectiveness of non-pharmacological interventions on RRBs, (Horner, Carr, Strain, Todd, & Reed, 2002) and other subsequent reviews (Patterson, Smith, & Jelen, 2010; John T. Rapp & Timothy R. Vollmer, 2005) showed that skill-based behavioural interventions (especially environmental enrichment and differential/alternative reinforcement), were effective in reducing RRBs. However, even if RRBs are frequently targeted by behavioural interventions, they are rarely included as outcome measures when evaluating the effectiveness of comprehensive interventions (Howlin, Magiati, & Charman, 2009). A systematic review has been published more recently (Zarafshan, Salmanian, Aghamohammadi, Mohammadi, & Mostafavi, 2017), but its scope is limited to pre-school children with ASD. Also, the search was run in April 2014 and no meta-analysis was performed. The review includes 15 studies, published from 1987 to 2013. Interventions were evaluated as completely

heterogeneous, but results showed positive effects from almost all the techniques. Interventions focused on positive reinforcement (tangible or social) and improvement in communication skills and sensory problems were the most effective.

Even if a substantial amount of other systematic reviews and meta-analysis summarizes treatments directed towards subjects with ASD (Barahona-Corrêa, Velosa, Chainho, Lopes, & Oliveira-Maia, 2018; Gates, Kang, & Lerner, 2017; Healy, Nacario, Braithwaite, & Hopper, 2018; Su Maw & Haga, 2018; Wolstencroft et al., 2018), they are not focused specifically on RRBs. Thus, the goal of this study was to conduct a systematic, exhaustive, and up-to-date systematic review and meta-analysis of controlled interventions aimed specifically at RRBs. To avoid methodological limitations found in other reviews, no limitations to the age of the sample and to the type of intervention were set.

Methods

Search strategy and inclusion criteria.

A systematic search was performed to identify original studies assessing the clinical efficacy of randomized controlled intervention trials in people with ASD, focusing on RRBs. Two investigators conducted a two-step literature search. As a first step, Web of Knowledge SM database by Thomson Reuters® (including Web of Science, MEDLINE®, KCI – Korean Journal Database, Russian Science Citation Index and SciELO Citation Index) was searched, including all databases and using the "topic" function of the database, which allowed the title, abstract, keywords, and other fields to be analysed at the same time. The hyper-inclusive following search string was used:

((autis* OR (developm* AND disorder) OR asperger OR Kanner OR ASD OR PDD)) AND ((RCT OR trial OR observational OR open label OR prospective OR longitudinal OR randomized OR cohort)).

The search was run in January 2020 and extended inception to December 2019, limited to English language and human studies only. Search records were imported into EndNote X9.3 and duplicates were removed. Titles and abstracts obtained were examined by two reviewers, while a third reviewer was consulted to resolve discrepancies.

Selection procedure

Studies meeting the following criteria were included:

a) original articles, published in a peer reviewed scientific journal, written in English;

b) studies including subjects with a PDD or ASD diagnosis, including previous classification (ex. infantile autism);

c) study design: randomized controlled trials (RCT) comparing at least two different interventions directed to ASD subjects or one treatment and placebo (or equivalent, ex. sham stimulation for transcranial magnetic stimulation, waiting list of treatment as usual (TAU) for psychotherapy);

d) studies reporting at least one behavioural outcome.

Studies meeting the following criteria were excluded:

a) review, meta-analysis, case report, congress abstracts, and articles in languages other than English;

b) studies with retrospective observational design or longitudinal observational design without a comparison group, open-label studies, non-randomised trials;

c) studies failing to report a behavioural measure (ex. biomarkers and imaging were not considered clinical outcome measures)

d) studies with less than 10 subjects in the intervention arm

Studies meeting inclusion criteria were then assessed further for eligibility by full text reading by two authors, and any doubt was resolved by a third author. The second step involved the implementation of a manual search of the reference lists of the retrieved articles and relevant reviews on the topic to identify additional studies not identified in the electronic search. To achieve a high standard of reporting, we followed the PRISMA guidelines and compiled the PRISMA checklist (Liberati et al., 2009).

Data extraction. Study selection and data extraction were performed independently by two investigators. Cases of inconsistency and disagreement were double-checked and discussed with a third author. The following variables were recorded on a standardized Microsoft Excel spreadsheet:

- Author
- Year of publication
- Study location
- Study design (randomized versus placebo, randomized versus active treatment, cross-over randomized trial)
- Length of follow-up
- Type of intervention

- Pharmaceutical: if the treatment involved the use of over the counter or prescription drugs (e.g. antipsychotics, antidepressants, stimulants…)
- Nutraceutical: if the intervention involved the use of dietary supplements or functional foods (e.g. omega-3, vitamins, gluten-free diet…)
- Psychotherapy: if the intervention used a manualized or non-manualized form of psychotherapy treatment technique (e.g. cognitive behavioural therapy, psychodynamic treatments, acceptance and commitment therapy…)
- Educational: if the treatment used a psychoeducational program, or any program that helped delivering information about the diagnosis (including possible causes and symptoms), management (including associated risks/side-effects) and prognosis, and how affected individuals can stay well (Jones, 2018)

- Sample size and number of patients assigned to each arm (intervention versus control)
- Sample age and gender (female proportion)
- Psychiatric comorbidities other than intellectual disability (excluded, acknowledged and unclear)
- Diagnostic tool

- Intelligence quotient (IQ) assessment, included the range of assessed IQ (inclusion range if the former data was not available), mean IQ of the sample, and IQ assessment tool.
- aim/target of the study (core symptoms, problem behaviours, adaptive functioning, cognitive functioning, and medical and psychiatric comorbidities). Aim was coded 'unclear' when the article did not clearly report a specific target symptom or domain for the intervention, or when the aims where too numerous or vaguely defined.

Appraisal of methodological quality. Assessment of study quality was performed independently by two investigators using the Cochrane Risk of Bias Tool (Higgins et al., 2011). Discrepancies were solved after consultation with a third reviewer. We considered the following domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other sources of bias. According to Cochrane's tool, each domain could be judged to be at "low risk", 'unclear risk' or 'high risk of bias" (Higgins et al., 2011).

Selection of studies for the systematic review and meta-analysis: after data extraction, we included in our systematic review the studies reporting a clear aim directed at reducing core symptoms of autism or RRBs in specific. We excluded trials where RRBs were measured, but not reported as a specific aim. The decision was taken to increase

the overall quality of RRB assessment and to reduce the possibility of selective reporting of outcomes. In fact, apart from the measures recommended by experts in the field and described in the introduction, changes in RRBs are often measured with multi-domain assessments, which comprise a small number of items dedicated to RRBs. Descriptive statistics and graphs were obtained with IBM SPSS 26.0 and Microsoft Excel.

All the studies providing specific RRB scores at baseline and endpoint, or as change from baseline, were subsequently included in the meta-analysis. Change in scores between baseline and the longest follow-up assessment were considered as measure of the primary outcome. Valid outcome measures were extracted as raw scores (means and SDs) or raw differences. In case raw measures were not available, we extracted computed differences, test statistics, effect sizes and significance levels. For any outcome measure extracted, we specified the time endpoint and the daily dosage (for pharmacological/nutraceutical interventions) or the weekly intensity of treatment (for psychoeducational and psychological interventions).

In case of multiple follow up points, since ASD core symptoms are generally considered longstanding features, we selected the longest follow up assessment to obtain an estimate of long-term effect. Studies with multiple dosage/intensity of treatment were only five (Bolognani et al. 2019, Chugani et al. 2016, Lemonnier et al. 2017, Mazahery et al. 2019, Wood et al. 2019): meta-regression of multiple dosages was of limited use

considering the low number of studies, so we assessed them case-by-case and we chose the same maximization approach (in this case selecting the maximum dosage/intensity authorized by FDA). When available, we gave priority to the outcome measures suggested by experts in the field, discussed in the Introduction (Lawrence Scahill et al., 2015). When more than one recommended outcome scale was used within a study, we computed the average effect size and standard error of the measures with the method suggested by (Borenstein, 2009).

Effect sizes were calculated as Hedges' g according to a random effects model, using DerSimonian Laird method (DerSimonian & Laird, 1986). Effect sizes were interpreted according to the common guidelines (i.e. 0.2, "small"; 0.5, "medium"; 0.8 "large"). Heterogeneity between the effect sizes was assessed using the I_2 statistic, with a value greater than 50% indicative of high heterogeneity. Subgroup analysis was performed according to the intervention type (pharmacological, psychotherapy and education, nutraceutical, and other types). Publication bias was assessed through visual inspection of the funnel plot, Egger's test analysis and cumulative analysis.

This systematic review and meta-analysis was conducted according to Assessing the Methodological Quality of Systematic Reviews (AMSTAR) guidelines (see Appendix for Amstar Checklist). Meta-Analysis was performed with Stata 16.0 (Statacorp.2019).

Results

Systematic review

The electronic search generated 125,195 items, from which 5,468 full-text publications were analysed for detailed screening. The manual search yielded additional 88 studies. In total, we identified 565 randomized controlled interventions in ASD. 80 studies (14.2%) reported a clear aim directed specifically at RRBs or at both ASD core symptom domains and were then included in the systematic review. The selection process is reported in the PRISMA flow diagram (Figure 1). See Appendix (Table 1a) for a complete list of the studies included in our systematic review.

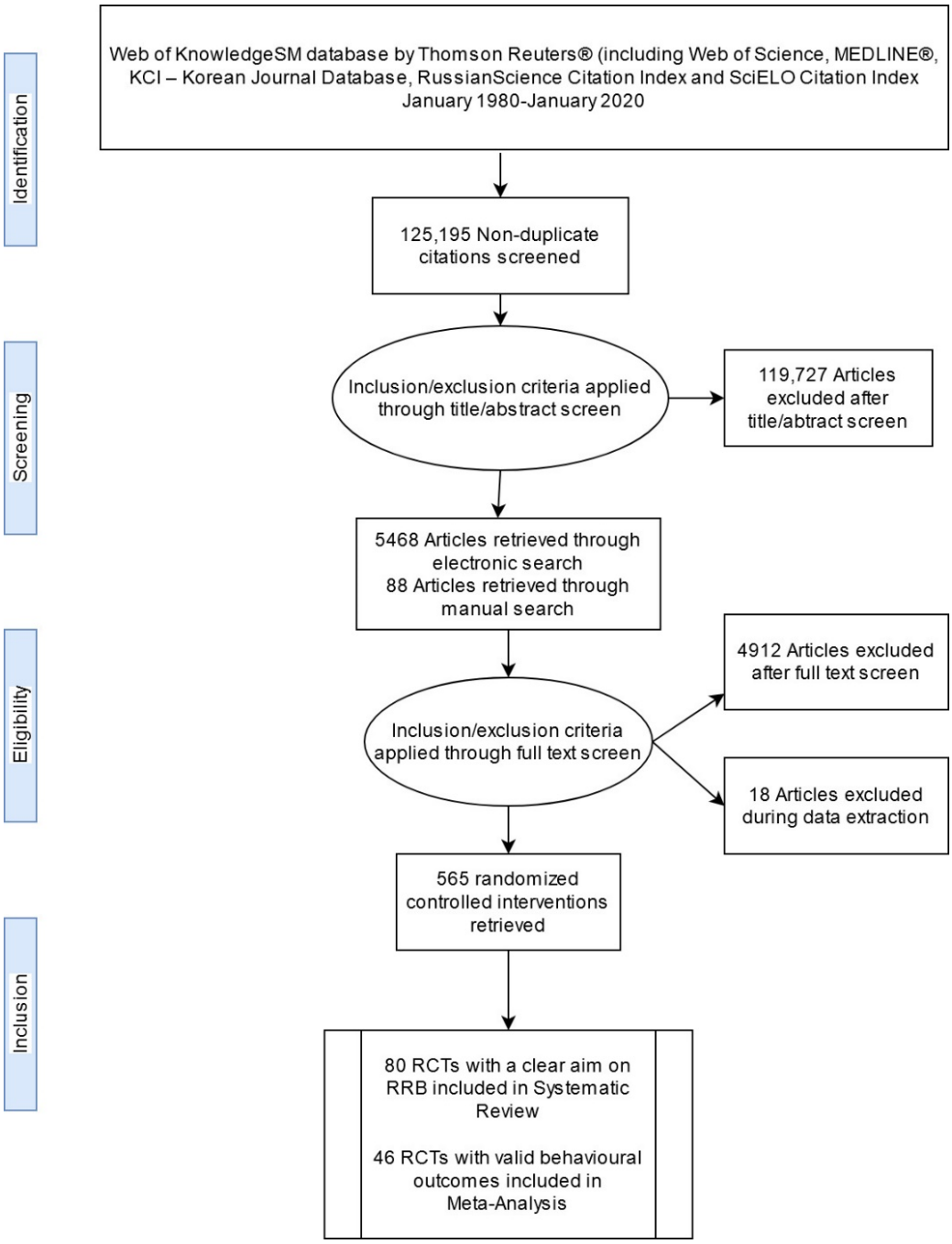


Figure 1. Prisma flow diagram of the selection procedure

Included studies were published between 1992 and 2019. No overlapping datasets were found. Our database included 3114 subjects assigned to the active treatment.

Sample sizes ranged from 10 to 319 individuals, with a mean of 39.92 (SD = 42.21) subjects assigned to treatment. Each study included on average 19% female participants (range 0%–50%, unclear in 3 studies). Maximum follow-up duration varied from a single administration (1 day) to 116 weeks, with a mean follow-up of 16 weeks (SD = 16.5).

The active intervention (Figure 2) was pharmacological in 36 (45%) studies, based on psychotherapy or psychoeducational techniques in 15 (18.8%) studies. Complementary (nutraceuticals, diets, etc) and other forms of interventions were adopted in 29 (36.3%) studies.

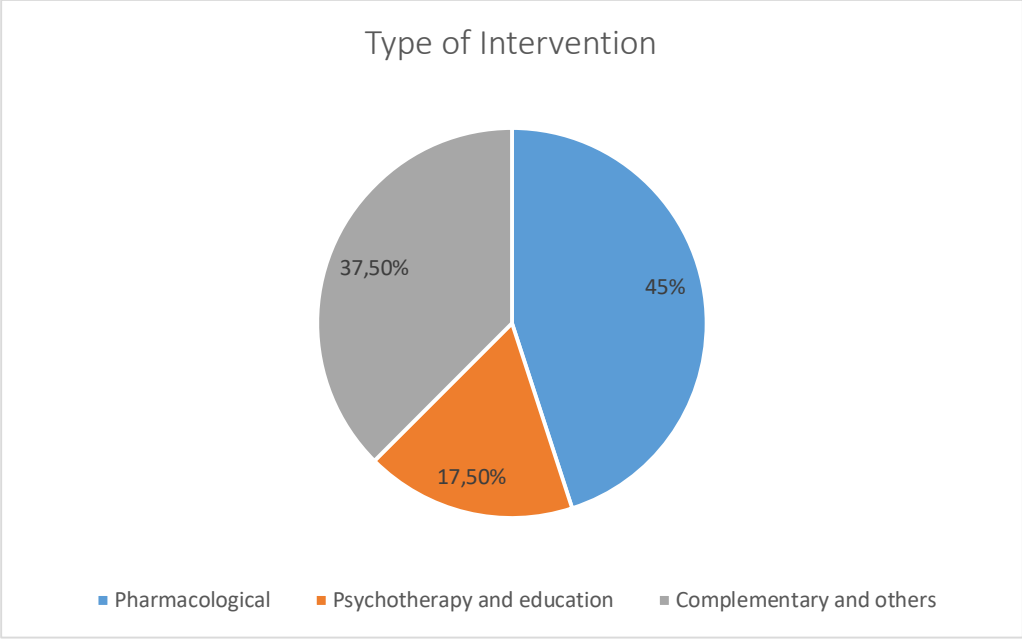


Figure 2. Pie chart of intervention types

In total, 66 (82.5%) studies recruited only children and 10 (12.5%) only adults, while 4 (4.9%) studies included both children and adults. Mean age of the sample was 10.52

years (SD = 7.52). Psychiatric comorbidities, excluding intellectual disability (ID), were excluded in 21 (26.3%) studies, reported in 10 (12.5%) studies and unclear in 50 (61.3%) studies.

Most interventions were conducted in the United States (38; 47.5% of the studies).

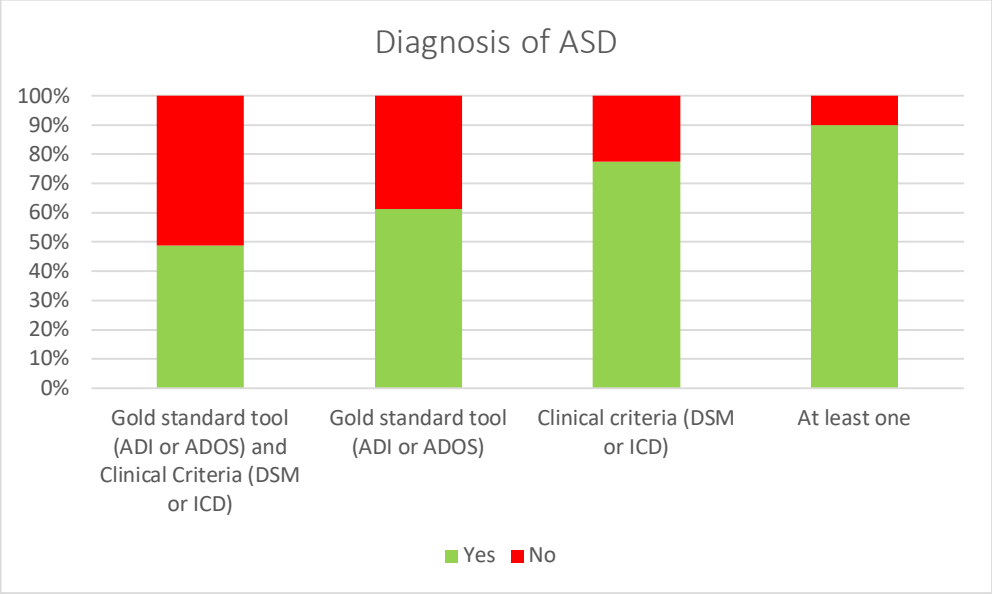


Figure 3. Use of diagnostic tools for ASD

Diagnostic tool of ASD (Figure 3) was clearly reported in 72 (90%) studies, while in 8 (10%) studies it was unclear. Specifically, 62 (77.5%) studies reported that the sample was diagnosed with clinical criteria (DSM or ICD), 49 (61.3%) studies included at least one gold-standard instrument for diagnosis (ADI or ADOS), while 39 (48.8%) adopted both clinical criteria and at least one gold-standard instrument for diagnosis.

IQ assessment of the included subjects was unclear in 38 (47.5%) studies, while in 42 (52.5%) studies it was specified. However, 7 studies only provided minimum IQ

required for recruiting, resulting in 35 (43.8%) studies providing a meaningful IQ assessment. One study (1,3%) recruited only subjects with ID, 22 (27.5%) recruited only individuals without ID and in 15 (18.8%) both ASD people with and without ID were recruited.

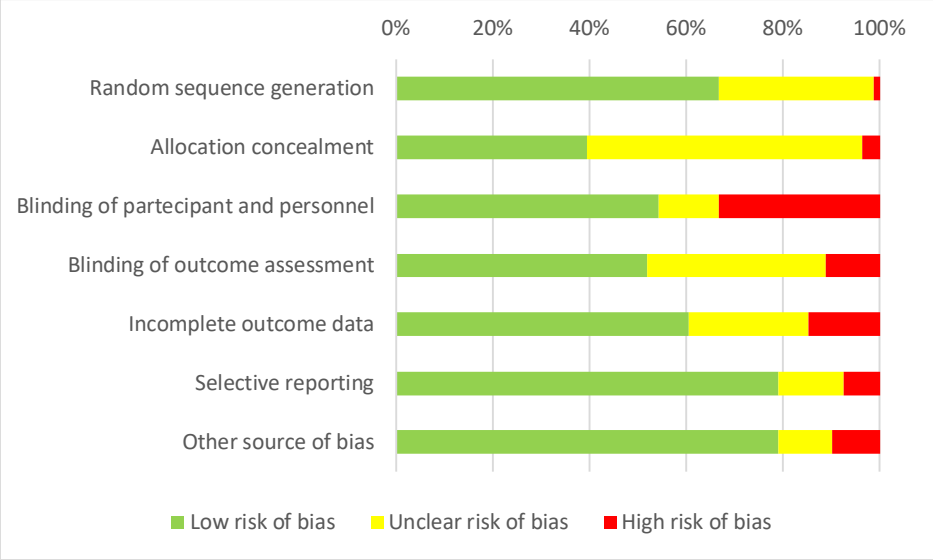


Figure 4. Risk of bias of the included RCT studies

The risk of bias relative to RCT studies (n=80) is presented in Figure 4. In total, 14 (17.5%) studies were rated as good quality, 37 (46.3%) scored as fair and 29 (36.2%) had poor quality of reporting according with the methodology reported in (Penson, Krishnaswami, Jules, & McPheeters, 2013). The complete list of RCT risk of bias evaluation can be found in the Appendix (Table 2a).

Meta-analysis

Effect of randomized controlled interventions on RRBs

We performed an exploratory meta-analysis including all the randomized controlled studies reporting valid outcome scores at baseline and at endpoint. Characteristics of the included studies are described in the Appendix (Table 1a). The general impact of interventions is synthesized in Figure 5.

All the outcome scales used to assess RRBs in this sample indicate clinical improvement with a reduction of scores. As consequence, negative effect sizes indicate clinical improvement. Also, since effect sizes in this sample are reflecting the difference of pre/post scores between intervention and placebo arms, negative values of the effect size favour intervention group over control group.

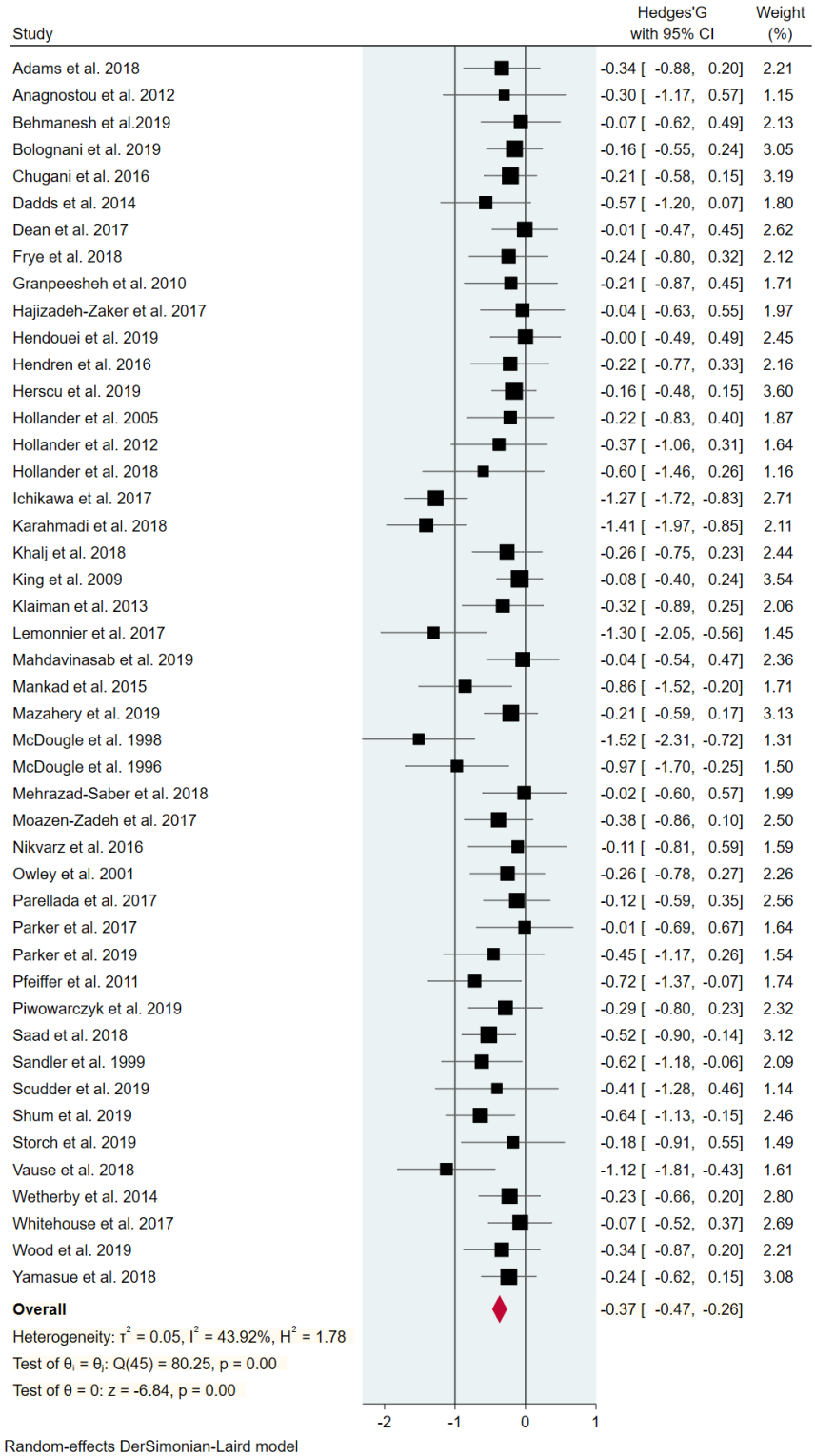


Figure 5. Forest plot: overall effect of interventions aimed at RRBs. Negative effect sizes indicate improvement and favour intervention over control.

Meta-analysis (46 studies, 1339 subjects) showed an overall small-to-medium effect of included interventions on RRB in subjects with ASD (g = 0.37, 95% CI -0,26 to -0,47). There was moderate, but significant heterogeneity in the complete sample of studies (I_2 = 43.92%, p < 0.01), which we expected in this overall analysis, since we compared very different interventions. We performed exploratory meta-regression with four separate moderators (mean age of the sample, length of follow up, sample size of the intervention group, female ratio) and building simple models (i.e., female ratio and mean age). We did not find any significant predictive model of heterogeneity (see Appendix, Figure 13a-d), so we proceeded with subgroup and sensitivity analysis for further evaluation.

Regarding instruments for the evaluation of RRBs, 17 (37%) studies used Aberrant Behaviour Checklist (ABC), 10 (22%) studies used Repetitive Behaviour Scale (RBS) and 11 (24%) studies employed CY-BOCS. Other outcome tools were Social Responsiveness Scale (SRS) in 9 (20%) studies, Autism Diagnostic Observation Schedule (ADOS) in 3 (7%) studies, Gilliam Autism Rating Scale (GARS) in 2 studies and PDD Behaviour Inventory (PDDBI) in 1 study. In summary, 23 (50%) studies used an outcome tool recommended in literature, 7 (15%) studies used more than one recommended scale and 16 (35%) studies used an outcome tool that was not recommended. Please note that studies which only employed the use of ad-hoc and not validated outcome assessments were excluded from our analyses.

Visual inspection of the forest plot revealed that 5 studies (Ichikawa et al. 2017, Karahmadi et al 2018, Lemonnier et al. 2017, McDougle et al. 1998, Vause et al. 2018) appeared as possible outliers. We performed cumulative analysis ordered for ascending effect size magnitude, showing that the exclusion of outliers did not bring a substantial change to overall effect size and its significance (g = -0.26, 95% CI -0.18 to -0.33, see Figure 11a and 11b in Appendix). However, sensitivity analysis revealed that with the exclusion of these five outliers, heterogeneity (I_2) was reduced to zero. Since four of these five studies were pharmacological interventions, we first performed a subgroup analysis based on intervention type, to carefully inspect the issue.

Subgroup analyses

Type of intervention

Subgroup analysis based on type of intervention is depicted in Figure 6. Subgroups were defined as studies including pharmacological interventions, psychotherapy/psychoeducation interventions and complementary (diets, oral supplements, others) interventions. Meta-analysis displayed a small, but beneficial effect in the three subgroups, with no significant difference based on intervention type (p = 0.16).

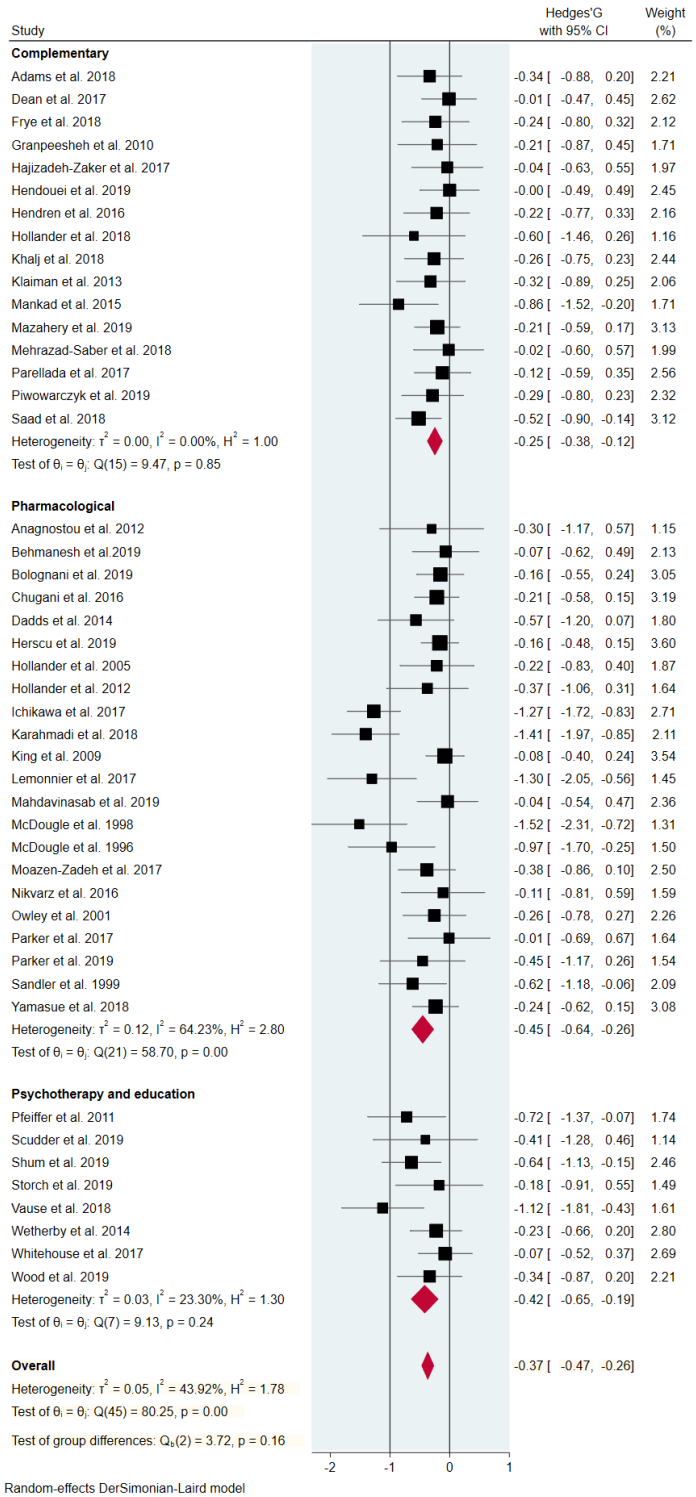


Figure 6. Forest plot: subgroup analysis by intervention type. Negative effect sizes indicate improvement and favour intervention over control.

Evidence of a small, but beneficial effect was displayed for complementary interventions (g = 0.25, CI -0.12 to -0.38), while a small to moderate beneficial effect was found for interventions based on medications (g = 0.45, CI -0.26 to -0.64), psychotherapy and/or psychoeducation (g = 0.42, CI -0.19 to -0.65). There was low heterogeneity for complementary interventions (I_2 = 0%, p = 0.85) and psychotherapy/psychoeducation interventions ((I_2 = 23.30%, p = 0.24). Heterogeneity was high and significant in the pharmacological subgroup (I_2 = 64.23%, p = < 0.01).

With the exclusion of the five abovementioned outliers, since four studies were pharmacological trials and one was a psychotherapy trial, we obtained changes in the overall effect size of the pharmacological subgroup, comparable to those evidenced in the complete sample. In the pharmacological subgroup, effect size was reduced, but still significant (g = -0.23, CI -0.12 to -0.35) and heterogeneity was notably reduced (I_2 = 0%, p = 0.89). Changes in the psychotherapy/psychoeducation subgroup were less marked (g = -0.34, CI -0.13 to -0.54, I_2 = 0%) (see Appendix, Figure 12a).

Focused analyses

We performed analyses of subset of studies investigating the effect of specific treatments, if five or more studies were available for each subset. Evidence of a small, but beneficial and significant effect was displayed for the use of SSRIs towards RRBs (g = 0.24, CI -0.01 to -0.47). Heterogeneity was low (I_2 = 22.33%) (

Figure 7)

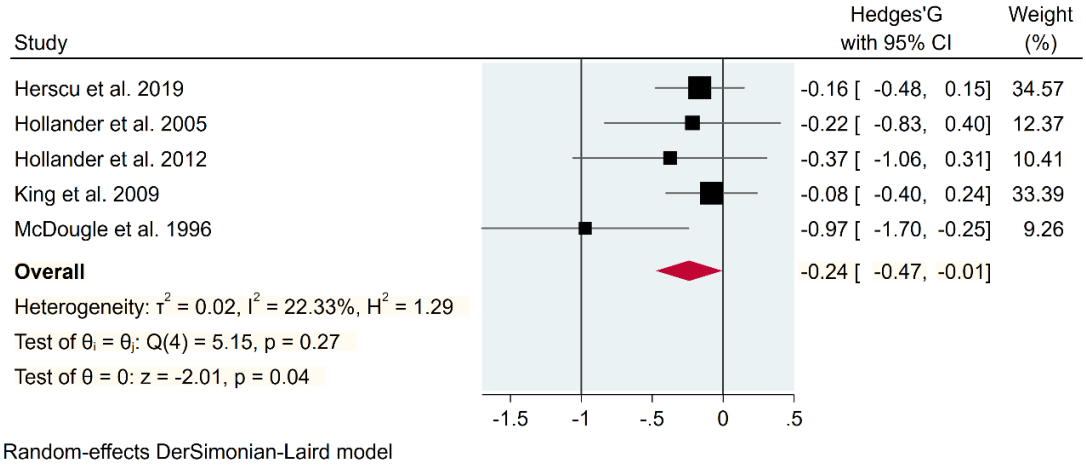


Figure 7. Forest plot: effect of SSRIs. Negative effect sizes indicate improvement and favour intervention over control

Meta-analysis of the overall effect of pituitary neuropeptides (oxytocin and vasopressin) also revealed a modest, beneficial effect on RRBs (g = 0.29 CI -0.03 to -0.56). Heterogeneity was low (I_2 = 0%) (

Figure 8)

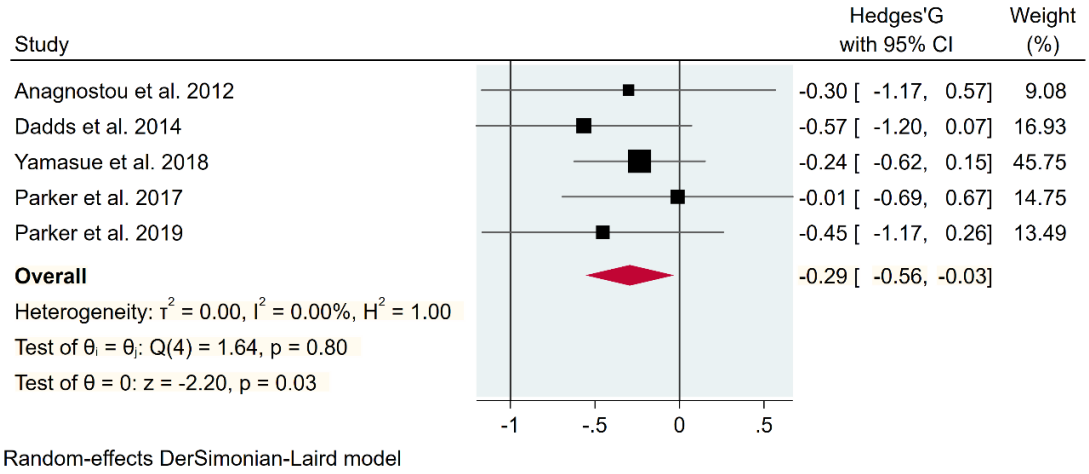


Figure 8. Forest plot: effect of pituitary neuropeptides (oxytocin and vasopressin). Negative effect sizes indicate improvement and favour intervention over control

Overall summary effect of nutraceutical supplements (vitamins, omega-3, folinic acid, acetylcysteine, etc) was also modest, but significant (g = -0.24, CI -0.10 to CI -0.38).

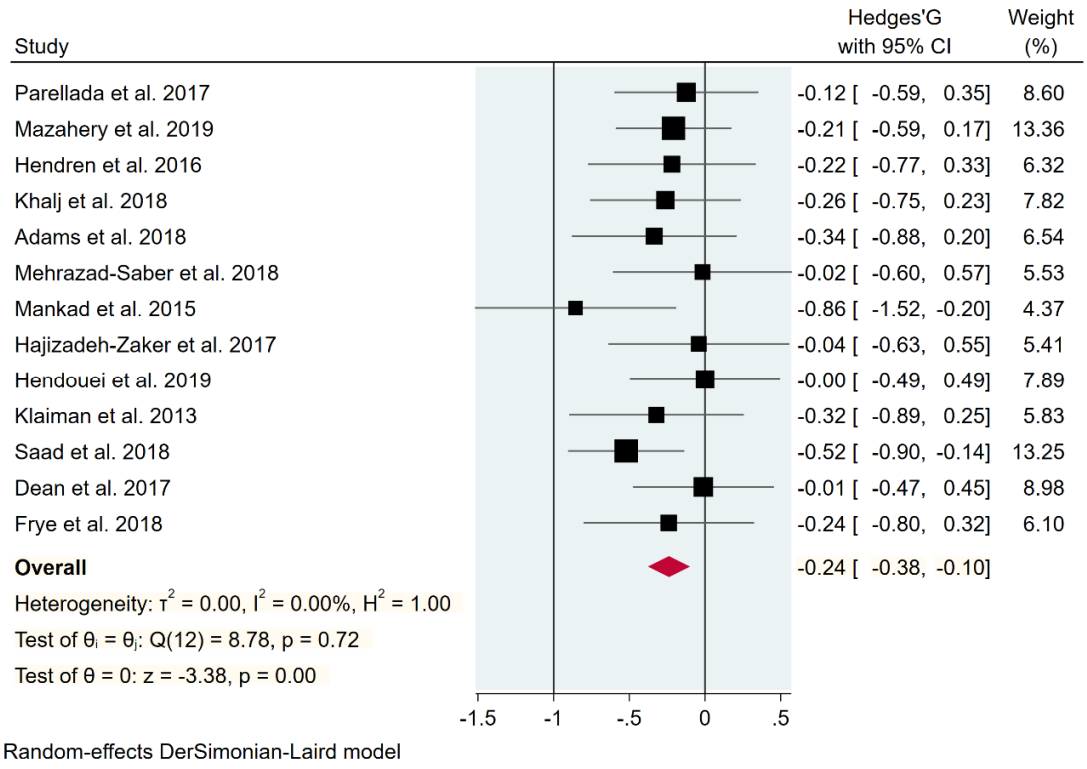
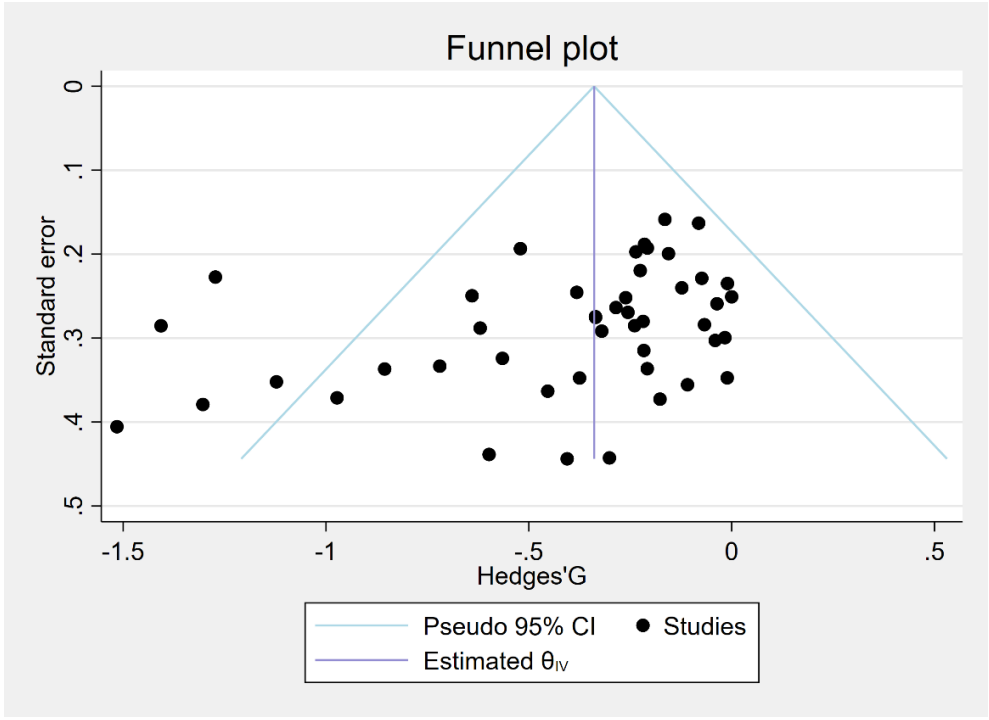


Figure 9. Forest plot: effect of nutraceutical supplements. Negative effect sizes indicate improvement and favour intervention over control

Heterogeneity was low (I_2 = 0%) (Figure 9). As visible from the forest plots, this result is very similar to the overall complementary subgroup effect size, which included only other three studies not assessing oral supplements (Granpeesheh et al. 2010, hyperbaric treatment; Hollander et al. 2018, trichuris suis ova; Piwowarczyk et al. 2019, gluten free diet).

Publication bias



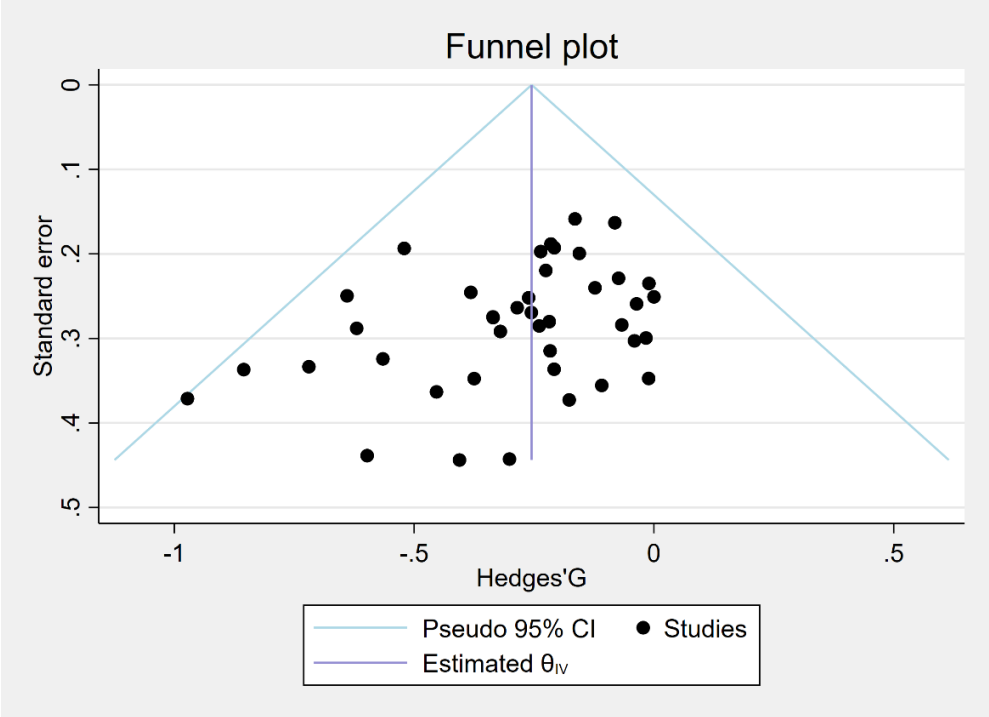


Figure 10. Funnel plot of the included studies (including and excluding outliers)

Visual inspection of the funnel plot revealed an evident asymmetry of included studies, consistent with forest plots inspection. A second funnel plot, produced after the exclusion of the outliers, still showed a slight asymmetry and a wider dispersion of studies favouring the intervention. Egger's regression test for funnel plot asymmetry performed on the whole sample revealed a significant asymmetry (β_1 = -1.86, z = -2.50, p = 0.0124). The same test was non-significant after exclusion of outliers (β_1 = -1.07, z = -1.74, p = 0.08) (see Appendix, Figure 14a and 14b).

To assess the impact of small studies on the overall effect size, we also performed cumulative analysis ordered for ascending sample sizes (number of patients in the intervention arm of each study). We repeated the analyses with and without the five outliers (see Appendix, Figure 15a and 15b). Cumulative forest plots did not display a remarkable change of direction or magnitude caused by smaller studies, even if we must underline that the range of sample sizes was limited (10 to 78).

Discussion

Key results:

- Among 525 RCTs targeting ASD, only 14% clearly defined both core domains of autism (or RRBs specifically) as a specific target.
- IQ assessment was unclear in nearly 50% of the sample and other psychiatric comorbidities (excluding ID) were not clearly disclosed in 61% of the studies
- Subgroup meta-analysis of changes in RRBs based on the type of intervention (pharmacological, psychotherapy/psychoeducation, complementary) did not reveal a significant difference between the three subgroups.
- Effect sizes of interventions directed at RRBs were, on the whole, small to moderate.
- The absence of studies reporting a worsening of RRBs could be related to publication bias, but also to the relatively short mean duration of follow up (4 months), which could not detect relevant changes in the natural course of these symptoms (thus underlining their stability over time).
- Risk of bias, measured with the Cochrane Risk of Bias tool, was generally medium to high.

Restricted and repetitive behaviours (RRBs) are a core symptom domain of autism spectrum disorder (ASD) and they constitute a critical adaptive challenge for subjects in the spectrum, with a strong impact on learning, social adaptation, family stress and quality of life (S. L. Bishop, Richler, Cain, & Lord, 2007). This review is the first attempt in literature to synthesize evidence from randomized controlled trials (RCTs) within this field of research, performing a systematic review and meta-analysis which comprehensively includes any type of intervention aimed specifically at RRBs, without limitations in timespan and age of the sample.

Systematic Review

We included 80 studies (3114 subjects) in the systematic review. Thus, among 525 RCTs targeting ASD analysed during our full-text screen, only 14% of the studies clearly defined RRBs or both core domains as their primary or secondary target. This is in line with previous considerations: despite being acknowledged as core feature of ASD, RRBs are still receiving less attention, if compared to the domain of social interaction and communication (Richler et al., 2010).

With the scope of giving a context to the results of our meta-analysis, it is important to underline key characteristics and limitations of the included studies:

- Sample sizes were generally small, with a mean of 40 subjects in the intervention arm. Thus, even if meta-analysis increases statistical power pooling results from otherwise low-powered studies, it is very important to interpret our results with prudence.

- Mean follow-up duration of the assessments was about 4 months: given the lifelong duration and acknowledged stability of ASD core symptoms over time (Matson & Horovitz, 2010), and the heterogeneity of developmental trajectories of RRBs over time in different patients (Boyd et al., 2012), any intervention targeted at RRBs should be assessed for an appropriate period of time and with multiple time-points.
- Regarding potential risk of bias, only 14 (17.5%) studies were rated as good quality.
- Along the same line, it is difficult to interpret results to a full extent if we consider that IQ assessment was unclear in nearly 50% of the sample and other psychiatric comorbidities (excluding ID) were not clearly disclosed in 61% of the studies. RRBs are not specific to ASD, being present in other psychiatric and neurological disorders (e.g. Tourette Syndrome, Parkinson's Disease, OCD) (M. Lewis & Kim, 2009) and they are influenced by ability and cognitive level, with different phenotypes (i.e. lower level and higher level order) and developmental trajectories (Lam et al., 2008). Without a precise description of psychiatric comorbidities and a characterization of the cognitive functioning of the sample, we are not able to fully evaluate the efficacy of treatments and to identify which subtypes of RRBs are most likely to be reduced in different subjects. Thus, we are also not able to provide precise prognostic indicators and reliable clinical recommendations.

- Finally, even if we did not set exclusion criteria based on age of the sample, mean age in our overall analysis was approximately 10 years. This value reflects how most studies (82% in our sample) in ASD populations are still targeted at children and adolescents (Stagg & Belcher, 2019). If we also consider the abovementioned changes in RRBs during development in ASD subjects, it is evident how results are difficult to extend to adults with ASD.

Most of the studies clearly reported at least one valid diagnostic tool (clinical criteria or gold-standard instruments), but we must underline that ADOS or SRS were also employed to assess changes in outcomes in 12 (26%) studies included in our meta-analysis. Both ADOS and SRS were originally designed as diagnostic tools, so their use to evaluate efficacy of treatments is controversial, since they might not detect subtle improvements (Lord et al., 1989).

Meta-Analysis

46 studies (1339 subjects) in our sample had valid RRBs outcome scores, measured at baseline and endpoint, and were thus included in the meta-analysis. Of note, 65% of these studies used an outcome tool recommended in literature to assess RRBs as outcomes of interventions (ABC, RBS, CY-BOCS) (Lawrence Scahill et al., 2015), with 15% using more than one recommended scale. Evaluation of RRBs with multiple measures, including at least one clinician-based scale and one self-report (or parent-report for children and adolescents), would probably increase sensitivity, specificity, and accuracy of measurement, as seen in other psychiatric assessments (Catalan et al., 2020). We excluded studies which employed only ad-hoc and not validated scales, since results are less reliable and more difficult to compare (Bolte & Diehl, 2013).

Exploratory analysis of the overall effect size of interventions aimed at RRBs was small-to-medium, but significant, with moderate and significant heterogeneity (which was plausible when comparing very different types of interventions). Meta-regression performed on the whole sample failed to predict heterogeneity, using four moderators (mean age of the sample, length of follow up, number of subjects in the intervention group, female ratio). These results might be influenced by the relatively small number of studies and missing data in the variables. Also, to pose a cautious hypothesis, most of the heterogeneity in the overall analysis was probably not correlated to intrinsic characteristics of the sample, but mostly because of some studies finding smaller effect sizes and other studies finding larger effect sizes in the same direction. Unfortunately, meta regression also did not provide significant models of prediction in

the subgroup analyses, most probably due to the small number of studies in each subgroup and for the prevalence of small sample sizes.

Subgroup analysis based on the type of intervention (pharmacological, psychotherapy/psychoeducation, complementary) did not reveal a significant difference between the three subgroups. Even if all effect sizes were on the whole modest, this result is rather surprising, especially when comparing pharmacological treatments with complementary interventions. This could be partly explained by the fact that many complementary treatments were in fact administered as adjuvant therapies, in combination with an atypical antipsychotic. Thus, differences in changes between the groups could be influenced by a different response to medications, and not only to complementary interventions alone. Also, our results are confirmed by previous research (Zhou et al., 2020) which showed the same beneficial, but modest effect on RRBs in smaller subgroups of studies employing medications. The modest effect size of the pharmacological subgroup could be related to:

- The abovementioned absence of effective medications for core symptoms of ASD, comprising RRBs (Pandina et al., 2020)
- If the hypothetical cause of RRBs is attributed to the maintenance of a 'steady-state of arousal' (i.e., increasing activity in moments of hypo-stimulation), the use of a sedative medication could be counter-productive.

Heterogeneity was higher in the pharmacological subgroup, probably caused by the presence of outliers, which are discussed in a separate section of this discussion.

The relative paucity of psychotherapy/psychoeducational treatments included in our synthesis could be explained by the lack of randomized trials employing non-pharmacological interventions to target RRBs (Zarafshan et al., 2017). In fact, psychotherapy and psychoeducational trials could be more expensive and more time-consuming than pharmacological trials, since they could require weekly (or more intensive) sessions led by a professional. Another related issue of importance is the operator-dependent fidelity of treatment. Also, focusing to our specific domain of symptoms, our previous research underlined how psychotherapy and educational trials are mostly aimed at the communication core domain of ASD (Provenzani et al., 2020) and less to RRBs.

Focused analyses on a smaller number of studies also showed a modest, but significant effect of SSRIs and pituitary neuropeptides (oxytocin and vasopressin) with effect sizes comparable to previous findings focused on specific treatments. Given the wider scope of our review, and also considering that meta-analyses regarding specific type of drug treatments in ASD have already been published (Fallah et al., 2019; Hirsch & Pringsheim, 2016; Ooi, Weng, Kossowsky, Gerger, & Sung, 2017; Rossignol & Frye, 2011; Sturman, Deckx, & van Driel, 2017), we will not discuss these results in detail. However, a possible and interesting explanation of the common (even if modest) effect of these medications on RRBs could be related not only to sedation, but also to reduction of anxiety (Russell, Frost, & Ingersoll, 2019) given its possible correlations with RRBs.

Publication bias is an acknowledged issue in meta-analysis of intervention studies in ASD subjects (Carrasco, Volkmar, & Bloch, 2012). Since we did not search (and thus included) grey literature in our synthesis, our conclusions could only be limited to studies published in peer-reviewed journals (see "limitations" section). Inspection of the funnel plot revealed possible presence of publication bias, with a visible difference in the dispersion of studies with larger effect sizes, compared to studies with smaller effect sizes. This difference was substantially modified when we removed the outliers. The absence of studies reporting a worsening of RRBs after treatment could be related to publication bias, but also to the relatively short mean duration of follow up, which could not detect relevant changes in the natural course of these symptoms. Thus, trials which employed an ineffective intervention did not show a worsening in RRBs, but only their stability. Cumulative analysis did not reveal a substantial change of overall effect caused by smaller studies, but we must underline the limited range of the sample sizes included, which limited the strength of this evidence.

Outliers

For a more detailed analysis of heterogeneity in the sample, we include a discussion of the five studies with larger effect size, since their exclusion (even if it did not changed substantially the modest effect sizes) led to considerable reduction of I_2 values, change of shape of the funnel plot and non-significance of the Egger test for funnel plot asymmetry.

Ichikawa et al. 2017 evaluated the efficacy of aripiprazole in 47 patients with ASD, IQ of the sample was unclear. RRBs were assessed with a modified version of ABC scale (ABC-J) and the compulsion scale of CY-BOCS, but the primary endpoint was irritability. Potential overall bias of the study was estimated as medium. A potential explanation of the large effect size could be related to the use of an antipsychotic drug, more likely to induce sedation compared to other treatments. Also, we underline the prominent evaluation of lower-order RRBs, in a sample without a clear assessment of IQ, that would be helpful to interpret the results.

Karahmadi et al 2018 evaluated the use of the glutamate modulator Memantine, subjects in the intervention arm were 30. RRBs were assessed with GARS scale, which is not recommended in literature for evaluation of repetitive behaviours as outcomes of treatment. IQ assessment was unclear, but the potential overall bias was estimated as low. Of note, subjects in intervention and control arm received memantine along with Applied Behaviour Analysis and unspecified “previous common medications” not disclosed in the article.

The study by Lemonnier et al. 2017 reports the results of a trial with bumetanide, a drug whose primary indication is the treatment of hypertension and oedema. Subjects assigned to active treatment were 52. Sedation is acknowledged as possible side effect of the drug, in fact asthenia and fatigue were reported in 33% of the sample. Effects

were assessed with SRS scale, which is not comprised in the instruments recommended in literature for the assessment of RRBs-related outcomes. IQ assessment was unclear, and the potential overall bias was estimated as medium.

McDougle et al. 1998 reported the results of a trial comparing risperidone to placebo, subjects assigned to active treatment were 15. A modified CY-BOCS scale was used to assess RRB, evaluating only repetitive behaviour (and not repetitive thoughts) since 52% of the sample was represented by non-verbal subjects. Mean IQ of the sample was 54.6. Potential bias of the study was estimated as low. As in Ichikawa et al., the large effect size could be related to the higher chance of sedation associated with an antipsychotic drug, compared to other treatments. Also, the limited assessment of motor stereotypies in a low-functioning sample (where "lower-order" RRBs are, according to literature, probably most prominent and thus detectable) could lead to an overestimation of the effect.

Vause et al. 2018 evaluated function-based CBT, subjects assigned to the active treatment were 18. Mean IQ of the sample was 93,61. Symptoms were assessed with RBS and CY-BOCS scale. Of note, authors decided to merge RRBs and OCD symptoms, thus introducing theoretical ambiguity to diagnosis and evaluation of outcomes, especially in a sample with a prevalence of high-functioning subjects.

In summary, three of the five outliers were studies evaluating the efficacy of a drug which had sedation as a potential side effect. Two studies assessed outcomes with a scale not recommended by experts. Sample sizes were small in the five studies, leading to wider confidence intervals and smaller relative weights: their exclusion produced a reduction of overall effect sizes, but they remained significant. IQ assessment was unclear in three of the five studies, thus interpretation of their results (and related effect sizes) is difficult, given the wide variability of RRBs in relation to functioning and cognitive abilities.

Strengths and limitations

As far as we know, this is the first attempt to summarize the efficacy of interventions aimed at RRBs without limitations regarding intervention type, timespan of search and age of the sample. Also, even if we examined a wide range of different interventions, the selection of one specific outcome (improvement of RRBs), linked to one core symptom domain, increased clearness of purpose and homogeneity of our findings.

However, the results from our synthesis must be approached cautiously, due to intrinsic characteristics of the included studies (small sample sizes, high proportion of missing assessments of IQ and psychiatric comorbidities, low mean age of the sample, short term follow ups). Also, some limitations related to our methodological choices must be taken into account.

First, we only included randomized controlled trials published in peer-reviewed journals, excluding case reports, observational studies, open-label, and not-

randomized longitudinal controlled studies. While this approach is supposed to increase the quality of included interventions, it reduces the total number of studies and increases the specificity of the therapeutic setting to experimental frameworks, thus excluding naturalistic observations and limiting the generalisation of results. Also, we did not search for unpublished studies, which could possibly increase publication bias.

Second, our inclusion criteria implied the selection of studies which addressed RRBs (or both core domains of ASD) as a clear aim, primary or secondary. Even if this decision is likely to reduce selective reporting and is supposed to increase the quality of assessments of RRBs, it reduces the total sample of studies and the inclusion of incidental findings. It is also important to note that, given the heterogeneous manifestations of RRB symptoms, it is not always possible to make a clear-cut diagnostic distinction between repetitive behaviours and other ASD and comorbid symptom domains (e.g. irritability, hyperactivity, OCD).

Ultimately, we excluded from our meta-analysis measures of effectiveness based on ad hoc measures and not validated tools. Even if the use of ad hoc measures reduces the possibility of comparison between studies and their psychometric properties are rarely evaluated in the studies which employ them, it is possible that they represent innovations in assessment and evaluations of ASD-related symptoms (Bolte & Diehl, 2013).

Conclusions

Our updated and comprehensive synthesis revealed that randomized controlled studies in ASD subjects aimed at RRBs are often limited by intrinsic characteristics, such as small sample sizes, short follow up period and missing assessment of IQ. Risk of bias is generally medium to high. Pharmacological and complementary interventions are more represented than psychotherapy/psychoeducational programs, probably because the latter are more frequently aimed at the social communication core domain of autism. Meta-analysis showed that there is a modest, but significant beneficial effect of the three subtypes (pharmacological, psychotherapy and education, complementary) of interventions towards RRBs. Overall effect sizes in the three subgroups were similar, but this result must be interpreted cautiously since some interventions (especially in the complementary therapies group) were sometimes administered in addition to pharmacological therapies. Results were influenced by the presence of five outliers, which exclusion did not change drastically the magnitude of the overall effect, nor the significance of the results. Potential publication bias was present as grey literature was not analysed.

Conflicts of interest

The author declare that there is no conflict of interest and that this study did not receive any external funding.

References

Ahearn, W. H., Clark, K. M., Gardenier, N. C., Chung, B. I., & Dube, W. V. (2003). Persistence of stereotypic behavior: examining the effects of external reinforcers. *J Appl Behav Anal, 36*(4), 439-448. doi:10.1901/jaba.2003.36-439

Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985). Psychometric characteristics of the aberrant behavior checklist. *Am J Ment Defic, 89*(5), 492-502.

Anagnostou, E., Chaplin, W., Watner, D., Silverman, J. M., Smith, C. J., Zagursky, K., . . . Hollander, E. (2011). Factor Analysis of Repetitive Behaviors in Autism as Measured by the Y-BOCS. *Journal of Neuropsychiatry and Clinical Neurosciences, 23*(3), 332-339. Retrieved from <Go to ISI>://WOS:000295018300011

Antezana, L., Factor, R. S., Condy, E. E., Strege, M. V., Scarpa, A., & Richey, J. A. (2019). Gender differences in restricted and repetitive behaviors and interests in youth with autism. *Autism Research, 12*(2), 274-283. doi:10.1002/aur.2049

APA. (2013). Neurodevelopmental Disorders. In *Diagnostic and Statistical Manual of Mental Disorders, 5th edition*. Washington, DC.

Arnott, B., McConachie, H., Meins, E., Fernyhough, C., Le Couteur, A., Turner, M., . . . Leekam, S. (2010). The Frequency of Restricted and Repetitive Behaviors in a Community Sample of 15-Month-Old Infants. *Journal of Developmental and Behavioral Pediatrics, 31*(3), 223-229. doi:10.1097/DBP.0b013e3181d5a2ad

Barahona-Corrêa, J. B., Velosa, A., Chainho, A., Lopes, R., & Oliveira-Maia, A. J. (2018). Repetitive Transcranial Magnetic Stimulation for Treatment of Autism Spectrum Disorder: A Systematic Review and Meta-Analysis. *Front Integr Neurosci, 12*, 27. doi:10.3389/fnint.2018.00027

Barrett, S., Prior, M., & Manjiviona, J. (2004). Children on the borderlands of autism: differential characteristics in social, imaginative, communicative and repetitive behaviour domains. *Autism, 8*(1), 61-87. doi:10.1177/1362361304040640

Beggiato, A., Peyre, H., Maruani, A., Scheid, I., Rastam, M., Amsellem, F., . . . Delorme, R. (2017). Gender differences in autism spectrum disorders: Divergence among specific core symptoms. *Autism Res, 10*(4), 680-689. doi:10.1002/aur.1715

Bishop, S. L., Richler, J., Cain, A. C., & Lord, C. (2007). Predictors of perceived negative impact in mothers of children with autism spectrum disorder. *Am J Ment Retard, 112*(6), 450-461. doi:10.1352/0895-8017(2007)112[450:POPNII]2.0.CO;2

Bishop, S. L., Richler, J., & Lord, C. (2006). Association between restricted and repetitive behaviors and nonverbal IQ in children with autism spectrum disorders. *Child Neuropsychology, 12*(4-5), 247-267. doi:10.1080/09297040600630288

Bodfish, J., Symons, F., & Lewis, M. (1999). The repetitive behavior scale. *Western carolina center research reports*.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders, 30*(3), 237-243. doi:10.1023/a:1005596502855

Bolte, E. E., & Diehl, J. J. (2013). Measurement tools and target symptoms/skills used to assess treatment response for individuals with autism spectrum disorder. *J Autism Dev Disord, 43*(11), 2491-2501. doi:10.1007/s10803-013-1798-7

Borenstein. (2009). Multiple Outcomes or Time-Points within a Study. In *Introduction to Meta-Analysis* (pp. 225-238).

Boyd, B. A., McBee, M., Holtzclaw, T., Baranek, G. T., & Bodfish, J. W. (2009). Relationships among repetitive behaviors, sensory features, and executive functions in high functioning autism. *Research in Autism Spectrum Disorders, 3*(4), 959-966. doi:10.1016/j.rasd.2009.05.003

Boyd, B. A., McDonough, S. G., & Bodfish, J. W. (2012). Evidence-Based Behavioral Interventions for Repetitive Behaviors in Autism. *Journal of Autism and Developmental Disorders, 42*(6), 1236-1248. doi:10.1007/s10803-011-1284-z

Carrasco, M., Volkmar, F. R., & Bloch, M. H. (2012). Pharmacologic treatment of repetitive behaviors in autism spectrum disorders: evidence of publication bias. *Pediatrics, 129*(5), e1301-1310. doi:10.1542/peds.2011-3285

Catalan, A., Salazar de Pablo, G., Vaquerizo Serrano, J., Mosillo, P., Baldwin, H., Fernandez-Rivas, A., . . . Fusar-Poli, P. (2020). Annual Research Review: Prevention of psychosis in adolescents - systematic review and meta-analysis of advances in detection, prognosis and intervention. *J Child Psychol Psychiatry*. doi:10.1111/jcpp.13322

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J. A., & Baird, G. (2005). Outcome at 7 years of children diagnosed with autism at age 2: predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *J Child Psychol Psychiatry, 46*(5), 500-513. doi:10.1111/j.1469-7610.2004.00377.x

Chen, Y.-H., Rodgers, J., & McConachie, H. (2009). Restricted and Repetitive Behaviours, Sensory Processing and Cognitive Style in Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 39*(4), 635-642. doi:10.1007/s10803-008-0663-6

Cohen, I. L., Tsiouris, J. A., Flory, M. J., Kim, S. Y., Freedland, R., Heaney, G., . . . Brown, W. T. (2010). A large scale study of the psychometric characteristics of the IBR Modified Overt Aggression Scale: findings and evidence for increased self-destructive behaviors in adult females with autism spectrum disorder. *J Autism Dev Disord, 40*(5), 599-609. doi:10.1007/s10803-009-0908-z

Cuccaro, M. L., Shao, Y. J., Grubber, J., Slifer, M., Wolpert, C. M., Donnelly, S. L., . . . Pericak-Vance, M. A. (2003). Factor analysis of restricted and repetitive behaviors in autism using the Autism Diagnostic Interview-R. *Child Psychiatry & Human Development, 34*(1), 3-17. doi:Doi 10.1023/A:1025321707947

Cunningham, A. B., & Schreibman, L. (2008). Stereotypy in autism: The importance of function. *Research in Autism Spectrum Disorders, 2*(3), 469-479. doi:10.1016/j.rasd.2007.09.006

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials, 7*(3), 177-188. doi:10.1016/0197-2456(86)90046-2

Eilam, D., Zor, R., Szechtman, H., & Hermesh, H. (2006). Rituals, stereotypy and compulsive behavior in animals and humans. *Neurosci Biobehav Rev, 30*(4), 456-471. doi:10.1016/j.neubiorev.2005.08.003

Ejiri, K., & Masataka, N. (2001). Co-occurences of preverbal vocal behavior and motor action in early infancy. *Developmental Science, 4*(1), 40-48. doi:https://doi.org/10.1111/1467-7687.00147

Elison, J. T., Wolff, J. J., Reznick, S., Botteron, K. N., Estes, A. M., Gu, H., . . . Infant Brain Imaging Study, I. (2014). Repetitive Behavior in 12-Month-Olds Later Classified With Autism Spectrum Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 53*(11), 1216-1224. doi:10.1016/j.jaac.2014.08.004

Esbensen, A. J., Seltzer, M. M., Lam, K. S. L., & Bodfish, J. W. (2009). Age-Related Differences in Restricted Repetitive Behaviors in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 39*(1), 57-66. doi:10.1007/s10803-008-0599-x

Fallah, M. S., Shaikh, M. R., Neupane, B., Rusiecki, D., Bennett, T. A., & Beyene, J. (2019). Atypical Antipsychotics for Irritability in Pediatric Autism: A Systematic Review and Network Meta-Analysis. *J Child Adolesc Psychopharmacol, 29*(3), 168-180. doi:10.1089/cap.2018.0115

Fulceri, F., Narzisi, A., Apicella, F., Balboni, G., Baldini, S., Brocchini, J., . . . Calderoni, S. (2016). Application of the Repetitive Behavior Scale-Revised--Italian version--in preschoolers with autism spectrum disorder. *Research in Developmental Disabilities, 48*, 43-52. doi:10.1016/j.ridd.2015.10.015

Gates, J. A., Kang, E., & Lerner, M. D. (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. *Clin Psychol Rev, 52*, 164-181. doi:10.1016/j.cpr.2017.01.006

Healy, S., Nacario, A., Braithwaite, R. E., & Hopper, C. (2018). The effect of physical activity interventions on youth with autism spectrum disorder: A meta-analysis. *Autism Res, 11*(6), 818-833. doi:10.1002/aur.1955

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ, 343*, d5928. doi:10.1136/bmj.d5928

Hiller, R. M., Young, R. L., & Weber, N. (2014). Sex differences in autism spectrum disorder based on DSM-5 criteria: evidence from clinician and teacher reporting. *J Abnorm Child Psychol, 42*(8), 1381-1393. doi:10.1007/s10802-014-9881-x

Hirsch, L. E., & Pringsheim, T. (2016). Aripiprazole for autism spectrum disorders (ASD). *Cochrane Database Syst Rev, 2016*(6), Cd009043. doi:10.1002/14651858.CD009043.pub3

Honey, E., McConachie, H., Turner, M., & Rodgers, J. (2012). Validation of the repetitive behaviour questionnaire for use with children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 6*(1), 355-364. doi:10.1016/j.rasd.2011.06.009

Horner, R. H., Carr, E. G., Strain, P. S., Todd, A. W., & Reed, H. K. (2002). Problem behavior interventions for young children with autism: a research synthesis. *J Autism Dev Disord, 32*(5), 423-446. doi:10.1023/a:1020593922901

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *Am J Intellect Dev Disabil, 114*(1), 23-41. doi:10.1352/2009.114:23;nd41

Hus, V., Pickles, A., Cook, E. H., Jr., Risi, S., & Lord, C. (2007). Using the autism diagnostic interview--revised to increase phenotypic homogeneity in genetic studies of autism. *Biol Psychiatry, 61*(4), 438-448. doi:10.1016/j.biopsych.2006.08.044

Hutt, S. J., & Hutt, C. (1970). *Direct Observation and Measurement of Behavior*: Charles C. Thomas.

Iverson, J. M., & Wozniak, R. H. (2007). Variation in vocal-motor development in infant siblings of children with autism. *Journal of Autism and Developmental Disorders, 37*(1), 158-170. doi:10.1007/s10803-006-0339-z

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*(2), 197-209. doi:10.1901/jaba.1994.27-197

Jiujias, M., Kelley, E., & Hall, L. (2017). Restricted, Repetitive Behaviors in Autism Spectrum Disorder and Obsessive-Compulsive Disorder: A Comparative Review. *Child Psychiatry Hum Dev, 48*(6), 944-959. doi:10.1007/s10578-017-0717-0

Kamp-Becker, I., Ghahreman, M., Smidt, J., & Remschmidt, H. (2009). Dimensional structure of the autism phenotype: Relations between early development and current presentation. *Journal of Autism and Developmental Disorders, 39*(4), 557-571. doi:10.1007/s10803-008-0656-5

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217-250.

Kim, S. H., & Lord, C. (2010). Restricted and Repetitive Behaviors in Toddlers and Preschoolers with Autism Spectrum Disorders Based on the Autism Diagnostic Observation Schedule (ADOS). *Autism Research, 3*(4), 162-173. doi:10.1002/aur.142

Klinger, L. G., Dawson, G., & Renner, P. (2003). Autistic disorder. *Child psychopathology, 2*, 409-454.

Lam, K. S. L., & Aman, M. G. (2007). The repetitive behavior scale-revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(5), 855-866. doi:10.1007/s10803-006-0213-z

Lam, K. S. L., Bodfish, J. W., & Piven, J. (2008). Evidence for three subtypes of repetitive behavior in autism that differ in familiality and association with other symptoms. *Journal of Child Psychology and Psychiatry, 49*(11), 1193-1200. doi:10.1111/j.1469-7610.2008.01944.x

Langen, M., Kas, M. J., Staal, W. G., van Engeland, H., & Durston, S. (2011). The neurobiology of repetitive behavior: of mice…. *Neurosci Biobehav Rev, 35*(3), 345-355. doi:10.1016/j.neubiorev.2010.02.004

Langen, M., Schnack, H. G., Nederveen, H., Bos, D., Lahuis, B. E., de Jonge, M. V., . . . Durston, S. (2009). Changes in the Developmental Trajectories of Striatum in Autism. *Biological Psychiatry, 66*(4), 327-333. doi:10.1016/j.biopsych.2009.03.017

Leekam, S. R., Prior, M. R., & Uljarevic, M. (2011). Restricted and Repetitive Behaviors in Autism Spectrum Disorders: A Review of Research in the Last Decade. *Psychological Bulletin, 137*(4), 562-593. doi:10.1037/a0023341

Lewis, M., & Kim, S.-J. (2009). The pathophysiology of restricted repetitive behavior. *Journal of Neurodevelopmental Disorders, 1*(2), 114-132. doi:10.1007/s11689-009-9019-6

Lewis, M. H., Tanimura, Y., Lee, L. W., & Bodfish, J. W. (2007). Animal models of restricted repetitive behavior in autism. *Behavioural Brain Research, 176*(1), 66-74. doi:10.1016/j.bbr.2006.08.023

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 healthcare interventions: explanation and elaboration. *BMJ, 339*, b2700. doi:10.1136/bmj.b2700

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Arch Gen Psychiatry, 63*(6), 694-701. doi:10.1001/archpsyc.63.6.694

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *J Autism Dev Disord, 19*(2), 185-212. doi:10.1007/bf02211841

Lovaas, I., Newsom, C., & Hickman, C. (1987). Self-stimulatory behavior and perceptual reinforcement. *Journal of Applied Behavior Analysis, 20*(1), 45-68. doi:10.1901/jaba.1987.20-45

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: evidence from a large sample of children and adolescents. *J Autism Dev Disord, 42*(7), 1304-1313. doi:10.1007/s10803-011-1356-0

Matson, J. L., & Horovitz, M. (2010). Stability of Autism Spectrum Disorders Symptoms over Time. *Journal of Developmental and Physical Disabilities, 22*(4), 331-342. doi:10.1007/s10882-010-9188-y

May, T., Cornish, K., & Rinehart, N. J. (2016). Gender Profiles of Behavioral Attention in Children With Autism Spectrum Disorder. *Journal of Attention Disorders, 20*(7), 627-635. doi:10.1177/1087054712455502

Mayes, S. D., Calhoun, S. L., Aggarwal, R., Baker, C., Mathapati, S., Anderson, R., & Petersen, C. (2012). Explosive, oppositional, and aggressive behavior in children with autism compared to other clinical disorders and typical children. *Research in Autism Spectrum Disorders, 6*(1), 1-10. doi:https://doi.org/10.1016/j.rasd.2011.08.001

Ooi, Y. P., Weng, S. J., Kossowsky, J., Gerger, H., & Sung, M. (2017). Oxytocin and Autism Spectrum Disorders: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Pharmacopsychiatry, 50*(1), 5-13. doi:10.1055/s-0042-109400

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *J Clin Child Adolesc Psychol, 34*(3), 523-540. doi:10.1207/s15374424jccp3403_8

Ozonoff, S., Macari, S., Young, G. S., Goldring, S., Thompson, M., & Rogers, S. J. (2008). Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism, 12*(5), 457-472. doi:10.1177/1362361308096402

Pandina, G., Ring, R. H., Bangerter, A., & Ness, S. (2020). Current Approaches to the Pharmacologic Treatment of Core Symptoms Across the Lifespan of Autism Spectrum Disorder. *Child Adolesc Psychiatr Clin N Am, 29*(2), 301-317. doi:10.1016/j.chc.2019.12.004

Patterson, S. Y., Smith, V., & Jelen, M. (2010). Behavioural intervention practices for stereotypic and repetitive behaviour in individuals with autism spectrum disorder: a systematic review. *Developmental Medicine and Child Neurology, 52*(4), 318-327. doi:10.1111/j.1469-8749.2009.03597.x

Penson, D., Krishnaswami, S., Jules, A., & McPheeters, M. L. (2013). Effectiveness of hormonal and surgical therapies for cryptorchidism: a systematic review. *Pediatrics, 131*(6), e1897-e1907. doi:10.1542/peds.2013-0072

Provenzani, U., Fusar-Poli, L., Brondino, N., Damiani, S., Vercesi, M., Meyer, N., . . . Politi, P. (2020). What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. *Autism, 24*(2), 274-284. doi:10.1177/1362361319854641

Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy I: a review of behavioral assessment and treatment. *Res Dev Disabil, 26*(6), 527-547. doi:10.1016/j.ridd.2004.11.005

Rapp, J. T., & Vollmer, T. R. (2005). Stereotypy II: A review of neurobiological interpretations and suggestions for an integration with behavioral methods. *Research in Developmental Disabilities, 26*(6), 548-564. doi:10.1016/j.ridd.2004.11.006

Ratto, A. B., Kenworthy, L., Yerys, B. E., Bascom, J., Wieckowski, A. T., White, S. W., . . . Anthony, L. G. (2018). What About the Girls? Sex-Based Differences in Autistic Traits and Adaptive Skills. *Journal of Autism and Developmental Disorders, 48*(5), 1698-1711. doi:10.1007/s10803-017-3413-9

Richler, J., Huerta, M., Bishop, S. L., & Lord, C. (2010). Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. *Development and Psychopathology, 22*(1), 55-69. doi:10.1017/s0954579409990265

Rogers, S. J., & Ozonoff, S. (2005). Annotation: what do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *J Child Psychol Psychiatry, 46*(12), 1255-1268. doi:10.1111/j.1469-7610.2005.01431.x

Rojahn, J., Barnard-Brak, L., Medeiros, K., & Schroeder, S. R. (2016). Stereotyped behaviours as precursors of self-injurious behaviours: a longitudinal study with infants and toddlers at risk for developmental delay. *Journal of Intellectual Disability Research, 60*(2), 156-166. doi:https://doi.org/10.1111/jir.12224

Rojahn, J., Matlock, S. T., & Tassé, M. J. (2000). The Stereotyped Behavior Scale: Psychometric properties and norms. *Research in Developmental Disabilities, 21*(6), 437-454. doi:10.1016/S0891-4222(00)00057-3

Rossignol, D. A., & Frye, R. E. (2011). Melatonin in autism spectrum disorders: a systematic review and meta-analysis. *Dev Med Child Neurol, 53*(9), 783-792. doi:10.1111/j.1469-8749.2011.03980.x

Russell, K. M., Frost, K. M., & Ingersoll, B. (2019). The relationship between subtypes of repetitive behaviors and anxiety in children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 62*, 48-54. doi:https://doi.org/10.1016/j.rasd.2019.03.006

Rynkiewicz, A., Schuller, B., Marchi, E., Piana, S., Camurri, A., Lassalle, A., & Baron-Cohen, S. (2016). An investigation of the 'female camouflage effect' in autism using a computerized ADOS-2 and a test of sex/gender differences. *Mol Autism, 7*, 10. doi:10.1186/s13229-016-0073-0

Scahill, L., Aman, M. G., Lecavalier, L., Halladay, A. K., Bishop, S. L., Bodfish, J. W., . . . Dawson, G. (2015). Measuring repetitive behaviors as a treatment endpoint in youth with autism spectrum disorder. *Autism, 19*(1), 38-52. doi:10.1177/1362361313510069

Scahill, L., Dimitropoulos, A., McDougle, C. J., Aman, M. G., Feurer, I. D., McCracken, J. T., . . . Vitiello, B. (2014). Children's Yale-Brown Obsessive Compulsive Scale in Autism Spectrum Disorder: Component Structure and Correlates of Symptom Checklist. *Journal of the American Academy of Child and Adolescent Psychiatry, 53*(1), 97-107. doi:10.1016/j.jaac.2013.09.018

Scahill, L., Riddle, M. A., McSwiggin-Hardin, M., Ort, S. I., King, R. A., Goodman, W. K., . . . Leckman, J. F. (1997). Children's Yale-Brown Obsessive Compulsive Scale: reliability and validity. *J Am Acad Child Adolesc Psychiatry, 36*(6), 844-852. doi:10.1097/00004583-199706000-00023

Schroeder, S. R., Marquis, J. G., Reese, R. M., Richman, D. M., Mayo-Ortega, L., Oyama-Ganiko, R., . . . Lawrence, L. (2014). Risk Factors for Self-Injury, Aggression, and Stereotyped Behavior Among Young Children At Risk for Intellectual and Developmental Disabilities. *Ajidd-American Journal on Intellectual and Developmental Disabilities, 119*(4), 351-370. doi:10.1352/1944-7558-119.4.351

Shao, Y., Raiford, K. L., Wolpert, C. M., Cope, H. A., Ravan, S. A., Ashley-Koch, A. A., . . . Pericak-Vance, M. A. (2002). Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *Am J Hum Genet, 70*(4), 1058-1061. doi:10.1086/339765

Soorya, L., Kiarashi, J., & Hollander, E. (2008). Psychopharmacologic interventions for repetitive behaviors in autism spectrum disorders. *Child and Adolescent Psychiatric Clinics of North America, 17*(4), 753-+. doi:10.1016/j.chc.2008.06.003

South, M., Ozonoff, S., & McMahon, W. M. (2005). Repetitive behavior profiles in Asperger syndrome and high-functioning autism. *J Autism Dev Disord, 35*(2), 145-158. doi:10.1007/s10803-004-1992-8

Stagg, S. D., & Belcher, H. (2019). Living with autism without knowing: receiving a diagnosis in later life. *Health Psychology and Behavioral Medicine, 7*(1), 348-361. doi:10.1080/21642850.2019.1684920

Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., & Weiss, B. H. (1999). Can Autism Be Diagnosed Accurately in Children Under 3 Years? *Journal of Child Psychology and Psychiatry, 40*(2), 219-226. doi:https://doi.org/10.1111/1469-7610.00435

Sturman, N., Deckx, L., & van Driel, M. L. (2017). Methylphenidate for children and adolescents with autism spectrum disorder. *Cochrane Database Syst Rev, 11*(11), Cd011144. doi:10.1002/14651858.CD011144.pub2

Su Maw, S., & Haga, C. (2018). Effectiveness of cognitive, developmental, and behavioural interventions for Autism Spectrum Disorder in preschool-aged children: A systematic review and meta-analysis. *Heliyon, 4*(9), e00763. doi:10.1016/j.heliyon.2018.e00763

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W., . . . Tuff, L. (2006). Investigating the structure of the restricted, repetitive behaviours and interests domain of autism. *J Child Psychol Psychiatry, 47*(6), 582-590. doi:10.1111/j.1469-7610.2005.01537.x

Troyb, E., Knoch, K., Herlihy, L., Stevens, M. C., Chen, C. M., Barton, M., . . . Fein, D. (2016). Restricted and Repetitive Behaviors as Predictors of Outcome in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 46*(4), 1282-1296. doi:10.1007/s10803-015-2668-2

Turner-Brown, L. (2020). Understanding and Addressing Restricted and Repetitive Behaviors in Children with Autism. In *Clinical Guide to Early Interventions for Children with Autism* (pp. 61-77).

Turner, M. (1999). Repetitive behaviour in autism: a review of psychological research. *J Child Psychol Psychiatry, 40*(6), 839-849. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10509879

Watt, N., Wetherby, A. M., Barber, A., & Morgan, L. (2008). Repetitive and stereotyped behaviors in children with autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 38*(8), 1518-1533. doi:10.1007/s10803-007-0532-8

Wigham, S., Rodgers, J., South, M., McConachie, H., & Freeston, M. (2015). The Interplay Between Sensory Processing Abnormalities, Intolerance of Uncertainty, Anxiety and Restricted and Repetitive Behaviours in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 45*(4), 943-952. doi:10.1007/s10803-014-2248-x

Wink, L. K., Erickson, C. A., & McDougle, C. J. (2010). Pharmacologic Treatment of Behavioral Symptoms Associated With Autism and Other Pervasive Developmental Disorders. *Current Treatment Options in Neurology, 12*(6), 529-538. doi:10.1007/s11940-010-0091-8

Wolff, J. J., Boyd, B. A., & Elison, J. T. (2016). A quantitative measure of restricted and repetitive behaviors for early childhood. *Journal of Neurodevelopmental Disorders, 8*. doi:10.1186/s11689-016-9161-x

Wolstencroft, J., Robinson, L., Srinivasan, R., Kerry, E., Mandy, W., & Skuse, D. (2018). A Systematic Review of Group Social Skills Interventions, and Meta-analysis of Outcomes, for Children with High Functioning ASD. *J Autism Dev Disord, 48*(7), 2293-2307. doi:10.1007/s10803-018-3485-1

Yerys, B. E., Wallace, G. L., Harrison, B., Celano, M. J., Giedd, J. N., & Kenworthy, L. E. (2009). Set-shifting in children with autism spectrum disorders Reversal shifting deficits on the Intradimensional/Extradimensional Shift Test correlate with repetitive behaviors. *Autism, 13*(5), 523-538. doi:10.1177/1362361309335716

Yu, Y., Chaulagain, A., Pedersen, S. A., Lydersen, S., Leventhal, B. L., Szatmari, P., . . . Skokauskas, N. (2020). Pharmacotherapy of restricted/repetitive behavior in autism

spectrum disorder:a systematic review and meta-analysis. *Bmc Psychiatry, 20*(1). doi:10.1186/s12888-020-2477-9

Zandt, F., Prior, M., & Kyrios, M. (2007). Repetitive behaviour in children with high functioning autism and obsessive compulsive disorder. *J Autism Dev Disord, 37*(2), 251-259. doi:10.1007/s10803-006-0158-2

Zarafshan, H., Salmanian, M., Aghamohammadi, S., Mohammadi, M. R., & Mostafavi, S.-A. (2017). Effectiveness of Non-Pharmacological Interventions on Stereotyped and Repetitive Behaviors of Preschool Children With Autism: A Systematic Review. *Basic and Clinical Neuroscience, 8*(2), 95-104. doi:10.18869/nirp.bcn.8.2.95

Zhou, M. S., Nasir, M., Farhat, L. C., Kook, M., Artukoglu, B. B., & Bloch, M. H. (2020). Meta-analysis: Pharmacologic Treatment of Restricted and Repetitive Behaviors in Autism Spectrum Disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*. doi:10.1016/j.jaac.2020.03.007

Appendix

Table 1a: List of included studies

(studies with symbol * are also included in meta-analysis)

Author	Year	Treatment	Comparison	Intervention Type	Maximum Follow Up (weeks)	Subjects in intervention arm	Overall Bias	Study Location
*Adams et al. **	2018	Vitamin-mineral supplement	TAU	Complementary and others	52	37	Fair	United States
Adams et al.	2011	Vitamin-mineral supplement	Placebo	Complementary and others	12	72	Poor	United States
Amatachaya et al.	2014	tDCS	Placebo	Complementary and others	1	20	Poor	Thailand
Amatachaya et al.	2015	tDCS	Placebo	Complementary and others	1	20	Fair	Thailand
*Anagnostou et al. **	2012	Oxytocin	Placebo	Pharmacological	6	10	Fair	United States
*Behmanesh et al. **	2019	Propentofylline	Placebo	Pharmacological	10	33	Good	Iran

*Bolognani et al. **	2019	Balovaptan	Placebo	Pharmacological	18	148	Fair	United States
*Chugani et al. **	2016	Buspirone	Placebo	Pharmacological	48	109	Fair	United States
*Dadds et al. **	2014	Oxytocin	Placebo	Pharmacological	13	19	Fair	Australia
DaWalt et al.	2018	Multi-family group psychoeducation	Waiting list	Psychotherapy and education	13	19	Poor	United States
De Veld et al.	2017	Theory of Mind training	Waiting list	Psychotherapy and education	8	74	Poor	Netherlands
*Dean et al. **	2017	N-acetylcysteine	Placebo	Complementary and others	26	48	Fair	Australia
Dollfus et al.	1992	Bromocriptine	Amisulpride	Pharmacological	4	18	Fair	France
Dunn-Geier et al.	2000	Secretin	Placebo	Pharmacological	3	47	Good	Canada

Elder et al.	2006	Gluten and Casein free diet	TAU	Complementary and others	6	15	Poor	United States
Eslamzadeh et al.	2018	Atomoxetine	Placebo	Pharmacological	8	21	Poor	Iran
Fazlioglu et al.	2008	Sensory integration therapy program	Waiting list	Psychotherapy and education	12	15	Poor	Turkey
*Frye et al. **	2018	Folinic acid	Placebo	Complementary and others	12	23	Fair	United States
Gordon et al.	1993	Desipramine; Clomipramine	Placebo	Pharmacological	10	14	Fair	United States
*Granpeesheh et al. **	2010	Hyperbaric oxygen therapy (HBOT)	Placebo	Complementary and others	15		Poor	United States
*Hajizadeh-Zaker et al. **	2017	L-Carnosine	Placebo	Complementary	10	25	Good	Iran

Hardan et al.	2019	Memantine	Placebo	Pharmacological	12	319	Fair	Multiple locations
*Hendouei et al. **	2019	Resveratrol	Placebo	Complementary and others	10	35	Good	Iran
*Hendren et al. **	2016	Vitamin B12	Placebo	Complementary and others	8	28	Fair	United States
*Herscu et al. **	2019	Fluoxetine	placebo	Pharmacological	14	78	Fair	United States
*Hollander et al. **	2005	Fluoxetine	Placebo	Pharmacological	8	39	Fair	United States
*Hollander et al. **	2018	Trichuris suis Ova	Placebo	Complementary and others	28	10	Poor	United States
*Hollander et al. **	2012	Fluoxetine	Placebo	Pharmacological	12	21	Fair	United States

Hollander et al.	2003	Oxytocin	Placebo	Pharmacological	1	15	Fair	United States
*Ichikawa et al. **	2017	Aripiprazole	Placebo	Pharmacological	8	47	Fair	Japan
*Karahmadi et al. **	2018	Memantine	Placebo	Pharmacological	12	30	Good	Iran
Kent et al.	2013	Risperidone	placebo	Pharmacological	6	61	Good	United States
Kerley et al.	2017	Vitamin D3	Placebo	Complementary and others	20	22	Fair	Ireland
*Khalaj et al. **	2018	Palmitoylethanolamide	Placebo	Complementary and others	10	35	Fair	Iran
*King et al. **	2009	Citalopram	Placebo	Pharmacological	12	73	Fair	United States
*Klaiman et al. **	2013	Tetrahydrobiopterin	Placebo	Complementary and others	16	23	Fair	United States

Kong et al.	2018	Acupuncture	TAU	Complementary and others	16	30	Poor	China
*Lemonnier et al. **	2017	Bumetanide	Placebo	Pharmacological	20	<65	Fair	France
Lomas Mevers et al.	2019	Multidisciplinary intervention for encopresis	waiting list	Pharmacological	10	10	Poor	United States
Lu et al.	2018	Massage plus routine rehabilitation intervention	TAU	Complementary and others	12	20	Fair	Taiwan
*Mahdavinasab et al. **	2019	Baclofen	Placebo	Pharmacological	10	32	Good	Iran
*Mankad et al. **	2015	Omega-3 fatty acids	Placebo	Complementary and others	24	19	Fair	Canada
Mazahery et al.	2019	Vitamin D and Omega-3 fatty acids	placebo	Complementary and others	48	88	Good	New Zealand

McDougle et al. *	1998	Risperidone	Placebo	Pharmacological	12	15	Good	United States
McDougle et al. *	1996	Fluvoxamine	Placebo	Pharmacological	12	15	Fair	United States
Mehrazad-Saber et al. *	2018	L-Carnosine	Placebo	Complementary and others	8,7	25	Fair	Iran
Moazen-Zadeh et al. *	2017	Simvastatin	Placebo	Pharmacological	10	35	Fair	Iran
Moharreri et al.	2017	Naltrexone plus risperidone	Placebo plus risperidone	Pharmacological	10	30	Poor	Iran
Ni et al.	2017	Intermittent Theta Burst Stimulation (iTBS)	Placebo	Complementary and others	1	25	Poor	Taiwan
Nikvarz et al. *	2016	Memantine	Risperidone	Pharmacological	8	18	Poor	Iran
Overwater et al.	2019	Everolimus	placebo	Pharmacological	48	15	Good	Netherlands

*Owley et al. **	2001	Secretin	Placebo	Pharmacological	8	56	Good	United States
Owley et al.	1999	Secretin	Placebo	Pharmacological	8	10	Good	United States
*Parellada et al. **	2017	Omega-3	Placebo	Complementary and others	8	77	Poor	Spain
*Parker et al. **	2019	Intranasal arginine vasopressin (AVP)	Placebo	Pharmacological	4	17	Fair	United States
*Parker et al. **	2017	Oxytocin	Placebo	Pharmacological	4	17	Fair	United States
*Pfeiffer et al. **	2011	Sensory integration treatment	Fine motor treatment	Psychotherapy and education	6	37	Poor	United States
*Piwowarczyk et al. **	2019	Gluten-free diet (GFD)	TAU	Complementary and others	24	33	Poor	Poland

Remington et al.	2001	Clomipramine; Haloperidol	Placebo	Pharmacological	7	36	Poor	Canada
Roberts W et al.	2001	Secretin	Placebo	Pharmacological	6	32	Poor	Canada
Rogers et al.	2019	Early Start Denver Model (ESDM)	TAU	Complementary and others	116	55	Poor	United States
*Saad et al. **	2018	Vitamin D3	Placebo	Complementary and others	17	60	Good	Egypt
Sampanthavivat et al.	2012	Hyperbaric treatment	Pressurized room air	Complementary and others	2	30	Fair	Thailand
*Sandler et al. **	1999	Secretin	Placebo	Pharmacological	4	28	Poor	United States
Schohl et al.	2014	PEERS®	Waiting list	Psychotherapy and education	14	34	Poor	United States

*Scudder et al. **	2019	Parent–child interaction therapy (PCIT)	Waiting list	Psychotherapy and education	18	13	Fair	United States
*Shum et al. **	2019	PEERS®	Waiting list	Psychotherapy and education	42	38	Fair	Hong Kong
Sizoo et al.	2017	CBT	MBSR (Mindfulness Based Stress Reduction)	Complementary and others	26	27	Poor	Netherlands
Srinivasan et al.	2015	Rhythm intervention; Robotic intervention	TAU	Complementary and others	10	24	Poor	United States
*Storch et al. **	2019	Family-based exposure-focused treatment (FET)	TAU	Psychotherapy and education	20	14	Fair	United States

Thomeer et al.	2012	Comprehensive psychosocial intervention	Waiting list	Psychotherapy and education	5	17	Poor	United States
*Vause et al. **	2018	Functional behavior-based cognitive behavior therapy (Fb-CBT)	TAU	Psychotherapy and education	24	19	Fair	Canada
Voigt et al.	2014	Omega-3 Fatty Acids	Placebo	Complementary and others	26	24	Fair	United States
Wang et al.	2017	Levetiracetam	TAU	Pharmacological	26	35	Poor	China
*Wetherby et al. **	2014	ESI (Early social interaction) individual	ESI (Early social interaction) group	Psychotherapy and education	39	82	Fair	United States

*Whitehouse et al. **	2017	iPad-based application (TOBY-Therapy Outcomes By You)	TAU	Psychotherapy and education	26	41	Poor	Australia
Whiteley et al.	2010	Gluten and Casein-free diet	TAU	Psychotherapy and education	52	38	Poor	Denmark
Wood et al.	2015	CBT	Waiting list	Psychotherapy and education	16	19	Fair	United States
*Wood et al. **	2019	CBT	TAU	Psychotherapy and education	17	148	Fair	United States
*Yamasue et al. **	2018	Intranasal oxytocin	Placebo	Pharmacological	8	53	Good	Japan

Table 2a: Cochrane Risk of Bias tool for randomized controlled studies (RCT)

Study name	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other source of bias	Overall bias level
Adams et al. 2011	Low	Low	Low	Low	Low	Low	High	Fair
Adams et al. 2018	Unclear	Unclear	High	Unclear	Unclear	Low	High	Poor
Amatachaya et al. 2014	Low	Unclear	Unclear	Unclear	Low	Low	Unclear	Poor
Amatachaya et al. 2015	Low	Unclear	High	Low	Low	Low	Unclear	Fair
Anagnostou et al. 2012	Low	Low	Low	Low	Unclear	Low	High	Fair
Behmanesh et al. 2019	Low	Low	Low	Low	Low	Low	Low	Good
Bolognani et al. 2019	Low	Low	Unclear	Low	Unclear	Low	Low	Fair
Chugani et al. 2016	Unclear	Unclear	Low	Unclear	High	Low	Low	Fair
Dadds et al. 2014	Unclear	Unclear	Low	Low	Low	Low	Low	Fair
DaWalt et al. 2018	Unclear	Unclear	High	High	Low	Low	Low	Poor
De Veld et al. 2017	Low	Unclear	High	High	Low	Low	Unclear	Poor
Dean et al. 2017	Low	Low	Low	Unclear	Unclear	Low	Low	Fair
Dollfus et al. 1992	Unclear	Unclear	Low	Low	Low	High	Low	Fair
Dunn-Geier et al. 2000	Low	Low	Low	Low	Low	Low	Low	Good
Elder et al. 2006	Low	Unclear	High	Low	High	High	Low	Poor
Eslamzadeh et al. 2018	Unclear	Low	Unclear	Unclear	Unclear	Unclear	Low	Poor
Fazlioglu et al. 2008	Unclear	Unclear	High	Unclear	Unclear	Low	Low	Poor

Frye et al. 2018	Low	Low	Low	Low	Low	Low	Unclear	Fair
Gordon et al. 1993	Unclear	Unclear	Low	Unclear	High	Low	Low	Fair
Granpeesheh et al. 2010	Low	Unclear	Low	Low	High	Unclear	High	Poor
Hajizadeh-Zaker et al. 2017	Low	Low	Low	Low	Low	Low	Low	Good
Hardan et al. 2019	Low	Low	Low	Unclear	Unclear	Low	Low	Fair
Hendouei et al. 2019	Low	Low	Low	Low	Low	Low	Low	Good
Hendren et al. 2016	Low	Low	Low	Low	High	Low	Low	Fair
Herscu et al. 2019	Low	Low	Low	Unclear	Low	Low	Low	Fair
Hollander et al. 2005	Unclear	Unclear	Low	Low	Low	Low	Unclear	Fair
Hollander et al. 2018	Low	Low	Unclear	Unclear	Unclear	Low	Low	Poor
Hollander et al. 2003	Unclear	Unclear	Low	Unclear	Low	Low	Low	Fair
Hollander et al. 2012	Unclear	Unclear	Low	Low	High	Low	Low	Fair
Ichikawa et al. 2017	Low	Low	Low	Low	Low	Low	Unclear	Fair
Karahmadi et al. 2018	Low	Low	Low	Low	Low	Low	Low	Good
Kent et al. 2013	Low	Low	Low	Low	Low	Low	Low	Good
Kerley et al. 2017	Low	Unclear	Low	Unclear	Low	Low	Low	Fair
Khalaj et al. 2018	Unclear	Unclear	Low	Unclear	Low	Low	Low	Fair
King et al. 2009	Low	Unclear	Low	Low	Low	Low	High	Fair
Klaiman et al. 2013	Unclear	Unclear	High	Low	Low	Low	Low	Fair
Kong et al. 2018	Unclear	Unclear	High	High	High	Unclear	Low	Poor
Lemonnier et al. 2017	Low	Low	Low	Low	Unclear	Unclear	High	Fair

Lomas Mevers et al. 2019	Low	Unclear	Unclear	Low	Unclear	Low	Low	Poor
Lu et al. 2018	Low	Unclear	High	Unclear	Low	Low	Low	Fair
Mahdavinasab et al. 2019	Low	Low	Low	Low	Low	Low	Low	Good
Mankad et al. 2015	Unclear	Low	Low	Unclear	Low	High	Low	Fair
Mazahery et al. 2019	Low	Low	Low	Low	Low	Low	Low	Good
McDougle et al. 1998	Low	Low	Low	Low	Low	Low	Low	Good
McDougle et al. 1996	Unclear	Unclear	Low	Low	Low	Low	High	Fair
Mehrazad-Saber et al. 2018	Unclear	Unclear	Low	High	Low	Low	Low	Fair
Moazen-Zadeh et al. 2017	Low	Low	Low	High	Low	Unclear	Low	Fair
Moharreri et al. 2017	Low	Unclear	Low	Low	Unclear	Low	Unclear	Poor
Ni et al. 2017	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low	Poor
Nikvarz et al. 2016	Unclear	Unclear	High	High	Low	Low	Low	Poor
Overwater et al. 2019	Low	Low	Low	Low	Low	Low	Low	Good
Owley et al. 1999	Low	Low	Low	Low	Low	Low	Low	Good
Owley et al. 2001	Low	Low	Low	Low	Low	Low	Low	Good
Parellada et al. 2017	Unclear	Unclear	Low	Low	Unclear	Unclear	Low	Poor
Parker et al. 2017	Low	Low	Low	Low	Unclear	Unclear	Low	Fair
Parker et al. 2019	Low	Low	Low	Unclear	Unclear	Low	Low	Fair
Pfeiffer et al. 2011	Low	Unclear	High	Low	Low	High	Low	Poor
Piwowarczyk et al. 2019	Low	High	High	Unclear	Unclear	Low	Low	Poor
Remington et al. 2001	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Low	Poor

Roberts et al. 2001	Low	Unclear	High	Low	Low	High	Low	Poor
Rogers et al. 2019	Low	Unclear	High	Unclear	Unclear	Low	Low	Poor
Saad et al. 2018	Low	Low	Low	Low	Low	Low	Low	Good
Sampanthavivat et al. 2012	Low	Low	Low	Low	High	Low	Low	Fair
Sandler et al. 1999	Unclear	Unclear	Low	Low	Unclear	Unclear	Low	Poor
Schohl et al. 2014	Unclear	Unclear	High	High	High	Low	Low	Poor
Scudder et al. 2019	Unclear	Unclear	High	Unclear	Low	Low	Low	Fair
Shum et al. 2019	Low	Unclear	High	Unclear	Low	Low	Unclear	Poor
Sizoo et al. 2017	High	High	High	High	Unclear	Low	Low	Poor
Srinivasan et al. 2015	Low	Unclear	High	Unclear	High	Low	Low	Poor
Storch et al. 2019	Low	Unclear	High	Unclear	Low	Low	Low	Fair
Thomeer et al. 2012	Unclear	Unclear	High	Unclear	High	Unclear	Low	Poor
Vause et al. 2018	Low	Low	Unclear	Unclear	Low	Low	Low	Fair
Voigt et al. 2014	Low	Unclear	Low	Low	Low	High	Unclear	Fair
Wang et al. 2017	Low	Unclear	High	High	Low	Low	Low	Poor
Wetherby et al. 2014	Low	Unclear	High	Unclear	Low	Low	Low	Fair
Whitehouse et al. 2017	Unclear	High	High	Unclear	Unclear	Low	High	Poor
Whiteley et al. 2010	Low	Unclear	High	Unclear	High	Unclear	Low	Poor
Wood et al. 2015	Low	Unclear	High	Low	Low	Low	Low	Fair
Wood et al. 2019	Low	Unclear	Unclear	Low	Low	Low	Low	Fair
Yamasue et al. 2018	Low	Low	Low	Low	Low	Low	Low	Good

AMSTAR 2

Checklist per la valutazione critica di revisioni sistematiche di trial controllati randomizzati e/o di studi non randomizzati sull'efficacia degli interventi sanitari

1. I quesiti di ricerca e i criteri di inclusione della revisione comprendono gli elementi del PICO?

☒ Sì	☐ No

Per rispondere **Sì**, devono essere presenti:

- ☒ popolazione
- ☒ intervento
- ☒ gruppo di confronto
- ☒ outcome
- ☐ durata del follow-up (*opzionale, ma raccomandato*)

2. La revisione sistematica dichiara esplicitamente che i metodi sono stati definiti prima della sua conduzione, motivando tutte le violazioni significative del protocollo?

☐ Sì	☒ Sì, in parte	☐ No

Per rispondere **Sì, in parte**, deve essere riportata la redazione di un protocollo che include tutti i seguenti elementi:

- ☒ quesito(i) di ricerca
- ☒ strategia di ricerca
- ☒ criteri di inclusione/esclusione degli studi
- ☒ valutazione del rischio di bias

Per rispondere **Sì**, inoltre il protocollo dovrebbe essere registrato e devono essere specificate:

- ☐ una meta-analisi/piano di analisi statistiche per la sintesi, se appropriata, *e*
- ☐ una strategia per esplorare le cause di eterogeneità
- ☐ una motivazione per ogni violazione dal protocollo

3. Gli autori motivano la scelta del disegno degli studi inclusi nella revisione?

☒ Sì	☐ No

Per rispondere **Sì**, la revisione deve motivare uno dei seguenti criteri:

- ☒ l'inclusione solo di RCT, *oppure*
- ☐ l'inclusione solo di NRSI, *oppure*
- ☐ l'inclusione sia di RCT che NRSI

AMSTAR 2

4. Gli autori hanno effettuato una ricerca sistematica della letteratura?		
☐ **Sì**	☒ **Sì, in parte**	☐ **No**

Per rispondere **Sì, in parte** devono essere presenti tutti i seguenti elementi:

- ☒ la ricerca è stata effettuata in almeno 2 database (rilevanti rispetto al quesito di ricerca)
- ☒ vengono riportate le parole chiave e/o la stringa di ricerca
- ☒ vengono giustificate le restrizioni applicate alla ricerca (es. lingua)

Per rispondere **Sì**, devono inoltre essere presenti tutti i seguenti elementi:

- ☐ sono state analizzate le voci bibliografiche degli studi inclusi
- ☐ è stata effettuata una ricerca nei registri di trial
- ☐ sono stati consultati esperti del campo
- ☐ è stata effettuata, se rilevante, una ricerca nelle fonti di letteratura grigia
- ☐ la ricerca bibliografica è stata effettuata entro 24 mesi dal completamento della revisione

5. La selezione degli studi è stata effettuata da almeno due autori in maniera indipendente?	
☒ **Sì**	☐ **No**

Per rispondere **Sì**, deve essere presente uno tra i seguenti metodi:

- ☒ almeno due revisori indipendenti erano concordi sulla selezione degli studi eleggibili da includere
  oppure
- ☐ due revisori hanno selezionato un campione degli studi eleggibili raggiungendo l'accordo per almeno l'80% degli stessi, mentre i rimanenti sono stati selezionati da un terzo revisore

6. L'estrazione dei dati è stata effettuata da almeno due autori in maniera indipendente?	
☒ **Sì**	☐ **No**

Per rispondere **Sì**, deve essere riportato uno tra i seguenti metodi:

- ☒ almeno due revisori hanno raggiunto l'accordo su quali dati estrarre dagli studi inclusi
  oppure
- ☐ due revisori hanno estratto i dati da un campione degli studi eleggibili, raggiungendo l'accordo per almeno l'80% degli stessi, mentre i rimanenti sono stati selezionati da un terzo revisore

7. Gli autori forniscono l'elenco degli studi esclusi giustificando le motivazioni?		
☐ **Sì**	☐ **Sì, in parte**	☒ **No**

Per rispondere **Sì, in parte**:

- ☐ viene fornito l'elenco di tutti gli studi potenzialmente rilevanti esclusi dalla revisione dopo lettura integrale

Per rispondere **Sì,** inoltre:

- ☐ deve essere giustificata l'esclusione di ogni singolo studio potenzialmente rilevante

Shea BJ, Reeves BC, Wells G et al. AMSTAR 2: strumento per la valutazione critica di revisioni sistematiche di trial randomizzati e/o di studi non randomizzati sull'efficacia degli interventi sanitari. Evidence 2020;12(2): e1000206.
Versione italiana a cura della Fondazione GIMBE.

8. Gli autori descrivono con sufficiente livello di dettaglio gli studi inclusi?

☐ Sì	☒ Sì, in parte	☐ No

Per rispondere **Sì, in parte** devono essere descritti tutti i seguenti elementi:
- ☒ popolazione
- ☒ interventi
- ☒ confronti
- ☒ outcome
- ☒ disegni di studio

Per rispondere **Sì**, devono inoltre essere presenti tutti i seguenti elementi:
- ☐ descrizione dettagliata della popolazione
- ☐ descrizione dettagliata dell'intervento (inclusa la dose se rilevante)
- ☐ descrizione dettagliata del confronto (inclusa la dose se rilevante)
- ☐ descrizione del setting dello studio
- ☐ descrizione delle tempistiche di follow-up

9. Gli autori hanno utilizzato un metodo adeguato per analizzare il rischio di bias dei singoli studi inclusi nella revisione?

RCT

☒ Sì	☐ Sì, in parte	☐ No	☐ Inclusi solo NRSI

Per rispondere **Sì, in parte**, il rischio di bias[1] deve essere stato valutato rispetto a:
- ☒ allocazione non occultata, *e*
- ☒ assenza di *blinding* di partecipanti e valutatori degli outcome[2]

Per rispondere **Sì**, inoltre il rischio di bias deve essere analizzato rispetto a:
- ☒ sequenza di assegnazione non randomizzata, *e*
- ☒ selezione dei risultati riportati da multiple misurazioni o analisi di un outcome specifico

NRSI

☐ Sì	☐ Sì, in parte	☐ No	☐ Inclusi solo RCT

Per rispondere **Sì, in parte**, il rischio di bias deve essere analizzato:
- ☐ rispetto ai fattori confondenti, *e*
- ☐ rispetto al bias di selezione

Per rispondere **Sì**, inoltre il rischio di bias deve essere analizzato rispetto a:
- ☐ metodi utilizzati per l'accertamento di esposizioni ed outcome, *e*
- ☐ selezione dei risultati riportati da multiple misurazioni o analisi di un outcome specifico

10. Gli autori riportano le fonti di finanziamento degli studi inclusi nella revisione?

☐ Sì	☒ No

Per rispondere **Sì**:
- ☐ devono essere riportate le fonti di finanziamento dei singoli studi inclusi nella revisione[3]

[1] Risk of bias (RoB)
[2] Non necessario per gli outcome oggettivi (es. mortalità per tutte le cause)
[3] Il punteggio viene assegnato anche nel caso in cui i revisori abbiano indagato questo aspetto anche se gli autori degli studi non lo hanno esplicitamente riportato.

AMSTAR 2

11. Se è stata condotta una meta-analisi, gli autori hanno utilizzato metodi appropriati per la combinazione statistica dei risultati?

RCT

☒ **Sì**	☐ **No**	☐ **La meta-analisi non è stata condotta**

Per rispondere **Sì** devono essere riportati tutti i seguenti dettagli:

- ☒ motivazione della meta-analisi come strumento appropriato per la sintesi dei dati, *e*
- ☒ utilizzo di tecniche appropriate per la combinazione pesata degli studi gestendo adeguatamente l'eterogeneità, se presente, *e*
- ☒ valutazione delle cause di eterogeneità

NRSI

☐ **Sì**	☐ **No**	☐ **La meta-analisi non è stata condotta**

Per rispondere **Sì** devono essere riportati tutti i seguenti dettagli:

- ☐ motivazione della meta-analisi come strumento appropriato per la sintesi dei dati, *e*
- ☐ utilizzo di tecniche appropriate per la combinazione pesata degli studi gestendo adeguatamente l'eterogeneità, se presente, *e*
- ☐ combinazione statistica delle stime d'effetto e aggiustamento per i fattori confondenti, oppure combinazione di dati grezzi se le stime aggiustate non sono disponibili, *e*
- ☐ sintesi delle stime d'effetto separatamente per RCT e NRSI se inclusi entrambi nella revisione

12. Se è stata condotta una meta-analisi, gli autori analizzano il potenziale impatto del rischio di bias dei singoli studi nei risultati della meta-analisi o nelle altre sintesi delle evidenze?

☒ **Sì**	☐ **No**	☐ **La meta-analisi non è stata condotta**

Per rispondere **Sì**:

- ☐ sono stati inclusi solo RCT a basso rischio di bias, *oppure*
- ☒ se le stime aggregate sono basate su RCT e/o NRSI a rischio di bias variabile, è stato analizzato il possibile impatto del rischio di bias nelle stime cumulative dell'effetto

13. Gli autori tengono in considerazione il rischio di bias nei singoli studi quando interpretano/discutono i risultati della revisione?

☒ **Sì**	☐ **No**

Per rispondere **Sì**:

- ☐ sono stati inclusi solo RCT a basso rischio di bias, *oppure*
- ☒ se sono stati inclusi RCT o NRSI con moderato o alto rischio di bias, la revisione include una discussione del potenziale impatto sui risultati

14. Gli autori spiegano e discutono in maniera soddisfacente ogni eterogeneità osservata nei risultati della revisione?	
☒ **Sì**	☐ **No**

Per rispondere **Sì**:
- ☐ nessuna eterogeneità significativa nei risultati, *oppure*
- ☒ se è presente eterogeneità sono state indagate le cause e discusso il suo impatto sui risultati della revisione

15. Se è stata effettuata una meta-analisi, gli autori hanno esplorato adeguatamente il bias di pubblicazione e discusso il potenziale impatto sui risultati della revisione?		
☒ **Sì**	☐ **No**	☐ **La meta-analisi non è stata condotta**

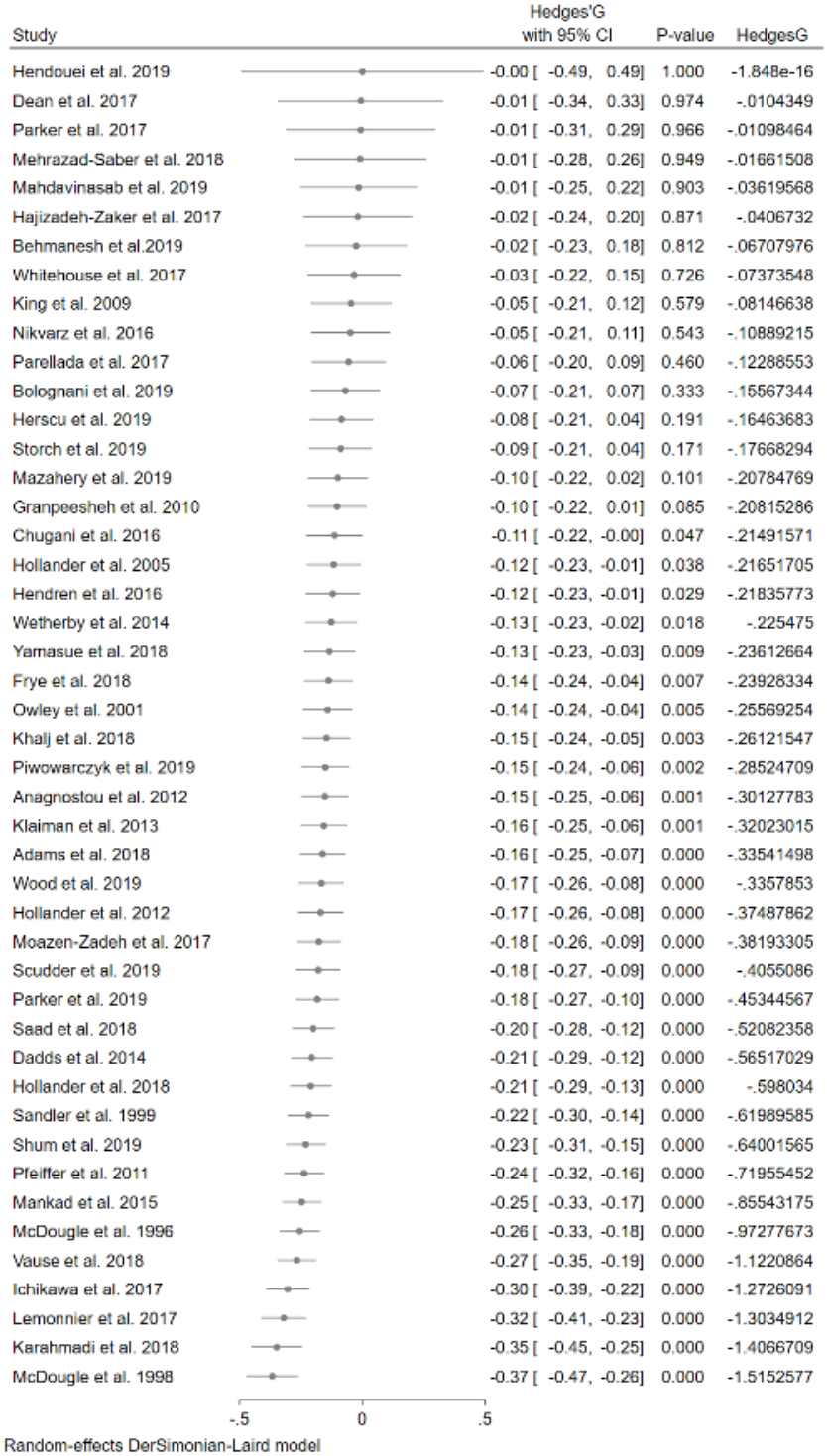
Per rispondere **Sì**:
- ☒ sono stati eseguiti test statistici o utilizzate modalità grafiche per stimare il bias di pubblicazione, discutendone l'eventuale esistenza e il suo impatto

16. Gli autori hanno riportato ogni fonte potenziale di conflitto di interessi, includendo anche eventuali finanziamenti ricevuti per condurre la revisione?	
☒ **Sì**	☐ **No**

Per rispondere **Sì**, gli autori
- ☒ riportano di non avere conflitti di interessi, *oppure*
- ☐ riportano le fonti di finanziamento e le modalità per gestire i potenziali conflitti di interesse

Figure 11a and 11b. Cumulative analysis for outliers, ordered for ascending effect size

a) Complete sample



b) Subgroup cumulative analysis

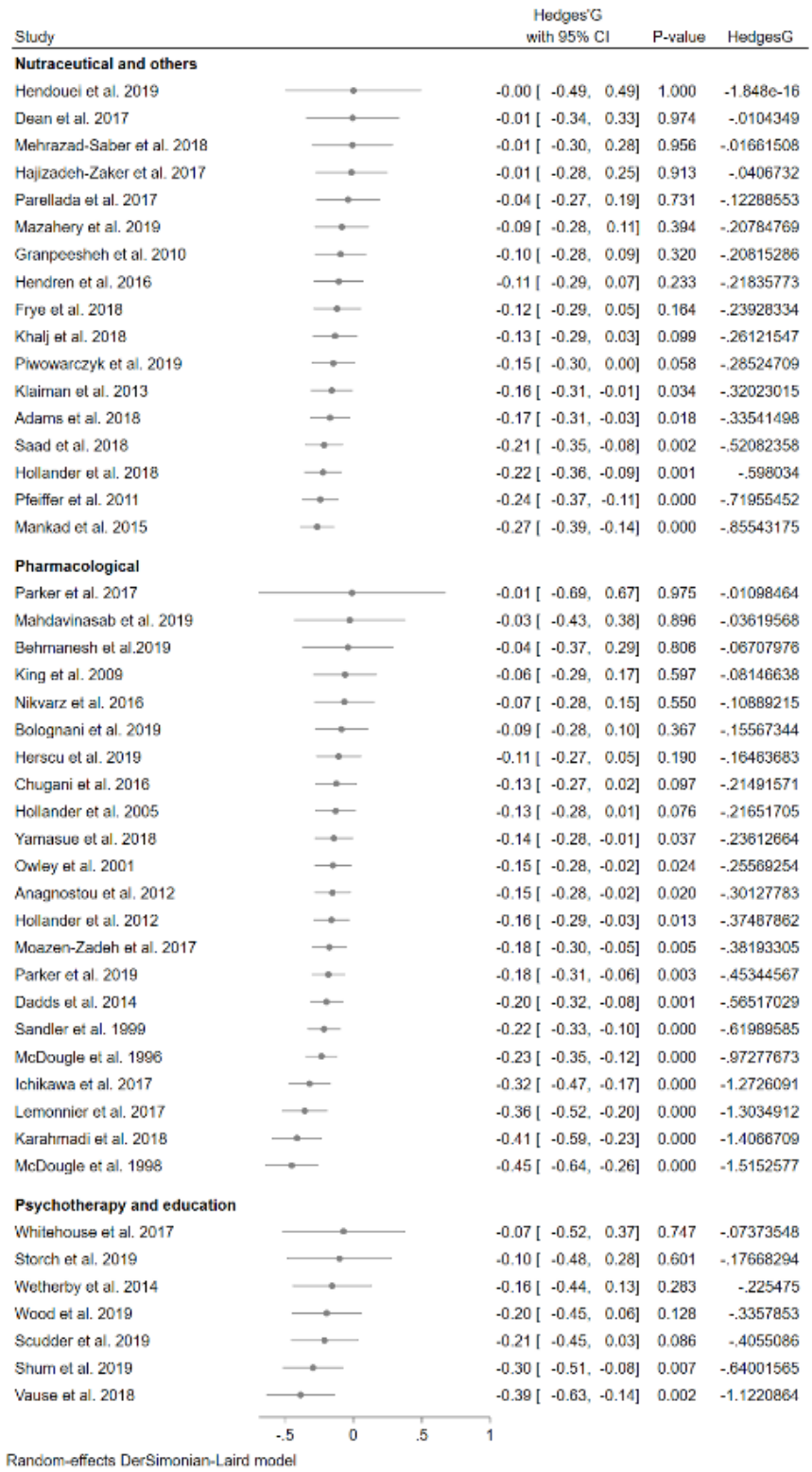


Figure 12a. Forest plot with exclusion of outliers: subgroup analysis and overall effect

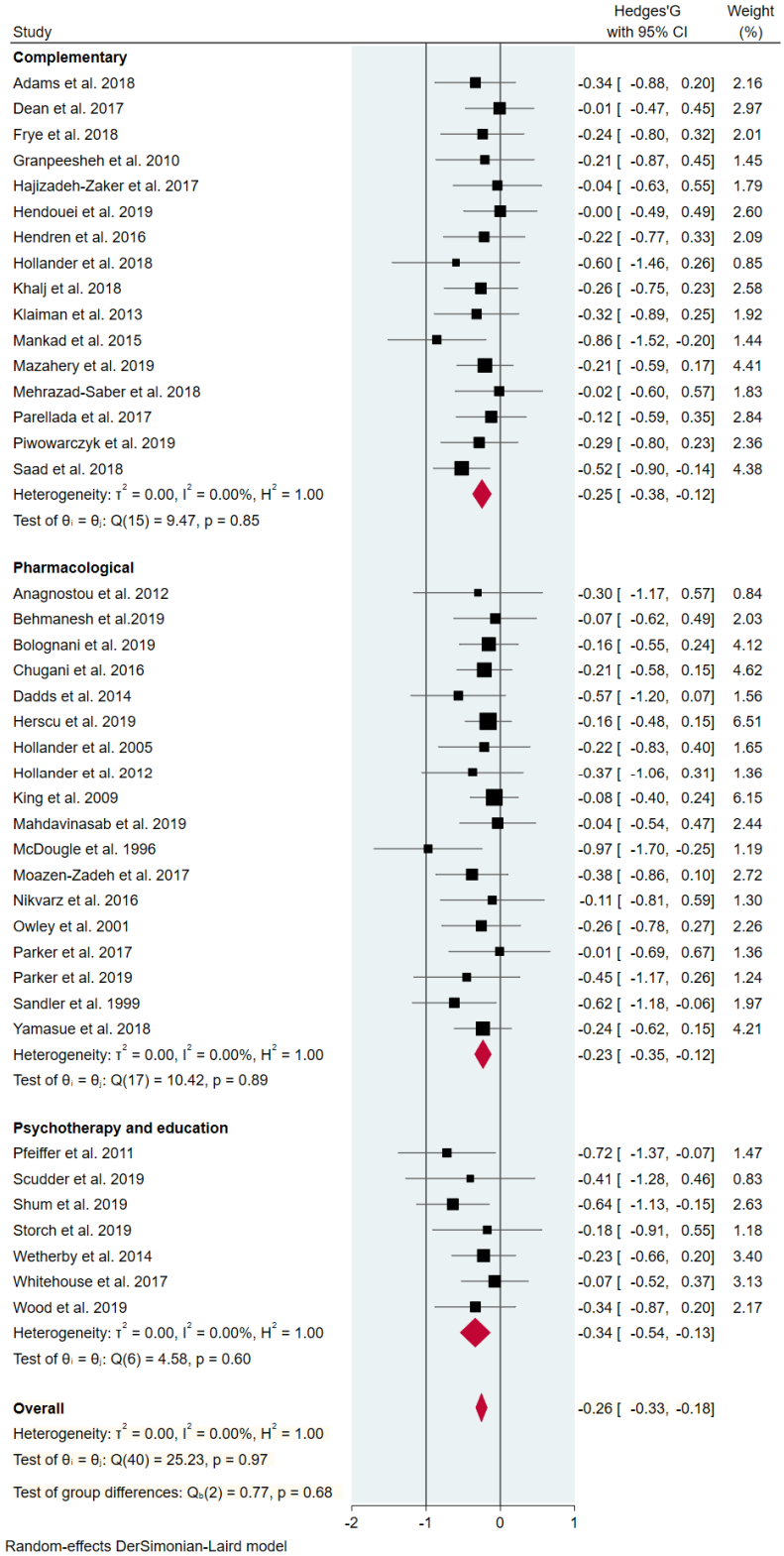


Figure 13a-d. Meta-regression results for the overall sample

a) Moderator: mean age of the sample

Random-effects meta-regression
Method: DerSimonian-Laird

Number of obs = 46
Residual heterogeneity:
tau2 = .05384
I2 (%) = 43.43
H2 = 1.77
R-squared (%) = 1.55
Model F(1,44) = 2.19
Prob > F = 0.1457

_meta_es	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-.010793	.0072862	-1.48	0.146	-.0254774	.0038914
_cons	-.2584991	.0906458	-2.85	0.007	-.4411837	-.0758145

Test of residual homogeneity: Q_res = chi2(44) = 77.78 Prob > Q_res = 0.0013

b) Moderator: female ratio of the sample

Random-effects meta-regression
Method: DerSimonian-Laird

Number of obs = 44
Residual heterogeneity:
tau2 = .05685
I2 (%) = 45.04
H2 = 1.82
R-squared (%) = 1.94
Model F(1,42) = 1.13
Prob > F = 0.2936

_meta_es	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	-.7929744	.7456183	-1.06	0.294	-2.297693	.7117443
_cons	-.2261473	.1418312	-1.59	0.118	-.5123742	.0600795

Test of residual homogeneity: Q_res = chi2(42) = 76.42 Prob > Q_res = 0.0009

c) Moderator: number of subjects in intervention arm

Random-effects meta-regression
Method: DerSimonian-Laird

Number of obs	46
Residual heterogeneity:	
tau2	.05395
I2 (%)	43.13
H2	1.76
R-squared (%)	1.34
Model F(1,44)	1.40
Prob > F	0.2435

_meta_es	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
nintervention	.0035551	.0030073	1.18	0.243	-.0025057	.0096159
_cons	-.4844076	.1128356	-4.29	0.000	-.7118129	-.2570023

Test of residual homogeneity: Q_res = chi2(44) = 77.38 Prob > Q_res = 0.0014

d) Moderator: maximum follow up

Random-effects meta-regression
Method: DerSimonian-Laird

Number of obs	46
Residual heterogeneity:	
tau2	.05602
I2 (%)	44.28
H2	1.79
R-squared (%)	0.00
Model F(1,44)	0.65
Prob > F	0.4248

_meta_es	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
weeks	.0038769	.0048122	0.81	0.425	-.0058214	.0135753
_cons	-.4261429	.0905841	-4.70	0.000	-.6087032	-.2435826

Test of residual homogeneity: Q_res = chi2(44) = 78.96 Prob > Q_res = 0.0009

Figure 14a and 14b. Egger test for funnel plot asymmetry

a) Complete sample

```
Random-effects meta-regression                     Number of obs  =        46
Method: DerSimonian-Laird                          Residual heterogeneity:
                                                              tau2 =  .04188
                                                           I2 (%) =   37.36
                                                               H2 =    1.60
                                                  R-squared (%) =   23.41
                                                   Wald chi2(1)   =      6.25
                                                   Prob > chi2    =    0.0124
```

_meta_es	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_meta_se	-1.856003	.7426588	-2.50	0.012	-3.311587	-.400418
_cons	.1342986	.2054737	0.65	0.513	-.2684225	.5370196

```
Test of residual homogeneity: Q_res = chi2(44) = 70.24   Prob > Q_res = 0.0072

Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.86
      SE of beta1 =     0.743
                z =     -2.50
       Prob > |z| =    0.0124
```

b) Outliers excluded

```
Random-effects meta-regression                     Number of obs  =        41
Method: DerSimonian-Laird                          Residual heterogeneity:
                                                              tau2 =         0
                                                           I2 (%) =      0.00
                                                               H2 =      1.00
                                                  R-squared (%) =      0.00
                                                   Wald chi2(1)   =      3.04
                                                   Prob > chi2    =    0.0812
```

_meta_es	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_meta_se	-1.073115	.6153337	-1.74	0.081	-2.279147	.1329169
_cons	.0139668	.159454	0.09	0.930	-.2985573	.3264908

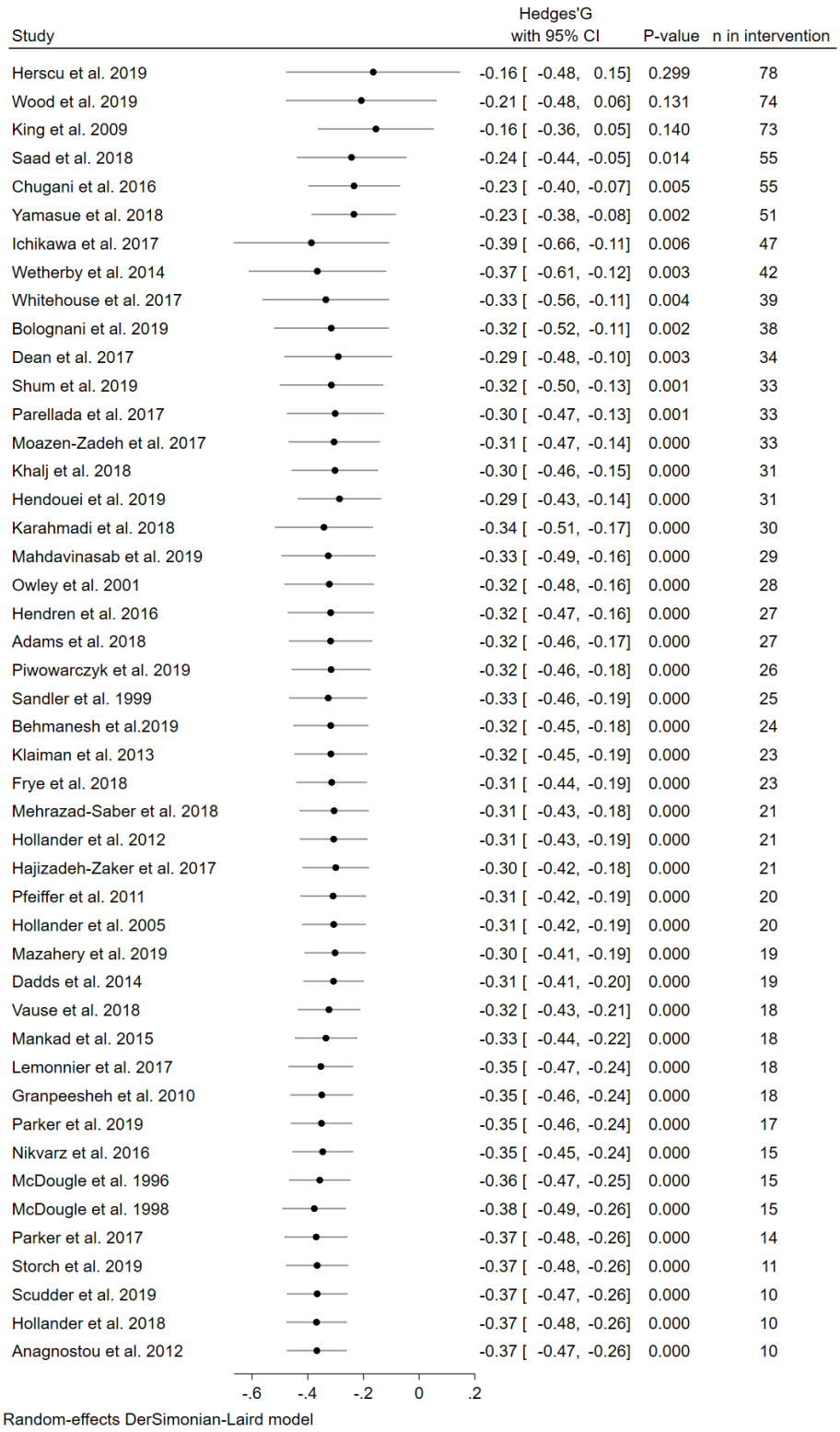
```
Test of residual homogeneity: Q_res = chi2(39) = 22.19   Prob > Q_res = 0.9860

Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.07
      SE of beta1 =     0.615
                z =     -1.74
       Prob > |z| =    0.0812
```

Figure 15a and 15b. Cumulative analysis by ascending sample sizes

a) Complete sample



b) Outliers excluded